PERSONAL NUMERACY STANDARDS OF ENTRANTS TO TEACHER EDUCATION COURSES

Rosemary Callingham
Ian Hay

University of Tasmania, Launceston

Abstract

There is continuing concern about prospective teachers’ knowledge and understanding of mathematics in an increasingly technological world. The Australian Institute of Teaching and School Leadership (AITSL) (2011) has recognised the importance of teachers’ literacy and numeracy and has set a standard for pre-service teacher education courses against which universities are expected to provide evidence. As a result, universities are increasingly attempting to measure their pre-service teacher education students’ achievement in literacy and numeracy. This paper reports on the outcomes of a numeracy test which all students undertook in their first semester of the course.

Background

Teaching is a demanding, rigorous, exciting and rewarding profession, and initial teacher education is the foundation on which quality teaching is built...Learning to teach well is intellectually demanding, it takes time, and it needs Australia’s most able and well-prepared people. (Mackay, 2011, p. 1).

Mackay’s comment was made against a background of growing interest in teacher quality. The single greatest in-school influence or factor on students’ engagement, learning, and achievement is the quality of the teacher (Hattie, 2009). Those individuals wishing to become teachers must be perceived by that profession, the public, and the school community to be people who are able, intelligent and worthy (Rienstra, 2010). Teaching is such an important profession it needs to attract and hold the “best and brightest” into the teaching profession for the long term benefit of society (Lovat, 2006). This issue is not unique to Australia. The Organisation for Economic Co-operation and Development (OECD) has also argued that teachers matter and that attracting, developing, and retaining effective teachers was an important priority across 25 OCED countries (OECD, 2005).

The importance of teachers’ knowledge is recognized in the standards set by the Australian Institute for Teaching and School Leadership (AITSL) (2011) for entry into teacher education programs. Clauses 3.1 and 3.2 of the National Program Standards in AITSL (2011) state:

All entrants to initial teacher education will successfully demonstrate their capacity to engage effectively with a rigorous higher education program and to carry out the intellectual demands of teaching itself. To achieve this, it is expected that applicants’ levels of personal literacy and numeracy should be broadly equivalent to those of the top 30 percent of the population. (AITSL, 2011, p. 13)

Clause 3.2 states that:

If students are admitted with scores below these requirements the provider (university) will need to establish satisfactory additional arrangements to ensure that all students are supported to achieve the required standard before graduation (AITSL, 2011, p. 13).
In response to these standards and concerns about teachers’ background mathematical knowledge many universities have started to give competency tests in mathematics or numeracy to their pre-service teachers.

**Teachers’ Mathematical Knowledge**

In 1987, Shulman described the different types of knowledge that teachers need to operate successfully in the classroom. Content knowledge was one of seven knowledge types that he identified. His work has led to considerable activity in the mathematics education field that explored the nature of teachers’ knowledge. Hill and her colleagues developed measures of teachers “mathematical knowledge for teaching” (Hill, Schilling, & Ball, 2004). This knowledge went beyond that of content alone and addressed the more complex understanding of mathematics drawn on by teachers in order to develop mathematical understanding in their school students. Callingham et al (2011) measured a complex construct of teachers’ mathematical knowledge that included mathematics content knowledge, mathematics pedagogical content knowledge and beliefs about teaching mathematics. They showed that pre-service teachers espouse appropriate beliefs about teaching mathematics but found the pedagogical aspects of teaching mathematics more difficult than the mathematics content. There is growing evidence that mathematics content knowledge alone—being able to “do the maths”—is insufficient for effective teaching (Baumert et al, 2010; Blömeke, Suhl, Kaiser, & Döhrmann, 2012; Hill, Rowan, & Ball, 2005). Measures such as the number and level of mathematics courses taken during teacher education courses were not related to school students’ learning outcomes (Mewborn, 2001).

Moran (2008) noted that many dimensions of teacher effectiveness, especially those associated with successful practice in high needs schools, are not reliably indicated by tests of academic ability. Similarly, there is literature that argues that in terms of teacher preparation USA teacher entrance tests (Berliner, 2005; Nichols & Berliner, 2010) and grade point average (GPA) (D’Agostino & Powers, 2009) have limited value in predicting who is going to be an effective teacher, which has more to do with the person’s ability to relate to others, deal with multiple issues simultaneously, problem solve and adapt subject content knowledge to the learning needs of the students. It is acknowledged, however, that mathematics content knowledge is closely associated with effective mathematics pedagogy.

