AN INVESTIGATION OF THE ENGLISH LANGUAGE DEMANDS OF MATHEMATICAL TEXTS ON DATA HANDLING USED IN INTERMEDIATE PHASE MATHEMATICS

A thesis submitted in fulfilment of the requirements for the degree of

MASTERS IN EDUCATION

of

RHODES UNIVERSITY

by

ANGELA PETA JONES

June 2022

ACKNOWLEDGMENTS

At the start of this journey, I realized just how much there was to learn and verse 5 from Proverbs 3 kept me focussed, 'Trust in the Lord with all your heart and lean not on your own understanding.'

It has been a journey of immersion, wonder, control, escape, frustration, encouragement.

A journey of a thousand steps, with many who cheered along the way, with words, deeds, messages of support, interest, gentle but firm nudges forward – for you I am grateful.

There are 3 who stand out, on whose shoulders I stood, sometimes firmly, sometimes precariously balanced. Propelling me forward academically, spiritually, lovingly, with their daily check-ins. To you 3, I am most grateful.

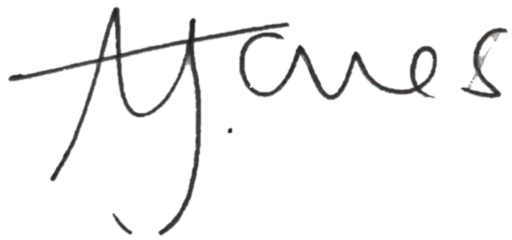
My supervisor, Dr. Pam Vale without whose great wisdom, quick thinking, patience, and equal love of coffee -I would not have enjoyed the journey as much as I did!

Her co-supervisor – Mr Mittens, my tuxedo cat, who spent most of this thesis next to my laptop, or in the box of articles, diligently checking up on my progress.

And Jack Jones – my amazing husband who never stopped listening to my academic rambling, who made me explain it all to him every step of the way – and who doesn't ever stop believing in me!

DECLARATION

I, Angela Jones, have read and understood the University's policy on plagiarism. This is my own work and, where I have drawn on the work of others, I have referenced appropriately. This work has not been submitted to fulfil the requirements of a degree at any other university.



22 June 2022

ABSTRACT

In the Intermediate Phase the majority of South African learners are transitioning from learning in their mother tongue to learning in English as well as from learning to read to reading to learn, and this is a major challenge. Textbooks are a key mediating artefact in the learning of mathematics and they present a challenge to the learner in terms of the language comprehension demands The data handling sections of mathematics textbooks are particularly dense in text. This is an important part of the mathematics curriculum as it is the beginning of statistical literacy learning. We need to be able to question, evaluate claims based on data, create arguments we can defend and use data meaningfully, it is thus crucial that learners acquire statistical literacy. This research sets out to examine the text in the data handling sections of four Intermediate Phase Mathematics book series in order to answer the following research question: What are the language comprehension demands of English mathematical texts on data handling that are used in South African Intermediate Phase Mathematics? The theories framing the study are Vygotsky's sociocultural theory and Cummins' second language acquisition theory. It is an interpretivist mixed method case study that takes the form of a document analysis.. The findings indicate that many units in the books analysed have a higher readability level than the grade level and will thus present a challenge to learners in terms of their ability to access the mathematical content. An analysis of the linguistic complexity revealed that the features contributing most to the complexity of the texts included words with seven or more letters, prepositional phrases, infinitives, complex verbs and complex/compound sentences. An examination of the non-textual elements revealed that most of them are accurate, connected, concise, contextual and these add to the comprehensibility although there were a few which could be possibly be distractors. It is hoped that the empirical findings of this study, will sensitise educators and publishers involved with the design of textbooks and workbooks to the type of language currently found and that they might give attention to the needs of English language learners when developing these texts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND CONTEXT

This study explores the language comprehension demands of English mathematical texts on data handling, in three Intermediate Phase textbook series and one workbook series. Although a small section of the curriculum, the area of data handling is of great importance. This content area is the branch of mathematics where tools are used to gather, organise and represent data. Large amounts of information confront people on a daily basis. As critical humans, we need to be able to make sense of this data. We need to be able to question, evaluate claims based on data, create arguments we can defend and use data meaningfully. It is crucial that learners acquire statistical literacy, and the textbooks are a key mediating artefact in the teaching and learning of this. For this reason I investigate the complexity of the language used in the mathematical texts and the non-textual elements used in the textbooks.

The study focusses on the South African context, where English, although not the mother tongue of the majority of children, is by Grade 4 the language of learning and teaching for 90% (Spaull, 2016) of learners. I draw on Vygotsky's (1986) theory of learning and his emphasis on the importance of language and Jim Cummins' (1981) language acquisition theory as the theoretical framing for the study.

In this chapter I describe South Africa's performance in mathematics and literacy in international comparative studies in order to start to outline the South African education context. Thereafter, to expand on this picture of the South African context, I discuss the challenges that English language learners face in transitioning from Foundation Phase [FP] (Grades R – 3) to the Intermediate Phase [IP] (Grades 4-6). The IP is the focus of this study. Having established the educational context, I then move on to argue for the focus on textbooks and data handling as the particular focus of this study. After providing this rationale for the study, I present the research questions and an outline of the thesis.

1.1 The South African Context

International and local systemic evaluations of education in South Africa indicate that South Africa needs to pay attention to the state of its education system. Both the Progress in International Reading Literacy Study (PIRLS) and the local Department of Basic Education (DBE) systemic evaluations, the Annual National Assessments, found the literacy attainment of South African learners poor (Mullis et al., 2007).

The Trends in International Mathematics and Science Study (TIMMS) 2019 is the seventh assessment cycle of mathematics and science conducted at the fourth and eighth grades in 64 countries. According to the TIMMS 2015 results, Grade 4 South African children placed 48th out of 49 countries in Mathematics. South African Grade 4 children scored 376 points, which is significantly lower than the centrepoint of 500 (Mullis et al., 2016). In 2019, South Africa assessed fifth and ninth grade students to better match their curricula to the TIMSS levels and to maintain trend measurement. The fourth grade mathematics assessment included three content areas—number, which included pre-algebra (50%); measurement and geometry (30%); and data (20%). The majority of TIMSS 2019 mathematics items assessed students' application and reasoning skills (Mullis et al., 2020). South African Grade 5 learners' average was 374, two points less than 2015, again much lower than the TIMMS scale centerpoint of 500 points (Mullis et al., 2020). These results indicate that South African Grade 5 learners were placed 56th out of 58 countries (Mullis et al., 2020).

There are many forms of literacy, the one this research focusses on is reading literacy – the ability to read text, make sense of it and action what is required. Reading literacy is defined by Mullis et al. (2009) as "the ability to understand and use those written language forms required by society and/or valued by the individual" (p. 11). The PIRLS assessment, administered by the International Association for the Evaluation of Educational Achievement [IEA], is focussed on assessing the reading literacy of students in their fourth year of schooling (Mullis et al., 2009).

It is acknowledged that in the classroom the teacher is able to mediate the use of the textbooks and workbooks. The more linguistically complex the text, the more that this is a requirement for learners to be able to make sense of the work presented. During COVID-19 lockdowns children would have been working more independently as they were either not attending school and were doing their work away from the teacher, or they were attending on a rotational schedule. In these times the reading literacy of the learners would have been a determining factor as to whether they would be able to make sense of the content in the textbooks and workbooks if they were working independently from them. There is evidence that this was the case in Makhanda schools. Vale and Graven (under review) conducted research with Intermediate Phase mathematics educators about their strategies to continue mathematics teaching and learning during the COVID-19 lockdowns. Twenty-one of the 25 teacher participants mentioned issuing extra homework, either through work packs, the DBE workbooks or textbooks, that learners needed to work on independent of the teacher. The language demands presented by these texts are important to consider when selecting work for learners to do independently as if the reading literacy of the learners is insufficient to comprehend the text, they will not be able to access the mathematical content presented. As a sample, this research examines the language demands of data handling texts from two textbooks from the national catalogue (DBE, 2021), the national workbooks issued by the DBE and a free online textbook promoted as a resource during COVID-19.

