A Comparison of Text Difficulty in Systemic Assessment Using Lexile Theory

In South Africa, the use of systemic assessments is controversial, raising issues of fairness. However, an argument can be made that they yield valuable insights into the current levels of literacy and education in the country. National systemic assessments are used to gauge scholastic progress across schools and provinces, while international tests provide some measure of comparability across countries. This article investigates the Lexile framework as an educational tool for gauging the reading difficulty of texts used in national and international assessments conducted in South Africa. The results of the Lexile analyses showed that the reading difficulty of the Annual National Assessments (ANA), conducted in 2014, varied between grades and did not match the same grade level in the Progress in International Reading Literacy Study (PIRLS). We argue that by using the Lexile framework during the designing phase of assessment, and selecting texts that are level appropriate for the learners, the assessment process may be enhanced.

Keywords: Lexile, ANA, text complexity, text difficulty, PIRLS, text length

Introduction

Systemic assessment, as a component of educational system, exerts an influence on both teaching and literacy development. For teaching and assessment to be in alignment, we propose that three features need to be in place: the acknowledgment of the teacher as a professional, a model for teaching and learning, and a transparent assessment process. In the complex South African context with its rich linguistic landscape, the assessment of reading comprehension faces particular challenges and difficulties with regard to the above three features.

The use of standardised and systemic tests in the South African context is controversial. From a systemic perspective, such assessments should serve both as overall indication of skills gained, but in addition serve as diagnostic tools for improving literacy levels (Chisholm and Wildeman 2013; Department of Basic

Education 2011a). There are underlying assumptions that assessments can provide both accurate indications of reading comprehension levels and that as well such assessments can be used as diagnostic tools.

Some authors, such as le Cordeur (2014), emphasise that systemic type and standardised type tests such as the Annual National Assessment (ANA) reduce creativity, higher order thinking and leads to teaching to the test. le Cordeur (2014:154) advocates a "holistic approach to assessing the quality of teaching" and proposes that there should be both quality and variety in both external and internal assessments. However, systemic tests and standardised assessments are often seen by educational departments as the ideal way to gauge progress in the schools within their responsibility. Therefore it is likely that such assessments will continue to enjoy prestige. At the same time, teachers are required to devise their own assessments to evaluate the reading comprehension levels of learners for school based assessment, and also as assessment to inform teaching and learning.

Whether assessments are used internally, by teachers to assess reading comprehension or externally by monitoring agents, these assessments should have evidence of their validity in the South African context but also be comparable to international standards. In this article we investigate the use of a tool, the Lexile framework, for gauging the reading difficulty of a text. Such a tool has many applications and is useful for comparing both internal and external assessments. It also places the power in the hands of the teacher, so that teachers can gauge the difficulty of their texts, as well as the texts found in other assessments.

Problem Statement

Over the past decade, an emphasis has been placed on improving assessment practises in basic education, internationally as well as nationally. The South African government has prescribed a guiding principle in schools with the intention of enhancing assessment

planning. The Department of Basic Education (DBE) has two policies that outline the needs and implementation of assessment in schools, namely: *The National Protocol for Assessment for Grades R - 12* and *The Curriculum Assessment Policy Statement (CAPS)*. The National Protocol for Assessment for Grades R – 12 mainly focuses on management for school assessment (DBE 2012a). It also provides a standardised reporting and recording process for the framework of the CAPS. The CAPS framework contains programmes and promotion requirements for all the subjects (DBE 2012a). These policy guidelines have outlined frequency and periods within the teaching schedule when the assessment activities are to be conducted. The policies also include exemplars on how these assessment activities should implemented.

The ideal is that the abovementioned policies should inform decision making concerning systemic assessment which then informs policy (Khosa 2010). In this article we take South Africa's systemic assessments as a starting point. We then focus on the concept of assessing and monitoring learners' progress in schools and finally propose that good assessment practises have the potential to enhance and support the function for professional agency and development. As educational researchers we can support the development of professional agency by facilitating efforts to enrich classroom instructions and create opportunities for teachers to engage with and challenge educational structures.

