The development and efficacy of an undergraduate numeracy assessment tool

Andrea Knowles
University of Otago
<andrea.knowles@otago.ac.nz>

Chris Linsell
University of Otago
<chris.linsell54@gmail.com>

Boris Baeumer
University of Otago
<boris.baeumer@otago.ac.nz>

Megan Anakin
University of Otago
<megan.anakin@otago.ac.nz>

This paper describes the development and efficacy of an online tool for assessing the numeracy of undergraduate students. The tool was designed to be easy to administer, provide immediate feedback to students on whether they had the required level of numeracy, and to be consistent with other measures of adult numeracy. When used with students taking a mathematics or statistics course, we found a significant correlation of $r = 0.45$ between their numeracy score and final mark in their enrolled course. Students who had a numeracy score less than our threshold had a 30.6% probability of failing their course, whereas students who had a numeracy score of at least our threshold had a probability of failing of only 8.0%.

We define *numeracy*, in an undergraduate university context, as having the knowledge, skills, and confidence to use mathematical tools in a range of disciplinary contexts. Tertiary educators may expect students entering their programmes to have the prerequisite numeracy to successfully complete their quantitative courses. However, student performance does not necessarily align with these expectations (Parsons, 2010). Students lacking numeracy skills are less likely to continue with a course when they are faced with difficulties with quantitative material (Matthews et al., 2009). Large scale numeracy assessment tools such as the Literacy and Numeracy Test for Initial Teacher Education (LANTITE) (Australian Council for Educational Research, 2016) and the Literacy and Numeracy for Adults Assessment Tool (LNAAT) (Tertiary Education Commission [TEC], 2008), have been developed to provide detailed feedback to individuals about their numeracy competency. Such tools are aimed at measuring the level of numeracy demonstrated by an individual rather than establishing if that person has a sufficient level of numeracy to be successful in a particular situation. Therefore, we sought to develop an *undergraduate numeracy assessment* (UNA) tool that could be used specifically for identifying if students have the prerequisite level of numeracy to enable them to be successful in their quantitative courses.

Background

The New Zealand Ministry of Education (2009) cautions us on using educational assessment as a sole means of assessing numeracy capability because high school students with high levels of success in formal qualifications may often present with low levels of numeracy. Since expectations from lecturers about students' mathematical competence does not necessarily align with numeracy entry levels (Parsons, 2010), high school leavers who are not identified by their teachers as having problems with numeracy may be identified subsequently in adulthood (Bynner & Parsons, 2006). Furthermore, the teaching of mathematical and statistical knowledge within courses of a quantitative nature does not necessarily link directly to a students' mathematical qualification (Gnaldi, 2006; Taylor et al., 1998).

2021. In Y. H. Leong, B. Kaur, B. H. Choy, J. B. W. Yeo, & S. L. Chin (Eds.), *Excellence in Mathematics Education: Foundations and Pathways (Proceedings of the 43rd annual conference of the Mathematics Education Research Group of Australasia)*, pp. 243-250. Singapore: MERGA.

We built upon descriptions of students' numeracy difficulties that were generally anecdotal or restricted to mathematical content (Taylor et al., 1998). We identified important underlying numeracy constructs for undergraduate students that included proportional reasoning, understanding of rational numbers, and multiplicative thinking (Galligan & Hobohm, 2015; Linsell & Anakin, 2012; Linsell et al., 2017). These constructs can be found in the large-scale numeracy assessment tools, such as the LANTITE and LNAAT. However, there are limitations when using these tools to assess the numeracy of undergraduate university students. First, students with high attainment take longer to answer questions than students with low attainment (TEC, 2017). Thus, students and education practitioners may feel that the time taken to complete a robust adaptive test across a six-step progression may be arduous or unnecessary. Second, assessment feedback provided to a student describes individual strategies, strengths, and knowledge (Hall & Zmood, 2019; TEC, 2008) but not a level of numeracy competency. Third, the New Zealand TEC has aligned numeracy progression benchmarks in the LNAAT to levels of the mathematics and statistics in the New Zealand Curriculum and to National Certificate of Educational Achievement (NCEA) standards for numeracy assessment (Thomas et al., 2014). A LNAAT score of 605 (Step 5) approximates to the NCEA numeracy standard as required for university entrance. However, further work is needed to confirm whether LNAAT is well aligned and represents numeracy competencies that adults require to be successful in society. Further study is also needed to investigate numeracy competency, to predict success in quantitative courses at the university level. One way to address the limitations of the large-scale assessments is to carefully frame assessment items. We define framing in three ways. First, assessment items need to be encased in appropriate and meaningful contexts (Mason et al., 2009). Second, items must allow for authentic user responses. Third, items must assess conceptual knowledge alongside procedural fluency (Hiebert & Carpenter, 1992). With well framed assessment items, educators may be able to establish a student's numeracy competence and predict their readiness to succeed in quantitative courses.

