AFA06011

Future Teachers' Developing Numeracy and Mathematical Competence As Assessed By Two Diagnostic Tests.

Karoline Afamasaga-Fuata'i
University of New England

Paul Meyer
National University of Samoa

Naomi Falo
National University of Samoa

Perenise Sufia
National University of Samoa

This paper reports the findings of the first year of a longitudinal mathematics project established to monitor the development of student teachers' numeracy and mathematical competence. The sample was a group of pre-service student teachers enrolled in the foundation and primary education programs. Data presented here is from the Second Mathematics Diagnostic Test (MDT2) administered in October 2005. Student responses from the two tests were analysed using the dichotomous Rasch Measurement Model. Results from the first diagnostic testing (MDT1) showed that no one achieved mastery level defined at 80% success rate. A comparison of student performance on both MDT1 and MDT2 was conducted to (a) identify any developmental trends for the group between testing and (b) determine the extent of the impact of normal mathematics content and mathematics education courses on students' numeracy and competence levels. Findings show that there was a significant difference between the performances of the students in the two tests. The paper presents and discusses performance of sub-groups and selected individuals. Findings will inform the development of customized enrichment programs to re-dress students' numeracy needs before exit. Implications for primary mathematics and mathematics education courses are provided.

Keywords: numeracy, mathematical competence, primary mathematics, misconceptions, diagnostic testing, mathematics pedagogical content knowledge

The preparation of competent and quality teachers is vital to the success of educational reforms in schools in almost all countries. According to Shulman's (1986) seminal conceptualisations of the diversity of knowledge required for teaching, three out of seven categories focus on content knowledge – subject matter knowledge (e.g. mathematics content knowledge), pedagogical content knowledge (PCK) and curriculum knowledge (CK). For Shulman, PCK consists of

> "the ways of representing the subject which makes it comprehensible to others … [it] also includes an understanding of what makes the learning of specific topics easy or difficult; the conceptions and preconceptions that students of different ages and backgrounds bring with them … If those preconceptions are misconceptions, which they often are, teachers need knowledge of the strategies most likely to be fruitful in reorganising the understanding of learners (Shulman, 1986, p.9).

Shulman further defines subject matter knowledge (SMK) as more than knowledge of facts or concepts, it requires knowledge of both the substantive structure (organising principles and explanatory frameworks) and syntactic structure (nature of enquiry in the discipline, and how new knowledge is introduced and accepted in that community). The transformation of SMK and CK into PCK is an important focus of teacher education. PCK is what "conceptualises the link between knowing something for oneself and being able to enable others to know it" (Huckstep, Rowland &

Thwaites, 2003). It is therefore argued that the quality preparation of pre-service teachers of mathematics include the development of student teachers' competence in the content knowledge of the curriculum to be taught in order to effectively design learning activities for their students.

Background of the Numeracy and Mathematics Competence Problem

In Samoa, much has been said and written about the mathematical competence of school leavers entering post-secondary institutions through analyses of national results from primary literacy tests, school certificate and senior school certificate examinations which record students' performances at various stages of schooling up to the end of secondary level (Afamasaga-Fuata'i, 2004, 2003, 2002a, 2002b, 2001, 1998a, 1997). On the other hand, post-secondary institutions routinely conduct testing programs to monitor students' achievement levels while at the departmental level, Samoa's Ministry of Education, Sports and Culture (MESC) continually regulate curricular reforms consequential to periodic reviews of the primary and secondary curricula (MESC, 2003a, 2003b; Afamasaga-Fuata'i, 1998b; Department of Education (DOE), 1995). From all of these various sources, one of the recurring themes, amongst many others, concerns the requisite numeracy capabilities and mathematical competence of students entering different tertiary programs such as science, commerce, vocational or teacher education – most of them are under-prepared to effectively cope with mathematics. Consequently, remediation and/or bridging courses are established to upgrade students' mathematical competence (Afamasaga-Fuata'i, 1998a, 1998b, 1997); thus the concern over student numeracy and mathematics performance at all levels unavoidably becomes a perennial issue, but it is particularly critical in relation to the preparation of future teachers. Through numerous publications, MESC mandates that teachers have deep understanding of their subject matter, curriculum, and pedagogy, as well as knowledge of their students and how they learn (MESC, 2003a) as the case is worldwide with other international teaching standards (NCTM, 2005; Committee for the Review of Teaching and Teacher Education (2003a, 2003b) cited in Mays, 2005; Center for Education, 2001; Darling-Hammond, 2000; Thomas, 2002; Jones, Mooney, & Harries, 2002; Rowland, Barber, Heal, & Martyn, 2003).

Restructured Primary Teacher Education Programs

At the institutional level in Samoa, the university, which is the only in-country provider of teacher education in 2003, reviewed its foundation and teacher education programs mainly to enhance the quality of graduating teachers. In the subsequent restructuring of primary pre-service programs (NUS Calendar, 2004, 2005), two previously compulsory content courses (bridging mathematics/statistics) which aimed at further developing students' mathematics/statistics content knowledge to higher levels than they did at school became reduced to electives in the Foundation Education Program (FEP) (i.e. entry level to the teacher education programs), disappeared from the Diploma in Education (Primary)(DEdP) while in the revised Bachelor of Education (Primary) (BEdP), two higher level mathematics/statistics content courses (with foundation mathematics/statistics as pre-requisites) became electives. Upon successful completion of the FEP, primary student teachers progress directly into either the DEdP or BEdP. However, whilst both the DEdP and BEdP each have two compulsory mathematic methods courses which develop student teachers' pedagogical skills in teaching others to learn and understand primary mathematics, the elective status of the content courses combined with students' preferred choices can potentially result in some primary student teachers exiting the program without necessarily upgrading their content knowledge of mathematics (e.g. as DEdP graduates). This potential scenario was raised by Burgess & Bicknell (2003) as a major concern amidst similar program restructuring in a New Zealand institution. Numerous studies have also demonstrated the necessity of developing both the content knowledge alongside and in parallel to pedagogical content knowledge (Shulman, 1986; Ma, 1999; Van Dooren, Verschaffel, & Onghena, 2002). Subsequently, in recognition of the university's vital role in the preparation of quality teachers, MESC's mandate for teachers to have expert knowledge of the content of the primary mathematics curriculum, and informed by the research literature on preservice teachers' content knowledge for effective teaching and learning, a Longitudinal Mathematics Project (Project) was established in 2005 to monitor and provide baseline and longitudinal data on the numeracy capabilities and mathematics competence (NAMC) with respect to the prescribed primary and early secondary mathematics curricula, of primary student teachers that enrol in the 1-year FEP and Primary Education Programs namely the 2-year DEdP and 3-year BEdP. Doing so would provide empirical data to base the assessment and certification of student teachers' competence with the content knowledge of the primary and early secondary mathematics curricula, independently of their formal preservice programs. Consequently, the main objective of the Project was to administer mathematics diagnostic tests to: (a) identify at-risk students upon entry with the aim of providing remediation early in the programs to rectify misconceptions, (b) provide empirical data of student errors and misconceptions to support remedial instruction and, (c) inform curricular and pedagogical development within relevant mathematics content and methods courses to ensure beginning teachers' numeracy, mathematical and pedagogical needs are satisfactorily met upon exit from the university's pre-service program. As Ma (1999) in her comparative study of Chinese and American elementary teachers demonstrated, primary teachers need to have *a profound understanding of fundamental mathematics;* without it, she found teachers could not promote students' learning that went beyond the teachers' own level of understanding. This paper therefore focuses on the numeracy capabilities and mathematical competence of the 2005 cohort of foundation and primary education students in the pre-service teacher education program. The major reason for this focus is because teachers' content knowledge of the curriculum and subject matter generally influence their selection of learning activities and pedagogical mediation of mathematical meaning in the classroom (Shulman, 1986; Thompson, 1992; Aubrey, 1997; Rowland, Barber, Heal, & Martyn, 2003; Ball, Lubienski, & Mewborn, 2001).

The next sections briefly discuss relevant definitions and methodology before presenting data followed by a discussion of findings, conclusions and ending with some implications for pre-service teacher education and educational policy.

Numeracy and Mathematical Competence

The numeracy capabilities and mathematical competence of student teachers is conceptualised as their ability to successfully solve items that are based on the content of MESC's prescribed Primary and Early Secondary Mathematics (PESM) curricula. '*Numeracy*' or 'quantitative literacy' is defined in this paper based on the definitions by Cockcroft (1986) and Steen (2001), For example, Cockcroft explains that to be numerate means

> *to imply the possession of two attributes. The first of these is an 'at-homeness' with numbers and an ability to make use of mathematical skills which enable an individual to cope with the practical mathematical demands of his everyday life. The second is ability to have some appreciation and understanding of information which is presented in mathematical terms, for instance in graphs, charts or tables or by reference to percentage increase or decrease*" (Cockcroft, 1986).

Steen, on the other hand, points out that

> "*there appears to be reasonable consensus among individuals of widely differing perspectives on the natural growth of numeracy from the basic arithmetic of grade school through the more sophisticated numerical reasoning of measurement, ratios, percentages, graphs, and data needs of modern life*" (Steen, 2001).

In contrast to the more practical, concrete, and contextual view of numeracy, '*mathematics competence*' as used in this paper encompasses the ability of student teachers to effectively apply the abstract knowledge and skills prescribed by MESC's PESM curricula, to solving items in a mathematics diagnostic test. "Being numerate does not necessarily mean learning more mathematics topics; numeracy is more to do with applications and interpretations of mathematics. … Whereas mathematics asks students to rise above the context, quantitative literacy is anchored in the messy contexts of real life" (Steen, 2001). In comparison, prescribed curriculum statements are useful to plan teaching sequences but "they are only 'frameworks' – teachers need in-depth knowledge of mathematical concepts and processes so as to enrich them" (Bobis, Mulligan & Lowrie, 2004). Therefore, it is argued that pre-service teachers should be competent and should have deep understanding of the content of the relevant mathematics curriculum in order for them to effectively develop their pedagogical mathematics content knowledge. Clearly, given the 2004 restructuring in effect at the university, there is vested interest and much educational value to be reaped by future students, in ensuring that primary student teachers' mathematical and numeracy needs are diagnosed early with follow-up remediation of misconceptions with a view of certifying student teachers' numeracy and mathematical competence before exit. In order to assess student teachers' competence and ability levels with respect to the subject matter knowledge of MESC's PESM curricula, a Mathematics Diagnostic Test that was field-tested and used in a similar project at an Australian regional university with different cohorts of primary student teachers was selected (May, 2005). The next section provides more details of the diagnostic test.