When educators identify themselves as researchers, and as lifelong learners they are able to acquire additional knowledge and skills such as proficiency with psychometric tools. In doing so, educators will be able to further demonstrate the usefulness of concepts or tools. For example, the Lexile Framework is an assessment tool that provides educator's with skills that will support their role as a professional.

Literature Review

The Lexile framework was developed and designed by MetaMetrics with the aim of matching a reader's ability to the difficulty of a text (Wright and Linacre 1994). The Lexile measure is a tool that makes use of an algorithm to calculate the difficulty of a text and therefore provide some clarity about the particular text's readability by a reader at a particular level of current proficiency. The Lexile measure is used for international and national assessments.

The Lexile Framework is an increasingly popular teaching tool in its founding country, the United States of America (USA) and is used by learners, teachers, parents, librarians and school administrators in states and districts level (Copeland and Liben 2013). On a state level, a study had been done in Texas and Florida, where Lexile measures were used in systematic evaluation. The Texas Assessment Knowledge and Skills (TAKS) is a standard-based testing method in the state that evaluates performance and progress of learners (Texas Education Agency 2004). By using the Lexile measure an accurate level of the learner's readability is measured. It also serves as a predictor of how a learner is going to perform in the TAKS by linking the text to a Lexile measure. Subsequently, the Lexile measure assists in assessment development to allocate the most accurate text for a test. Alaska Elementary school also makes use of the Lexile measures for part of reading and testing programme that help learners to increase their reading ("Lexile Measures helps Alaska Elementary School" n.d.). At the beginning of each year, learners participate in a scholastic reading test in order to establish a baseline Lexile measure. The learners receive their Lexile measure and set a reading count for the year. This initiative motivates the learners to reach their reading counts by the end of the year and the progress is monitored by their teachers and librarians. In 2001 grade 3 learners scored 56% proficiency rate in reading. After three years of the school's reading initiative, the same learners scored 79% proficiency rate in

reading in the grade 6 assessment ("Lexile Measures helps Alaska Elementary School" n.d.). In the USA, the Lexile measure is being used in systematic evaluation and monitoring mainly to be able to measure a learner's readability which will indicate a specific reading level (Lennon and Burdick 2014). The information can be used to provide suitable reading materials for a learner that target their reading ability needs. Additionally, the use of Lexiles also indicate whether the learners are within a band that they are expected to be in.

Internationally, Korea utilises the Lexile framework in the English-Lectio Quotient (E-LQ) Assessment by reporting on a reader's Lexile measure (Fitzgerald et al. 2015). The United Kingdom also makes use of the framework by reporting a Lexile measure in the Granada Learning (GL) Assessments (Morgan 2009). The GL Assessment makes use of the Lexile framework to enhance the Progress in English (PIE) assessment by giving a detailed grading of the texts used by schools and publishers. A total of 400 000 UK learners complete the PIE tests each year and receive Lexile measures.

The Test of English as a Foreign Language (TOEFL) is a test that is written by learners from countries where English is not spoken in their daily life (Educational Testing Services (ETS) 2015). The TOEFL utilises the Lexile measure in three steps, firstly by measuring the learner's readability level in the beginning of the program. Secondly, by prescribing reading materials according to the text difficulty as measured by the Lexile framework. Thirdly, by providing a reader's Lexile measure at the end of a program to measure progress (ETS 2015).

The Lexile framework has attracted a demand from the international market as many organisations are incorporating the tool into their assessments. However, very little evidence of research done on the use of Lexile measures in Africa and to be more

particular South Africa. This paper explores the possibility of using the Lexile framework as a tool for professional development and language assessment in a developing context.