Development of Assessment Tool

Our aim was to produce a dependable assessment tool that was easy to administer, gave immediate feedback to students on whether they had the required level of numeracy, and that was consistent with other measures of adult numeracy. We decided that an online assessment would be necessary for facilitating marking and giving immediate feedback to students. We had previously used the LNAAT for investigating numeracy of undergraduates (Linsell & Anakin, 2012; Linsell et al., 2017). The LNAAT has been aligned with other measures of numeracy (Thomas et al., 2014) and we therefore decided to benchmark our tool against this.

We wanted to determine whether students had a particular level of numeracy, rather than measure what level of numeracy students had. Therefore, it was unnecessary to set questions that could be answered with lower levels of numeracy than our requirement. Our previous work (Linsell et al., 2017) had indicated that Step 6 of the LNAAT numeracy scale was necessary for success in undergraduate quantitative courses. Furthermore, detailed examination of the responses of students to the LNAAT numeracy questions suggested to us that a score of 740 was necessary, considerably higher than the 690 threshold for Step 6 (Casey & Knowles, 2018). Step 6 includes requirements for students to:

- solve addition and subtraction problems involving fractions, using partitioning strategies;

- solve multiplication or division problems with decimals, fractions and percentages, using partitioning strategies;
- use multiplication and division strategies to solve problems that involve proportions, ratios, and rates;
- know the sequences of integers, fractions, decimals and percentages, forwards and backwards, from any given number.

Our assessment consisted of 20 questions on the topics of fractions, decimals, ratios and proportions, and percentages. Students were required to answer five questions, which covered a range of sub-topics, in each topic.

Using a question format similar to that of the LNAAT, our assessment made use of meaningful contexts, previously unseen by the students, to determine whether the students could use mathematical tools to solve problems. This use of contexts ensured that conceptual knowledge (Hiebert & Carpenter, 1992), rather than just procedural knowledge, was required to solve the problems. Contexts were chosen that reflected the experiences of undergraduate students but that were not specific to any particular academic subject. Figure 1 shows an example of a question that requires students to make use of their knowledge of operating with fractions (this sample question is for illustrative purposes only and was not used in any assessments). The format for this question was multiple-answer, while other questions made use of numeric answers, fractions (both proper and mixed), multi-choice and drag-and-drop formats.

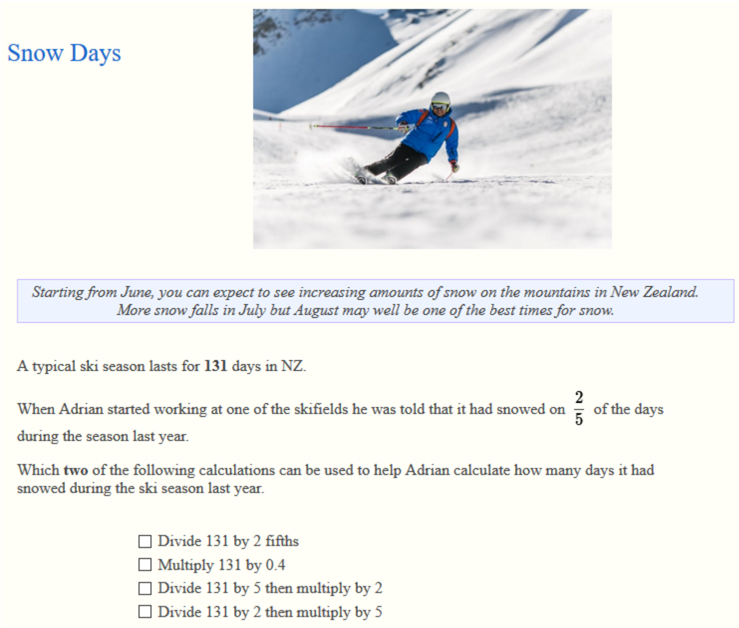

Figure 1. Snow Days question employing multiple answer format.