Being able to adapt subject knowledge, such as mathematics, to school students’ learning needs, is predicated on knowing appropriate mathematics. Concern about teachers’ understanding of mathematics is not new. In 1977, at the first meeting of the Mathematics Education Research Group of Australasia (MERGA), Brown (1977) indicated that entrants to primary teacher training had low levels of mathematics skill. One reason that he gave for this was the increasing number of women entering the profession. Such an explanation would not be accepted today. Mays (2005) found that a
small but important proportion of primary pre-service teachers held serious misconceptions about basic mathematical topics such as decimals. Lowrie, Logan, and Scriven, (2012) expressed concern about teachers’ understanding of geometry and measurement. Ma (1999) suggested that one factor in the superior performance of Chinese students in mathematics in international tests was that Chinese teachers had a deep understanding of relevant mathematical content whereas American teachers struggled with basic procedures in some instances.

Such a body of research raises questions about the nature of the mathematics courses taught during pre-service teacher education. These generally assume some basic knowledge of mathematics, and it is this mathematical knowledge of one cohort of pre-service teacher entrants at one university that is the subject of this report.

Methodology

In 2010, new undergraduate (BEd (primary), BEd (early childhood)) and postgraduate (MTeach) pre-service teacher education courses were introduced at the University of Tasmania. At the same time, off-campus provision was included through online delivery for every unit in the degrees. Because of concerns about the quality of entrants to teacher education, it was also decided to develop a test of competence in literacy and numeracy which all incoming pre-service teachers (PST) would undertake. Only the test of numeracy competence is reported here.

Test Preparation

All members of the mathematics education team at the time contributed to the test preparation. Examples of test items used at other universities or for other purposes were considered, as well as new items developed. Because a decision had been made for ease of large-scale delivery that the tests would be delivered online, and that this also had to be available for distance students in a cost effective way, the test conditions had to be unstandardised. Hence some item types could not be used. For example, items that only required computational skills were not used because under the non-standardised test conditions these become trivial. For similar reasons items that could be found directly on the web were not used. Items requiring open answers were either changed to a multiple choice format or not used.

A pool of about 60 items was developed that all addressed mathematics content knowledge; there was no attempt to measure any pedagogical content knowledge because the students were at too early a stage in their course. The purpose of this test was only to provide an indicator to staff and students of numeracy competence and it was felt that an automatically scored test would be a sufficiently good indicator without the additional workload attached to manual marking. Hence all items were of a multiple choice (MCQ) format. A final selection of 40 items was chosen across all areas of mathematics. The breakdown of mathematics content areas was Number and Algebra, 18 items;
Measurement and Geometry, 12 items; Statistics and Probability, 10 items. The test was delivered using Qualtrics survey software (www.qualtrics.com) which supports a wide variety of item types, including some novel interactive versions, but for auto-scoring MCQ was the easiest.

The most time consuming part of the item preparation was the development of graphics and specialist mathematical text. Many of the items were available only in hard copy and where diagrams were needed most had to be developed from scratch. In addition, to include a graphic, the picture or diagram has to be prepared exactly as needed in a .jpg format, imported into a specified Qualtrics graphics library and used from there. This procedure was also needed for fractions questions or anything that needed specialist mathematical symbols, because text editors do not deal well with mathematical text. Word files, for example, with mathematical equations prepared in Equation Editor lost this formatting when imported unless imported as a graphic. At an advanced mathematics level specialist programs are available but none of these translates easily to html or other file types used for web-based delivery. This issue does lead to additional expense and time when mathematics items are being prepared.

Test delivery
The test was attached to two first year units, Personal and Professional Numeracy, which were taught in the BEd and the MTeach degree. The content was identical but delivered through different units for administrative reasons. For the purposes of the test of numeracy competence, however, the two groups were considered to be a single cohort.