The PIRLS achievement scale summarizes fourth grade students' performance answering questions designed to measure their reading comprehension across two overarching purposes for reading—literary and informational purposes, as well as a range of comprehension processes. The results in 2006 and 2011 ranked South African learners as last (Robertson, 2017) and in 2016 South African learners were again placed at the bottom of 50 countries.

The Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) is a consortium of education ministries, policy-makers and researchers who, in conjunction with UNESCO's International Institute for Educational Planning (IIEP), aims to improve the research capacity and technical skills of educational planners (Moloi & Strauss, 2005). By generating information from school surveys SACMEQ enables decision-makers to monitor general conditions of schooling and the quality of basic education. SACMEQ III, was conducted in 2007, with South Africa being one of 14 participating countries. For the purposes of SACMEQ, reading literacy was defined as: "the ability to understand and use those written language forms required by society and/or valued by the individual" (Ross et al. 2005, p. 74), while mathematics literacy was defined as "the capacity to understand and apply mathematical procedures and make related judgements as an individual and as a member of the wider society" (p. 78).The SACMEQ III showed that South Africa ranked 10th out of fourteen education systems for reading, and 8th for mathematics, performing below Tanzania, Kenya and Swaziland, which are less resourced countries (Spaull, 2013).

What these systemic evaluations point out is that learners are struggling with both reading literacy and numeracy. When considering the textbooks used for mathematics teaching and learning, this is significant as an inability to comprehend the text independently would mean that learners are unable to access the mathematical content in order to develop their mathematical knowledge unless the teacher mediates the text. If the textbook is a key mediating artefact in mathematics teaching and learning, the comprehensibility of the textbook is important to understand.

1.2 The transition to the Intermediate Phase

This study focuses on the Intermediate Phase (Grades 4-6). The choice to focus on this phase is because the transition to this phase is challenging for many learners. For this study, it is particularly important to consider the transition to a new language of learning and teaching that most learners in South Africa experience at this point in their education.

1.2.1 The four transitions

According to Sibanda (2016), in South Africa in Grade 4 the learners experience four significant transitions from the Foundation Phase (FP) to the Intermediate Phase (IP). The first transition is from learning through the medium of their home language in the FP, to learning through the medium of either English or Afrikaans in the IP. In the FP, most learners are taught in their home language, but in the IP English or Afrikaans are used as the language of learning and teaching for all subjects, including Mathematics.

The second transition is from reading descriptive, story-like texts with language close to the ordinary, everyday language of social interaction in the FP, to reading texts with more content-dense vocabulary in Grade 4 (Chall et al., 1990). Adams (2003) suggests that mathematical reading involves both the knowledge of the language of mathematics and also linguistic comprehension skills. This is even more important in the IP, as the language of mathematics becomes more challenging. Österholm (2006) suggests that mathematical texts are conceived as compact, technical and complex, and that it is even more difficult to comprehend this text when it contains symbols. According to Vilenius-Tuohimaa et al. (2008), Grade 4 learners' reading comprehension was connected strongly to their mathematical word problem performance.

The third transition is the shift is in comprehending from 'learning to read' to 'reading to learn' (DBE, 2008). Learners in the FP are developing the skill of reading but when they get to Grade 4 they are expected to know how to read and to read different content subjects, in English, and learn from what they read. According to Setati (2016) some skills like general reading skills have been known to make cross-linguistic transfers, except for vocabulary between English and isiXhosa because of their orthographic distance. According to Cummins (1979) academic and cognitive bilingualism can only be achieved on the basis of sufficiently developed first language skills. The Curriculum and Assessment Policy Statement [CAPS] (DBE, 2011) advocates' mother tongue education in the FP and the teaching of a First Additional Language

(FAL) from Grade 1. It is expected that the skills developed in the home language during the FP will transfer to the new language of learning and teaching in the IP.

The fourth transition is the movement from more concrete thinking in the FP to more abstract thinking in IP, since mathematical abstraction is particularly critical for learner progress and attainment in the IP (Sibanda, 2016).

1.2.2 The transition to learning in English

The right of all children "to receive education in the official language or languages of their choice" is upheld in the South African Constitution (Constitution of the Republic of South Africa, 1996, Section 29(2)). The benefits of home language instruction as well as bilingual education are emphasized in the national Language in Education Policy (Department of Education, 1997). This policy allows South African schools to choose their language of learning and teaching [LoLT] and makes provision for a bilingual approach by being supportive of language practices like code-switching. In Foundation Phase, most learners are taught in their home language and learn an additional language, often English. In the Intermediate Phase the language of learning and teaching becomes either English or Afrikaans, which learners have often only learned at an additional language level until that point. In this study, it is the transition to learning in English that is of concern as English books will be analysed. The current curriculum (DBE, 2011) supports teaching mathematics through the medium of the home language until the end of the FP. However, beyond Grade 3 it does not support teaching through the medium of the home language (where it is not English or Afrikaans). This is exemplified by the DBE's distribution of workbooks exclusively in English or Afrikaans from Grade 4 onwards.

The rationale behind this policy is that children should find the transition into literacy in English easier, if they are first literate in their home language, as supported by empirical evidence from South Africa (Taylor & Von Fintel, 2016). Cummins'(1979) 'developmental interdependence' hypothesis proposes that developing proficiency in a second language is to some extent a function of the type of competence already developed in the home language at the time of exposure to the second language is the basis of this rationale. The 'threshold hypothesis' suggests that there may be thresholds of language proficiency that the bilingual child needs to reach in order to avoid cognitive difficulties and allow beneficial characteristics to impact cognitive and academic functioning (Cummins, 1979). According to Cummins (2000) the learner should achieve requisite thresholds in both languages for cross-linguistic transfer to occur, but learners should cross a certain threshold in their second language proficiency for this crosslinguistic transfer to happen. Considering the performance on the international benchmarking studies reviewed in section 1.1, it is questionable whether learners are crossing this threshold prior to making the transition to learning in English.

Spaull (2016) notes that in 2013, the percentage of learners in FP learning in English was 23%, while in Grade 4 this percentage was as high as 90%. The start of the Intermediate Phase is an area of significant concern since it is the start of most learners' educational journey in English and learners enter the Intermediate Phase [IP] with vastly different levels of language proficiency. According to Robertson and Graven (2015), almost 80% of these Grade 4 learners have the double burden of mastering their LoLT, as well as using it to gain access to mathematics through this LoLT. This sets up a significant challenge for these learners as they move through the IP grades where they encounter an expanded range of curriculum areas and academic texts.

There are major implications for access to learning when there is dominance of a single language of power as the medium for education (Robertson & Graven, 2020). Especially when those who do not speak this language at home are often already marginalized in terms of social and economic access. Of particular interest to this study is the challenge encountered by learners who are working with texts in English where this is their additional language. According to Setati (2014) language engenders access to mathematical concepts and thus learning in an unfamiliar language would compromise that access to mathematical concepts.

English language proficiency is essential in the English Mathematics classroom, especially since there is greater emphasis on students' discursive skills in the mathematics classroom, as more student centred pedagogies have taken over from traditional teacher-centred pedagogies in mathematics education (Robertson & Graven, 2019). More explicit attention is necessary in assisting students in developing their English proficiency in the mathematical dialogues needed, for example, to defend their mathematical thinking and explain, reason and argue (Robertson & Graven, 2019). According to Prediger, Erath and Moser Optiz (2019) "language is a major learning medium used for communicative and epistemic purposes in mathematics classrooms", which makes it imperative that language "become a learning goal also, in mathematics classrooms" (p. 11).

In South Africa, the majority of Grade 4 learners are forced to read and learn in a language which they have not acquired sufficient basic vocabulary to communicate in (Sibanda, 2017).

The use of abstract concepts in the teaching of mathematics may not always be meaningful to the learner, even more so to the second language learner. For these learners, moving from informal spoken to formal written mathematical language, must happen on three levels "from spoken to written language, from main language to English, and from informal to formal language" (Setati, 2002, p. 10). According to Graven (2014) the majority of South Africa's students are being schooled through an as yet incompletely mastered second language, without attention being given to how their home language can be used as a resource for learning. This is a significant challenge to the effective learning of mathematics.