Mathematics Diagnostic Tests

Two versions of a written Mathematics Diagnostic Test (MDT1 and MDT2) developed by a group of mathematics educators to certify the numeracy needs and levels of different cohorts of primary students teachers (May, 2005), were used as one means of assessing and diagnosing students' numeracy capabilities and mathematical competence and understanding of the prescribed PESM curricula. The MDT tests consisted of five mental computation (MCT) and thirty items selected primarily from the TIMSS-R 1999 mathematics paper (Mullis, Martin, Gonzalez, Gregory, Garden, O'Connor, Chrostowski & Smith, 2000) and 3 items from the misconception literature. The TIMSS-R 1999 items were selected as these have been field tested and used already, with available data on their reliability and validity and success rates with 15 year olds from 38 different countries. Based on their collective extensive teaching, research and curriculum development experiences in Samoa, the authors undertook a validating exercise to ensure that items in the diagnostic test aligned with and reflected the spirit of the MESC's PESM curricula. Of the 38 items, 15 items were kept common in the two tests to provide a means of anchoring the performance of students between the two testings in accordance with the principles of test equating within item-response theory. Items covered five content areas matching those prescribed by the content of the *MESC PESM Curricula* (MESC, 2003a, 2003b) - fractions and number sense (FNS), measurement (MSR), algebra (ALG), geometry (GEO) and data presentation, analysis (DRA) and probability (PRB) - and five cognitive domains: *knowing, using routine procedures, investigating and problem solving, and mathematical reasoning and communication* (TIMSS, 2003). Unlike TIMSS-R 1999, all diagnostic items were left open-ended to provide the researchers with access to all the resultant errors (Mays, 2005). The diagnostic test does not purport to be a comprehensive assessment instrument of MESC's PESM but instead, it aims to assess a sample of its content with a range of items that students at the end of their first year of secondary schooling are most likely to solve successfully. As Aubrey (1997) states "If teaching involves helping others to learn then understanding the subject content to be taught is a fundamental requirement of teaching." It is therefore argued that a quality and competent teacher of primary mathematics is expected to demonstrate mastery of such a diagnostic test.

Methodology

Data was collected by administering MDT1 early in Semester 1 and MDT2 at the end of Semester 2 with student teachers undertaking their normal load of courses for the two semesters in between testing. A total of 140 students took MDT1 in March 2005 with 141 students completing MDT2 in October 2005. Students' responses were categorized as Correct (1), Incorrect (0) or Blank (B) and analysed using the Dichotomous Rasch Measurement Model (Rasch, 1980) and QUEST software (Adams & Khoo, 1993). Items are referenced (AAA##) to indicate the relevant content area (i.e. FNS, MSR, ALG, GEO, DRA, or PRB) and corresponding number (##) in the test. The MDT2 data was Rasch analysed initially with the 15 common items of the two tests anchored on MDT1 estimates in order to obtain meaningful and valid estimates of improvement by the second test (Bond & Fox, 2001).

Analysis of Data

The Rasch analysis provides statistics to determine the compatibility of the data with the Rasch Model in terms of *goodness-of-fit* statistics, and the *reliability* of the test to measure the latent trait (NAMC) as manifested through student responses to the items. These are further elaborated next using cohort statistics to establish the reliability of the diagnostic test as a measurement tool before discussing in detail the performance of the 46 students that sat both MDT1 and MDT2. Item estimates are presented first followed by person estimates.

Item Estimates

Reliability - With the 2005 MDT1 mean item estimate theoretically set at zero logit in accordance with the Rasch Model, and estimates of the 15 MDT1/MDT2 common items anchored on MDT1 values, the Rasch analysis generated a MDT2 cohort mean item estimate of 0.30 logits (Table 1, column 1a, column I). That is, overall, MDT2 item estimates were relatively higher on average compared to the MDT1 zero mean item estimate. In contrast, MDT2 showed more or less the same standard deviation of 1.73 logits compared to 1.76 in MDT1. Like the MDT1, MDT2 items also showed the same high reliability index of 0.97 confirming that the items separated out into a hierarchy along the logit continuum from low to high indicating the extent of the presence of the latent trait 'numeracy capabilities and mathematical competence' the test is purporting to measure.

Goodness-of-fit - Infit and outfit mean square values of MDT2 item estimates (see Table 1, column 1a, columns I) were around the expected value of 1.00 (0.99 and 1.09 respectively) indicating the overall fit of the MDT2 item data to the Rasch model. Further inspection of the item-fit map however showed that 3 items (PRB10, ALG14 and MSR15) of the 38 items were outside the 95% acceptable bounds. These 3 items were part of the 15 MDT1/MDT2 common items whose MDT1 estimates were used to anchor the MDT2 calibrations.

Table 1
Rasch Analysis Item and Person Statistics

```
                    1a
Item Estimates (Thresholds)
 all on all (N = 141 L = 38
Probability Level= .50)
------------------------------------------
Summary of item Estimates
==========================================
                              I      II
Mean                         0.30   0.34
SD                           1.75   1.80
SD (adjusted)                1.73   1.76
Reliability of estimate       .97   0.96

Fit Statistics
==========================================
Infit Mean Square       Outfit Mean Square
           I     II                I     II
 Mean     0.99  0.99     Mean     1.09  1.06
 SD       0.23  0.11     SD       0.69  0.61

Infit t                 Outfit t
           I     II                I     II
 Mean    -.16   -.08     Mean     -.01   .03
 SD      1.35   1.14     SD       1.18  1.05

   0 items with zero scores
   0 items with perfect scores
==========================================
         I – first round of analysis
```

```
                                   1b
Person-Ability Estimates
all on all (N = 141 L = 38
Probability Level= .50)
------------------------------------------
Summary of case Estimates
==========================================
                              I      II
Mean                         -.49   -.47
SD                           1.04   1.04
SD (adjusted)                0.93   0.93
Reliability of estimate      0.79   0.79

 Fit Statistics
==========================================
Infit Mean Square       Outfit Mean Square
          I     II                I      II
Mean     0.99  0.99     Mean     1.07   1.04
SD       0.20  0.21     SD       0.96   0.94

Infit t                 Outfit t
          I     II                I      II
Mean    -.05   -.04     Mean     .14    .13
SD       .93    .94     SD       .81    .74

   0 cases with zero scores
   0 cases with perfect scores
==========================================
II – second round of analysis
```

To improve the fit of the data to the model, a second analysis was conducted but this time with the above 3 items deleted from the anchor-list leaving 12 items anchored on their MDT1 estimates. With this second round of item and case calibrations (Table 1, column 1a, columns II), a subsequent item fit map showed all items (except GEO17) were within the 95% acceptable bounds confirming that 37 out of 38 items were behaving as expected by the Rasch Model. The exception was the same item that showed erratic behaviour in MDT1 (meaning that some able students got it wrong while some less able students got it correct) but because GEO17 tested a basic geometric fact (*selection of 2 similar triangles out of 4 given triangles*), it was theoretically appropriate that it be left in the analysis (Bond & Fox, 2001). As can be seen in Table 1 (column 1a, columns I and II), the respective statistics for item mean, standard deviation and reliability remained relatively the same ($\leq$ 5% difference). The subsequent goodness-of-fit statistics (infit and outfit mean squares of 0.99 and 1.06 respectively in Table 1, column 1a, columns II) were within the expected mean values of 1.00. Likewise, the infit and outfit *t* values of -0.08 and 0.03 were within the expected mean of 0 logits hence confirming the overall fit of the item data to the Rasch Model (Bond & Fox, 2001).

The Rasch analysis also provides an item analysis output that includes, amongst other statistics, the *frequency* of incorrect (B and 0) and correct (1) responses and *mean ability estimate* of each response. It is therefore possible under the Rasch Model to predict, using the mean ability estimates, which students are making which response. Table 2 shows the relevant statistics for misfit item GEO17 indicating that while 46.8% were successful, 51.1% got it incorrect and 2.1% baulked. As expected with mean ability for each response, there is a monotonic trend from B→0→1 suggesting that the least capable students (mean of -1.24 logits) baulked while the most capable students (mean of -0.34 logits) were successful with the mean ability of the incorrect response between extreme means. The erratic behaviour of GEO17 can be further viewed by considering kidmaps - another unique output of a Rasch analysis to be discussed later on. Next are the cohort's statistics for person ability estimates.

Table 2
Item Analysis Statistics – GEO17

Infit MNSQ = 1.36 Disc = .14

Categories	B	0	1
Count	3	72	66
Percent (%)	2.1	51.1	46.8
Mean Ability	-1.24	-.56	-.34
Thresholds		-.26	
Error		.00	

```
----------------------------------------------------------------
Item Estimates (Thresholds)
all on all (N = 141 L = 38 Probability Level= .50)
----------------------------------------------------------------
 5.0                            |
                                |                Harder Items
                                |
                                |      MSR27     FNS31
                                |      ALG14
 4.0                            |
                                |
                                |
                                |
                                |
                                |
 3.0                            |
                                |
                                |
                                |
                             X  |      FNS09     FNS24
                                |      PRB13     ALG18
 2.0                            |
                                |      MSR15
                                |
                             X  |      PRB10
                           XXX  |      DRA16     ALG38
                            XX  |
 1.0                         X  |      FNS08     ALG33    MSR34
                      XXXXXXXX  |
                       XXXXXXX  |
                           XXX  |      FNS06     FNS32
                       XXXXXXX  |------PRB25     FNS30 Mean 0.34
                       XXXXXXX  |      FNS26
                     XXXXXXXXX  |      ALG21     FNS35
  .0                      XXXX  |      ALG11     FNS22
                       XXXXXXX  |      GEO17
      Mean -0.47 XXXXXXXXXXX    |      MCT02
                   XXXXXXXXXXX  |
                  XXXXXXXXXXXX  |
                          XXXX  |      MCT04     MCT05
-1.0                   XXXXXXX  |      ALG07     ALG37
                           XXX  |      GEO20
                        XXXXXX  |      MCT03     ALG19
                            XX  |      GEO29
                         XXXXX  |
                      XXXXXXXX  |      DRA12     ALG28
-2.0                        XX  |      GEO23
                           XXX  |      ALG36
                           XXX  |
                                |      MCT01
                             X  |
                             X  |                  Easier Items
-3.0                         X  |
                                |
                                |
                                |AAA##  - common MDT1/MDT2 items
                                |
                                |
-4.0                            |
----------------------------------------------------------------
  Each X represents    1 students
================================================================
```

Note: The 12 common items are bolded and underlined in the item-person map.

Figure 1. Item-Person Map for the 2005 MDT2 Cohort.