Looking at other large scale studies such as the Programme in International Student Assessment (PISA) and Progress in Reading Literacy Study (PIRLS) the assessment frameworks differs from the Lexile Framework slightly. The PIRLS framework comprises of two major reading purposes a) literary experience and b) acquire and use of information (Mullis, Martin, Kennedy, Trong and Sainsbury 2009). The PIRLS specifically considers the comprehension processes in reading which is relative to its study. On the other hand PISA take into account the Common European Framework of Reference which describes second and foreign language learning. The PISA framework examines reading from the perspective of reading for private use, reading for public use, reading for work and reading for education (PISA 2015). All three frameworks addresses the needs and demands of reading literacy but in different ways. The Lexile Framework takes into account the scientific Rasch Theory, whereas PISA and PIRLS reflects at reading from a theoretical perspective. This paper explores the readability of the text used in the ANA and the PIRLS by using the Lexile framework.

The Lexile Framework

The Lexile framework for reading was developed by MetaMetrics and uses an algorithm to match a potential reader with a text (Wright and Linacre 1994). The algorithm is used to measure the difficulty of a text, for example a book, article or news clipping, and give some indicator of that particular text's readability. The Lexile measure is based on word frequency, semantic difficulty and syntactic complexity. It offers a good prediction of how difficult text is to comprehend (Lennon and Burdick 2014). The framework is an educational tool designed to assist teachers and parents to

gauge the reading ability and overall comprehension of learners. The Lexile framework places the reader and the text on the same scale and thereby matches reading ability to text ("Lexile Measures in the Classroom" 2008). The process of aligning the two variables on the same scale involves the transformation of raw scores to log-odds units of both item difficulty and reader proficiency (Tennant and Conaghan, 2007).

By using this framework potential readers are able to match their Lexile measure with a vast repertoire of reading resources, organised in roughly three categories: too easy (not challenging); just right; and complex reading. Moreover, the Lexile framework allows teachers or parents to monitor a learner's reading progress throughout their schooling years. The Lexile framework supports teachers and parents in managing a learner's reading comprehension by suggesting appropriate reading materials which may challenge the learner and encourage reading progress ("Lexile Measures at Home" 2008).

The inner workings of the Lexile framework is based on reading experts, Flesch, Carroll and Bormuth's work and was originally funded by the National Institute for Child Health Development (NICHD). The framework utilizes the Rasch item response theory to determine a common scale on which to match readers and text. Within the Lexile framework the text difficulty and reader ability is aligned on one scale (Lennon and Burdick 2014).

There are limitations that should be noted. One of the limitations include that the Lexile framework cannot measure all kinds of text. It can, however, measure short stories, books, interviews, newspaper articles and plays. Some of the texts it cannot measure include songs, questions and poetry because these texts lack conventional punctuation (Lennon and Burdick 2014). Furthermore, it cannot measure a learner's writing or any non-prose content such as pictures.

Methodology

By applying the Lexile framework we were able to compare the sets of texts used for national and international investigations into reading proficiency. Texts were selected from national systemic tests and the Progress in International Reading Literacy Study (PIRLS). For the national systemic tests, texts were selected based on different general education and training grade bands and exit levels (see Figure 1 below). The reading comprehension sections of English Home Language tests[1] were adapted to meet the Lexile framework requirements. This included removing non-prose elements and saving the document in the correct format.

[Insert Figure 1 here]

Passages selected from the international study, PIRLS, were aimed at grade 4 level, but were administered to grade 5 learners in South Africa, judged to be appropriate for the developing context. A sister project, prePIRLS, was developed for countries where there was a need to gather information about learners who are in the process of learning to read, but not at the expected level, prescribed internationally. An easier assessment was needed to determine reading literacy performance (Mullis et al. 2009). prePIRLS was administered to South African grade 4 learners. Two released passages were selected from PIRLS and prePIRLS respectively to determine the Lexile measure and these were compared with the national systemic tests. The passages selected from PIRLS[2] and prePIRLS[3] included informational and literary passages aimed at assessing reading comprehension.