1.2.3 Evidence of the effect of the transition to English from the Annual National Assessments

The DBE Annual National Assessments [ANA] results from 2011 to 2014 (DBE, 2012; 2013; 2014) also indicate poor performance in literacy and mathematics. While these assessments can be argued as outdated, and they have been critiqued, they are the only assessments available that report on each Grade of FP and IP, and are thus useful in showing the influence of the change in the language of learning and teaching in the IP. For this reason, I report on them in this section.

Table 1.1 shows the national Mathematics ANA results for the Foundation Phase from 2011 - 2014. There is no denotation of home language (HL) and first additional language (FAL) learners and we are thus unable to disaggregate the results by language. Learners in FP were assessed in their home language. There is some improvement evident across the years, with the lowest scores being achieved by the Grade 3s. With an average of 53% in 2013 and 56% in 2014, this means that Grade 3s are entering the Intermediate Phase with approximately 50% competency in Mathematics, even when being taught in their home language.

Table 1.1 ANA percentage scores in FP Mathematics (2011-2014)

Phase/ Grade		**2011**	**2012**	**2013**	**2014**
FP	1	63	68	60	68
	2	55	57	59	62
	3	28	41	53	56

Data derived from DBE (2012); DBE (2013c); DBE (2014).

Table 1.2 shows the average Intermediate Phase Mathematics achievement results. The scores are significantly lower than in the FP, with none of the results exceeding 45%. The major difference is the shift to teaching and assessing the learners now in exclusively English or Afrikaans, and not their home language.

Table 1.2 ANA percentage scores for IP Mathematics (2011-2014)

	2011	**2012**	**2013**	**2014**
Grade 4	28	37	37	37
Grade 5	28	30	33	37
Grade 6	30	27	39	43

Data derived from DBE (2012); DBE (2013c); DBE (2014).

Table 1.3 shows the percentage scores of Grade 1 – Grade 6 in Home Language and First Additional Language from 2011 to 2014. For 2012 – 2014 the averages range from 51% to 63% for Home Language in FP. There is again a drop evident in the IP where the averages range from 30% to 63%, with the majority of the averages being below 50%. The table doesn't distinguish between which languages are assessed, but it is evident that learners are not fully proficient in their languages, and this would include in the language of learning and teaching. Spaull and Hoadley (2017) report that by the end of the Grade 4 year, more than 50% of South African learners have not learnt to read fluently and with comprehension in any language. Reddy et al. (2016) also identified language as one of the main contributing factors to poor performance.

Table 1.3 ANA percentage scores in FP and IP Language (2012-2014)

Phase/ Grade		**2012**		**2013**		**2014**	
		HL	**FAL**	**HL**	**FAL**	**HL**	**FAL**
FP	1	58	-	60	-	63	-
	2	55	-	57	-	61	-
	3	52	-	51	-	56	-

IP	4	43	34	49	39	57	41
	5	40	30	46	37	57	47
	6	43	36	59	46	63	45

from DBE (2012); DBE (2013c);DBE (2014)

There are a number of reasons for dissatisfaction with the Annual National Assessments, and many of these revolve around language of assessment issues. As indicated, by Grade 4, 90% of South Africa's students are officially learning and being assessed through the medium of English (Spaull, 2016). As noted by Sibanda (2017) and Sibanda and Graven (2018) assessment in a language in which the majority of students are not yet proficient cannot be seen to provide valid evidence of students' competencies. Nevertheless, it does provide us with a picture of the change in performance from FP to IP and highlights the challenge in making the transition from FP to IP.

Robertson (2017) explains the phenomena evident in the IP Mathematics results:

> "The bigger range differential in the ANA scores for Mathematics may be partly explained by the fact of IP learners having to contend with what is, in effect, a treble assessment load: firstly, the IP move towards more challenging and abstract mathematical tasks; secondly, the fact that, irrespective of whether or not it is the learners' HL, by this stage English will, in most South African IP classes, be the main LoLT and thus the language of assessment; and thirdly, the strong likelihood that the assessment tasks will have been couched in less carefully-scaffolded language as regards vocabulary and syntactic structure." (p 36)

The poor performance of learners in their languages and mathematics points to the need to closely examine the mathematics textbooks as a key mediating artefact in the teaching and learning of mathematics. If the textbooks are not written in comprehensible language, there is little use in relying on these texts in the classroom and for learner independent work.

1.3 Data handling as a focus of this study

In this study, I focus in particular on the data handling sections of the textbooks and workbooks. According to Naidoo and Mkhabela (2017) data handling encompasses real-world situations and assists in developing critical thinking skills in learners and thus it plays an important role in mathematics education. The contexts for data handling activities should be selected to create an awareness of social, economic and environmental issues. Naidoo and Mkhabela (2017) add that despite data handling playing such an important role in mathematics education, global international assessments, as discussed in section 1.1, disclose that learners are not performing well in data handling.

The data handling section of the curriculum can be considered to be the content area that contributes to the development of statistical literacy. According to Watson (2006), statistical literacy is "the meeting point of the chance and data curriculum and the everyday world, where encounters involve unrehearsed contexts and spontaneous decision making based on the ability to apply statistical tools, general context knowledge and critical literacy skills" (p.11). For both Gal (2004) and Watson (2006), questioning claims in social contexts such as media reports is fundamental to statistical literacy.

Data handling is a complex mathematical topic and confronts learners with terminology that may have a different meaning when used in daily life. In textbooks and workbooks learners encounter this data as embedded in text which is informationally dense. Since most South African learners are not learning in their home language the focus of the research is on the language demands of the text used in data handling sections in textbooks and workbooks.

Data handling is a well-defined sequence of activities completed in order to solve a problem. The text would begin this problem solving cycle, thus the need to examine the texts in order to understand whether they facilitate or hinder children's access to mathematics. Data handling texts are also dense with pictures, graphs and charts. For learners to interpret the data properly, they need to understand the mathematical content, context and the graphics (Lowrie & Diezmann, 2007). They also need to be able to verbalise their interpretation based on information in the graphics and text.

In Grade 4, children should be able to answer simple questions based on data in graphs and charts. They should be able to read, interpret and represent data – collect, organise, analyse and understand information, from one or more sources, in order to answer questions (DBE, 2011). Higher order thinking is required when having to integrate information and process it in order to create new information or solve problems based on quantitative information. Learners need to understand that information is organized through the use of graphs and charts and that they provide a way of being able to compare data. The interpretation of these graphical representations is reliant on general communication as well as a specific language type. Across the Intermediate Phase, according to CAPS (DBE, 2011) the main progression in Data Handling across the grades is achieved by working with new forms of data representation and developing new analytic tools for interpreting and reporting data. By the end of Grade 6, learners should be able to represent data in pictographs, bar graphs and double bar graphs and should be able to critically interpret data represented in words, pictographs, bar graphs, double bar graphs and pie charts (DBE, 2011). In addition, they should be able to summarise data verbally and in writing, including drawing conclusions about the data and making predictions based on the data (DBE, 2011).

The typical data handling problem solving cycle would begin by posing a problem. The learner needs to translate the problem into a question that the data handling cycle/process must answer. This is dependent on comprehension of the problem. Once the problem has been posed, the data is collected, organised and represented and then interpreted. Conclusions are drawn or predictions are made in order to solve the problem. Learners should work through the full data cycle a few times a year – this involves collecting, organizing, representing, analysing, interpreting and reporting data. Some of these aspects of data handling can also be dealt with as discrete activities. How the child makes sense of the question posed by the data collected, may be more complicated due to the texts.

Generally a person works with a large amount of data and because it is impossible to work with all the data at once, will try to describe it, typically in two ways. Numerical values are used to describe data – measures of central tendency (the mean or median) or measures of distribution (percentiles, quintiles) and pictures in the form of graphs and charts are used to describe data. Data handling texts are thus dense in information and graphics. It is necessary to consider both the complexity of the language used and the comprehensibility of the graphics used. This research sets out to do both.

1.4 Textbooks as the focus of this study

The ability to make sense of the text and the accompanying graphics, in order to represent, analyse and interpret data, becomes even more important when the reliance on text and workbooks is heavy and the learner needs to be able to analyse and write down answers to questions about the data. In 2020 and 2021, during the COVID-19 school closures and the rotational attendance that followed, the text and visual and graphical representations in the textbooks and workbooks became more important. Texts could not be mediated as usual by the teacher, and it is assumed that learners needed to work independently with their textbooks and workbooks. Mathematical social interactions with the teacher or peers were likely scarce and the texts that learners were working from would likely be relatively unmediated by any more knowledgeable other.