Case Estimates

Reliability – Table 1 (column 1b, column 1) shows the MDT2 mean ability estimate was –0.49 in the first calibration and became finalised to -0.47 in the second calibration. While the negative mean value indicated MDT2 was difficult for the cohort, it was a relatively higher mean case estimate compared to the lower MDT1 mean case estimate of -0.93 logits in Semester 1, 2005. Compared to the variation of item difficulty measures (1.76 logits) around item mean difficulty estimate of 0.34, the MDT2 person standard deviation of 0.93 indicated lesser variation of ability measures around its mean ability estimate (-0.47). On the other hand, while the person reliability index is high at 0.79, it is not as reliable as the item separations, but nonetheless more than acceptable (Bond & Fox, 2001). Theoretically, the person reliability requires not only ability estimates well targeted by a suitable pool of items but also a large-enough spread of ability across the sample so that measures demonstrate a hierarchy of ability/development (person separation) on the construct (Fox & Jones, 1998; cited in Bond & Fox (2001, p. 32)). Comparatively speaking, the relatively lower person reliability index further corroborates the targeting problem indicated by the negative mean ability estimate of -0.47 logits namely that the students found the test difficult.

Goodness-of-fit - The case infit and outfit mean square values were around the expected value of 1.00 (0.99 and 1.04 respectively) as expected by the Rasch Model. Similarly for the infit and outfit *t* values, -0.04 and 0.13 were around the expected value of 0 (Table 1, column 1b, column II) hence confirming the overall fit of the person data to the model (Bond & Fox, 2001). The item-person map in Figure 1 visually displays the location of means and spread of items/persons along the common logit continuum.

Reliability of the Diagnostic Test as a Measurement Tool

Both the item and person reliability indices presented above corroborate that MDT2, like MDT1, is a reliable test to measure unidimensionally the NAMC variable as indicated by the relatively high reliability indices (0.96 and 0.76 respectively). This means that items towards the top end of the logit continuum in Figure 1 are cognitively and mathematically more demanding than those at the bottom end. Theoretically, it is argued that, students who are highly numerate and mathematically competent with the PESM curricula are expected to have ability estimates at the top end of the hierarchy with *more than an average probability* of successfully completing the majority of the items (i.e. ability levels higher than item estimates) while less capable students would be located towards the bottom end with *less than average probability* of being successful with most items (i.e. ability levels below item estimates). On the contrary, overall, the item-person map (IPM) in Figure 1 indicates that the majority of items (25/38) are above the average person ability estimate (of -0.47 logits) with 3 items (MSR27, FNS31 and ALG14) located beyond the ability of the cohort and 5 other items (FNS09, FNS24, PRB13, ALG18 and MSR15) above the ability levels of 99% (140/141) of the cohort. Clumped around the mean ability estimate of -0.47 logits is the majority of the cohort. The apparent misalignment

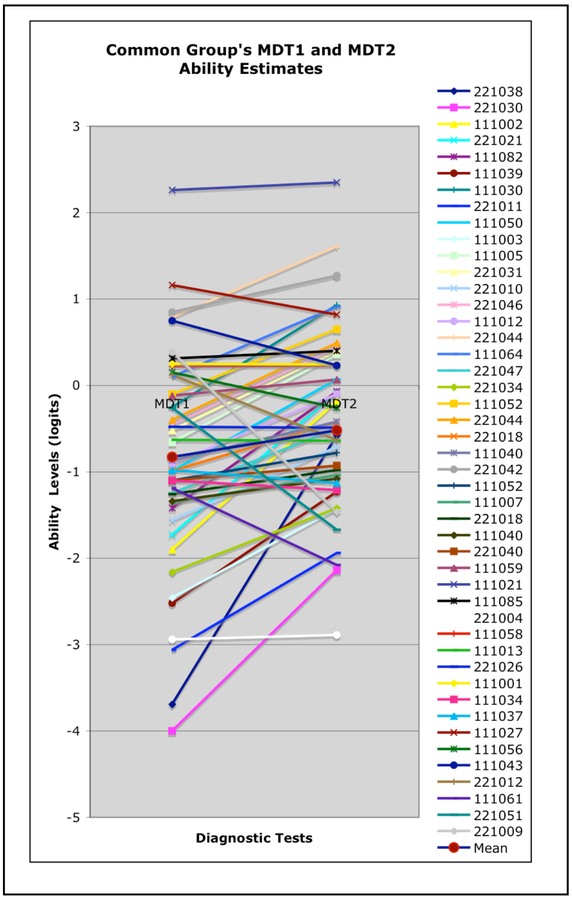

Figure 2. Common Group's MDT1 and MDT2 Ability Estimates.

between the item and person means, locations and overall distributions further corroborate that the 2005 MDT2 cohort, like the MDT1 cohort, found the test difficult.

The next sections discuss in more detail the performance of the group of students that sat both MDT1 and MDT2.

Student Performance in the Second Diagnostic Testing

Forty-six (46) of the 2005 cohorts took both MDT1 and MDT2 (to be referenced the 'common-group' CG). CG had group mean abilities of –0.83 and –0.52 for MDT1 and MDT2 respectively, demonstrating a positive ability change of 0.31 hence reflecting a general improvement by the second testing. A paired t-test showed a significant difference (*p=0.0009*) between the two performances. Shown in Figure 2 are the CG's paired results including test means to facilitate visual comparisons with students' ability estimates.

Table 3. Categories of Subgroups Relative to MDT1/MDT2 Means

Group	**Ability Change**	**MDT1 Mean**	**MDT2 Mean**	**Total Number**
CG	-1.85 to 3.11	-0.83	-0.52	46 (100%)
CG-Positive-Change	0.01 to 3.11	-0.83	-0.52	34 (74%)
CG-Negative-Change	-1.85 to -0.01	-0.83	-0.52	12 (26%)
Sub-groups	**Ability Change**	**MDT1 Mean**	**MDT2 Mean**	**Total Number**
CG-Above-0.31 (G1)	0.32 to 3.11	-	-	26 (56.5%)
❖ SG11	0.42 to 1.18	Above Mean	Above Mean	9
❖ SG12	0.45 to 1.59	Below Mean	Above Mean	8
❖ SG13	0.42 to 2.99	Below Mean	Below Mean	9
CG-Below-0.31 (G2)	0.29 to -1.99	-	-	20 (43.5%)
❖ SG21	-0.41 to 0.19	Above Mean	Above Mean	9
❖ SG22	-1.85 to -0.01	Above Mean	Below Mean	4
❖ SG23	-0.89 to 0.26	Below Mean	Below Mean	7

As Figure 2 and Table 3 show, 74% (34/46) of the students showed *positive improvement* by the second testing (CG-Positive-Change) with 26% (12/46) showing negative trends (CG-Negative-Change) suggesting *regression in performance* between the two tests. Whilst classifying student performance as "improvement" (positive gradient) or "regression" (negative gradient) is informative, a second analysis is necessary to determine whether (1) the change in ability estimates for each student was above or below the CG's mean ability change of 0.31, (2) the relative positioning of paired-ability-estimates were above or below each test mean, and based on resulting sub-groups from (1) and (2), determine whether (3) the performance of each sub-group between the two tests was significant.

Of the 46 students in the common group, 26 (56.5%) obtained ability changes greater than mean ability change of 0.31 logits (forming G1) by the second test with their performances in the two tests showing a significant difference (*p=0.0*). In contrast, 43.5% (20/46) had ability changes *below 0.31* (forming G2); there was also a significant difference (*p=0.007*) between the two performances. In accordance with criteria (1) and (2) above, 3 sub-groups per group were found as detailed in Table 3. The ability trends within each sub-group are discussed next as *one* way of making sense of the data. A case study per subgroup (SG11 to SG23) is presented and analysed in terms of *new correct items* in MDT2, *new errors* introduced in MDT2, and *persistent misconceptions* as evidenced by incorrect and baulked items in both tests.

To frame the analysis of the case studies and subsequent discussion of findings, the following three general questions are used:

(1) What are the most common areas of improvement?
(2) What are the most common misconceptions?
(3) What are the similarities and differences, if any, between the groups and subgroups?

Table 4 provides text descriptions of the test items. Besides the 15 common items, the rest used different numerical values in the two versions of the diagnostic test.

Table 4. MDT2 Item Text Descriptions.

Item	Text Descriptions	Item	Text Descriptions
MCT01	7 x 9 = ?	**GEO20**	Missing angle of a quadrilateral given 70°, 115° and 115° .
MCT02	What is 20% of 70?	ALG21	Find the value of x if 11 −10x=6x−52.
MCT03	6013 – 3987 = ?	FNS22	Write three fractions that are equivalent to $\frac{3}{7}$.
MCT04	$\frac{2}{5}+\frac{1}{4}=?$	**GEO23**	In the diagram, which angle has a measure closest to 45°?
MCT05	0.3 x 0.3 = ?	FNS24	Penny had a bag of marbles. She gave one quarter of them to Rebecca. She then gave a third of the remaining marbles to Jack. If Penny ended up with 32 marbles, how many did she start with?
FNS06	Write in ascending order 0.625, 0.25, 0.037, 0.5, 0.125.	**PRB25**	Eleven chips are labelled 2, 3, 5, 6, 8, 10, 11, 12, 14, 18 and 20 respectively. The eleven chips are placed in a bag and one is drawn out at random. What is the probability that the number on the chip is both even and a multiple of 3?2,3,5,6,8,10,11,12,14,18,20.
ALG07	An unknown number *n* is multiplied by 12 and then 11 is subtracted from the result. The final answer is 71. Write this as a mathematical expression.	FNS26	Shade 2/9 in a 9 x 4 grid.
FNS08	Write in ascending order $\frac{5}{7},\frac{3}{4},\frac{7}{10},\frac{4}{9}$.	MSR27	The length of the rectangle below is one and a half times its width. What is the ratio of the width to the perimeter?
FNS09	An athlete ran 4.5 kilometres in exactly 12 minutes. What was her average speed in metres/sec?	ALG28	Write in simplest form 4 × n × n × n × n.
PRB10	If a fair coin is tossed, the probability that it will land heads up is 1/2. A fair coin is tossed 4 times and it lands heads up each time. What is likely to happen when the coin is tossed a fifth time?	**GEO29**	In the diagram (of similar triangles), what is the length of the interval BD?
ALG11	If 6 times a number is 54, what is two thirds of the number?	FNS30	Sound travels at approximately 330 m/sec. A lighting strike was followed 5.2 seconds later by a clap of thunder. How far away did the lightning strike?
DRA12	The graph shows the time taken to travel to school by a group of students. How many pupils travel for more than 10 minutes to reach school?	FNS31	A pile of salt contains 600 individual crystals and has a weight of 7.8g. What is the average weight of a salt crystal?
PRB13	A sample of 200 voters is chosen at random from an electorate containing 6000 voters. When the sample was asked, it was found to contain 13 people who voted informally. How many informal votes would you expect to find in the complete electorate?	FNS32	Laura had $270 but spent five ninths of it. How much money did she have left?
ALG14	At a particular university, there is an average of 16 students to every professor. Write this as a mathematical equation.	**ALG33**	A club has 72 members with 16 more female members than male members. How many males and females are members of the club?
MSR15	A rectangular garden bed adjoins a building as shown in the diagram. The garden bed has a path on 3 sides. What is the area of the path?	**MSR34**	What is the area of the shaded rectangle?
DRA16	Two streets in a town have 30 houses (Orange St.) and 21 houses (Lime St.) respectively. This is represented in the pictogram below. How many houses are represented by the symbol?	FNS35	A fertilizer mix contains 250g of nitrate, 500g of phosphate and 750 g of potash. What is the ratio of the weight of the nitrate to the total weight of the fertilizer?
GEO17	Which two of the four triangles are similar?	**ALG36**	Extension of a geometric pattern
ALG18	An elevator starts at the first floor of a building. It travels up to the sixth floor, then down to the second floor and back up to the fifth floor. If the floors are 2.8 metres apart, how far did the elevator travel?	**ALG37**	Extension to 2 terms of a numeric tabular pattern based on ALG36
ALG19	If x = 5, what is the value of $\frac{7x-3}{3x-2}$?	**ALG38**	If we produced a figure with 50 rows, we would require 1275 blocks. Explain how to calculate the number of blocks required to construct a figure with 51 rows.