---

[1] The Annual National Assessment (ANA) Grade 3, 4, 5, 6 and 9 comprehension passages of 2014 were selected.
[2] PIRLS 2011 passages included Fly Eagle Fly and The Giant Tooth Mystery.
[3] prePIRLS 2011 passages included The Lonely Giraffe and Two Giant Dinosaurs.

The following features were compared: the *sentence length*; *word count*; *log word frequency*; and the *Lexile measure* which is represented by L, e.g. 800L. These statistics were computed via the Lexile website. Alongside sentence length, the program also calculated the total number of words in the text and the logarithm of the number of times a word appears in a text. Finally, the Lexile measure was calculated which shows the reading demand of the text as well as the semantic difficulty and syntactic complexity.

The above mentioned features were compared to the Curriculum and Assessment Policy Statement (CAPS) English Home Language for Grades 1 – 9 prescriptions or suggestions for teaching. By doing this comparison, we were able to determine how comparable the national systemic tests and international assessments are based on the Lexile framework. This comparison entailed some curriculum analysis by focussing on what learners should be able to do (and read) by each grade; the prescribed length of texts; and the factors literal comprehension, reorganisation, inferential, evaluation and appreciation (DBE 2012b). The following section discusses the analysis and results of the investigation.

Analysis and Results

All of the selected passages[4] were adapted to meet the Lexile framework requirements before submitting it on the Lexile website for analysis. The approach to analysis was to report on the word count and the Lexile measure for each of the passages. This section will start with the national systemic results before continuing to the international study passages.

For the intermediate phase (grades 4 – 6) and senior phase (grade 9) we needed to take into consideration what the curriculum stipulates for the length of the text per

---

[4] A total of seven passages were selected from the national systemic tests and PIRLS.

grade (see Table 1 below). There is a 50 word count increase per grade which should allow for greater text complexity.

[Insert Table 1 here]

It appears that most of the ANA selected passages exceeded the prescribed word count range of CAPS. Since there are no clear prescribed word count for grade 3 learners, it makes it difficult to gauge whether the number of words are too high. However, by comparing the 290 word count of the grade 3 passage to the 221 words of the grade 4 passage, it seems that the grade 3 passage was too lengthy. The grade 9 passage's word count was below the prescribed length, at 411 words. Figure 2 depicts the word count of the selected ANA passages.

[Insert Figure 2 here]

The grade 3 ANA passage had a word count of 290 and a Lexile measure calculated at 540L (see Figure 3). This measure is within the 330L to 700L range for the grade 3 band of the Lexile framework and seems to be appropriate for the learner cohort; this particular text was neither too easy nor too complex. The grade 4 ANA Englilsh comprehension passage's Lexile measure was estimated at 800L and had a word count of 221 which fits within the Lexile range of 445L – 810L for grade 4 but it exceeds the recommended word count based on the curriculum document. Nevertheless, the grade 4 Lexile measure was almost at the end of the 'just right' range.

However, some of the ANA passages may have been too complex, especially the grade 5 assessment. The grade 5 ANA passage had a word count of 254, and a

Lexile measure of 1030L thereby exceeding the range of 565 – 910L. Even though the grade 5 ANA passage was much shorter in length, it was more complex and used difficult concepts or terminology in the passage.

[Insert Figure 3 here]

In comparison to the grade 5, the English grade 6 ANA passage was calculated at 860L which is within an acceptable Lexile range (665L – 1000L) and had a word count of 313. The grade 6 passage did exceed the curriculum recommendation. Note that the grade 5 passage also exceeded the grade 6 Lexile band.

