A Preliminary Analysis of the Linguistic Complexity of Numeracy Skills Test Items for Pre Service Teachers

Lisa O'Keeffe
University of South Australia
lisa.okeeffe@unisa.edu.au

Language is frequently discussed as barrier to mathematics word problems. Hence this paper presents the initial findings of a linguistic analysis of numeracy skills test sample items. The theoretical perspective of multi-modal text analysis underpinned this study, in which data was extracted from the ten sample numeracy test items released by the Australian Council for Educational Research (ACER) in 2015. The initial data presented here maps the core content of these items to the Australian Curriculum and identifies the lexical density of each sample item. The findings indicate that the sample test items typically correlate to Year 6 of the national curriculum but each have high lexical density.

Skills testing for pre-service teachers (PSTs) was introduced in England in 2000. This was one of a number of education policy changes following on from the introduction of the national curriculum in 1988 (Machin & Vignoles, 2002). The intention of this testing was to 'raise the skills levels of the teaching profession' and hence "raise ... [students'] … professional standing and the profile of teachers and teaching" (Morris, 1999, p.3). However, the introduction of the testing was met with some criticism with many believing that their previous knowledge and qualifications were being undermined by the testing and referred to it as a "rite of intensification" (McNamara, Roberts, Tehmina & Brown, 2002, p.869). In the Australian context, the conversation around skills testing came to the forefront in 2015 following the *Action Now: Classroom Ready Teachers Report* (Teacher Education Ministerial Advisory Group (TEMAG) (Australian Government, Dept. of Education and Training, 2015). The report presents five key areas of suggested reform, the second of which is "Sophisticated and transparent selection for entry to teaching" (TEMAG, 2015). Within this section, the report makes direct reference to Initial Teacher Education (ITE) providers selecting entrants with a blend of skills, including literacy and numeracy. The report also recommends that Higher Education provides should use the proposed national testing to demonstrate "that all PSTs are within the top 30 per cent of the population in personal literacy and numeracy" (TEMAG, 2015, p.14). Following this, ACER announced that they had been tasked with trialling the national literacy and numeracy skills testing for PSTs. There appears to be a strong belief, by ACER in Australia and the Department for Education and Employment (1998) in the UK, that passing these 'tests' is an indicator of a pre-service teachers' literacy or numeracy capabilities as future professionals. Thus it is important to understand and hence investigate the content requirements of such assessments.

There is little doubt that the analysis of the mathematical structure of these test items would also be of benefit to the reader, however the author has deliberately focused on the initial linguistic analysis as this is the initial hurdle that faces the PSTs. Schleppegrell (2007) summarises recent research which integrates linguistic analysis and mathematics and highlights many of the challenges that language adds to mathematics education. She suggests that linguistic structures are often used differently in mathematics than in everyday language and this can present a challenge to students. This notion of linguistic challenge stems from the work of O'Halloran (2000) who argues that the grammatical structures and the multiple

2016. In White, B., Chinnappan, M. & Trenholm, S. (Eds.). Opening up mathematics education research (Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia), pp. 503–510. Adelaide: MERGA.

semiotic systems, such as everyday language, the mathematical register, symbols and visual representations, presented by mathematics require initial comprehension before one can begin to work through the mathematics. Added to this, Veel (1999) identifies a gap between the linguistic complexity, presented by teachers and text in mathematics compared with that used by the students. Veel (1999) makes particular reference to the high ratio of content words (also known as lexical words) to grammatical words, otherwise known as lexical density.

The findings of the *National Numeracy Review* (Council of Australian Governments, 2008) further support the author's decision to focus initially on linguistic analysis. The review referred to the work of Newman (1977), who found that students made seven categories of errors when solving word problems. These were reading, comprehension, transformation (to a mathematical representation), process skills, encoding, carelessness and lack of motivation. Newland's findings suggest that language and linguistic structure presents a considerable challenge to students. This is further supported by Zevenbergen (2001) who also identifies a number of literacy demands which students encounter when reading and interpreting mathematical text. She makes reference to the mathematical register versus everyday language, the lexical density in mathematical expression and also to the semantic complexity presented in the structure of word problems. This is reinforced by Martiniello (2008) who notes that often the most linguistically complex items were those that contain complicated grammatical structures that are central to comprehending the item or task. Martiniello also makes direct reference to multiple clauses (a clause is a grammatical structure that typically includes at least a subject noun phrase and a finite verb) embedded in item sentences and long nominal phrases and nominalisations.

