THE DEVELOPMENT OF ENGINEERING STUDENTS' ALGEBRA SKILLS: FINDINGS FROM AN ALGEBRA SKILLS TEST REFINED BY RASCH ANALYSIS

Tracy Craig and Anita Campbell

University of Cape Town

Most approaches to teaching calculus rely on proficiency in high school algebra. Given this strong dependence, one would expect correlation between high school algebra skills and performance in first year mathematics. However, three years of data from first year engineering students in an academic support programme found no such correlation. Were students gaining the necessary algebra skills while they studied calculus? Was the algebra test well designed? To learn more, the algebra test was refined using Rasch analysis. Students were given the refined algebra test at the beginning of their second year in 2013. A pre-test and post-test comparison suggested that weakness with algebra was not the main factor preventing students from progressing to second year mathematics.

INTRODUCTION

The Academic Support Programme for Engineering in Cape Town (ASPECT) offers mathematics, physics and communication courses to students who would otherwise not have been offered a place in an engineering degree based on their Grade 12 results and National Benchmark Test (NBT) results. The class averages for first year mathematics and physics are comparable with mainstream results yet only 50% of ASPECT students eventually graduate with an engineering degree. (ASPECT Review, 2009). Concern over these statistics generally and, more specifically, the need to understand the impact of multiple factors including algebraic proficiency on success rates were motivating factors for this study.

An externally developed algebra skills test had been written by the first-year ASPECT cohorts at the beginning of the academic year from 2009 to 2012 and had been found to have little diagnostic value, with respect to ultimate success in the course. Yet it is generally acknowledged (Trillan 1980; Kaput 1997) that algebra skills are important throughout mathematics, particularly in topics such as calculus (Wilson & MacGillivray 2007) which forms the majority of the first-year engineering syllabus. Research has shown that new university students do find algebra difficult (Kajander & Lovric 2005; Jennings 2009). Wilson and MacGillivray (2007) carried out a Rasch analysis on a tertiary-entry questionnaire assessing a variety of pre-tertiary (or indeed pre-secondary) mathematical skills. Their observations include that the symbolic algebra questions were among the hardest questions, with multiple steps, surds and inequalities adding to the difficulty.

Craig, T. & Campbell, A. (2013). The development of engineering students' algebra skills: findings from an algebra skills test refined by Rasch analysis. In Z. Davis & S. Jaffer (Eds.), *Proceedings of the 19th Annual Congress of the Association for Mathematics Education of South Africa*, Vol. 1. (pp. 59 – 69). Cape Town: AMESA.

We were curious as to why the algebra test had almost no diagnostic power. It seems natural that students with weak algebra skills would struggle with the symbolic requirements of calculus and hence would achieve a poor result for the course. Of course, the situation is never that simple. Success in first-year mathematics requires a lot more of the student than simply algebraic fluency; simply doing the calculus course work may provide sufficient practice in algebra to improve weak algebra skills, as determined by the algebra test. Alternatively, perhaps the algebra test was poorly designed. This article examines some of those concerns through three research questions:

1. Was the instrument well designed to measure algebra skills?
2. What is the relationship between students' algebra proficiency and their success in first year mathematics?
3. Was there improvement in algebra skills over the first year of study?

RASCH MEASUREMENT

The Rasch measurement model is centred on the requirement that measurement in fields such as education should retain the specificity and rigour of measurements in fields such as physics. A measure of temperature, for instance, can be determined in meaningful units as determined by a tool such as a thermometer. An assessment instrument in education should similarly produce measures which are not just ordinal (Student 1's result is larger than Student 2's result) but which can be summed and compared in ways traditional assessment results usually cannot be. Rasch measurement theory is discussed in detail in a number of publications (Andrich 1988; Rasch 1960/80; Wilson, 2005; Wright & Stone 1979, 1999; Dunne, Long, Craig & Venter 2012). It is beyond the scope of this article to delve into the depth of mathematical and statistical underpinnings of the model. The reader should look elsewhere for the theoretical and statistical foundations of Rasch measurement theory; in this article we simply discuss its use as a tool for determining the validity of a test.