AAA## - common MDT1/MDT2 items

Group G1- Ability Change Above 0.31

Only 9 of the 26 (34.6%) students (SG11) in G1 had individual ability estimates that were *both* located higher than the tests' mean abilities, see Figure 3. A paired t-test showed a significant difference (*p=0.0)* between the performances of the 9 students. Figure 4 shows 30.7% (8/26) students (SG12) with ability changes above 0.31 whereas individual estimates were below MDT1 mean and above MDT2 mean; a paired t-test indicated a significant difference between the two performances

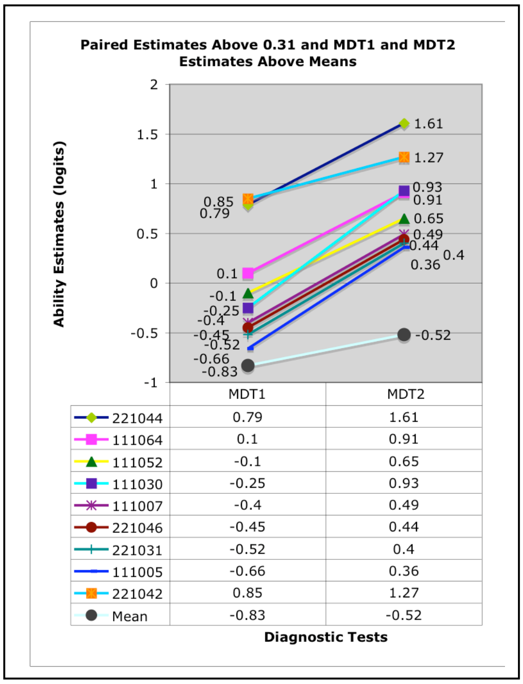

Figure 3. SG11 – Ability Estimates Above 0.31 and Estimates Both Above MDT1/MDT2 Means.

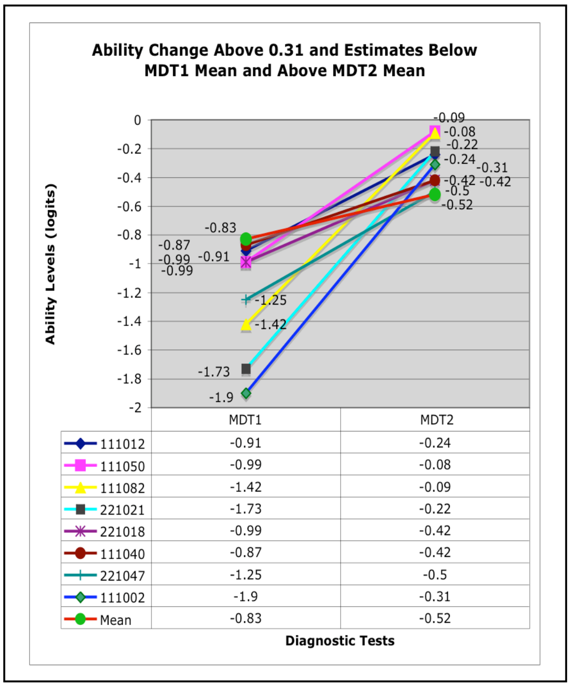

Figure 4. SG12 - Ability Estimates Above 0.31 and Estimates Below MDT1 Mean and Above MDT2 Means.

(*p=0.001*). On the other hand, Figure 5 shows 34.6% (9/26) students (SG13) had ability changes above 0.31, but first and second ability estimates were *below* test means; a paired t-test demonstrated a significant improvement by their second test (*p=0.004*) despite their below-mean performances.

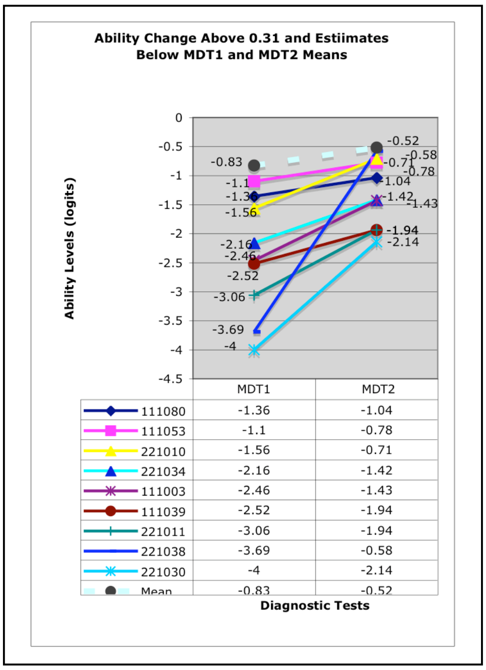

Figure 5. SG13 – Ability Change Above 0.31 and Estimates Below MDT1/MDT2 Means.

A case study of each of the subgroups SG11, SG12 and SG13 are presented next by utilising the diagnostic information illustrated on Rasch's kidmaps. A kidmap typically locates an individual's ability estimate in the middle column (marked XXX) along the logit continuum with correct items on the left and incorrect items on the right. The left and right dotted horizontal lines represent the upper and lower bounds respectively, of the student's ability estimate effectively separating the kidmap into items that the model predicts are more difficult (above upper bound) and less difficult (below lower bound) for the student with the items within the bounds representing those the model predicts the individual will have a 50% of being successful.

Group G1 - Subgroup SG11– Student 111030

A case study of Student 111030 with the highest ability change of 1.18 logits in SG11 (Figure 3)

```
                              MDT2
------------------------ K  I  D  M  A  P-------------------
   Candidate: 111030                    ability:    .93
   group:     all                       fit:        .83
   scale:     all                      % score:  65.63

----Harder Achieved -------------------Harder Not Achieved-
                            |   |
                            |   |
                            |   |
                            |   |  (MSR27) FNS31
                            |   |  ALG14
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   | (FNS09)  (FNS24)
                            |   |  PRB13    ALG18
                            |   |
                            |   |  MSR15
    ............................   PRB10
                            |   |
                  DRA16     |   | (ALG38)
                  ALG33     |   |
                  FNS08     |XXX|  MSR34
                            |   |
                  FNS32     .........FNS06................
                            |   |  PRB25    (FNS30) FNS26
                  ALG21     |   |
                  FNS35     |   |
                  ALG11     |   |  (FNS22)
                  GEO17     |   |
                            |   |  MCT02
          MCT05   MCT04     |   |
                  ALG37     |   |
                  ALG07     |   |
 GEO20    ALG19    MCT03    |   |
                  GEO29     |   |
                  DRA12     |   |
          ALG28   GEO23     |   |
                  ALG36     |   |
                            |   |
                  MCT01     |   |
                            |   |
                            |   |
                            |   |
 AAA## - new successes      |   | AAA## - new errors in MDT2
                            |   | (AAA##) – MDT2 blank items
                            |   |
-----Easier Achieved -------------------Easier Not Achieved
===========================================================
```

Figure 6. Student 111030'd MDT2 Kidmap.

provides a detailed view of a student's successes, errors, and persistent misconceptions within this subgroup. A comparison of his two kidmaps showed that in MDT2 (Figure 6), (1) he successfully completed MDT1 incorrect items (MCT05, ALG37, DRA12, and ALG36) in his MDT1 "Easier Not Achieved" quadrant; (2) he successfully completed a baulked MDT1 item (GEO17); and (3) he successfully completed 3 items (FNS08, DRA16, and ALG33) that were incorrect in MDT1. On the other hand, in MDT2, (1) he unsuccessfully completed PRB13 and FNS26 which were both correct in MDT1; (2) he continued to demonstrate difficulties successfully solving (a) multi-step word problems that require interpretation and complex procedures (FNS31, ALG14, ALG18, MSR15, and PRB10) shown in the "Harder Not Achieved" quadrant in Figure 6, indicating he had *less than average* probability of succeeding (above ability estimate), and (b) items (MSR34 and FNS06) in which he had an *average probability* (within his MDT2 ability band) and *higher than average probability* (below ability level) of being successful (PRB25 and MCT02) shown in his MDT2 "Easier Not Achieved" quadrant; and (3) he baulked from word problems (MSR27, FNS09, FNS24, ALG38, and FNS30) and knowledge item FNS22 as illustrated in Figure 6. Student 111030's demonstrated improvements since MDT1 included correctly multiplying decimals; ordering fractions; identifying similar triangles; determining a pictograph scale; interpreting data from a bar graph; and solving a simple 2-step word problem on extending a geometric and tabular patterns, and additively distributing amounts over two quantities. His new errors included solving a word problem on proportional reasoning and representing an equivalent fraction. An inspection of responses showed that he misrepresented percentage of an amount (20% of 70) as a percentage (i.e. 14%); incorrectly shaded the equivalent number of blocks (9 out of 36 squares) to represent a fraction (2/9); incorrectly calculated probability of selecting an even and multiple of 3 chip as 2/4; incorrectly arranged ascending decimals as 0.625, 0.25, 0.037, 0.5, 0.125; and incorrectly calculated shaded/nested areas within other geometric shapes. Moreover, his

```
                               MDT2
------------------------ K  I  D  M  A  P------------------------
    Candidate: 111002                        ability:   -.31
    group:     all                           fit:       1.15
    scale:     all                           % score:  46.67

--------Harder Achieved ---------------Harder Not Achieved ------
                              |   |
                              |   |
                              |   |   FNS31  (MSR27)
                              |   |
                              |   |
                              |   |   ALG14
                              |   |
                              |   |   PRB10
                  FNS24       |   |   FNS09
                              |   |   PRB13      ALG18
                              |   |
                  DRA16       |   |
                              |   |   ALG33    (ALG38)
                              |   |   FNS08    (MSR15) (MSR34)
                              |   |
..................FNS32.............  FNS06    (PRB25)
                              |   |   FNS26
                              |   |   ALG21    (FNS30)
                  FNS22       |   |   FNS35    (ALG11)
                              |XXX|   GEO17
                  MCT02       |   |
                              .................................
                  MCT05       |   |   MCT04
                              |   |   ALG37    (ALG07)
                  GEO20       |   |
                  MCT03       |   |   ALG19
                  GEO29       |   |
          ALG28   DRA12       |   |
                  GEO23       |   |
                  ALG36       |   |
                              |   |
                  MCT01       |   |
                              |   |
  AAA## - new successes       |   |(AAA##) – MDT2 Blank items
                              |   |
                              |   |AAA##  - new error in MDT2
                              |   |
----------Easier Achieved --------------Easier Not Achieved -----
=================================================================
```

Figure 7. Student 111002's MDT2 Kidmap.

responses to word problems indicated difficulties comprehending probability and quantitative descriptions such as in the elevator and salt crystal problems, reasoning with rates and proportions, and mathematically transforming his interpretations into algebraic expressions and/or equations such as in the professor:student problem. Overall, Figure 6 shows strong evidence that the student consistently had difficulties operating with fractions, rates, ratios, proportions, probability, measurement, and solving complex word problems.