In the context of this study, this literature all supports the case of basing the preliminary analysis of the mathematical text (the numeracy skills test sample items) on mathematical content analysis which integrates linguistic analysis. This is further supported, particularly in the context of a numeracy assessment, by the view of Aiken (1972), who suggested that students with higher reading abilities tend to do better than others, not because they are any better at mathematics but that they can read and better understand what is being asked of them. Therefore, this paper focuses on one aspect of the author's analysis of these sample questions and presents the initial findings from the linguistic analysis of the ten sample numeracy skills test items which were made available in July 2015. The actual pilot assessments, conducted in August and September of 2015, are not yet freely available and hence the ten sample questions are the best available measure for such an analysis.

Theoretical Framework

In their work on text structure and cohesion, Halliday and Hasan (1991) discuss the importance of cohesion in text and suggest analysis should be based on five key cohesive devices which are based on linguistic analysis. From a mathematics text perspective, the work of O'Halloran (2005) in particular highlights the need for utilising linguistic tools to better understand the structure of mathematics text. This view is supported by Morgan (1996) and O'Keeffe (2014) who both integrate language analysis in various forms into their work on mathematics text analysis.

O'Halloran's early work is based on that of Halliday (1973). Halliday's research provides the basis for much language analysis in different subject areas, focusing on the functional aspects of language (often referred to as systematic functional linguistics (SFL)). SFL, according to Halliday (1993) allows for the interpretation of the meaning of the language in the context of its function. That is, it examines how language is used, the purpose it serves, and how its form has evolved to serve its function. Drawing from and expanding on the work of Halliday (1973), O'Halloran (2005) highlights that discourse analysis of mathematical text requires the consideration of language, mathematical symbolism, and visual images. She discusses how the "meaning-making" process in mathematical discourse often integrates each of these semantic resources (language, mathematical symbols, and visual images) and suggests that translation between these resources (defined as intersemiosis) impacts on one's overall construction of meaning (O'Halloran, 2005). For example, a visual image in a mathematical text must be understood and interpreted in the context of the accompanying language and mathematical symbolism.

As noted by Halliday (1973), symbolism and images are often utilised to compensate for the inadequacy of language discourse to fully construe the intending meaning. Symbolism and visual image both add another dimension to mathematical text, in particular to support the making of connections and demonstrating relationships, and hence can often contribute to supporting the reader in concretizing abstract mathematical concepts (Hammill, 2010). This view is supported by Schleppegrell (2007) who agrees that ordinary language is often insufficient to construe the meaning required for mathematics. However, she also indicates that grammatical structures themselves such as dense phrases, nominalisations, logical connectors (words or phrases that connect two or more sentences/ideas), and verbs can all present a challenge for the mathematics student/reader. Halliday (1993) echoes this in regard to scientific English, adding that one of the potential sources of difficulty with scientific text is lexical density. Schleppegrell (2007) discusses how one would expect to find different clause/sentence structures in different registers and hence one can expect higher lexical density in academic text compared with informal text. She also notes how Halliday suggests we would expect informal English to have a lexical density of 2 on average and a school based text would typically have a lexical density of 6. To, Fan and Thomas (2013) add to this, noting that the higher the lexical density of a text the more difficult it is and suggest that a typical text would be expected to have a lexical density of between 3 and 6.