We used the RUMM2030 software (Andrich, Sheridan & Luo 2011) to conduct detailed analyses using the Rasch model on the algebra skills test. Results which are possible to glean from such an analysis include, but are not limited to, detailed quantitative statistics, a ranking of items in order of difficulty together with what accuracy level that has been determined, a similar ranking of students in order of proficiency, a relative positioning of the difficulty of the instrument in the context of the population, an indication of items whose solutions have been guessed, an indication of items which are too closely aligned and are therefore redundant, a measure of unidimensionality or multidimensionality (how many constructs the test is measuring), graphical depiction of how the proficiency quintiles chose multiple choice distractors and summary charts such as the person-item common axis.

For the purposes of this article we rely on the Rasch measurement model to

- determine whether our instrument is unidimensional (i.e. whether the items comprising the instrument are "pulling in the same direction")
- determine whether any items are anomalous items and why
- develop a ranking of items in order of difficulty in the context of the population being tested
- measure the algebra proficiency of students within the tested population.

THE ASSESSMENT INSTRUMENT

The algebra skills test used from 2009-2012 in ASPECT was designed by colleagues in the Department of Statistical Sciences. The items all aimed to assess algebraic skill in the contexts of words and symbols, symbols alone, and graphs. In early 2013 we applied the Rasch measurement model through the use of RUMM2030 software to analyse the instrument for unidimensionality and for suitability with the population within which we were using it.

At its simplest, unidimensionality refers to the instrument measuring one particular attribute, or construct of interest, where all items add something to that measure. In this case the attribute was understood to be algebraic proficiency. We ran the Rasch analysis on the entire database (four years' worth of data) and were immediately aware of the unsuitability of Item 24, discussed further below. We removed the item and reran the analysis, resulting in the statistics shown in Tables 1 and 2.

Table 1: Fit statistics

Measure	Statistic	Standard deviation
Item mean	0	1.326
Person mean	0.795	0.895
Item fit residual	−0.125	0.958
Person fit residual	−0.181	0.92
Chi Square	136.8, df = 112, p = 0.056	
Cronbach's Alpha	0.682	

As shown in Table 1, the item mean is calibrated at zero while the person mean is estimated to be at 0.795.

Table 2: Unidimensionality t-tests

Number of significant tests	Out of	Percentage of significant tests	Lower bound 95% confidence interval
11	195	5.64%	2.60%

That means the test is slightly too easy for the population, but not by much. A perfect test (impossible to achieve realistically) would have the person mean at zero as well. The standard deviations in the item and fit residuals are comfortably less than 1.4 implying a good fit. The Chi Square p value is 0.056, which is rather low, but still

higher than the required 0.05, implying that there is not a significant difference between the expected outcome according to the model and the observed outcome. Table 2 shows the statistics related to unidimensionality. The lower bound of the 95% confidence interval of the t-tests is less than 5% which informs us of acceptable unidimensionality. In simpler language: the test does satisfactorily measure a single construct of interest, in this case algebraic proficiency. We discuss below how Item 24 was removed from the analysis prior to the generation of the statistics in Tables 1 and 2. Removal of Item 15 (discussed below) improves all of the measures of fit residual, Chi square and unidimensionality, but we kept the problem in as a potential indicator of students with high proficiency.

The analysis drew attention to the presence of one "bad item," Item 24, shown in Figure 1. The problem requires effective use of linear functions, inequalities, grouping of constraints and, possibly, re-ordering of variables within an inequality.

The shaded region below shows the solutions to which of the following systems of inequalities?