Group G1 – Subgroup SG12 – Student 111002

A case study of Student 111002 with the highest ability change (1.59 logits) (see Figure 4) provides a detailed view of a student's success and persistent misconceptions within this subgroup. A comparison of her two kidmaps showed that in MDT2 (Figure 7), (1) she successfully completed MCT01 which was incorrect in MDT1; (2) she successfully completed MCT03, GEO29, and GEO20 which were within her MDT1 ability band; and (3) she successfully completed MCT05, DRA16, and FNS24 which were in her MDT1 "Harder Not Achieved" quadrant. On the other hand, in MDT2, (1) she unsuccessfully completed MDT1 correct item ALG19; (2) she was again unsuccessful with ALG37 and MCT04 shown within her MDT2 "Easier Not Achieved" quadrant; (3) she was consistently unsuccessful with items FNS06, FNS26, ALG21, FNS35, and GEO17 shown within her MDT2 ability band; (4) she was persistently unsuccessful solving "Harder Not Achieved" items FNS08, ALG33, ALG18, PRB13, FNS09, PRB10, ALG14, and FNS31; (5) she consistently baulked from MSR27 and ALG07; and (6) she baulked from MDT1 incorrect items ALG38, MSR15, MSR34, FNS30, PRB25, and ALG11.

Overall, Student 111002's successes since MDT1 included correctly solving basic computational problems such as multi-digit subtraction and multiplication of decimals and whole numbers; knowledge items such as finding unknown angle of a quadrilateral, and items on proportional reasoning with similar triangles, determining ratio of a pictograph scale; and operating with fractions. In contrast, her new error

```
                              MDT2
--------------------- K  I  D  M  A  P----------------------
 Candidate: 221038                          ability:   -.58
 group:     all                             fit:        .59
 scale:     all                             % score:  41.18

-------Harder Achieved -------------Harder Not Achieved ---
                            |   |  (MSR27)  (FNS31)
                            |   |
                            |   |
                            |   |
                            |   |  ALG14
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |
                            |   |  FNS24   (FNS09)
                            |   |
                            |   |
                            |   |  PRB10
                            |   | (PRB13)  (ALG18)
                            |   |  DRA16   (ALG38)
                            |   |  ALG33   (MSR34)
                            |   |  FNS08   (MSR15)
                            |   |
               FNS06        |   | (FNS32)
 ...............................   PRB25    FNS26   (FNS30)
                            |   |
                            |   | (ALG21)  (FNS35)
                            |   |  GEO17   (ALG11)  (FNS22)
                            |XXX|  MCT02
                            |   |
                            |   | (ALG07)  (GEO20)
               ALG37        |   | (MCT04)  (MCT05)
                            ...............................
                            |   | (MCT03)  (GEO29)  (ALG19)
                            |   |
                            |   |
              DRA12         |   |
      ALG28   GEO23         |   |
              ALG36         |   |
                            |   |
              MCT01         |   |
                            |   |
                            |   | (AAA##) – MDT2 Blank Items
AAA## - new successes       |   | AAA## - new error
                            |   |
                            |   |
------Easier Achieved --------------Easier Not Achieved ----
===========================================================
```

Figure 8. Student 221038's MDT2 Kidmap.

included evaluation of a rational algebraic expression while her persistent misconceptions included ordering fractions and decimals and representing equivalent fractions; reasoning and operating with fractions, rate, ratio, and proportions; reasoning with additively and multiplicatively related quantities; reasoning with probability, likely outcomes and expected number; solving a linear equation; extending, reasoning with, and generalising numeric tabular patterns; and solving word problems requiring interpretation and mathematical representation and/or application of complex procedures. The only algebraic items she could solve successfully were basic ones on extending a pictorial geometric pattern and simplifying a basic expression. Overall, the student demonstrated difficulties reasoning and operating with fractions, rates, probability, evaluating algebraic expressions and solving linear equations; generating and generalising numerical patterns; and solving more complex word problems.

Group G1 – Subgroup SG13 – Student 221038

A case study of Student 221038 with the highest ability change (3.11 logits) see Figure 5 provides a detailed view of a student's successes and persistent misconceptions within this subgroup. A comparison of her two kidmaps showed that in MDT2 (Figure 8), she successfully completed incorrect MDT1 items MCT01, ALG36, DRA12, ALG37, FNS06, and ALG28. On the other hand, in MDT2, (1) she made a new error when she unsuccessfully completed previously correct MDT1 item GEO17; (2) she persistently made errors when solving incorrect MDT1 items ALG14, FNS24, DRA16, FNS08, FNS26, and MCT02 shown in her MDT1 "Harder Not Achieved" quadrant; (3) she unsuccessfully attempted MDT1 ignored items PRB10 and ALG33 shown in her MDT2 "Harder Not Achieved" quadrant and PRB25 shown within her MDT2 ability band; (4) she baulked from incorrect MDT1 items PRB13, FNS32, ALG11, MCT04, MCT05, ALG07, and MCT03; and (5) in MDT1 as well, she consistently baulked from MSR27, FNS31, FNS09, ALG18, ALG38, MSR34, MSR15, FNS30, ALG21, FNS35, FNS22, GEO20, ALG19 and GEO29. In fact, her MDT1 performance (Figure 5) was quite poor as indicated by her low MDT1 ability estimate of -3.69

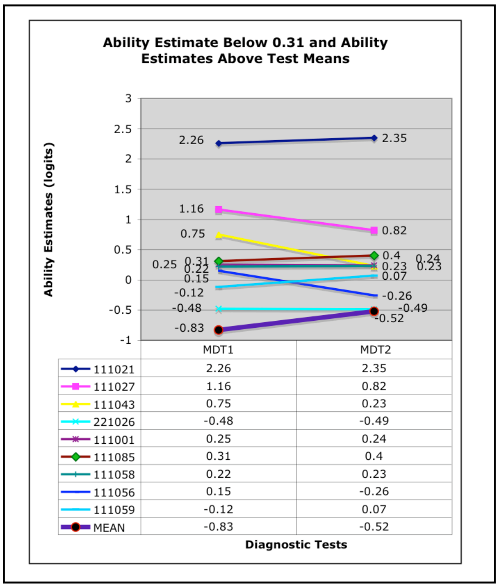

Figure 9. SG21 – Ability Change Below 0.31 and Estimates Above Means.

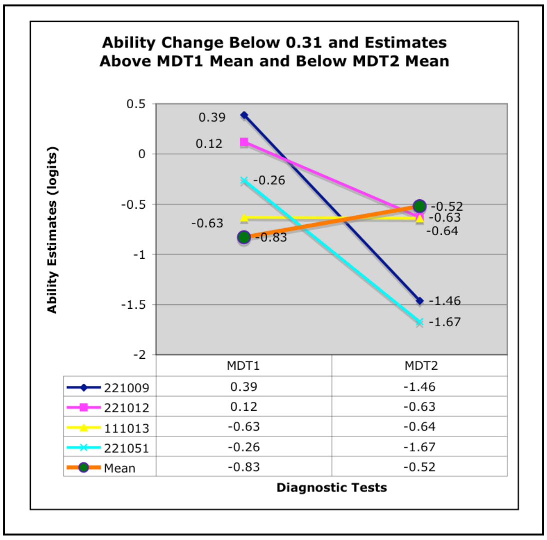

Figure 10. SG22 – Ability Change Below 0.31 and Estimates Above MDT1 Mean and Below MDT2 Mean.

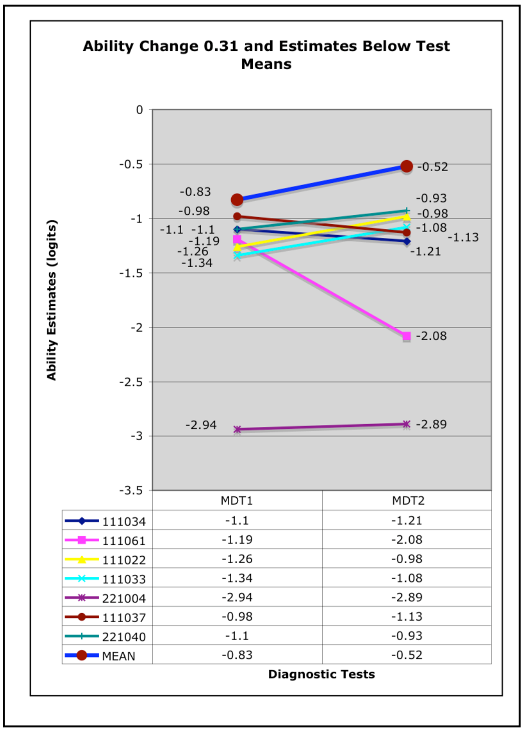

Figure 11. SG23 – Ability Change Below 0.31 and Estimates Below Test Means.

logits and the number of questions she actually completed successfully in each test (2 out of 25 attempts in MDT1 and 7 out of 17 attempts in MDT2) out of a total possible of 38. In addition, she baulked from 13 items in MDT1 and 21 items in MDT2. Overall, Student 221038's second performance may have improved 3.11 logits but her performance with incorrect and blank items shown in Figure 8 indicates that she is potentially at-risk of not achieving mastery level (>80% correct) of the NAMC construct. Her successes are limited to basic computations and ordering of decimals, basic algebraic, basic geometric and basic reading bar graph items.

The following sections present data from the second group G2's 3 subgroups – those with ability changes below 0.31 logits.