Research Design

This paper reports on the preliminary findings of an investigation of both the mathematical and linguistic complexity of the ACER 'numeracy skills test' sample items. A qualitative research design was utilised which concentrated on line by line item analysis. The initial analysis of the ten items, which will be referred to as *tasks* throughout the rest of this paper, was guided by two key research questions (RQ):

- RQ 1: What is the level of mathematical challenge (identified by Year Level correlation) presented in these questions?
- RQ 2: What is the level of initial linguistic challenge (identified by lexical density) presented in these questions?

O'Halloran's (2005, p. 160) multimodal model for mathematics discourse provides the overarching framework for the text analysis undertaken in this study. This framework is founded in Halliday's (1973) systemic functional grammar analysis and enables a simultaneous analysis of three key functioning properties of mathematics text: Language, Mathematical symbolism and Mathematical Visual-image. However, the initial analysis presented in this paper, an investigation of the lexical density, requires only the *grammatical* analysis of the *Content* section of this framework which is presented below.

Content	**Language**	**Mathematical Symbolism**	**Mathematical Visual Image**
	Grammar		
	Clause Complex, clause word group/phrase word	*Statements, Clause, expression, Components*	*Episode, Figure, Parts*

Figure 1: Extract from O'Halloran's (2005) Multimodal Model for Mathematics Discourse

Some elements of this framework are not relevant to include in the discussion of findings as they remain consistent throughout the text and hence do not add to the discussion. This is due to the nature of the text under analysis. For example the first three rows in O'Halloran's framework (omitted from Figure 1) refer to ideology, which is the perspective put forward by the text, genre and register; each of these remains consistent throughout the entire text. For example with ideology, the purpose of this text is for a non-specialist to read and complete a series of problems/questions. This tenor, which does not change, can be described as a hierarchical relationship of *specialist v's trainee*, where information is provided and unmodulated commands are issued by someone who is considered a specialist in the field. The grammatical analysis comprises three sections. The first explores the use of sentence structure (referred to as clause complex) and phrases (clause word group/phrase word) in the language presented. The second explores statements, sentences, expressions and components which combine to make-meaning which is reliant on mathematical symbolism. Finally, the third section explores the grammatical structure of visual images. This element of O'Halloran's (2005) framework is based on O'Toole's rank grammar, which was created for visual image in the Arts, and comprises three categorisations; episode, which is the defined as a combination of process which are represented such as the interplay of actions, figure which relates to the specific images or objects within an episode and parts are the elements that make up each figure.

Data Collection

The data was collected from the 10 sample numeracy skills test items provided by ACER (2015). The transcript of which comprises pages 9 to 16 of the ACER sample literacy and numeracy test for initial teacher education students document (available at ACER, 2015). Given the small sample size, all ten questions were included in the data collection and analysis process. RQ 1 required a mapping of the mathematical task presented in each task to the ACARA content descriptors. The mathematical concept which is central to each question was identified by the author and this was aligned to the ACARA content descriptors to allow for appropriate Year Level alignment. The initial data are intended to suggest the mathematical level and complexity of each task. Finally, O'Halloran's framework for content analysis was applied to address RQ 2. This included line by line analysis of the entire text, noting quantities for each element of the framework. This data was then combined to identify the lexical density of each of the ten tasks. Lexical density is defined as the ratio of lexical content words throughout each sentence to the number of independent clauses in each sentence.

Findings

The preliminary findings for the mathematical content analysis (extracts from each of the tasks are provided in Table 1) indicate that three of the tasks can be defined as one step tasks; that is they can be completed by the application of one skill in one step. Of the seven multi-step tasks five of these can be considered as having multiple solution paths. Table 1 presents an over view of this data along with the curriculum alignment recorded for each.