(a) $\left\{\begin{matrix} y \geq 1 \\ x - y \geq 1 \end{matrix}\right\}$

(b) $\left\{\begin{matrix} y \leq 1 \\ x - y \leq 1 \end{matrix}\right\}$

(c) $\left\{\begin{matrix} y \leq 1 \\ x - y \geq 1 \end{matrix}\right\}$

(d) $\left\{\begin{matrix} y \geq 1 \\ x - y \leq 1 \end{matrix}\right\}$

(e) I do not remember

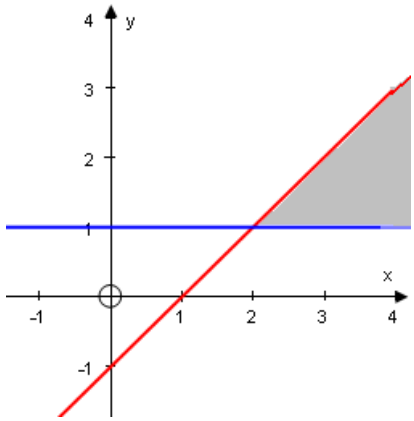

Figure 1: Item 24

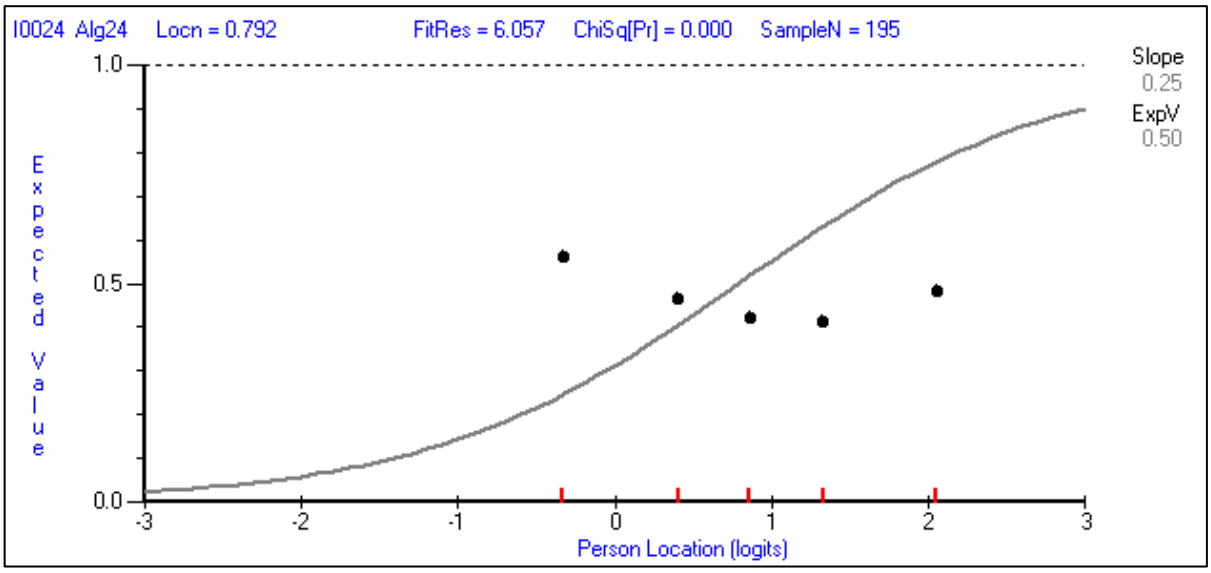

Figure 2: Item Characteristic Curve for Item 24

Figure 2 shows the Item Characteristic Curve for Item 24. As the proficiency quintile increases on the horizontal axis, so the probability of answering correctly (as shown by the dots) should approximately follow the given logistic curve. This pattern is not apparent for Item 24. In fact the lowest quintile attains the highest mean. This item therefore does not fit the expectation of the model, that learners of greater proficiency are more likely to get an item correct than learners of lower proficiency (Rasch 1960/80: 117).

The Multiple Choice Distractor curves in Figure 3 reveal that almost random guessing is occurring between distractors 1 and 4. We have a number of ideas of how the item could be made easier or split into multiple items. There is potential for deeper investigation here into students' facility with inequalities, but for the purposes of current analysis the item was withdrawn from the skills test as it did not add to the measure of algebraic proficiency for the population being tested.