Group G2 – Ability Change Below 0.31

Only 9 of the 20 (45%) students had individual ability estimates that were both higher that two mean estimates (SG21), see Figure 9. A paired t-test showed a significant difference (*p=0.04)* between the performances of the 9 students. In contrast, Figure 10 shows 4 of the 20 (20%) students (SG22) had ability changes that were less than 0.31 with individual estimates above MDT1 mean and MDT2 estimates *below* the MDT2 mean; a paired t-test indicated a significant difference between the two performances (*p=0.03*). On the other hand, Figure 11 shows 7 out of 20 (35%) students (SG23) whose ability changes

```
                          MDT2
------------------- K  I  D  M  A  P-----------------
 Candidate: 111021                   ability:   2.35
 group:     all                      fit:        .82
 scale:     all                      % score:  87.50

-----Harder Achieved ---------Harder Not Achieved --
                          |   | (MSR27) (FNS31)
                          |   |
                          |   |
                          |   |
                          |   |
                          |   |  (ALG14)
                          |   |
                          |   |
                          |   |
                          |   |
                          |   |
...........................
                          |   |
                          |   |
                          |XXX|   FNS09  (FNS24)
          ALG18           |   |  (PRB13)
                          |   |
                          ........MSR15................
                          |   |   PRB10
                          |   |
            DRA16         |   |  (ALG38)
            ALG33         |   |
   MSR34    FNS08         |   |
                          |   |
                          |   |
 FNS32 FNS30 FNS06        |   |
       FNS26              |   |   PRB25
             ALG21        |   |
     FNS35   FNS22        |   |
             ALG11        |   |
             GEO17        |   |
             MCT02        |   |
                          |   |
      MCT05  MCT04        |   |
             ALG37        |   |
             ALG07        |   |
 GEO20 ALG19 MCT03        |   |
             GE029        |   |
                          |   |
             DRA12        |   |
     ALG28   GEO23        |   |
             ALG36        |   |
                          |   |
             MCT01        |   |
                          |   |
                          |   |
                          |   |(AAA##) —Blank MDT2 items
AAA## - new successes     |   |
                          |   | AAA## - new errors
                          |   |
----Easier Achieved -----------Easier Not Achieved --
=====================================================
```

Figure 12. Student 111021's MDT2 Kidmap.

were less than 0.31 with *both* their first and second ability estimates *below* the two means; a paired t-test showed no significant difference between the two performances ($p>0.05$). Within group G2, 60% (12/20) showed regression in performance as indicated by negative ability changes with the rest showing differential positive improvements.

Group 2 - Subgroup SG21 – Student 111021

A case study of Student 111021 with the highest ability estimates in both tests, provides a detailed view of a student's success and persistent misconceptions within this subgroup. A comparison of her two kidmaps showed that in MDT2 (Figure 12), she successfully completed MDT1 incorrect items DRA16 and FNS32 shown within her "Easier Not Achieved" MDT1 quadrant. On the other hand, in MDT2, (1) she baulked from the MDT1 correct item FNS31 shown in her "Harder Achieved" MDT1 quadrant; (2) she unsuccessfully completed the MDT1 correct item FNS09 shown within her MDT1 ability band; (3) she was persistently unsuccessful with MSR15 and PRB25 shown within her "Easier Not Achieved" MDT1/MDT2 quadrants; (4) she unsuccessfully attempted the ignored MDT1 item PRB10 and baulked from the MDT1 incorrect item FNS24; and (5) she consistently baulked from word problem items MSR27, ALG14, PRB13, and ALG38. Demonstrated improvements since MDT1 consisted of correctly determining the ratio of a pictograph scale and operating successfully with fractions to determine balance left from an initial amount. New errors were word problems on rates such as expressing average speed (m/s) as 4500m:720sec and average weight of salt crystal. Error responses included interpreting likely outcomes of an event as a value (2.5), miscounting favourable outcomes (1/11 instead of 3/11), and incomplete procedure suggesting misinterpretation of the question. While new errors were related to rates such as speed, unit conversion and average weight, persistent misconceptions included difficulties identifying favourable outcomes and likely outcomes given an event, distinguishing between relevant and irrelevant information on a diagram, and solving word problems on rates, ratios, and operations with fractions. In contrast, the student consistently ignored word problems such as those involving interpretation and representation of multiplicative relationships, use of complex procedures, extending and generalising patterns, and making predictions based on sample proportions.

Evidence from her kidmap and actual responses indicate that whilst she had done well compared to her peers, she needs to focus on improving her interpretive and problem solving skills particularly regarding rates, fractions, generalising patterns and probability. Whilst she may have scored at mastery level (>80%), the Rasch model predicts that she had average to greater than average probability of successfully completing 7 of the incorrect/baulked items at or below her ability estimate with clearly less than average probability of succeeding with the difficult word problems at the high end of the logit continuum at the "Harder Not Achieved" quadrant.

Group 2 - Subgroup SG22 – Student 111013

A case study of Student 111013 with ability change of -0.16 (see Figure 10) provides a detailed view of a student's success and persistent misconceptions within this subgroup. A comparison of her two kidmaps showed that in MDT2 (see Figure 13), (1) she successfully completed MDT1 incorrect items (GEO29 and ALG19) shown within her MDT1 "Easier Not Achieved" quadrant and FNS26 within her MDT1 "Harder Not Achieved" quadrant; and (2) she successfully completed MDT1 ignored item FNS35 within her MDT1 ability band. On the other hand, in MDT2, (1) she unsuccessfully completed MDT1 correct items (a) MCT04 and ALG11 shown within her MDT2 ability band and (b) PRB25 and FNS32 shown within her MDT2 "Harder Not Achieved" quadrant; (2) she persistently had difficulties solving MCT02 and GEO17 shown within her MDT2 ability band, FNS30, FNS06, FNS08, ALG33, MSR15, DRA16, ALG18, FNS24, PRB13, FNS09, PRB10, and FNS31 shown within her MDT2 "Harder Not Achieved" quadrant; (3) she unsuccessfully completed ignored MDT1 items MSR34, ALG36, and MSR27; (4) she baulked from incorrect MDT1 item ALG14; and (5) she consistently baulked from items ALG37 and ALG38 in both tests. Demonstrated improvements since MDT1 included successes with computing ratio of a quantity to a total, representing an equivalent fraction, evaluating an algebraic rational expression, and applying proportional reasoning to similar triangles. Whereas new errors were incorrectly computing sum of fractions, operating with fractions,

```
                        MDT2
----------------- K  I  D  M  A  P-------------------
 Candidate: 111013                ability:   -.64
 group:     all                   fit:        .89
 scale:     all                   % score:  37.14

----Harder Achieved ----------Harder Not Achieved ---
                        |   |
                        |   |
                        |   |
                        |   |   MSR27   FNS31
                        |   |
                        |   |
                        |   |  (ALG14)
                        |   |
                        |   |
                        |   |
                        |   |
                        |   |
                        |   |   FNS09   FNS24
                        |   |   PRB13   ALG18
                        |   |
                        |   |   MSR15
                        |   |   PRB10
                        |   |
                        |   |   DRA16  (ALG38)
                        |   |   ALG33
                        |   |   FNS08   MSR34
                        |   |
                        |   |
                        |   |   FNS06   FNS30   FNS32
           FNS26        |   |   PRB25
           ALG21        |   |
...FNS35...FNS22............
                        |   |   ALG11
                        |   |   GEO17
                        |   |   MCT02
                        |XXX|
            MCT05       |   |   MCT04
                        |   |
            ALG07       ............................
   ALG19   MCT03        |   |   GE020  (ALG37)
           GEO29        |   |
           DRA12        |   |
  ALG28    GEO23        |   |
                        |   |   ALG36
                        |   |
                        |   |
           MCT01        |   |
                        |   |
                        |   |
                        |   |
                        |   |
                        |   |(AAA##)— MDT2 Blank Items
AAA## -  new successes  |   |
                        |   | AAA## - new errors
                        |   |
-----Easier Achieved ----------Easier Not Achieved --
=====================================================
```

Figure 13. Student 111031's MDT2 Kidmap.

and determining favourable outcomes of an event, persistent misconceptions were also related to fractions such as ordering decimals and fractions, and reasoning with fractions to determine initial amount; rates and number sense such as computing distance given speed and time, average speed in new units, average weight and summing distances; probability such as making predictions based on sample probability, and describing likely outcomes of an event; proportions such as additively distributing an amount; applying basic geometric properties; interpreting and identifying relevant information on diagrams and pictograph; identifying similar triangles, and computing percentage. Overall, the student consistently had difficulties solving word problems that require investigation, interpretation of relational statements, mathematical representation, pattern extension, generalisation, and communication of her reasoning.

Group 2 - Subgroup SG23 – Student 1110022

A case study of Student 111022 with ability change of 0.28 (see Figure 11) provides a detailed view of a student's success and persistent misconceptions within this subgroup. A comparison of her two kidmaps showed that, in MDT2 (Figure 14), she successfully solved incorrect MDT1 item DRA12. In comparison, her new errors included MDT1 correct items ALG19, GEO17, and PRB25. She continued to unsuccessfully complete (1) MDT1 incorrect items (MCT02, MCT04, ALG07, and GEO29) (shown within her MDT2 ability band), and (2) MDT1 'Harder Not Achieved' items (FNS22, FNS35, ALG21, FNS26, FNS06, FNS08, MSR34, ALG33, DRA16, ALG38, PRB10, MSR15, PRB13, ALG18, FNS09 and MSR27) (shown above her MDT2 ability band). Consistently also, she baulked from FNS31, ALG14, FNS30, and ALG11. Her demonstrated improvement since MDT1 included correctly reading data from a bar graph whereas her new errors included incorrectly evaluating a rational algebraic expression, identifying similar triangles, and identifying favourable outcomes and computing probability. Persistent misconceptions included determining percentage of a number, adding simple fractions, algebraically representing a relation, and applying proportional reasoning all of which she had average probability of being successful. While the rest of the incorrect items were all above her

```
                        MDT2
---------------- K  I  D  M  A  P--------------
 Candidate: 111022                    ability:   -.99
 group:     all                  fit:         .70
 scale:     all                  % score:  34.38
----Harder Achieved -------Harder Not Achieved ----
                        |   |
                        |   |
                        |   |
                        |   |  MSR27   (FNS31)
                        |   | (ALG14)
                        |   |
                        |   |
                        |   |
                        |   |  FNS09 (FNS24)
                        |   |  PRB13  ALG18
                        |   |
                        |   |  MSR15
                        |   |  PRB10
                        |   |  ALG38
                        |   |  DRA16
                        |   |  ALG33
                        |   |  FNS08  MSR34
                        |   |
          FNS32         |   |  FNS06
                        |   |  FNS26  (PRB25) (FNS30)
                        |   |  ALG21
                        |   |  FNS22  FNS35   (ALG11)
                        |   |
.........................      GEO17
                        |   |  MCT02
                        |   |
          MCT05         |   |  MCT04
          ALG37         |XXX|
                        |   |  ALG07
     GEO20  MCT03       |   |  ALG19
                        ........GEO29................
                        |   |
         DRA12          |   |
   ALG28  GEO23         |   |
          ALG36         |   |
                        |   |
          MCT01         |   |
                        |   |
                        |   | (AAA##) – MDT2 Blank Items
AAA##-new successes     |   | AAA## - new errors
                        |   |
                        |   |
--Easier Achieved -----------Easier Not Achieved ----
====================================================
```

Figure 14. Student 111022's MDT2 Kidmap.

ability estimate (less than average probability of being successful), she was successful with basic multiplication of single-digit whole numbers and simple decimals, basic operations with fraction of an amount, basic algebraic items, and basic geometric items as well as successfully completing a relatively more difficult 2-step 'operating with fractions' item. In contrast, her persistent misconceptions as evident from her error responses concerned multi-step operating with fractions, ordering fractions and decimals; evaluating algebraic expressions and solving a linear equation; reasoning with probability and likely outcomes; determining rates; and solving word problem items that involve interpreting quantitative relationships, mathematically representing interpretations, and applying complex procedures. Overall, the student had difficulties solving word problems on operations with fractions, rates, ratios, probability, and interpretation of additive and multiplicative relationships. Her kidmap shows 6 items she had average and more than average probability of being successful with and 22 items that were harder with less than average probability of getting them correct.