Table 1 Initial analysis of the sample ACER Tasks (Q1 -10)*

Extract from Question/Problem	Multiple Solution Paths	Key Concept	Curriculum Alignment
Q1 What percentage of the total operating expenditure on education in 2011–2012 was spent on the remaining aspects of the education budget?	×	Percentage is out of 100.	Year 6
Q2 For a 12-month 'Gym and Swim' membership, how much more does it cost to pay by monthly debit rather than upfront?	×	Extracting data from table. Value for money.	Year 6 Year 7
Q3 Which of these directions would lead Angela to Beckett Primary School? Select 'Yes' or 'No' for each set of directions.	✓	Read map and interpret directions.	Year 5
Q4 Which of these values is closest to the area of Beckett Primary School in square metres?	✓	Area and Grid reference. Portion of a quantity.	Year 5 Year 6
Q5 What is Alex's award for science? A High Distinction. B Distinction. C Credit. D Satisfactory	✓	Reasoning out of solution. Working out of solution: % of each (weighting)	Year 7/8 Year 8/9
Q6 These graphs show the percentage of children (11–13 years) playing sport at different times during school days and during non-school days. Select 'True' or 'False' for each statement.	×	Read and interpret line graph with percentages.	Year 6
Q7 What percentage of the total enrolments completed VET Certificates at the school in December 2011?	✓	Extracting data from table and express as a proportion.	Year 7
Q8 The table below contains some statements about the graph. Select 'True' or 'False' for each statement.	×	Interpreting bar graph with three data sets presented. Comparison of data.	Year 6
Q9 The weight of a box of stationery is 3.2 kilograms. What is the weight of 100 such boxes?	×	Multiplication of a decimal number by 100.	Year 6
Q10 About how many people lived in remote or very remote areas in Australia in 2011? A 11 000 B 44 000 C 110 000 D 440 000	✓	Calculating 2% of 22m	Year 7

* Calculators are allowed for Questions 1-8 only. For complete questions see: https://teacheredtest.acer.edu.au/files/Literacy_and_Numeracy_Test_for_Initial_Teacher_Education_students_-_Sample_Questions.pdf

At first glance (Table 1) the range across the Australian curriculum appears to be from Year 5 to Year 9, with a mode of Year 6. Of the questions aligned to Year 6, Q1 presents the weakest alignment, as other than understanding that percentage is out of 100 the procedural skills required are simply addition and subtraction of two digit numbers. Q5 also presents an interesting case, in regard to curriculum alignment. The question asks the reader to identify the grade band for the given student records. This question can be approached from two different perspectives: (a) by reasoning out the answer (which would align the interpretation expectation at Year 7) or (b) by mathematical calculation which would align to Year 9 content.

The second phase of data collection comprised an initial exploration of the linguistic complexity by exploring the lexical density presented in each of the tasks. Lexical density provides a "measure of the proportion of lexical items (i.e. nouns, verbs, adjectives and some adverbs)" in a text (Johansson, 2008, p. 61) defined by Halliday (1985, p. 64) as "the number of lexical items… as a proportion of the number of clauses". This analysis explores the lexical density, recognised as a significant factor of complexity of written language (Halliday, 1985), within and across each task of the sample of ten. As noted previously, the higher the lexical density, the higher the reading challenge presented. Tasks 1, 4, 5, 9 and 10 are singular modes of presentation; that is the task is presented as text only and hence does not require for translation between semiotic resources or includes images which do not including supporting full sentences (typically just labels as in Q 5). The lexical density of each statement and question clause in each task was calculated and is presented in Table 2.

Table 2 Lexical density (LD) for questions with single mode of presentation

Question/problem:	Q 1	Q 2**	Q 5**	Q 9	Q 10
Lexical Density	**8.3**	**10**	**10.2**	**4.5**	**6.75**

** Contains additional image/table however this does not contain any complete sentences.

Table 3 presents the data for the remaining five questions. These questions contain additional text which is included in either graphs, tables or figures which also has to be interpreted in the context of the task. Hence, it is relevant to consider the lexical density presented in these semiotic resources in conjunction with the main text in the task. This data is presented in Table 3 below. It is important to note that this analysis does not include an intersemiotic analysis of the interplay and potential resemiosis (which is an additional layer of "meaning-making" in which the integrated semiotic resources (inter-semiosis) are necessarily re-conceptualized or 're-semiotised' by the reader) between the modes of representation. Such analysis will be part of the next phase of data collection which is not discussed in this paper.