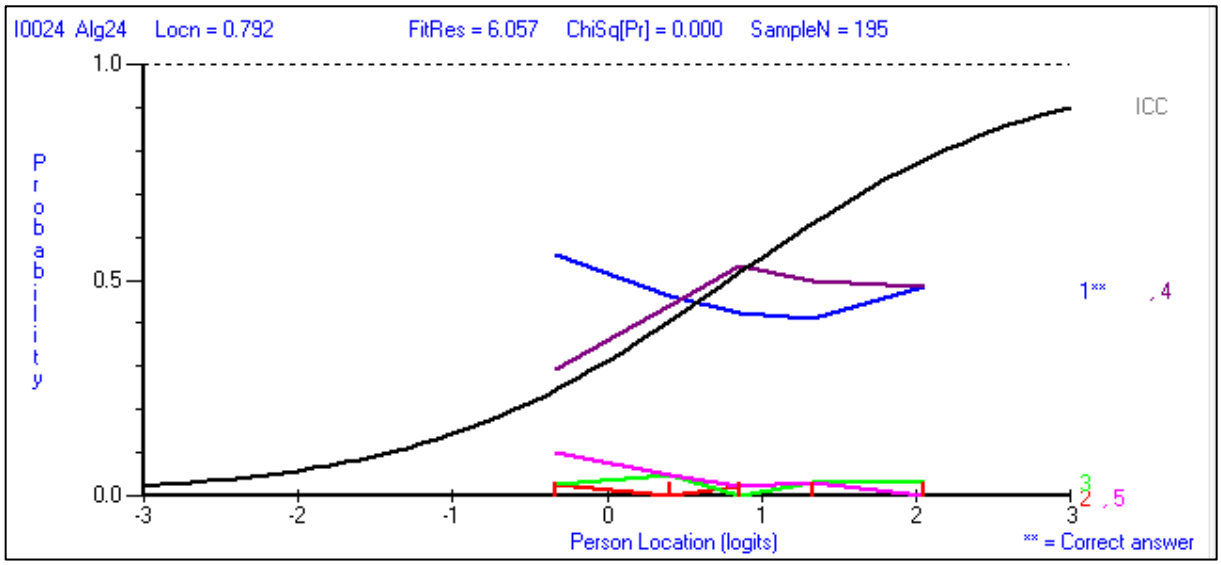

Figure 3 – Multiple Choice Distractor curves for Item 24

Other than that single problem, the test is well suited to the population. Figure 4 shows the item map generated by the software. A single common axis serves to represent the algebra proficiency of the students and the difficulty level of the items. The higher an item appears on the axis, the more difficult it is and the higher a student (represented in pairs by the × symbol) appears, the greater his or her proficiency. A student at the same level as an item has a 50% chance of successfully answering the question, a student higher than the item has a greater than 50% chance of answering the item successfully and a student lower on the axis than an item has a less than 50% probability of answering the item correctly.

In Figure 4 it is clear that the bulk of the students is shifted slightly higher up the axis than the bulk of the questions. This visual effect conveys the same information as the person mean in Table 1 being higher than the item mean. To decrease this shift of means, we plan in future designs of this test to include a few problems of slightly higher than current average difficulty.

```
RUMM2030      Project: TRY3    Analysis: TRY3N
Title:  ORIGINAL DATA WITH ONLY ALG24 REMOVED
Date: 5 Feb 2013 12:42:55 PM
Display: ITEM MAP
-----------------------------------------------------------------
LOCATION          PERSONS     ITEMS [uncentralised thresholds]
-----------------------------------------------------------------
  4.0                       |
                            |
                            |
                            |
                            |
  3.0                     x | I0015
                            |
                         xx |
                            | I0014
                      xxxxx |
  2.0                       | I0013
                      xxxxx |
                    xxxxxxx | I0016
                            | I0008
                  xxxxxxxxx | I0011
  1.0            xxxxxxxxxx | I0009
                xxxxxxxxxxx |
                  xxxxxxxxx | I0020
                   xxxxxxxx |
                     xxxxxx | I0026
  0.0              xxxxxxxx | I0019  I0004  I0027
                      xxxxx | I0017
                      xxxxx | I0025  I0028  I0002
                       xxxx | I0010  I0003
                         xx | I0018  I0021
 -1.0                    xx | I0007  I0022
                            |
                            | I0005
                          x | I0006  I0001  I0023
                            | I0029
 -2.0                       |
                            | I0012
                            |
                          x |
                            |
 -3.0                       |
-----------------------------------------------------------------
x = 2 Persons
-----------------------------------------------------------------
```