[Insert Figure 4 here]

An investigation of the ANA and the PIRLS and prePIRLS passages showed discrepancies in word count and Lexile measurement. The PIRLS literary and informational passages had a word count of 845 and 970 whereas the prePIRLS passages had a word count of 436 and 409 respectively (see Figure 4). These word counts are much higher than the ANA grade 4 (221) and grade 5 (254) word counts. The grade 5 ANA selected passage was a literary text with a word count of 254 and the PIRLS literary passage comprises 845 words. Based on these counts, it seems that learners in South Africa are expected to cope with a larger number of words and sentences when taking part in international assessments than when taking part in national assessments. This situation may account for some of the poor performance.

[Insert Figure 5 here]

The PIRLS literary and informational passages were calculated at 670L and 750L with word counts of 845 and 970, respectively. Both of these PIRLS passages are within the 450L – 810L range of the grade 4 band. The easier assessment, the prePIRLS, literary passage was calculated at 620L and the informational passage at 690L. These results fit well within the Lexile grade 4 band and are approximately 40L – 50L easier than the PIRLS passages.

There is approximately a 594 word count, and a 360L, difference between the grade 5 ANA and PIRLS literary passage. Based on the Lexile measurements (see Figures 3 and 5), the national systemic results had higher Lexile measures and were more complex when compared to the international assessment[5]. The next figure (Figure 6) depicts the mean sentence length for both assessments.

[Insert Figure 6 here]

The grade 5 ANA sentence length surpasses not only the grade 6 sentence length but also the international assessments' sentence length. The grade 4 ANA sentence length is also somewhat longer than the PIRLS and prePIRLS sentence lengths. The syntactic complexity of a text can be seen as a good predictor of the difficulty of a particular sentence (Lennon and Burdick 2014). As such, the mean sentence length of the grade 5 ANA test, may require the learners to store more information from the text in their short term memory compared to the other texts.

In conjunction with sentence length, the word frequency is used to calculate the Lexile measure (see Figure 7 below). The difficulty of the words is a continuum based

---

[5] Lexile measures and predefined grade bands are found in Appendix 1.

on the reader's exposure to it of which two categories of words exist, namely high-frequency and low-frequency words: the high-frequency words are the easiest and the low-frequency are seen as difficult (Monaco, Abbott and Kahana 2007). A text with high-frequency words means that the text possibly has a number of words that the readers may not have been exposed to.

[Insert Figure 7 here]

It is important to note that there will always be a difference in semantic difficulty between literary and informational texts. Based on the results from Figure 5 (above) the word frequency varies between 3.48, grade 6 ANA, to 3.76, PIRLS literary text. The average word frequencies of the national systemic and international assessments, match the word frequencies based on Hiebert's (2011) averages (see Table 2 below).

[Insert Table 2 here]

Recommendations

The application of the Lexile framework is useful for choosing texts. Checking the range of an existing text could assist teachers in the design of their own assessments. In an education environment such the South African one, the standards of reading can vary widely. Lexiles could support teachers in choosing texts for their learners. This would increase the chances that the texts are age appropriate and sufficiently challenging for learners. The list of books on the Lexile website also offers teachers the opportunity to choose from a wide variety of genres. Using Lexiles to choose or

evaluate texts could assist teachers with formative assessment design and gauging learner reading and comprehension levels would also contribute to summative assessment performance.

      With the goal of implementing an educational system that can compete internationally, ways need to be found to benchmark reading and therefore to choose texts which are diverse and well-targeted for age, interest and ability. Lexiles can contribute to benchmarking by providing a way to gauge the difficulty of a text. This of course should be used in conjunction with the teacher's own knowledge and skills. It also gives teachers the opportunity to develop their own skills in selecting materials and designing assessments. Essentially it is a tool which can be employed to enhance teaching practice and assessment design.

References

Chisholm L, Wildeman R. 2013. The politics of testing in South Africa. *Journal of Curriculum Studies* 45(1): 89-100.

Copeland M, Liben D. 2013. *Quick and Easy Access to Measures of text Complexity Using the Lexile Framework for Reading*. Metametrics, Inc.