Table 3 Lexical density (LD) for questions with more than one mode of presentation

Question/problem:	Q 3	Q 4	Q 6**	Q 7	Q 8
LD of main text	5.33	6	7.67	9.2	4
LD of associated Table:	10.67	-	10.33	5	12
LD of associated Image	9	9	-	-	10
Lexical Density*	**8.14**	**7**	**9**	**8**	**9.14**

* Note: Lexical density if calculated on the raw scores for each statement included in the task not on the average across the semiotic resources.

As evident in Tables 2 and 3 the lexical density of the individual statements and questions within each task tends to be high, with one exception (Q 9). The range of lexical density across all ten test items is from 4.5 to 10.2, all of which can be considered high in the context of Halliday's (1985) works which suggests lexical density should be between 3 and

6. Examining the lexical density of the ten questions collectively gives a score of 12.8 (based on raw scores not averages across the ten questions), again this figure is well outside of the range of between 3 and 6 noted by Halliday (1985). Also, worth noting is that the use of symbolic notation is limited across all ten questions, with only 4 symbols being used in total: %, $, m, and $\overset{N}{\uparrow}$. Opportunities for symbolic notation are not fully utilised, for example in question 4 when referring to square metres the symbolic notation is not used.

Discussion and Conclusion

In summary the mathematics content that these numeracy sample skills test items are primarily based on is year 6 of the Australian Mathematics Curriculum. Both ACER (2015) and the TEMAG (2015, p.14) report make direct reference to how the test is intended to identify that PSTs are in the "top 30 per cent of the population in personal [literacy and] numeracy". However, a test that is pitched at Year 6 content level, which allows the use of a calculator for the majority of the tasks, does not appear to align with the Australian Institute for Teaching and School leadership (AITSL) suggestion as to what the top 30% might look like when they refer specific score ranges in Mathematics Methods, Accounting and Further Mathematics directly (AITSL, 2014). There is also the added complexity of context which is presented through the mathematical language and representation utilised in these test items. The ten test items are all word problems, as expected for numeracy test items and Martiniello (2008) confirms how important language skills are for interpreting and solving mathematical word problems. She identifies the correlations between linguistic complexity and low scores in mathematics tests. While her work is primarily focused on students with English as an additional language the central idea is still relevant and supports Newman's (1970) findings.

The initial linguistic analysis presented in the sample test items indicates a high lexical density across all ten questions and across the sample of test items as a whole. In light of the work of Halliday (1985), who notes that lexical density adds to the complexity of written problems and may inadvertently be an additional barrier for many, such lexical density levels would contribute to additional complexity in regard to readability of the test items. Such a view is reinforced by Zevenbergen (2000) who suggests that high lexical density can add to difficulties in translating between the mathematical and non-mathematical register and between different modes of representation (semiotic resources). This is supported by O'Halloran (2005) who suggests that translation between these resources impacts on one's overall construction of meaning; which in the case of assessment items could hinder ones interpretation of the problem and hence incorrect answers may be as a result of mistranslation rather than lack of mathematical knowledge of a concept.

The purpose of these test items is to determine that PSTs have, what those setting the test have determined to be, adequate numeracy skills. However, given that test is intended to be taken either before or after a literacy test and the high lexical density presented in this data it is fair to say that the literacy demands of these ten tasks appears to present a considerable challenge. It is important to note that additional analysis of the language and the intersemiotics across the modes of representation will add to this discussion.

References

Australian Council for Educational Research. (2015). *Literacy and Numeracy Test for Initial Teacher Education Students.* Retrieved from https://teacheredtest.acer.edu.au/files/Literacy_and_Numeracy_Test_for_Initial_Teacher Education_students_-_Sample_Questions.pdf.

Australian Government, Dept. of Education and Training. (2015). Action Now, Classroom Ready Teachers - Report of the Teacher Education Ministerial Advisory Group (TEMAG). Retrieved from https://www.studentsfirst.gov.au/teacher-quality.