Figure 4: Person-item map

The decisions we made as a result of this analysis were to

- remove item 24 from the test (removed before Figure 4 was generated);

- retain a slightly problematic item, item 15. This item was answered poorly by the majority of the students, however, unlike item 24, it did add to the measure of algebraic proficiency and hence was retained;
- reorder the items in increasing order of difficulty.

The resulting algebra test is 28 items long and is available from the authors on request. Three sample questions are included in Appendix A.

APPLYING THE ALGEBRA TEST

First year ASPECT students were given the algebra test without prior warning in the first week of lectures from 2009 to 2012. In 2013, students in the new second year ASPECT vector calculus course were given the refined algebra test in the first week. Of the 85 students in the second year course, 30 had not taken the algebra test in their first year, mainly because those students began or completed their first mathematics course outside of the ASPECT programme. Students shaded their answers on multiple choice answer sheets which were scanned by a data capture service. Question papers were collected after each test to minimise the skewing of results from future students who might gain access to the test before writing it.

It should be pointed out that algebra skills are not taught as an early section of the first year course but rather embedded within other topics. Students are given summaries of algebraic formulae (such as logarithm laws) and required to achieve an average of 80% for 'skills tests', retaking similar tests if necessary. The skills tests topics include factorisation, trigonometry, logarithms and basic graphs.

RESULTS FROM PRE- AND POST-TESTS

Comparisons between the scores for the algebra test written by second year students in 2013 and the same test written in 2011 or 2012 showed some interesting trends. Of the 85 students who wrote the post-test in 2013, 39 improved on their previous score while 16 students obtained a lower score. The remaining 30 students who wrote the post-test did not write the pre-test. The difference in scores for improving students ranged from 2 to 27 percentage points with an average improvement of 12 percentage points. The students whose scores decreased had an average decline of 7 percentage points, ranging from a drop of less than 1 percentage point to a drop of 15 percentage points. Figure 5 shows the averages and ranges of pre- and post-test results for students who wrote both tests. The lowest, highest and average marks for the improving, declining, 'new' and non-continuing students are given in Table 3.

Categorising students' academic proficiency as improving or declining in such a binary manner as we have done here contrasts with the detailed continuum results of Rasch analysis. Certainly there is scope for further analysis and investigation, for instance, to take note of the key items for which there might be global improvement. For the purposes of this article, we consider it notable that significant numbers of students appear to have divergent progress in algebraic skill development.