To ensure authenticity of students' work when sitting the assessment in computer laboratories, we designed the assessment to make it unlikely that nearby students would be answering the same question, or that one student's answer would be useful to another student sitting the assessment later. The assessment used a number of levels of randomisation. In addition to randomising the order of questions, contexts were randomised (e.g., for multiplying fractions the context of recipes was randomised with the context of student allowances) and pictures accompanying the questions were changed accordingly, names of people, objects, places and courses were randomised (e.g., quantity of flour to quantity of sugar), and the numbers used in each question were randomised. When randomising numbers, it was important to select values that did not alter the level of difficulty of the question (e.g., in the Snow Days question only the fractions 2/5, 2/10, 3/5, 3/10 were used and the number of snow days was randomised between 131 and 139 excluding 135).

The platform we used was adapted and further developed from an online system for assessing first-year university students of mathematics and statistics at the University of Otago. Question presentation was simplified, fractional and drag-and-drop answer formats were added, and the reporting of feedback expanded. The development of the question bank and its benchmarking took multiple iterations of setting the test, analysing answers (e.g., too easy, too hard, misleading etc.), improving questions, and adding questions. The test was first administered in MATH151 *General Mathematics*, and the success rate for questions was found to vary between 28% and 89%. Possible reasons for the range of difficulty were identified and questions were revised. Next, two parallel versions of the test were developed and used in EMAT198 *Essential Mathematics for Teaching*. Again, questions that were particularly easy or hard were identified and modified if necessary. Students taking EMAT198 (n = 67) also sat a LNAAT assessment, which was used for benchmarking. There was a strong correlation of r=0.45 (p<0.001) between EMAT198 students' scores on UNA and their LNAAT results (see Figure 2). Regression showed that a LNAAT score of 740 corresponded with a UNA score of 14.

We combined all questions (modified if necessary) from iterations 2 and 3 for use in STAT115 *Introduction to Biostatistics* in the second semester. For this fourth iteration the success rate for questions was found to vary between 49% and 92%. This variation is likely to be due to general gaps in students' conceptual knowledge rather than assessment item difficulty. In total, there were five iterations of question development and improvement to develop a test for use in the following year.

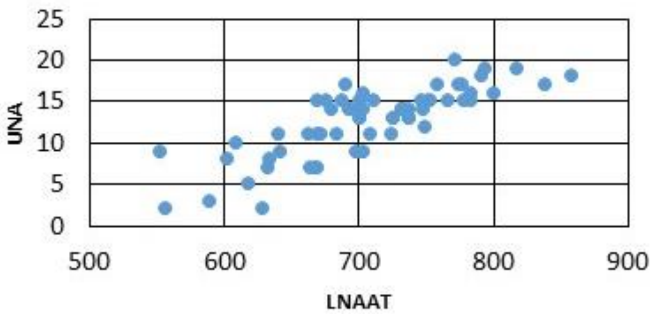

Figure 2. Correlation of UNA vs LNAAT assessment score in EMAT198 (n = 67)

Numeracy of Undergraduates

For students taking MATH151 *General Mathematics*, the UNA numeracy assessment was administered during tutorials in the third week of Semester 1 2019. The test was carried out under exam conditions. Of the 142 consenting students taking MATH151, 131 sat the UNA test, with the remaining 11 students not attending the tutorial in which the test was administered. Students scored between 1 and 20 on the 20-item test (M=13.3, SD=4.2) (see Figure 3). Sixty students (45.8%) scored less than our threshold score of 14 marks and 24 students (18.3%) scored less than 10 marks.

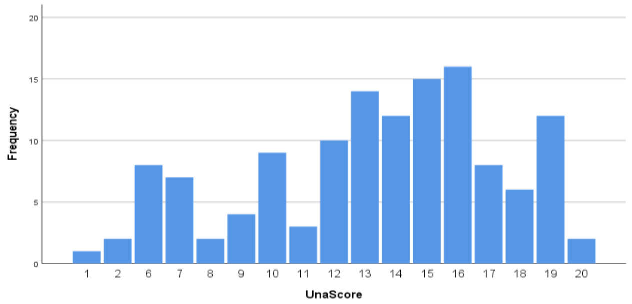

Figure 3. MATH151 distribution of students' scores (n = 131) on the 20 item UNA test

For students taking STAT115 *Introduction to Biostatistics*, the UNA numeracy assessment was completed by students in their own time in the first week of Semester 2 2019 and was unsupervised. However, students were encouraged to take the test to inform themselves of their numeracy needs and were given five marks towards their final grade in the course for taking the test. Of the 785 consenting students taking STAT115, 701 sat the UNA test, with the remaining 84 students opting not to do so, despite the inducements. Students scored between 0 and 20 on the 20-item test (M=14.9, SD=4.7) (see Figure 4). One hundred and eighty-eight students (26.8%) scored less than our threshold score of 14 marks and 90 students (12.8%) scored less than 10 marks.