The test was delivered to PSTs via a unique email link created within the software that meant that the test could be taken only once by each PST. Along with this link went detailed instructions about taking the test, including advice about scrolling through the questions and the amount of time available. A two day “window” was provided to students and they could log on at any time that was suitable for them within this period. Having logged on, however, they had to complete the test in one sitting. Once they had logged off, the software was set so that they could not re-access the test, and it automatically shut down after four hours of inactivity. Students were given 60 minutes to complete the 40 questions presented, and a score of 75 percent (30/40) was considered an appropriate score to demonstrate competence.

Sample
A total of 371 PSTs in the BEd (primary and early childhood), (179 on-campus; 192 online) and 168 students in the postgraduate MTeach (124 on-campus; 44 online) undertook the test. In the graduate entry MTeach, the unit to which the test was attached was being studied by every PST, including those taking secondary methods in subjects other than mathematics. The rationale for this was the
expectation in the Australian Curriculum that numeracy is a cross-curriculum priority that all teachers are expected to support and develop. The MTeach group was, therefore, a very mixed one.

A further breakdown of courses provided some additional information because there were still some students in “teach-out” courses and the units in which the test was embedded were also available to students in other courses across the university as general studies units. This breakdown is shown in Table 1. The total number of students shown in this breakdown is slightly less than the number who took the test because information about the course was not available for every student. No other demographic information was easily available, such as prior mathematics experience.

Table 1
Breakdown of Students in Courses

<table>
<thead>
<tr>
<th>Course</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEd(ECE)</td>
<td>59</td>
<td>10.9</td>
</tr>
<tr>
<td>BEd(Prim)</td>
<td>197</td>
<td>36.5</td>
</tr>
<tr>
<td>BEd(old)</td>
<td>26</td>
<td>4.8</td>
</tr>
<tr>
<td>BEd(Ins) (Distance old)</td>
<td>80</td>
<td>14.8</td>
</tr>
<tr>
<td>MTeach</td>
<td>168</td>
<td>31.2</td>
</tr>
<tr>
<td>GenStuds</td>
<td>9</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Findings

Test Quality

The items were scored right/wrong. The data were analysed using Winsteps (Linacre, 2009) software and the simple dichotomous Rasch model (Rasch 1960) to provide evidence of the measurement properties of the test. The summary statistics of item fit to the model are shown in Table 2. The ideal value for infit and outfit is 1 with the usual standardised $z$ values of ±2 being accepted. In all instances the values obtained from the test analysis were acceptable. The test also provided a Cronbach alpha reliability of 0.83.

Table 2.
Fit statistics for Test of Numeracy Competence

<table>
<thead>
<tr>
<th>Items</th>
<th>Mean measure (Logits)</th>
<th>Error Infit mean square</th>
<th>Standardised Infit $z$</th>
<th>Outfit mean square</th>
<th>Standardised Outfit $z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_i = 40$</td>
<td>0.00</td>
<td>0.16</td>
<td>0.99</td>
<td>0.00</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Performance of Pre-service Teachers

On the basis of raw scores, the proportion of students gaining mastery scores (≥30) was 84.5 percent in the MTeach and 55.0 percent in the BEd.
The overall mean person ability measure obtained from the Rasch analysis was 1.62 logits. This mean value was well above the item mean, constrained to 0.00, indicating that overall the test was relatively easy for the PST cohort. There were, however, a number of other analyses of interest, including any difference between the on-campus PSTs and those who were studying off-campus through online delivery. For these analyses, the BEd group and the MTeach group were considered separately.

Independent sample $t$-tests were conducted to determine if there was any significant difference between the mean scores on the basis of mode of study, on-campus or online. Results are shown in Table 3.

<table>
<thead>
<tr>
<th>Course</th>
<th>Mean score On-campus (logit)</th>
<th>Std Dev On-campus</th>
<th>Mean score Off-campus (logit)</th>
<th>Std Dev Off-campus</th>
<th>$t$</th>
<th>df</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEd</td>
<td>1.10</td>
<td>0.71</td>
<td>1.11</td>
<td>0.70</td>
<td>.148</td>
<td>368</td>
<td>.882</td>
</tr>
<tr>
<td>MTeach</td>
<td>2.72</td>
<td>0.79</td>
<td>2.66</td>
<td>0.97</td>
<td>.426</td>
<td>166</td>
<td>.670</td>
</tr>
</tbody>
</table>

The findings suggested that there was no difference in performance on the Test of Numeracy Competence on the basis of mode of delivery at either undergraduate or postgraduate level.