According to Harries and Spooner (2000), workbooks and textbooks play a central role in increasing understanding of mathematical concepts for both teachers and learners. Textbooks support learners' understanding of concepts, by presenting visual representations and symbols, which together with experiences and language, promotes learners' understanding of mathematics (Hoadley & Gallant, 2016; Liesbeck, 1984). They are useful curriculum tools in that they provide curriculum aligned activities that allow learners to practice mathematics and provide teachers with a guiding frame in preparing classwork and homework activities (Nicol & Crespo, 2006). Hoadley and Gallant (2016) add the monitoring and assessment function of textbooks.

Texts that to a native English speaker may be cognitively undemanding, could be very demanding for an English language learner (Cummins & Swain, 1986). This is especially true for mathematics, since mathematical language is challenging even for learners learning mathematics in their home language (Halliday, 1978). Since English language learners take longer to reach a requisite level of proficiency in academic language than English home language learners, they may not have reached the necessary threshold by Grade 4 (Cummins & Swain, 1986), and will in those cases be unable to engage meaningfully with the text provided in the book. There is thus an increased need to closely examine these texts and to consider their comprehensibility. The texts in focus for this research are both textbooks and workbooks.

The DBE have provided workbooks to all Grade 1-6 learners in public schools since 2011 (Hoadley & Galant, 2016). Grade 4-6 Mathematics learners in South Africa are only provided with these workbooks in either English or Afrikaans. The intentions of the DBE workbooks are: "to provide learners with worksheets to practice language and numeracy skills that have been taught in class…[t]hey are also meant to help teachers track the progress of learners and provide extra support if needed, they are a simple way to structure learning activities for learners" (DBE 2015). The DBE workbooks have garnered mixed opinions as to the use and success of them in terms of being an effective curriculum tool (Hoadley & Gallant, 2016). Hoadley and Gallant (2016) explain that these workbooks provide structured activities for the learners to practice and are a teaching tool. Fleisch et al. (2011) argue that the DBE workbooks are hard to follow, there is no sequence of concepts and learners are only exposed to one concept at a time. In other research by Mdluli and Ramsingh (2014) on the use of DBE workbooks in primary schools, the findings showed that teachers used the workbooks in varied ways. Some "seemed to think that the use of the workbooks alone was sufficient for the development of mathematical meaning and understanding" (p. 91) and the resource became the most important object in the mathematics classroom. Several teachers in this study used the workbook to the exclusion of using any other activities of resources and others used them as a source of homework, leaving learners to work through the content of the workbook independently (Mdluli & Ramsingh, 2014). Given the role that these workbooks, as well as textbooks, have come to play in many South African classrooms, there is need to interrogate their suitability.

As already argued, most children do not learn to read for meaning in any language by the end of Grade 3 and are consequently switching into learning through an additional language when they are not yet literate in their home language (Van der Berg et al., 2016). The language used in work and textbooks thus needs to be investigated. Setati (2002) explains the similarity between learning a language and learning mathematics: both require the learner to learn new terminology and symbols, and also to learn how to use different vocabulary in different contexts, as well as learning an awareness of grammar. This means that encountering texts in textbooks and workbooks in Mathematics is exceptionally complex for an English language learner. The language demands of these texts need to be carefully considered.

According to Halliday (1993), linguistic features that have a negative effect on learner comprehension include long phrases, complex sentences, syntactic ambiguity, special expressions, lexical density and more. According to Abedi (2006), these features slow learners down, their thinking can become confused and this could lead to not comprehending the particulars of the questions. Academic text requires different comprehension and production strategies from everyday oral language interactions (Cummins & Swain, 1986). Mediation of texts for English language learners is essential, but this is not always possible when the textbook or workbook becomes the teacher and the learner needs to be able to comprehend the mathematical text independently.

1.5 Research goal and research questions

The research goal:

To investigate the nature of the English language comprehension demands of the Intermediate Phase texts on data handling in textbooks and workbooks used in South Africa

Main research question:

What are the language comprehension demands of English mathematical texts on data handling that are used in South African Intermediate Phase Mathematics?

Sub-questions

1. What is the readability of mathematical texts on data handling that are used in Intermediate Phase Mathematics?
2. What is the linguistic complexity and what language features contribute to the linguistic complexity of mathematical texts on data handling that are used in Intermediate Phase Mathematics?
3. What is the comprehensibility of the non-textual elements that accompany the mathematical texts on data handling that are used in Intermediate Phase Mathematics?

1.6 Chapter summary and overview of the thesis

This chapter has outlined the purpose of this research and its location in the South African context. I have provided an overview of the state of South African education in terms of literacy and numeracy and have made clear the concern in this research for the English language learners who will be encountering mathematics textbooks and workbooks written in English from Grade 4. A rationale for the focus on the Intermediate Phase is provided. In addition, I provide an explanation for the focus on data handling and on textbook and workbook analysis. Finally, I have presented the research goal and research questions, as well as an overview of the structure of the study.

In Chapter 2 literature I present the theoretical framing of the study. The study is framed by Vygotsky's (1986) sociocultural theory and Cummins's (1981) theory of second language acquisition. Thereafter I present a review of the literature guided by the key concepts in this study. I review literature on the challenges of learning mathematics in an additional language and the challenge English language learners face when working with mathematical texts in English. I also review literature on reading comprehension, the particular demands of mathematical texts and visual and statistical literacy as well as reviewing studies on textbook analysis.

In Chapter 3 the methodology of the research study is outlined. This is an interpretivist study which uses a mixed method case study methodology. The sample is described in Chapter 3 and the instruments for data analysis are detailed. Finally, issues of validity and reliability are discussed.

In Chapters 4 and 5 the data is presented. Chapter 4 focuses on the readability and linguistic complexity of the texts in each of the books analysed. Chapter 5 focuses on the analysis of the non-textual elements used with the text in the books analysed.

In Chapter 6 I discuss the findings presented in Chapters 4 and 5 in order to answer the research questions. Finally, I discuss the practical implications of the findings of this study as well as the limitations of the study and potential avenues for future research.

CHAPTER 2

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

In this chapter I first present the theoretical framing of the study. This is a study that is framed by the sociocultural theory of Vygotsky and by Cummins' theory of second language acquisition. I then proceed to review the current literature on key aspects of this study.

2.1 Theoretical framework

The study is located within the theoretical perspective of the sociocultural theory. This framework serves to highlight the nature of the study as well as inform it. The study assumes that language is central to learning and particularly draws on Vygotsky's (1976) sociocultural view of language and learning. In addition, this study draws on Cummins' theory of second language acquisition.

2.1.1 Vygotsky's sociocultural theory

In this study language plays an important role. I draw on Vygotsky's (1989) theory on the social aspect of language development - he believed that the construction and change of the human mind occurs through the acquisition of speech and later, verbal concepts. This sociocultural theory's main feature is that learning and cognitive development, including language development, happen as a result of social interactions. The development of language is considered to be a major principle in Vygotsky's sociocultural theory as language is an important tool in the development of higher mental processes of learners (Vygotsky, 1989). The focus of this theory is that social and cultural processes shape cognitive development. Language is the tool that mediates learning processes and the word is the carrier of concepts. Textbooks thus become tools in the acquisition of mathematical concepts, especially in the IP, as teacher mediation of texts becomes less and less.

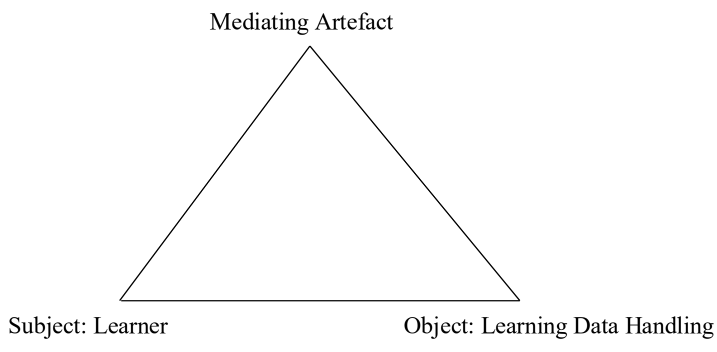


Figure 2.1 Vygotsky's mediation triangle

Figure 2.1 shows Vygotsky's notion of the 'mediated action' which is represented as a triangle containing three elements – subject, mediation artefact and object.