Australian Institute for Teaching and School Leadership (AITSL). (2014). *Initial Teacher Education Programs Literacy and numeracy standards.* Retrieved from http://www.aitsl.edu.au/docs/default-source/initial-teacher-education-resources/ite-standards-literacy-numeracy-vic.pdf?sfvrsn=6.

Aiken, L. (1972). Language factors in learning mathematics. *Review of Educational Research*, 5(3), 359-385.

Council of Australian Governments. (2008). *National Numeracy Review report*. Canberra: Commonwealth of Australia. Retrieved from http://www.coag.gov.au/reports/docs/national_numeracy_review.pdf.

Department for Education and Employment (DfEE). (1998). *Teaching: High Status, High Standards. Requirements for Courses in Initial Teacher Training*, London: DfEE.

Halliday, M. (1973). *Explorations in the Functions of Language*. London: Edward Arnold.

Halliday, M. (1985). *Spoken and written language.* Victoria: Deakin University Press.

Halliday, M. (1993). On the language of physical science. In Halliday, M. & Martin, J. (Eds.), *Writing science: Literacy and discursive power* (pp.54–68). London: Falmer.

Halliday, M. & Hasan, R. (1991). *Language, Context, and Text: Aspects of Language in a Social-semiotic Perspective*. Oxford: Oxford University Press.

Hammill, L. (2010). The interplay of text, symbols, and graphics in mathematics education. Transformative Dialogues: *Teaching & Learning Journal*, 3(3), 1-8.

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. Lund University, *Department of Linguistics and Phonetics: Working Papers*, 53, 61–79.

Machin, S., & Vignoles, A. (2006). *Education Policy in the UK*. London: School of Economics.

Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, 78(2), 333-367.

Mc Namara, O., Roberts, L., Tehmina, N., & Brown, T. (2002). Rites of passage in initial teacher training: ritual, performance, ordeal and numeracy skills test. *British Educational Research Journal*, 28(6), 863–878.

Morgan, C. (1996). The language of mathematics: Towards a critical analysis of texts. *For the Learning of Mathematics*, 16(3), 2–10.

Morris, E. (1999). *Teacher Training Agency (TTA) National Skills Tests: A guide for trainee teachers*. London: TTA.

Newman, N. A. (1977). An analysis of sixth-grade pupils' errors on written mathematical tasks. *Victorian Institute of Educational Research Bulletin*, (39), 31-43.

O'Halloran, K. L. (2000). Classroom discourse in mathematics: A multisemiotic analysis. *Linguistics and Education*, 10(3), 359-388.

O'Halloran, K. (2005). *Mathematical discourse, language, symbolism and visual images*. London: Continuum.

O'Keeffe, L. (2014). Mathematics textbook analysis: Supporting the implementation of a new mathematics curriculum. In K. Jones, C. Bokhove, G. Howson & F. Lianghuo (Eds.), *Proceedings of the International Conference on Mathematics Textbook Research and Development (ICMT) 2014*, (pp. 365-370), Southampton, University of Southampton.

Schleppegrell, M. J. (2007). The linguistic challenges of mathematics teaching and learning: A research review. *Reading and Writing Quarterly*, 23(2), 139–159.

To, V., Fan, S., & Thomas, D.P. (2013). Lexical density and Readability: A case study of English Textbooks. *The International Journal of Language, Society and Culture*, 37(7), 61-71.

Veel, R. (1999). Language, knowledge and authority in school mathematics. In F. Christie (Ed.), *Pedagogy and the shaping of consciousness: Linguistic and social processes* (pp. 185–216). London: Continuum

Zevenbergen, R. (2000). "Cracking the code" of mathematics classrooms: School success as a function of linguistic, social, and cultural background. In Boaler, J. (Ed.), *Multiple perspectives on mathematics teaching and learning*. (pp. 201-224). Westport, CT: Ablex Publishing.

Zevenbergen, R. (2001). Mathematical literacy in the middle years. *Literacy Learning: the Middle Years*, 9(2), 21-28.