There was interest in identifying any differences that might exist between undergraduate and postgraduate entry PST performances. For this analysis an additional interest was in identifying whether there were differences among the sub-groups undertaking the test within the BEd course. An independent samples $t$-test between undergraduate and postgraduate groups showed a statistically significant difference at the 0.05 level ($t = 22.96$, $df = 537$, $p = .000$) in favour of postgraduate entry students. To explore this finding further, a one-way ANOVA was carried out using a Bonferroni post-hoc analysis. The ANOVA showed no statistically significant difference among sub-groups within the undergraduate entry PSTs, but significant differences between the postgraduate entry group and every other sub-group. All of these analyses suggested that the postgraduate entry students behaved as a unique group. They appeared to have a different skill set from students who had not yet achieved a degree.

**Further Analyses**

Data were also available from the two other assignments undertaken in the *Personal and Professional Numeracy* course, which enabled further analyses to be undertaken to explore whether the differences identified from the Test of Numeracy Competence results were replicated in the other assignments. In addition, consideration was given to identifying any differences among the three outcome results available—the two assignments and the test—to identify any systematic variations among the
assessments. The two written assignments were very different from the test itself. The first required PSTs to “teach” some mathematical concept to another person and report on the results; the second expected critical analysis of a media article in terms of its numeracy use.

Results showed similar patterns to those of the test scores between On-campus and Online PSTs for the two assignments. There was no statistically significant difference between On-campus and Online PSTs with the exception of the second assignment for students enrolled in the BEd, which showed a statistically significant difference at the .05 level in favour of on-campus PSTs. This second assignment difference in the BEd group may have reflected additional support provided to on-campus students in choosing a particular media article, or tutors’ familiarity with the material when marking if on-campus students had sought extra help with the assignment.

Table 4.
Differences between Oncampus and Online Students on Assignments by Course

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Course</th>
<th>Mean score</th>
<th>t</th>
<th>df</th>
<th>p</th>
<th>Mean score</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignment 1</td>
<td>On-campus</td>
<td>57.7</td>
<td>1.48</td>
<td>369</td>
<td>.139</td>
<td>62.1</td>
<td>0.47</td>
<td>166</td>
<td>.143</td>
</tr>
<tr>
<td>Assignment 1</td>
<td>Online</td>
<td>58.7</td>
<td></td>
<td></td>
<td></td>
<td>59.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assignment 2</td>
<td>On-campus</td>
<td>60.7</td>
<td>2.27</td>
<td>369</td>
<td>.024*</td>
<td>62.1</td>
<td>0.27</td>
<td>166</td>
<td>.786</td>
</tr>
<tr>
<td>Assignment 2</td>
<td>Online</td>
<td>58.4</td>
<td></td>
<td></td>
<td></td>
<td>62.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Difference is significant at the .05 level

A significant difference was identified between the performances of students on both assignments when these were considered by undergraduate and postgraduate entry (Assignment 1: \( t = 4.2, \, df = 537, \, p = .000 \); Assignment 2: \( t = 3.0, \, df = 537, \, p = .003 \)). As with the numeracy test, an ANOVA with a Bonferroni post hoc test was carried out to identify where these differences were located. For Assignment 1, there were statistically significant differences between the MTeach and the new BEd students – BEd(ECE) and BEd(Prim) but not between the MTeach and the “old” BEd and BEd(Ins) students. For Assignment 2, the only statistically significant difference was between the MTeach and BEd(ECE) students.

These findings suggest that one reason for the observed difference between the MTeach cohort and other groups is likely to be because of skills acquired during a degree. Students in the “teachout” courses – BEd(old) and BEd(Ins) had been undertaking their course for some time, and were studying the unit as part of a completion plan, whereas the BEd(ECE) and BEd(Prim) were new to university study.

Comparisons between the Three Different Outcome Measures
A final set of analyses was undertaken to consider relationships among the different assessments: the Test of Numeracy Competence, Assignment 1 – Teaching and Assignment 2 – Media Analysis.
The first of these analyses was a correlation between the three assessment marks. This analysis indicated the associations among the marks on the basis of rank order. The correlations were moderate but were all highly statistically significant. Results are shown in Table 5.