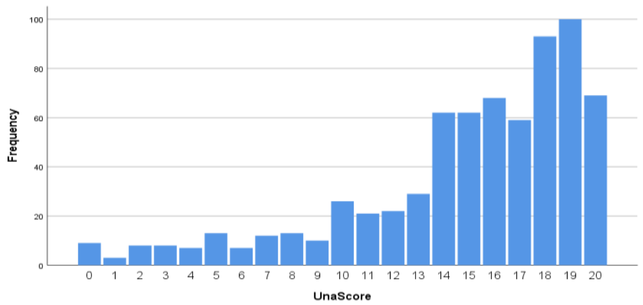

Figure 4. STAT115 distribution of students' scores (n = 701) on 20 item UNA test

As can be seen from Figures 3 and 4, the distribution of scores for STAT115 students sitting the test independently is rather different to that for MATH151 students sitting under exam conditions. Not only did a smaller proportion score less than our threshold score, but a much higher proportion scored 18 or more on the 20-item test. This difference could be accounted for by the variation in testing procedures rather than any differences between cohorts of students. The numeracy and attainment of the two cohorts is explored further in the next section.

Numeracy and Attainment

Overall, there was a strong and significant correlation of r=0.45 (p<0.001) between UNA numeracy score and the final mark of students in MATH151 and STAT115. Students who had a numeracy score less than our threshold of 14 marks had a 30.6% probability of failing their course, whereas students who had a numeracy score of at least our threshold had a probability of failing of only 8.0%. However, a much clearer picture is obtained by examining the attainment in MATH151 and STAT115 courses separately.

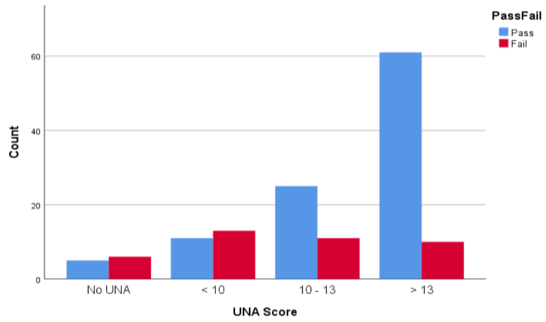

Figure 5. MATH151 students' attainment (n = 131) on course vs UNA score

For MATH151 there was a strong and significant correlation of r=0.41 (p<0.001) between UNA numeracy score and the final mark in the course. Of the students scoring less than 10 marks, 54% failed MATH151 (see Figure 5) with a mean score of 41% (M=41, SD=32). Similarly, 31% of students scoring 10 to 13 marks failed MATH151 with a mean score of 55% (M=55, SD=28). Only 14% of students scoring 14 or more marks failed MATH151 with a mean score of 71% (M=71, SD=26). It was interesting to note that the students who did not attend the tutorial and therefore did not sit the UNA test had a similar failure rate to those students who scored less than 10 marks. The failure rate (54%) for students scoring less than 10 marks or not sitting the UNA test was 3.9 times as high as the rate (14%) for students who achieved at least our threshold score of 14 marks.

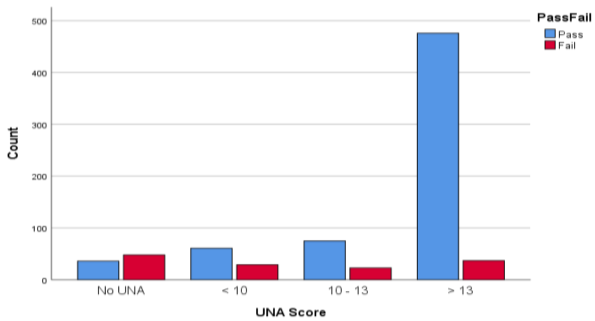

Figure 6. STAT115 students' attainment (n = 701) on course vs UNA score

For STAT115 there was a strong and significant correlation of r=0.46 (p<0.001) between UNA numeracy score and the final mark in the course. Of the students scoring less than 10 marks 32% failed STAT115 (see Figure 6) with a mean score of 56% (M=56, SD=17). Similarly, 24% of students scoring 10 to 13 marks failed STAT115 with a mean score of 62% (M=62, SD=20). Only 7% of students scoring 14 or more marks failed STAT115 with a mean score of 76% (M=76, SD=17). It was extremely interesting to note that the students who chose not to sit the UNA test had a failure rate even higher than those students who scored less than 10 marks. The failure rate (44%) for students scoring less than 10 marks or not sitting the UNA test was 6.3 times as high as the rate (7%) for students who achieved at least our threshold score of 14 marks.