The next sections discuss some of the general issues emerging from the cohort's item-person map and 6 case studies.

Discussion

The discussion is framed around the three questions: (1) *What are the most common areas of improvement?* (2) *What are the most common misconceptions?* (3) *What are the similarities and differences between the groups and subgroups?* First before considering the case studies, according to the cohort's item-person map, the misalignment between the items/persons distributions resulted in 8 items being located clearly above 99% of the cohort implying that as a group of prospective teachers, the empirical evidence demonstrated the need to provide targeted and specific remediation if ability estimates are to be improved by the final round of testing. A content analysis of the 8 most difficult items indicated that they are word problems involving multiplicative quantitative relationships including ratios, rates, proportions and identifying relevant information in a complex diagram. Cognitively, these items require thinking and reasoning multiplicatively, interpreting and comprehending quantitative descriptions, and transforming subsequent interpretations mathematically to find solutions. In contrast, at the bottom of the item-person map were 3 persons whose ability estimates indicated that the test did not have sufficient 'easiest items' to accurately and adequately indicate the range of their numeracy capabilities and mathematical competencies.

(1) *What are the most common areas of improvement?*

Findings from the cohort's item-person map demonstrated that there is a distinct hierarchy of items ranging from basic computational items that use routine procedures such as addition, subtraction, substitution, and evaluation, to those that use 1-2 step procedures and through to progressively more complex word problems that require multiple steps such as interpretation and comprehension of relational descriptions, reasoning with the interpretations and mathematically representing the results algebraically, and using complex procedures to find solutions. Collective evidence from the item-person map, 6 case studies and kidmaps suggest that students tended to have *more than average probability* of being successful with items towards the bottom level of the hierarchy with progressively *less chances* of being successful with items that have higher estimates. Whilst individual improvements, new errors and persistent misconceptions vary from individual to individual, the clear pattern emerging is that operating with fractions; reasoning about probability; distinguishing between relevant and irrelevant information in nested geometric shapes, bar graph and pictograph; and making sense of relational descriptions that involve multiplicative and additive relationships, rates, and ratios appear to be more challenging for students. Basically, items that are presented with fewer words (i.e. computational problems or simple diagrams) appear less challenging while student performances with knowledge items such as those on similar triangles, sum of interior angles of a quadrilateral, estimating size of angle, reading a bar graph, simplifying an algebraic expression, and equivalent fractions vary depending on the recency of their mathematical experiences with them. Overall, the most common areas of improvements appeared to be towards the bottom end of the hierarchy of items with computational problems, knowledge items, interpreting diagrams and bar graph, and applying basic geometric and algebraic properties.

(2) *What are the most common misconceptions?*

Findings as presented through cohort's item-person map and six kidmaps suggested that when students have difficulties with knowledge items, such as those on ordering fractions - they tended to use either the numerator and/or the numerator while it was the number of decimal places for decimal fractions. On the other hand, of the four mental computation items, the most difficult was the percentage of a whole number followed by the addition of fractions, multiplication of decimals and multi-digit subtraction whereas multiplication of single-digit whole numbers was the least difficult. Understanding and applying the concept of probability appeared quite challenging particularly distinguishing between likely, favourable and unfavourable outcomes of events, and making predictions for larger samples. Although average weight and ratio of width:perimeter appeared as relatively straightforward rates, students found these two items the most challenging; the first one due to students' tendency to simply multiply/divide given quantities and the second one because of students' struggles to synthesise the multiplicative relationship between length and width with perimeter. Converting units for computed rates (average speed) appeared difficult for students as was interpreting word problems consisting of a mix of fractional distributions and whole numbers. The transformation of descriptions of stops and starts of an elevator into total distance was hard for most students and so were the extension and generalisation of numerical patterns. Asking students to provide algebraic expressions or equations for co-relationships appeared complicated if the relationship was multiplicative as evident from error responses to the 'student:professor' item compared to that for the 'equation of a relational description'. Overall, there was evidence that students in the common-group had specific and basic misconceptions related to fractions, decimals, geometry, and measurement as suggested by their error responses to computational and knowledge items. In addition, they tended to baulk from, or were unsuccessful with items that required investigation, interpretation of relational statements, reasoning with and synthesising information, and use of complex procedure; that is, solving items that require critical and higher order problem solving skills. Studies by other researchers (Shulman, 1986; Thompson, 1992; Aubrey, 1997; Rowland, Barber, Heal, & Martyn, 2003; Ball, Lubienski, & Mewborn, 2001) reported that preservice teachers do not always have the conceptual understanding of the mathematics content they will be expected to teach. Van Dooren et al. (2002) in their study of the impact of preservice teachers' content knowledge on their evaluation of students' strategies for solving arithmetic and algebra word problems, offered further evidence that demonstrated the "influence of future teachers' content specific knowledge and skills and their pedagogical knowledge in a particular curricular domain on certain aspects of their future teaching in that domain" (p. 345). Furthermore, the work by the multiplicative conceptual field group (Harel & Confrey, 1994) and Saxe, Taylor, McIntosh, & Gearhart, (2005) provide some additional insights into ways students' difficulties in interpreting, operating and reasoning with multiplicative quantitative relationships (which includes reasoning and operating with fractions, rates, ratios and proportions) could be remediated.

(3) *What are the similarities and differences between the groups and subgroups?*

While a paired t-test for the common-group showed a significant difference between the group's two performances, further testing at the subgroup level indicated significant differences in performances with 5 of the 6 subgroups. For group G1, it can be plausibly inferred that there was a *significant improvement* given that (a) all ability changes were above the mean ability change, *and* (b) trends shown in Figures 3, 4 and 5 were all positive. In contrast, for group G2, the first two subgroups showed significant differences between their performances implying different scenarios for the different subgroups and as shown in Figures 9 to 11. For subgroup SG21, the significant difference means there was either performance improvement (positive gradients), or performance regression (negative gradients). Unlike the mixed positive/negative patterns of SG21, the significant difference in performances for SG22 can be interpreted as significant regression in performance, a possibility that is also numerically and visually supported by Figure 10. With subgroup SG23, there was no significant difference between the performances, and as shown in Figure 11, all ability estimates were below mean values suggesting that these students are potentially more at-risk of not meeting numeracy and mathematical competence standards. Whilst using ability change as an organising criterion of the data was useful, more information is needed to compile a more complete picture of student misconceptions such as those detailed and presented on individual kidmaps and actual written solution strategies.

On the basis of ability changes and graphs in Figures 3, 4, 5, 9, 10, and 11, it is suggested that students who are in subgroups SG13, SG22, and SG23 and regression cases of SG21 are potentially at-risk if *at least one* of the following three criteria is *not true*:

- *Significant difference in sub-group performance between the two tests;*
- *Ability change was less than average ability change; and*
- *Second estimate was lower than the second test's mean estimate.*

An interesting example is that of the top SG21 student (ID 111021, -0.09 ability change) with ability estimates above both tests' means - her kidmap (Figure 12) and error analysis suggested there were specific areas and general areas that she should specifically focus on in order to further enhance her NAMC levels by the third testing.

In summary, triangulating data from different sources such as the item-person map, kidmaps, and actual error responses provides substantiated empirical evidence to (1) inform recommendations for the cohort as a whole as well as (2) guide the development of customized remedial programs to specifically address each individual's identified misconceptions. Finally, findings from the case studies suggest that there was generally substantive sub-group improvement and/or substantive regression since the first diagnostic test in 5 of the 6 subgroups. The cohort and the common group found word problems the most difficult. Focussed and targeted remediation as suggested by Van Dooren et al (2002) for solving word problems flexibly by using *arithmetic* or *algebraic* multiple methods should form part of student teachers' repertoire of strategies. In the study reported here and like Schmidt's study of preservice teachers (1994) (cited in Van Dooren et al., 2002), the empirical evidence from the two diagnostic tests on student teachers' poor problem solving skills raises "*concerns about the ability of these preservice teachers' to support their future students in solving word problems.*" Similar concerns are also raised in this study inclusive of the wide-ranging content areas that these student teachers demonstrated persistent misconceptions with such as fractions, ratio, rates, proportions, measurement and probability.

Conclusions

The 3 main conclusions relate to (1) general improvements, (2) persistent misconceptions, and (2) potential at-risk students.

(1) General Improvements – With the apparent hierarchy of items, from the bottom end with basic computational problems using whole numbers through to visual representations of data, quantitative data expressed as whole numbers, fractions, and decimals, variously combined and applied in context, towards the middle with relational descriptions that are additive, and becoming increasingly descriptive and multiplicative towards the top end with the most difficult items, general improvements were more noticeable at the bottom and becoming less so towards the middle and top end. Students appear to cope better when relationships are described additively, performing relatively better as words become simple or non-existent (i.e. basic computation). In contrast, as relational descriptions become multiplicative, students tend to struggle to access the mathematics embodied in problem statements.