The mediating artefacts take the form of signs and tools. The difference between signs and tools is the ways they position human behaviour. The function of a tool is to serve as a conductor of human influence on the object of activity, in this study the text is the tool, which the learner uses in order to engage with the mathematical activity. A tool is externally focused on leading to changes in objects – the way mastery over nature is an aim of human activity. In order for internalization, which is the outward sign that learners need, has been developed into an internal sign that adults produce, human behaviour breaks away from its biological development and produces new psychological processes which are culturally based. Written language is a system of signs that represent the sounds and words of spoken language, which, in turn, are signs for real objects and associations between them. The intermediate link of spoken language, disappears, and written language is altered into a system of signs that symbolize objects and their relationships.

In this study, the subject is the learner and the object is the learners' learning of data handling. In the classroom, the teacher is usually the key mediator, the 'more knowledgeable other', who uses language and other mediating artefacts as the means to facilitate teaching and learning. This behaviour or human action and psychological functions are mediated by technical or psychological tools and signs. Vygotsky (1986) refers to using language as one of mans' greatest tools of cognitive development. Private speech becomes internal speech and eventually thought process as language is a tool of learning. Scaffolding, the strategy that supports learner development and provides structures to get to independent learning is the role of the teacher and instruction is crucial to cognitive development in the classroom.

From this Vygotskian perspective, textbooks and workbooks can be considered to be a key mediating artefact in mathematics teaching and learning. These texts are often mediated by a teacher in a classroom situation, but during COVID-19, learners were isolated from their classrooms during lockdown and would likely have been independently working with texts. In addition, in the IP learners are increasingly encouraged to work independently when using a textbook. For the purposes of this research, the textbook is viewed as one of the key mediating artefacts in the learning of mathematics. This study concerns itself with assessing the suitability of this mediating artefact for South African learners.

2.1.2 Cummins' theory of second language acquisition

In addition to Vygotsky's (1989) sociocultural theory, I draw on Cummins' (1981) second language acquisition theory. As argued in Chapter 1, the majority of South African learners are learning through the medium of English without this being their home language. It is therefore pertinent to include consideration of the acquisition of a second language and for this reason Cummins' theory is selected as a framing theory in this research. According to Cummins (2008), a theory is a manner of viewing phenomena that has degrees of use, dependant on its purpose, the communication with the intended audience, and the value for practice of following through on its implications. The reason for drawing on this theory for the purposes of this research is because the majority of the learners in South Africa who are learning mathematics in English, are not learning it through their home language. They are still in the process of developing this language and it is important to understand how this development happens if the textbooks and workbooks are to be appropriately appraised.

Two underlying proficiencies in second language acquisition, as introduced by Cummins in 1979 and 1981, are Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP). This theoretical framework assigns a central role to the interaction between socio-cultural, linguistic and school program factors in the development of these proficiencies (Cummins, 1979).

The theoretical intent of the distinction between BICS and CALP was initially to qualify the claim made by Oller (1979) that all individual differences in language proficiency could be accounted for by a single underlying factor, which he called 'global language proficiency.' The argument used by Cummins (1979) was that a single aspect known as general or global language proficiency could not incorporate all aspects of language performance or use, as phonology development slowed down after a few years, yet academic language development continued into adulthood.

As the acronym implies, basic interpersonal conversational skills refers to 'everyday' conversational language used in general, at home and around the playground as Gibbons (1991) refers to 'playground language' which includes language which "enables children to make friends, join in games and take part in a variety of day-to-day activities that develop and maintain social contacts" (p. 3). This mostly happens in face-to-face situations and is greatly dependant on gesture, body language and visual context.

According to Cummins, in Oliver (2002), BICS is social language - it occurs in context-embedded communication and refers to the communicative ability of a person which helps him/her function in daily interactive exchanges. It can involve writing, but is less formal that academic language using language for different reasons. Greeting, through talking or waving, asking for directions through pointing and answering a question through gestures or actions are all basic social skills. Hill (2009) highlighted the development of logical, pragmatic skills through BICS, as it involves conversational skills and adjusting language based on the situation. Generally learners with good social language skills have better interpersonal skills. BICS generally takes about 2 years to develop, according to Collier (1987) and Cummins (1984). Conversational language is developed through primary discourses, according to Cummins (2001), and these are developed through interactions in the home and represent the language of early socialization.

The development of academic language takes longer, it develops through social interaction from birth, as BICS does, but becomes differentiated in that it reflects primarily the language learners acquire at school which they have to use appropriately, if they want to proceed successfully. CALP, or cognitive academic language proficiency, can be defined as "the extent to which an individual has access to and command of the oral and written academic registers of schooling" (Cummins, 2000, p. 67). This language is specific to the social context of schooling and entails knowledge about language forms and rules (Cummins in Olivier, 2002).

The way language is used in the school or work setting is more formal and conforms to some societal standard which allows more self-correction than social language according to Everhart (2010). CALP has two components, academic language and academic context. Academic language is defined by Krashen and Brown (2007) as complex syntax, academic vocabulary and a complex style of discourse, and academic content, as the content of the subject. For example, History or Mathematics would have difference styles of discourse.

Academic language proficiency as defined by Bailey (2007) refers to students' understanding and ability "to use general and content-specific vocabulary, specialized or complex grammatical structures – all for the purpose of acquiring new knowledge and skills, interacting about a topic, or imparting information to others". Learners with well-developed CALP have a better chance at seeking information, comparing, ordering, classifying, analysing, inferring, justifying and persuading, solving problems, synthesizing, and evaluating information for their learning (O'Malley & Pierce, 2002 in Everhart, 2010) and thus a better chance of learning new knowledge through listening and reading, and expressing themselves through discussions and written work.

There is a gap of a few years between the development of BICS and that of CALP. According to Cummins (1999), grade norms in academic aspects of English required a period of 5-7 years to develop. Secondary discourses are acquired in social institutions and involve the acquisition of functions of language appropriate to the context and specialized vocabulary. These can be oral or written and are present in the social life of literate and non-literate cultures. An example of non-literate cultures' academic language is the language of marital or burial rituals, which have been passed down within generations, here academic language proficiency refers to the access and command of specialized vocabulary and functions of language, characteristic of the institution of schooling. Even language specific to careers and hobbies can be classified as secondary discourses with academic language, for example, professional chefs and amateur cooks have the same vocabulary which someone outside of the field would not understand. The reason that secondary discourse acquisition is fundamentally connected to schooling, is that the degree of proficiency people acquire in understanding and using the language, determine their life choices directly (Cummins, 1999). This second language acquisition theory introduced the concepts BICS and CALP, which have implications for policy and practice.

Developmentally, BICS and CALP are not necessarily different as all children learn basic concepts through interactions at home. Their knowledge of the world is shaped through these conversations. In the same way, academic language is developed through discussion about conceptual issues, and is an important way of deepening conceptual understanding and developing critical literacy (Cummins, 1999). This being said, there are clear differences in the acquisition and patterns of development between BICS and CALP. The nature of language proficiency needs to be understood as the difference in conceptual understanding can sometimes be misinterpreted, which often leads to learners who are highly proficient in BICS, being misconceived as being proficient in CALP and can lead to failure. Differences in the rate of acquisition needs to be considered so that the academic language of these children is not compromised.

Initially, the difference between BICS and CALP was expanded into two overlapping continua (Cummins, 1981a) highlighting the range of cognitive demands and contextual support in a language task/activity, either context-embedded/context reduced and cognitively demanding/cognitively undemanding. The internal and external dimensions of contexts were noted to reflect that context is made up of prior knowledge and experiences we 'bring' to the task as well as the visual or graphic support within the task. It was argued that second language learners needed context embedded, cognitively demanding tasks in order to acquire academic language. It was also noted that the terms ‘context-embedded’ and ‘cognitively demanding’ cannot be absolute, as learners’ prior knowledge will be different (Coelho, 2004; Cummins, 1981a). However cognitive skills are involved in most forms of social interaction, as the meaning being communicated is supported by contextual/interpersonal cues or linguistic cues, which needs to be interpreted. The term ‘context-reduced’ was used instead of ‘decontextualized’ as all literacy and language practices are contextualized, however the range of contextual academic language support between face-to-face interactions and textbook reading, is vast.