Discussion and Conclusions

We used assessment items from UNA with students enrolled in EMAT198 to reliably calibrate using regression analysis against the LNAAT test to map a threshold score of 14 on UNA with the LNAAT adult progression at Step 6 and a score of 740. This score is higher than the 605 (Step 5) benchmark which corresponds to NCEA Level 1 numeracy assessment (Thomas et al., 2014) that is required for university entrance. Results from 832 students enrolled in mathematics and statistics courses within this study, using a UNA benchmark score of 14, indicate a significant correlation between UNA score and final examination result, demonstrating its suitability across a range of undergraduate courses with quantitative material. Furthermore, the cost and management of large-scale assessment (Brumwell et al., 2018; Hall & Zmood, 2019) can be mitigated by the provision of a well framed, 20 item assessment, which identifies a particular level of numeracy competence (Galligan & Hobohm, 2015) rather than a description of a learners' strategies, strengths, and knowledge (TEC, 2008) making it both time and financially advantageous. The importance of presenting questions in real-life contexts (Norton, 2006; Mason et al., 2009) is widely understood. Furthermore, UNA uses familiar adult contexts to assess the use of conceptual knowledge rather than procedural fluency (Hiebert & Carpenter, 1992).

In describing how the UNA was developed, we also demonstrated the efficacy of the UNA to identify whether students had a particular level of numeracy rather than measure what level of numeracy students had. This decision allows us to not only analyse the data but consider appropriate actions to take as a result (Blaich & Wise, 2011). The next steps are to examine how other disciplines, such as commerce, health sciences, and humanities, may use the UNA. Expanded use of the UNA may assist lecturers to question and examine their expectations about their students' mathematical competence and its alignment with numeracy entry levels (Parsons, 2010). Additionally, educators may find the UNA convenient for identifying the number of students who are likely to experience conceptual difficulties in their course. The UNA also provides an alternate source of numeracy feedback to educators that is consistent with other measures of adult numeracy such as the LNAAT. Educators may use results from the UNA to suggest that identified students seek numeracy support. To this end, students may be more likely to continue with the course and complete it successfully.

Further areas to address include: developing a larger bank of questions in the context of students' specific disciplines (e.g., nursing, pharmacy, business); and the process and potential issues (e.g., resources, time) in scaling up the use of UNA across an institution. We anticipate that educators will find the UNA useful for identifying if students have the prerequisite level of numeracy to enable them to be successful in their quantitative courses and that it will be a dependable assessment tool that is easy to administer, provides immediate feedback to students, and is consistent with other measures of adult numeracy.

Acknowledgements

The authors would like thank John Shanks for providing expertise with technical support. We are grateful for a University Teaching Development Grant from the University of Otago's Committee for the Advancement of Learning and Teaching for providing funding that supported the development of the UNA tool.

References

Australian Council for Educational Research. (2016). *Literacy and Numeracy Test for Initial Teacher Education Students: Assessment Framework.* https://teacheredtest.acer.edu.au/files/Literacy-and-Numeracy-Test-for-Initial-Teacher-Education-Students-Assessment-Framework.pdf

Blaich, C. & Wise, K. (2011). *From gathering to using assessment results: Lessons from the Wabash National Study*. [Report]. Urbana, IL: National Institute for Learning Outcomes Assessment. http://www.bu.edu/provost/files/2015/09/From-Gathering-to-Using-Assessment-Results_Lessons-from-the-Wabash-Study-C.-Blaich-K.-Wise1.pdf

Brumwell, S., Deller, F., & Hudak, L. (2018). The case for large-scale skills assessment. *Driving Academic Quality: Lessons from Ontario's Skills Assessment Projects*, 53-64.