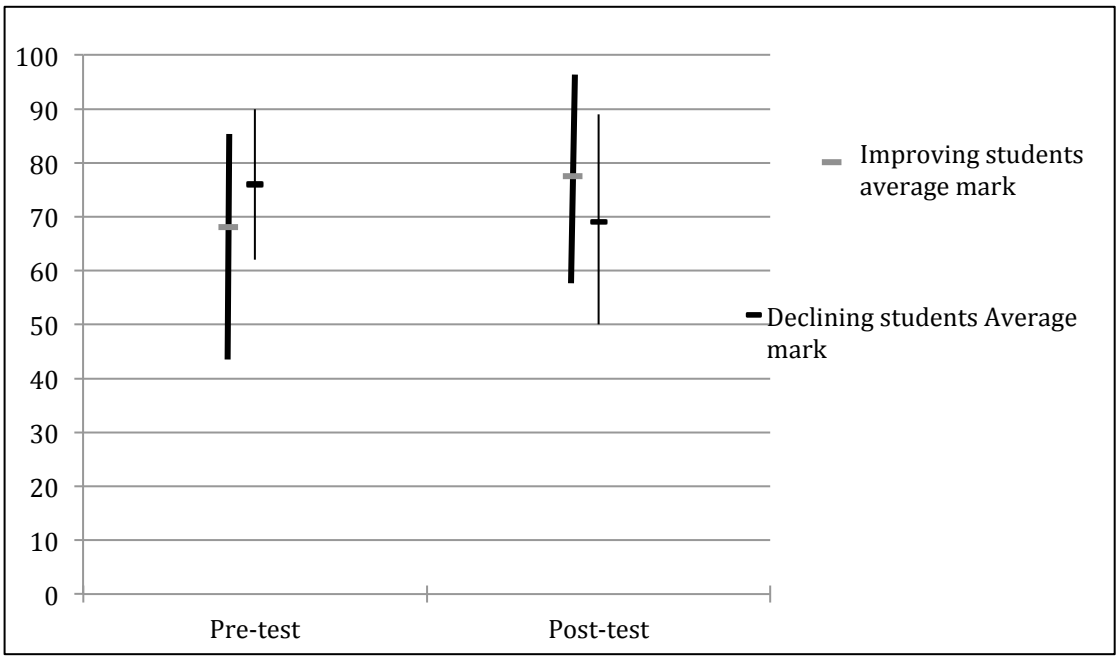

Figure 5: Pre- and post-test averages and ranges for improving/declining students

The overall class average in the post-test (74%) was higher than the pre-test average of 68%. None of the students scored less than 50% in the post-test, whereas 2 students scored less than 50% in the pre-test. Interestingly, these 2 students both increased their scores by 27 percentage points to post-test scores of 68% and 75%.

Table 3: Comparison of pre-test and post-test marks

	Pre-test results			Post-test results		
	Average (%)	Lowest mark (%)	Highest mark (%)	Average (%)	Lowest mark (%)	Highest mark (%)
Improving students (N = 39)	**65**	41	83	**77**	57	96
Declining students (N = 16)	**76**	62	90	**69**	50	89
New students (N = 30)	**-**	**-**	**-**	**73**	50	86
Non-continuing students (N = 16)	**69**	52	86	**-**	**-**	**-**
Overall	**68**	41	90	**74**	50	96

DISCUSSION AND CONCLUSIONS

For the 55 students that wrote the pre-test and post-test, only 12 students (22%) had scores below 60% in the pre-test and all of these improved in the post-test. For these students, starting with weaker algebra skills (as determined by the skills test) did not stop their progress into second year mathematics and their algebra skills did improve over the first year course. Students who did not progress into second year mathematics had pre-test scores ranging from 52% to 86%, with 4 students out of 16 (25%) scoring below 60%. As these results are comparable with the results of the students who progressed, we can infer that weakness with algebra was not a major factor preventing them from progressing to second year mathematics.

The 16 students who showed a decline in algebra test scores are a cause for concern, particularly the 3 students whose marks fell to below 60%. It would be interesting to track the performance of these students in their second year course to see if their basic algebra skills cause them significant difficulties in tests and exams.

The algebra tests were given to students at the start of the year, as early as possible. Should we, as suggested by Jennings (2010), perhaps give the algebra test after the students have had 2-3 weeks of study to refresh themselves after the 3 month vacation? The benefit of giving the algebra test unannounced is that it potentially gives a better indication of students' long term recall of algebra concepts, which might be an indication of concepts which have been well understood and are in the long term memory of the students. On the other hand, algebra fluency improves with practice and perhaps after 2-3 weeks of increasing fluency with both new and familiar algebra-rich mathematics, the algebra test performance of some students would be substantially improved compared to taking the test 'cold' without preparation time to revise algebra topics.

A limitation of the test being slightly too easy for the population, as measured by the person-mean estimate of 0.795, is that optimum information may not be gleaned in applying this test. As well as considering running the test slightly later in the term, further investigations include redesigning the test by adding slightly harder problems, as discussed earlier. In addition, in 2013 we aim to test the first-year students' algebra with the current instrument before and after their first semester as well as to assess their performance in algebra-heavy tests during the first semester to see if more interventions are needed for students whose algebra skills are not robust.