This finding suggested that the three tasks were tapping the same domain but that students responded differently to each task, indicating that the three assessments provided associated but varied information about students’ numeracy.

Table 5
Correlation matrix for three tasks

<table>
<thead>
<tr>
<th></th>
<th>Ass 1 mark</th>
<th>Ass 2 mark</th>
<th>Test score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ass 1 mark</td>
<td>1</td>
<td>.436**</td>
<td>.228**</td>
</tr>
<tr>
<td>Ass 2 mark</td>
<td>.436**</td>
<td>1</td>
<td>.235**</td>
</tr>
<tr>
<td>Test score</td>
<td>.228**</td>
<td>.235**</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: ** indicates significance at the .01 level (two tailed)

A linear regression analysis was conducted to determine the extent to which results on any of the three tasks predicted the others, that is each of the three assessment tasks was tested as the dependent variable predicted by the other two tasks. In every case the two remaining tasks were highly significant predictors of the target outcome. In other words, success on Assessment task 1 was predicted by success on Assessment task 2 and by the Test score, success on Assessment task 2 was predicted by success on Assessment task 1 and by the Test score, and the Test score was predicted by success on both Assessment task 1 and Assessment task 2.

This finding indicates that all three variables were confounded, as different domains of students’ numeracy. The nature of the marking in the two assignments ruled out any attempt at scaling the three assessments together using Rasch techniques to determine the extent to which these three outcome measures were actually measuring the same construct.

Discussion

There were important differences between the two student cohorts, MTeach and BEd, and some indication that this may be due to prior university experience. There are implications here for pre-service teacher courses for assessment and course development.

Postgraduate entry PSTs are entering a new domain of study, If, however, graduate attributes desired by universities at the end of the program of study are taken seriously, these students already have a well-developed set of generic skills. The results from the Test of Numeracy Competence would suggest that this is the case. In particular, even in a new domain, it seems reasonable to expect higher order outcomes from graduate entry students. As a result of these findings, changes were made to the
standards expected for MTeach students in their assessments in the *Personal and Professional Numeracy* unit, even though the content delivered is identical to that of the undergraduate course.

The assessment program in the *Personal and Professional Numeracy* units appeared to measure students’ numeracy in different but related ways. There was little evidence of either inconsistent marking or cheating for any task, including the Test of Numeracy Competence, despite the non-standardised conditions and the unknown quantity of students studying online. Although not as rigorous as an invigilated test, the measure did provide some indication of students’ strengths and weaknesses in terms of numeracy. At this point in time, the test appears “good enough” for the purpose of identifying students who have significant deficits, or those for whom an additional mathematics curriculum unit would be beneficial.

It is worth noting the history of the test since 2010. After the first administration, the test was decoupled from the unit in response to the lecturer’s concerns about administration. It has become a stand-alone test, supported by an administrative assistant. It no longer has “teeth” because it does not contribute to a unit assessment, although successful students do receive a certificate. The number of students attempting the test has fallen, and enrolment in the follow up elective unit designed for those who had done poorly dropped from 120 in 2011 to 35 in 2012. No one “owns” the test, and there is serious consideration being given to re-embedding it in the first year mathematics unit, *Personal and Professional Numeracy*. This history suggests that pre-service teachers will not undertake monitoring or improvement of their personal numeracy unless there is some incentive. Such an incentive will need to be external to the university, such as not entering the teaching profession unless personal numeracy competence is shown, or internal, such as not graduating without a demonstration of personal numeracy competence.

Further discussion is needed among teacher education providers and policy makers about the efficacy of tests such as the Test of Numeracy Competence reported here. If the nature of teachers’ mathematical knowledge is complex, going beyond straightforward content knowledge, then it may be better to have some “hurdle” test at the end of the teacher education course, if this is desired. In addition, some consideration should be given to the different cohorts of pre-service teachers. Post-graduate entry students appear to have different levels of competency from those entering university for the first time. Further research is needed to identify how best to develop teachers’ mathematical competency in ways that are consistent with their intended career paths.

**References**