Bynner, J., & Parsons, S. (2006). *New light on literacy and numeracy: Summary report*., London, UK: National Research and Development Centre for Adult Literacy and Numeracy. https://dera.ioe.ac.uk/22309/1/doc_3186.pdf

Casey, B., & Knowles, A. (2018). *Building Academic Numeracy: Improving undergraduate student outcomes with proactive numeracy learning support*. Wellington: Ako Aotearoa. https://ako.ac.nz/assets/Knowledge-centre/RHPF-s1701-Building-Academic-Numeracy/RESEARCH-REPORT-Building-Academic-Numeracy.pdf

Galligan, L., & Hobohm, C. (2015). Investigating students' academic numeracy in 1st level university courses. *Mathematics Education Research Journal, 27*(2), 129-145.

Gnaldi, M. (2006). The relationship between poor numerical abilities and subsequent difficulty in accumulating statistical knowledge. *Teaching Statistics, 28*(2), 49-53.

Hall, J., & Zmood, S. (2019). Australia's Literacy and Numeracy Test for Initial Teacher Education Students: Trends in Numeracy for Low-and High-Achieving Students. *Australian Journal of Teacher Education, 44*(10).

Hiebert, J., & Carpenter, T. P. (1992). Learning and Teaching with Understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York, NY: MacMillan.

Linsell, C., & Anakin, M. (2012). Diagnostic Assessment of Pre-Service Teachers' Mathematical Content Knowledge. *Mathematics Teacher Education and Development, 14*(2), 4-27.

Linsell, C., Casey, B., & Smith-Han, C. (2017). Numeracy of Undergraduate Business School Students. In A. Downton, S. Livy, & J. Hall (Eds.), *40 years on: We are still learning!, Proceedings of the 40th annual conference of the Mathematics Education Research Group of Australasia* (pp. 699). Melbourne, Australia: MERGA.

Mason, J., Stephens, M., & Watson, A. (2009). Appreciating mathematical structure for all. *Mathematics Education Research Journal, 21*(2), 10-32.

Matthews, K. E., Adams, P., & Goos, M. (2009). Putting it into perspective: mathematics in the undergraduate science curriculum. *International Journal of Mathematical Education in Science and Technology, 40*(7), 891–902.

Ministry of Education (March, 2009). *How well do qualifications predict literacy and numeracy* (Fact sheet). Ministry of Education, Wellington, New Zealand: Ministry of Education https://www.educationcounts.govt.nz/__data/assets/pdf_file/0017/36017/How-Well-do-Qualifications-Predict-Literacy-and-Numeracy-Fact-Sheet.pdf

Norton, S. (2006). Pedagogies for the engagement of girls in the learning of proportional reasoning through technology practice. *Mathematics Education Research Journal, 18*(3), 69-99.

Parsons, S. J. (2010). Mathematics Support in a University College and Research into Students' Experiences of Learning Mathematics and Statistics. In C. M. Marr & M. J. Grove (Eds.). *Responding to the mathematics problem: The implementation of institutional support mechanisms.* (pp. 59-53). Birmingham, UK: Maths, Stats & OR Network.

Taylor, J. A., Galligan, L., & Van Vuuren, P. (1998). *Perceived mathematical proficiencies required for first year study at the University of Southern Queensland*. [Report] Office of Preparatory and Continuing Studies (OPACS) Toowoomba, Australia: University of Southern Queensland. https://eprints.usq.edu.au/4591/2/Taylor_Galligan_VanVuuren_1998_PV.pdf

Tertiary Education Commission. (2008). *Learning progressions for adult numeracy. Tertiary Education Commission.* Wellington, New Zealand: Tertiary Education Commission. https://ako.ac.nz/assets/Knowledge-centre/ALNACC-Resources/Learning-progressions/Learning-progressions-numeracy.pdf

Tertiary Education Commission. (2017). *Guidelines for using the literacy and numeracy for adults assessment tool.* Wellington, New Zealand: Tertiary Education Commission. https://www.tec.govt.nz/assets/Forms-templates-and-guides/0869ebeb24/Guidelines-for-using-the-Literacy-and-Numeracy-for-Adults-Assessment-Tool.pdf

Thomas, G., Johnston, M., & Ward, J. (2014). *Alignment of Literacy and Numeracy Measures Research for the Tertiary Education Commission.* [Report]. Wellington, New Zealand: NZ Maths Technology. https://tec.govt.nz/assets/Reports/fa1264ae8f/Alignment-of-Literacy-and-Numeracy-Measures