These two theoretical constructs have an influence on this study as in South Africa, children enter school after having spent about five or six years in the process of acquiring their mother tongues. Learners generally come to school with prior knowledge in their home language and, according to Brown (2012), most learners from rural areas may have had little to no exposure to the English language until they start school.

As Cummins’ (1981) research shows that the child requires between five and seven years to acquire sufficient Cognitive Academic Language Proficiency (CALP) to perform well on academic tasks. The acquisition of Basic Interpersonal Communication Skills (BICS) takes about two years. The BICS/CALP distinction should never be lost sight of, as well as progression from the one to the other, with the latter requiring a substantive cognitive leap from learners.

Theoretically, the BICS and CALP distinction has evolved since 1981 to include the addition of discrete language skills, as a component of language that is separate from conversational fluency and academic language proficiency according to Cummins (2001). These skills encompass the learning of rule-governed aspects of language, for example, phonology, grammar, spelling, where acquisition of the general case allows generalization to other cases directed by the particular rule. Discrete language skills can be learned apart from academic language proficiency according to Cummins, Brown and Sayers (2007), illustrated by the fact that some children who are able to ‘read’ English, have a limited understanding of the words they are able to decode.

There are many studies that support the different time periods needed to acquire peer-appropriate conversational second language proficiency as compared with meeting grade expectations in Canada (Klesmer, 1994), Europe (Snow & Hoefnagel-Hohle, 1978), Israel (Shohamy et al., 2002), and the United States (Hakuta et al., 2002; Thomas & Collier, 2002). In South Africa, there are 11 official languages of which learners need to be proficient in two (DBE, 2011). The Threshold Hypothesis, which Cummins proposed in 1979, explains the cognitive effects of bilingualism in children. Cummins (1979) suggests that English second language learners, need to cross a threshold in home language proficiency, so that competencies developed in their home language transfers to the second or target language. For this transfer to happen, learners should be competent in both the home and additional languages.

In this study, the distinction between BICS and CALP is important. Learners can be expected to arrive in Grade 4 with some measure of BICS in English – if not learning through the medium of English in FP, they are required to learn English at an additional language level in the FP. However, in order to transition to learning through the medium of English, they need to have a measure of CALP. The CALP required in Mathematics is distinct from the CALP that may be required in other subjects. In order to develop this CALP, there is a threshold level of home language proficiency that is required, according to Cummins. As is evident in the ANA results for language, this threshold may not have been reached for many learners. All of this needs to be kept in mind when appraising textbooks and workbooks for their suitability for South African learners.

2.2 Literature review

In this section I review past and current literature in relation to the key concepts in this study. I open with a review of literature about the use of English as the language of learning and teaching in South Africa and the challenges of learning mathematics in an additional language. I then review current studies on children's reading comprehension, before moving to considering the particular demands of mathematical texts and thereafter considering some recent studies on the analysis of textbooks. Finally, as data handling is the section of the mathematics curriculum that is directed towards the acquisition of statistical literacy, I close the chapter with a consideration of recent literature about statistical literacy.

2.2.1 The challenges of learning mathematics in an additional language

The Language in Education Policy (DoE, 1997) supports moves towards multilingual education, with the underlying principle to maintain home languages while providing access to the acquisition of an additional language. An additive approach to bilingualism is seen in the orientation of the policy. The rights of the individual, to choose the language of learning and teaching is emphasized in the policy as vested in the individual, to be exercised within the framework of the education system's obligation to promote multilingualism. This policy was written to counter disadvantages which could result from mismatches between home languages and languages of learning and teaching.

The language(s) of learning and teaching in a public school must be (an) official language(s) (DoE, 1997). Of the 11 official languages in South Africa, the Language in Education Policy (1997) mandates School Governing Bodies to select the language of learning and teaching in their school. Many of these select English which contradicts the advocacy of additive bilingualism and forces many learners to learn all subjects, including mathematics, in English. The fact that mathematical assessments can only be accessed by having an understanding of the language of learning, teaching and assessment challenges their fairness (Sibanda & Graven, 2018).

There is a large amount of research confirming South African learners' consistent poor performance in literacy and mathematics (DBE, 2013, 2014; Reddy et al., 2015; Spaull, 2016; Pretorius & Spaull, 2015). Sibanda (2017) highlights the limited exposure to English in most young South African learners which hampers these learners attaining the 'threshold' suggested by Cummins (1979) and limits the linguistic skills transferring from the home language to English. Many learners thus enter the IP with very limited proficiency in English. This challenge is compounded by teachers using English for teaching and assessment, which, particularly in mathematics, is complex. Spaull (2016) notes that by Grade 4 learner competence in English is not well developed yet, but neither is the mother tongue competence well developed, thus Grade 4 becomes a critical stage of transition and extensive learner support is needed.

South African learners may grapple with learning mathematical concepts in a second language as language is essential for mental development as it provides the concepts for thinking and thus a means for conveying ideas and asking questions (Vygotsky, 1989). As explained previously, there is a big difference between teaching English for basic interpersonal conversation skills and teaching English for cognitive academic language use and in the teaching of mathematics. As Setati and Adler (2000) note: "Learners… have to cope with the new language of mathematics [both informal BICs/and formal CALP] as well as the new language in which mathematics is taught" (p. 247). If teachers are not aware of the differences between "the surface or conversational aspects of children's language and the deeper aspects of proficiency that are more closely related to conceptual and academic development"

(Cummins, 1994, p. 37), there is a risk that they may take a learner's BICS as evidence of overall linguistic proficiency. They may then fail to adequately provide for the ongoing development of that learner's CALP. Language acquisition practices would need to be additive, rather than subtractive, within the learner context. Focussed teacher talk and increased learner interaction is the key to improving the link between basic communication and academic language.

Graven (2014) suggests that South African children experience significant challenges in their mathematics learning as most are being taught through a second language, which has not been mastered yet, without thought being given to using their home language as a learning resource. This may compromise the learners' capacity to participate in the learning process and restricts their opportunities for deepening conceptual understanding of mathematical ideas (Robertson & Graven, 2019).

The current curriculum, CAPS (DBE, 2011), supports teaching mathematics through the medium of the home language until the end of the FP (Grade 3). However, beyond Grade 3 it does not support teaching through the medium of the home language (where it is not English or Afrikaans). This is exemplified by the DBE's distribution of workbooks exclusively in English or Afrikaans from Grade 4 onwards. The reality is that most of the learners are learning in English as a second language and have to learn in and read texts in a language in which they have had little to no exposure and are not adequately proficient in.

A recent study by Sibanda (2017) reveals that the learners' low performance in mathematics assessments was largely due to poor proficiency in reading and comprehending the problems, in the majority of questions, in the English language. Her study focus was on Grade 4 learners and the transitions they faced, particularly from 'learning to read' in Grade 3, to 'reading to learn' from Grade 4 (Chall et al., 1990). According to Graven (2014) the Annual National Assessments (ANAs) and outcomes of other assessments indicate that many learners lack grade appropriate foundational mathematical understanding, in the case of the majority of South Africa's IP learners, which may be caused by a language deficiency rather than a mathematical challenge.

Prediger et al. (2019) report on research that has established that native speakers of the language of learning and teaching and assessment, perform better that non-native speakers on mathematics assessments. According to Robertson and Graven (2019) there are major deficiencies in education systems in which the majority are forced to learn mathematics in a second language, but which provide no support for the development of academic language proficiency. These students are under threat to ‘grow’ mathematically, as they are learning mathematics through a language other than their home language.

Robertson and Graven (2019) report that many South African children are exposed to English only in the school context. The 2015 TIMMS results confirm this as only 31% of Grade 5 learners spoke the language of learning and teaching at home (Reddy, 2016). The other 69%, whose home language was different to the language used at school, received significantly worse scores in the assessment. Because of this language disparity in classrooms, language needs to become part of the aim of mathematics teaching, as Prediger et al. (2019) remind us that language is important for epistemic and communicative purposes and is a “major learning medium” in mathematics classrooms. South Africa is not the only country to experience the phenomenon of learners learning mathematics in their second language, according to the UNESCO Global Education Monitoring Report (2016), “as much as 40% of the global population does not have access to an education in a language that they speak or understand.”