DBE (Department of Basic Education). 2011a. *Annual national assessments: Report to basic education portfolio committee*. Pretoria, South Africa: Government Printing Works.

DBE (Department of Basic Education). 2011b. *Curriculum and Assessment Policy Statement Grades R – 3: English Home Language*. Pretoria, South Africa: Government Printing Works.

DBE (Department of Basic Education). 2012a.*The National Protocol for Assessment for Grades R -12*. Pretoria, South Africa: Government Printing Works.

DBE (Department of Basic Education). 2012b. *Intermediate Phase Grades 4 – 6: English Home Language*. Pretoria, South Africa: Government Printing Works.

ETS (Educational Testing Service). 2015. *Understanding TOEFL Primary: Reading and Listening Score Reports*. Retrieved 18 June 2016 from https://www.ets.org/s/toefl_primary/.../understand_reading_listening_score_reports.pdf

Fitzgerald J, Stenner AJ, Sanford-Moore EE, Koons H, Bowen K, Kim KH. 2015. The relationship of Korean students' age and years of English-as-a-foreign language (EFL) exposure with English-reading ability. *Reading Psychology* 36: 173-202.

Hiebert F. 2011. *Looking "Within" the Lexile for More Guidance: Word Frequency and Sentence Length*. Retrieved 10 June 2016 from http://textproject.org/library/frankly-freddy/looking-within-the-lexile-for-more-guidance-word-frequency-and-sentence-length/

Khosa G. 2010. *The South African national education evaluation system: What will make it work?* Occasional Paper 1: JET Education Services.

le Cordeur M. 2014. Constantly weighing the pig will not make it grow: do teachers teach assessment tests or the curriculum? *Perspectives in Education* 32(1): 142-157.

Lennon C, Burdick H. 2014. *The Lexile® Framework as an Approach for Reading Measurement and Success*. MetaMetrics, Inc.

Lexile Measure help Alaska Elementary School Foster strong reading habits, increase students reading proficiency. [Brochure]. n.d. Retrieved 22 July 2016 from www.lexile.com/m/cms_page_media/casestudy

Lexile Measures at Home [Brochure]. 2008. Retrieved 22 July 2016 from http://cdn.lexile.com/m/cms_page_media/135/Lexiles-at-Home_2.pdf

Lexile Measures in the Classroom [Brochure]. 2008. Retrieved 22 July 2016 from http://cdn.lexile.com/m/cms_page_media/135/Lexiles-in-the-Classroom.pdf

Monaco JD, Abbott LF, Kahana MJ. 2007. Lexico-semantic structure and the word-

frequency effect in recognition memory. *Learning and Memory* 14: 204–213.

Morgan D. 2009. *GL Assessment Personalises Literacy With Metametrics' Lexile Framework for Reading*. Press Release on 20 March 2009. Speech retrieved 14 August 2016 from https://www.gl-assessment.co.uk

Mullis IVS, Martin MO, Kennedy AM, Trong KL, Sainsbury M. 2009. *PIRLS 2011 Assessment Framework*. Chestnut Hill, MA: Boston College.

PISA (Programme for International Student Assessment). 2015. *Draft Reading Literacy Framework*. Retrieved 25 May 2016 from www.pisa.co.za/framework/readingliteracy

Tennant, A, Conaghan PG. 2007. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research* 57(8): 1358–1362.

Texas Education Agency. 2004. *Texas Assessment of Knowledge & Skills: Information Booklet. Exit Level English Language Arts Revised*. Texas Education Agency, Student Assessment Division: Texas.

Wright BD, Linacre JM. 1994. *The Rasch Model as a Foundation for the Lexile Framework* [White Paper]. Retrieved 13 May 2016, from https://metametricsinc.com/research-publications/rasch-model-foundation-lexile-framework/

Table 1. Length of Texts for Home Language

Table 2. Averages for Word Frequency and Sentence Length

Figure 1. Lexile Measures for National Systemic Tests and International Assessments