(2) Persistent Misconceptions - Firstly, students appear to have persistent misconceptions representing equivalent fractions, ordering, operating with, and applying fractions in context such as when used to measure probability, ratio and proportion. Secondly, findings suggest that students tend to struggle with probability problems particularly distinguishing between likely outcomes, favourable outcomes, probability value and expected values. Thirdly, there is evidence of weak or forgotten knowledge of basic geometric properties of similar triangles, quadrilaterals, rectangles and angle types. Fourthly, findings reveal that students have difficulties interpreting relational statements, transforming interpretations algebraically and generalising patterns. Fifthly, students tend to avoid items that involve the use of complex procedures and critical analysis and synthesis. Student errors demonstrate weak or no critical problem solving skills in relation to effectively interpreting and analysing given information, representing, organising and synthesizing information with relevant knowledge in order to identify appropriate procedures to generate correct responses. Findings also show that students struggle to distinguish between relevant and irrelevant measurements and synthesising information in visually presented data when the relational description is multiplicative and/or the diagram is complex, that is, geometric shapes are nested within other shapes.

(3) Potentially At-risk Students – Subgroup analysis of data suggest that possibly 3.5 of the 6 subgroups may be potentially at-risk. However, detailed analysis of kidmaps and error responses suggest differently. That is, there is sufficient empirical evidence indicating that each student in the common-group can greatly benefit by a thorough analysis of the information displayed by his or her two kidmaps.

Subsequently, there is an urgency to remediate individual misconceptions if students are to improve their more recent ability estimate.

The tests were to diagnose misconceptions. Findings suggest a range of specific ones and more general ones related to critical problem solving skills, requiring specific remediation and specific teaching within the current mathematics and mathematics education courses if future students are to be competent problems solvers and confident teachers of mathematics. Since student teachers are expected to teach numeracy and primary mathematics in the future, it is imperative that identified misconceptions are specifically remediated before exit. Aspiring to become quality and effective teachers of primary mathematics means being proficient problem solvers who are competent and numerate at mastery level with the primary and early secondary mathematics curricula

Implications

The nature and type of misconceptions identified over the two semesters imply that teacher educators need to (a) specifically address student teachers' misconceptions before they exit, (b) use the empirical error data as bases for designing activities to remediate content-specific mathematics misconceptions, and (c) utilise empirical error responses as a rich source of authentic data for pedagogical class discussions. Conclusions from the study further imply that critical problem solving skills should be explicitly and routinely taught as part of regular classroom teaching, not as an add-on component at the end of a topic. These suggestions have implications for:

(1) *the department of education* in terms of the need to (a) explicitly include the development of critical problem solving skills as an integral part of national curricula in all subject areas right throughout primary and secondary levels, and (b) provide continuing professional development for practicing teachers to share and exchange ideas on effective ways to improve classroom practices with the ultimate goals of *cultivating students' critical problem solving skills* and *improving primary and secondary students' numeracy and mathematical competence levels*;
(2) *providers of teacher education programs* in terms of the need to develop policies on minimum entry competencies for pre-service teachers and policies on monitoring and certifying primary student teachers' numeracy and mathematical competence before exit;
(3) *teacher educators* to use the empirical evidence gathered proactively and constructively to remediate preservice teachers' content knowledge and pedagogical content knowledge of the primary mathematics curriculum; and
(4) *individual student teachers* themselves to use their kidmaps diagnostically to guide their own self-study and to seek out specific assistance.

Ideally, critical problem solving should be an underlying process that underpins mathematics courses at all levels and mathematics education courses at tertiary level. Further, implications for content and curriculum educators at university include the need to collaborate more effectively in sharing ideas and resources for remediation in explicit ways to enhance both students' *content knowledge and critical problem solving skills*.

References

Adams, R. J., & Khoo, S. (1993). Quest: The interactive test analysis system. ACER.

Afamasaga-Fuata'i, K., (2004). Mathematics education in Samoa: From past and current situations to future directions. *Journal of Samoan Studies*, Volume 1,125-140.

Afamasaga-Fuata'i, K., (2003). Numeracy in Samoa: From Trends & Concerns to Strategies. A Paper Presented at the Samoa's Principal Conference, Department of Education, Samoa, EFKS Hall, January 28-30, 2003.

Afamasaga-Fuata'i, K., (2002a). A Samoan Perspective on Pacific Mathematics Education. *Pasefika Mathematics Education*. Proceedings of the 25th Annual Conference of the Mathematics Education Research Group of Australasia (MERGA-25). July 7 – 10, 2002 (pp. 1-13). University of Auckland, New Zealand.

Afamasaga-Fuata'i, K., (2002b). The State of Mathematics Teaching & Learning in Schools. A Paper Presented at the Science Faculty Mathematics Teachers Seminar/Workshop (August 26), Education Faculty Seminar Series (September 13), National University of Samoa; and to the Director of Education and Assistant Director Curriculum & Assessment, MESC (September 26).

Afamasaga-Fuata'i, K., (2001). New Challenges to Mathematics Education in Samoa. *Measina 2000 Volume 1* (pp. 90 - 97). The Institute of Samoan Studies.

Afamasaga-Fuata'i, K., (1998a). The Development of Pre-Degree Courses At the National University of Samoa. In Bridging Mathematics Network Eighth Annual Conference. Conference Proceedings. The University of South Queensland, Toowoomba, Australia. 43-47.

Afamasaga-Fuata'i, K., (1998b). Innovations in Primary Mathematics Teacher Education in Samoa. In S. Muralidhar, (ed). *A Study on Innovations in Mathematics Teacher Education in the Region.* A Report Prepared for the Education Department of the Human Resource Development Division of the Commonwealth Secretariat December 1998. 14–29.

Afamasaga-Fuata'i, K., (1997). Mathematics Education in Samoa. *Journal of the Faculty of Arts*. National University of Samoa. 94-102.

Aubrey, C. (1997). *Mathematics Teaching in the early years: an investigation into teachers' subject knowledge.* London, Washington D.C.: The Palmer Press.

Ball, D. L., Lubienski, S.T., & Mewborn, D. S. (2001) Research on Teaching Mathematics: The Unsolved Problem of Teachers' Mathematical Knowledge. In V. Richardson, (Ed.), *Handbook of Research on Teaching.* Fourth Edition. American Research Association.

Bobis, J., Mulligan, J., & Lowrie, T., 2004, *Mathematics for Children. Challenging children to think mathematically.* Pearson Prentice Hill, Australia.

Bond, T. & Fox, C. (2001). *Applying the Rasch Model Fundamental Measurement in the Human Sciences.* Lawrence Erlbaum Associates.

Burgess, T., & Bicknell, B. (2003). Subject matter knowledge in preservice teacher education. In L. Bragg, C. Campbell, G. Herbert, & J. Mousley, (Ed.), Mathematics Education Research: Innovation, Networking, & Opportunity. Proceedings of the 26th Annual Conference of the Mathematics Education Research Group of Australasia (pp. 176-182), Sydney, MERGA.

Center for Education (CFE). (2001). *Knowing and Learning Mathematics for Teaching.* Mathematics Teacher Preparation Content Workshop Program Steering Committee. Mathematical Sciences Education Board, National Research Council, USA.

Cockcroft, W. H. (1986). *Mathematics Counts.* London: Her Majesty's Stationery Office.

Darling-Hammond, L. (2000). Teacher quality and student achievement. A review of state policy evidence. *Education Policy Analysis Archives*, 8(1). Retrieved February 25, 2005 from http://epaa.asu.edu/epaa/v8n1.

Department of Education (DOE). (1995). *Education Strategies: 1995 – 2005.* Education Policy and Planning Development Project, Apia, Samoa.

Harel, G., & Confrey, J., (Eds.). (1994). *The development of multiplicative reasoning*. State University of New York Press, Albany, New York.

Huckstep, P., Rowland, T., & Thwaites, A. (2003). Primary teachers' mathematics content knowledge: what does it look like in the classroom?. Retrieved on February 16, 2005 from http://www.leeds.ac.uk/educol/documents/00002534.htm.

Jones, K., Mooney, C., & Harries, T. (2002) Trainee primary teachers' knowledge of geometry for teaching. Proceedings of the British Society for Research into Learning Mathematics, 22(2), 95-100. Retrieved on February 2, 2005 from http://eprints.soton.ac.uk/12596/.

Ma, L. (1999). *Knowing and teaching elementary mathematics.* Mahwah, NJ: Lawrence Erlbaum Associates.

Mays, H. (2005). Mathematical knowledge of some entrants to a pre-service education course. In M. Coupland, J, Anderson & T. Spencer (eds). *Making Mathematics Vital.* Proceedings of the

Twentieth Biennial *Conference* of the Australian Association of Mathematics Teachers (AAMT), (pp. 43-52). January 17 – 21, 2005. University of Technology, Sydney, Australia.

Ministry of Education, Sports and Culture (MESC). (2003a). A review of the secondary single curriculum. MESC.

Ministry of Education, Sports and Culture (MESC). (2003b). Primary mathematics curriculum. MESC.

Mullis, I., Martin, M., Gonzalez, E., Gregory, K., Garden, R., O'Connor, K., Chrostowski, S. & Smith, T. (2000). *TIMSS 1999 International Mathematics Report.* Retrieved on August 27, 2005 from http://timss.bc.edu/timss1999/math_achievement_report.html.

National Council of Teachers of Mathematics (NCTM). (2005). *2000 Principles and Standards.* Retrieved on August 28, 2005 from http://my.nctm.org/standards/document/index.htm.

National University of Samoa (NUS). (2004). University Calendar.

National University of Samoa (NUS). (2005). University Calendar.

Rasch, G. (1980). Probabilistic models for some intelligence and attainment test. (expanded version). Chicago: The University of Chicago Press.

Rowland, T., Barber, T., Heal, C., & Martyn, S. (2003). *Prospective primary teachers' mathematics subject knowledge: Substance and consequence.* Retrieved on February 24, 2005 from http://www.skima.maths-ed.org.uk/IoEupdateBSRLM1103.pdf.

Saxe, G. B., Taylor, E. V., McIntosh, C., & Gearhart, M., 2005, Representing fractions with standard notation: A developmental analysis. *Journal of Research in Mathematics Education,* 36(2): 137-157.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher* 15(2), pp.4-14.

Steen, L. A. (2001). Mathematics and Numeracy: Two Literacies, One Language. Retrieved on June 30, 2005 from http://www.stolaf.edu/people/steen/Papers/numeracy.html.

Thomas, C. S. (2002). *Data analysis knowledge among preservice elementary teachers*. Retrieved on February 16, 2005 from http://www.stat.auckland.ac.nz/ ~iase/publications/1/10_23_ct.pdf.

Trends in International Mathematics and Science Study (TIMSS). (2003). *Highlights From the Trends in International Mathematics and Science Study: TIMSS 2003*. Retrieved on February 2, 2003 from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005005.

Van Dooren, W., Verschaffel, L., & Onghena, P. (2002). The impact of preservice teachers' content knowledge on their evaluation of students' strategies for solving arithmetic and algebra word problems. *Journal of Research in Mathematics Education* 33(5), 319-351.