Terminology in text can be ambiguous and the use of words to describe processes and visual images to represent data also has an effect on the learners’ understanding (Friel et al., 2001). Reading the data, reading between the data and reading beyond the data would be more complicated in a second language (Friel et al., 2001). In addition, rhetorical information is often implied and not explicitly stated in mathematical discourse (Flick & Anderson, 1980), leading to students with English as a second language struggling to understand the whole meaning of a paragraph, even though they may understand individual sentences (Flick & Anderson, 1980).

Abedi (2010) notes that if learners are to perform well in mathematics, they should possess competence in both everyday language and mathematics specific language. This calls for the teacher to mediate mathematical texts and assessments taking into consideration the manifested disadvantage that limited linguistic exposure places on the learners.

2.2.2 The demands of mathematical texts

Constructing a mental model or the representation of a problem, seems specific to mathematical problem solving (Lucangeli et al., 1998). This mental construction in word problems specifically, is dependent on four abilities. The first two involve reading comprehension skills, being able to transform each sentence in the text, into a mental representation and the integration of different pieces of information into a single representation of the problem

(Lucangeli et al., 1998). Being able to plan the steps towards the solution and the execution of the plan, are the other two abilities needed.

In this research I view the language proficiency of the learner as either allowing access to the mathematical concept, or as presenting a barrier to access. The complexity of the text itself influences whether the learner is able to make sense of the problem in the real world. The following model clarifies that position. When a problem is posed in a real world context, the learner needs to translate this into a mathematical problem in order to formulate a mathematical solution. The gathered information must then be interpreted in order to make the decisions about the calculation to be performed. Once the mathematical solution is calculated, these results need translation back into the real world context. Below is a visual representation of this process.

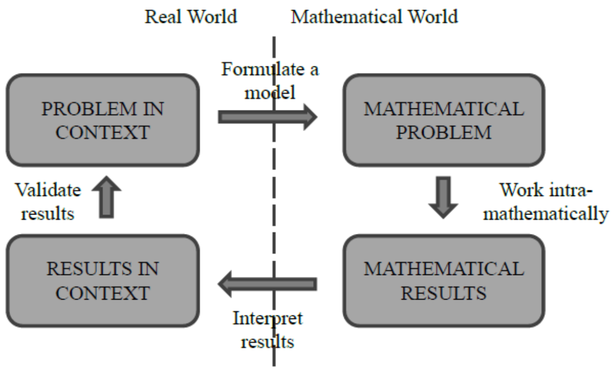


Figure 2.1. Problem solving in mathematics (PISA Governing Board, 2010, p. 6)

When working with a text, there is an additional layer to this, as shown in the figure below. This mental construction is dependent on reading comprehension skills – being able to transform each sentence in the text, into a mental representation and the integration of different pieces of information into a single representation of the problem (Lucangeli et al., 1998). It is language proficiency in reading and interpreting the text (including visual representations) that allows access to the mathematical task.

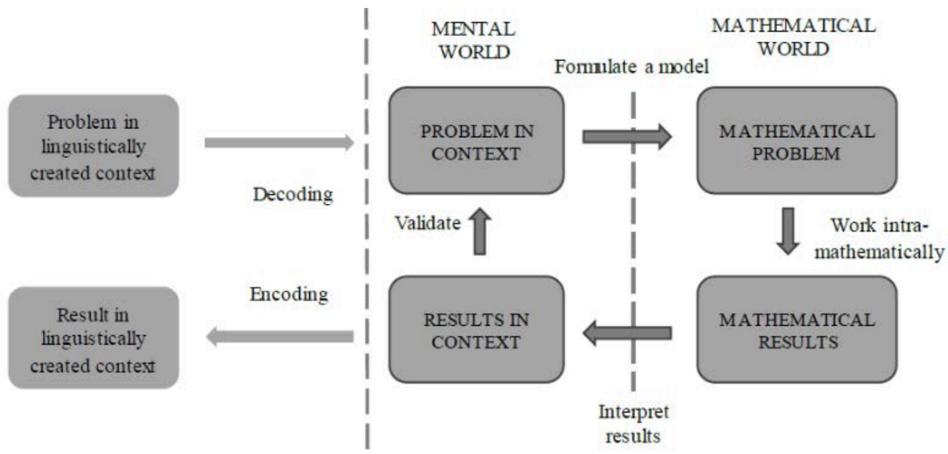


Figure 2.2 Problem-solving when a context is presented in text (from Vale, 2018, p.90)

This is pertinent in this research as 80% of learners are learning mathematics in a second language in South Africa and are needing to decode problems in a linguistically created context and then need to 'translate' them to the mathematical world, before solving them and interpreting the results in context, the added obstacle is encoding the result in the linguistically created context.

In Martiniello (2009) there is an illumination of the role of linguistic complexity of text and interaction with forms of visual representation. In order to solve maths word problems successfully needs decoding of all of these modalities. The learners need to understand the item's verbal language first, the more complex the text, the more difficult it is to understand, leading to misinterpretations and incorrect solutions (Mestre, 1988). Secondly, they need to know mathematics specific vocabulary, that has both specialized vocabulary (tables, data, pictograph) as well as syntactic structures (greater than, the same as) that are typical of mathematical discourse (Abedi et al., 1997; Halliday, 1978; Lemke, 2003; Mestre, 1988). The last action of the learners is the interpretation of non-linguistic mathematical symbols and their sentence structure to make sense of mathematical concepts as well as make sense of visual displays like graphs.

2.2.3 English language learners working with mathematical texts in English

As stated in the above writing the research is concentrated on the language demands of mathematical texts in a second language environment and how the work of Vygotsky (1978) addresses language being a tool for learning. According to Vygotsky (1978), tools assist learners in accomplishing an activity and learners are active participants in constructing their knowledge using language as a tool. The language used in textbooks is thus an important psychological tool in learners' acquisition of mathematical concepts. This is usually mediated by a teacher, however learners are increasingly expected to use texts independently as they progress through the IP grades. Therefore, the mediation from the teacher is gradually withdrawn as learners gain independence. However, the provision of a workbook or textbook on its own does not guarantee that learners will make meaning of mathematical concepts or make mathematical connections (Uttal et al., 1997). Mathews et al. (2014) suggest that while providing a resource like a workbook may help address some teaching and learning needs, it is not a surety that learner performance and understanding of mathematics will improve. This is particularly the case in this study as learners enter the IP with different levels of language proficiency.

In Cummins's work, as unpacked earlier in the chapter, the distinction between basic interpersonal communicative skills (BICS) and cognitive academic language proficiency (CALP) is made. It is pointed out that even under optimal conditions it takes learners considerably longer to acquire CALP in their additional language than it does for them to acquire conversational everyday fluency, or BICS, in an additional language (Cummins, 2008). As previously noted in this chapter, his research findings indicate that it may take as much as 5–7 years to develop CALP in an additional language. Cummins's linguistic interdependence hypothesis highlights how important learners' home language is as an already present foundation of linguistic and conceptual understanding that could be transferred to the additional language. As he points out "conceptual knowledge developed in one language helps to make input in the other language comprehensible" which can then form the basis for developing collective underlying competences across both languages (Cummins 2000, p. 39). These two aspects resonate with the South African language context as CAPS (2011) advocates multilingualism and mother tongue education in the FP, yet all IP support material is printed in English and Afrikaans.

Transferring the mathematical skills developed in the home language to English is complicated for English language learners (Vale, 2018). English language learners need to first unpack the language, before unpacking the meaning of mathematics, and this makes mathematical sense-making more cognitively demanding (Robertson & Graven, 2020). These learners need to move from their developing proficiency in informal spoken English to becoming proficient in formal written mathematical language (Setati, 2002). This needs to occur at three levels: from spoken to written English, from the mother tongue to English and from informal to formal language (Setati, 2002). Mathematics, with its own complex language, is for this reason a challenge for learners, and their mathematics teachers need to help them to navigate this.