In ASPECT we seek to support students academically as well as helping with challenges which are environmental and affective. We are aware that a suite of parameters influences any one student's chances of success in a course and through a degree. We must emphasise that we are not intending to falsely simplify a complex situation. As lecturers of mathematics however, we are most interested in specifically mathematical characteristics which support or hinder success in mathematics courses. First-year engineering mathematics is heavily algebraic, which prompted our scrutiny of the role algebraic proficiency plays in students' success. Our conclusions from the

analysis discussed in this article lead us to draw the conclusions that the algebra assessment instrument is measuring the attribute for which it was intended, that students' algebra proficiency does not play a major role in course success, and that algebra skills, in general, do improve over the first year of the degree.

REFERENCES

Andrich, D. (1988). *Rasch models for measurement*. Beverly Hills, CA: Sage Publications.

Andrich, D., Sheridan, B. & Luo, G. (2011). *RUMM 2030 software and manuals*. Perth: University of Western Australia

ASPECT Review (2009). Cape Town: University of Cape Town.

Dunne, T., Long, C., Craig, T. S. & Venter, E. J. (2012). Meeting the requirements of both classroom-based and systemic assessment: The potential of Rasch measurement theory. *Pythagoras*, 33(3), Art.#19, 16 pages. http://dx.doi.org/10.4102/pythagorasv33i3.19.

Jennings, M. (2009). Issues in bridging between senior secondary and first year university mathematics. In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd Annual Conference of the Mathematics Education Research Group of Australasia*. MERGA32 (pp. 273-280).

Jennings, M. (2010). A changing cohort—what do high school students remember and what does it mean for university lecturers? *CULMS Newsletter*, 2, 3-10. http://www.math.auckland.ac.nz/CULMS/wp-content/uploads/2010/08/CULMS-No2-Complete-v2.pdf.

Kajander, A. & Lovric, M. (2005). Transition from secondary to tertiary mathematics: McMaster University experience. *International Journal of Mathematical Education in Science and Technology*, 36(2-3), 149-160.

Kaput, J. J. (1997). *Teaching and learning a new algebra with understanding*. Unpublished manuscript. http://www.educ.fc.ul.pt/docentes/jponte/DA/da-textos/kaput_99algund.pdf.

Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests* (Expanded edition with foreword and afterword by B.D. Wright). Chicago, WC: Chicago University of Chicago Press.

Trillan, A. S. (Ed.). (1980). *Teaching basic skills in college: A guide to objectives, skills assessment, course content, teaching methods, support services, and administration*. San Francisco: Jossey-Bass Publishers.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. London: Lawrence Erlbaum.

Wilson, T. M. & MacGillivray, H. L. (2007). Counting on the basics: mathematical skills among tertiary entrants. *International Journal of Mathematics Education in Science and Technology,* 38(1), 19-41.

Wright, B. D. & Stone, M. H. (1979). The measurement model. In B.D. Wright, & M.H. Stone (Eds.), *Best test design*. Chicago (pp. 1-17). WC: MESA Press.

Wright, B. D. & Stone, M. H. (1999). *Measurement essentials*. Wilmington, DE: Wide Range, Inc.

APPENDIX A: SELECTED ITEMS FROM THE ALGEBRA SKILLS TEST

Item 2

$$\frac{2-x}{2} =$$

(a) $1-x$ (b) $-x$ (c) $2-\frac{x}{2}$ (d) $1-\frac{x}{2}$ (e) None of the above

Item 14

$$\frac{\sqrt[3]{p}}{\sqrt[5]{p}} =$$

(a) $p^{3/5}$ (b) $p^{5/3}$ (c) $p^{2/15}$ (d) $p^{8/15}$ (e) p^{-1}

Item 28

The number of solutions to the equation $\sqrt[3]{x} = x$ is

(a) 3 (b) 2 (c) 1 (d) 0 (e) I do not know how to do this