In Sibanda's (2020) research on the impact of language on the transition period between Grade 3 and Grade 4, she argues that authors should ensure that the readability of the text corresponds to the reading level of the target learners. This transition is not only complicated by learner language abilities, but also by the change in the nature of the texts, from less challenging narrative texts to more challenging, expository texts (Sibanda, 2020). These texts, including mathematical texts, make more demands on learners than the narrative presentation as they are dense in technical, academic language.

Robertson and Graven (2019) draw attention to the challenges learners have when they are denied access to the linguistic resources of their home language as tools for learning mathematics and are forced instead to learn mathematics through an additional language in which they as yet have little fluency. They also point out that when learners learn mathematics through a language that is not their home language, their capacity of 'thriving' mathematically is hampered. In many cases these learners are denied the opportunity to develop logical mathematical thinking in any language, which ultimately leads to exclusion from learning. Most IP learners in South Africa are only exposed to English at school, for the 6 hours in the classroom, and in some classrooms, there is more isiXhosa spoken than English. As Morrow (1994) points out, epistemological access is being denied to children who are being linguistically challenged as their access to successful participation in practice is hampered by the language.

The importance of literacy was discussed earlier in this chapter, according to Robertson and Graven (2019) early definitions of becoming literate highlighted mastering the skills of reading and writing, Hammond et al. (1992) described literacy as being a broader process of development, incorporating different forms of literacy. Various content areas would require different kinds of literacy forms and different "literate actions" (Castanheira et al. 2001, p.354).

2.2.4 Reading comprehension of mathematical texts

As Connor (2016) states, reading comprehension is important in the information-based world and learners who do not develop adept reading skills are at a disadvantage. Weaker language and social-cognitive development, together with ineffective instruction are linked to weaker literacy development.

In the study conducted by Connor (2016) reading comprehension was conceptualised as a complex activity that requires the reader "to call on the coordination of cognitive, regulatory, linguistic, and text-specific processes, including decoding [and encoding] of text, which are developing over time and that have reciprocal and interacting bootstrapping effects on one another" (p. 2). Encoding and decoding, fluency with word and text, orthographic knowledge and the structure of text, are all text specific processes enabled by writing and reading what has been written. Connor (2016) suggests that when learners read with discernment, a situation model is built. This model is a coherent mental representation of the texts' meaning.

The ability to evaluate and regulate mental operations is part of metacognition – this is an awareness of the cognitive process and the ability to manipulate this process. The self-monitoring of learner comprehension encourages recognition of non-comprehension of the text and using strategies such as rereading or questioning the text for clarity. This allows a learner to make sense and assimilate the knowledge. Connor (2016) suggests that learners with weaker academic language skills, spend less time on these strategies to clarify comprehension, than stronger learners.

In a study on the grammatical forms in mathematics texts, Abel and Exley (2007) present findings that show that different subject areas have literary demands that are not supported in the Language subjects. Teachers need to be aware of the literacy demands of discrete subject areas. According to Unsworth (1997) and Wyatt-Smith and Cummings (2003) different subjects have distinctive language structures and grammatical features that learners must move between in order to be successful, relying on their ability to move among complex 'curriculum literacies' rather than a common curriculum literacy. This highlights the demands of mathematical texts from learners.

In many studies the grammatical challenges presented by worded texts reveal that underlying language and structural patterns hinder interpretation of mathematical discourses (MacGregor, 2002; Kalogeropoulos, 2005; Monroe & Panchyshyn, 2005; O'Halloran, 2005; Parkin & Hayes, 2006). This research highlights the difficulty learners may experience when faced with having to make meaning from worded maths texts and transferring this into numerical number stories (Zevenbergen, 2001). Graphical density of sentences, with long difficult words, complex and lengthy verb and noun groups, containing more content and fewer grammatical words, are all linguistic challenges that confuse young learners (Parkin & Hayes, 2006). In addition, worded mathematical texts containing language features such as homophones and homonyms, which create lexical density and the use of symbols and complex semantic structures are misinterpreted by learners often (Zevenbergen, 2001; Monroe & Panchyshyn, 2005).

According to Unsworth (1997), Derewianka (1998) and O'Halloran (2005) when learners have not been supported in their learning of mathematics texts, understanding and developing distinctive grammatical structures that are essential to the mathematics learning area, they will be less likely to participate successfully at school. This support for mathematics specific vocabulary needs to be scaffolded, not only by the teacher, but also the tools for learning (Zevenbergen, 2002).

In a 2006 study done by Parkin and Hayes, an analysis of worded maths texts in secondary school textbooks was done by using functional Halliday's functional grammar as a tool. Findings indicated that using student familiar language to identify the grammatical structure of the text was valuable for scaffolding the language demands of mathematics texts. It is thus important for teachers to prepare learners properly in dealing with the linguistic demands of mathematics texts (Zevenbergen, 2002; O'Halloran, 2005).

2.2.3 Visual literacy: Interpreting visual elements in mathematics texts

Mathematical texts in data handling present an additional challenge. They are often accompanied by visual elements, like images, charts and graphs that similarly require interpretation. It is relevant therefore to consider the visual literacy of learners. Learning to interpret and create charts and graphs are part of the learning outcomes for the data handling section of the curriculum, and thus learners are still learning how to make sense of them. There are also additional images used in the texts that require interpretation.

Many studies have reported that visual representations could attract learners' attention and produce educational outcomes (e.g. Evans et al., 1987; Levie and Lentz, 1982; Levin and Mayer, 1993; Pettersson, 1990), but other contradictory studies, mainly on literacy and science textbooks, have described how visual images may have been negative in the contribution to learner comprehension text (Watkins et al., 2004). I take the position that student understanding of content can be assisted by visual representations since they provide concrete images of associated concepts (Levin and Mayer 1993). This can allow students to see patterns, images and visualized objects that can be experimented with and manipulated, which are not easily seen only from texts (Arcavi, 2003). This is however dependent on the quality of the visual representations. Kim (2012) highlights the importance of ensuring that visual representations, that Kim refers to as 'non-textual elements', are accurate, connected, and concise and related to problems or situations, so that students are supported in developing mathematical knowledge. If these conditions are met, non-textual elements can provide many different opportunities to learn (Kim, 2012), dependent on the visual literacy of the learner.

According to Messaris (1994) the constructs of visual literacy is underpinned by social cognitive theory, which suggests that the cognitive process is affected by individuals' observations, combining visual images with pre-existing schemata, allowing the reader to supplement the text with the visual, to enhance understanding. According to Semali (2003, p. 97), to be considered literate, "students must be taught to read visual images in addition to connected text".

Interpretation or synthesis are skills close to the top of Bloom's hierarchy of skills, which raises these abilities to higher order thinking skills (Bloom, 1984). New connections extend learner comprehension as the reader synthesizes the visual representation and its connection to what is read in the text, this extends pre-existing schemata or creates new schemata, creating a wider base for future learning. Viewing is the ability to understand a visual graphic, communicating a thought or concept using graphics and extending language skills.

Since the acquisition of mathematical knowledge is visually mediated into symbolic artefacts (Sfard, 2008) and mathematically abstract phenomena is being represented, Cook (2006) argues that graphics appearing in mathematical texts are crucial. Both Presmeg (2006) and Sfard (2008) agree that the primary modality in mathematics is visual.

The complex process of making mathematical meaning from visual displays is emphasized by O'Halloran (2005) – learners cannot see what is intended unless they know what it is they are looking for and they understand the conventions being used, the visual must thus be treated in context. Cook (2006) similarly agrees, proposing that guidance is important when learning from visuals and learners need to use prior knowledge to select relevant information to developmental models.

There are also challenges presented by visual images and the ability of the learner to solve problems is affected by the strength of the linkage between information in illustrations and the accompanying text (Berends & van Lieshout, 2009). Mayer and Moreno (2003), suggest that visuals and texts use different perceptual and cognitive pathways and they need to be explicitly linked. Without this, a lower retention and transfer of knowledge can occur due to the bigger cognitive load arising from divided attention and double processing of information.

Since mixed technological media has been used acceptably at schools, visual literacy has been recognized as a basic arts standard (NCTE, 2006) and necessary skill. The increasing use of technology solidifies visual literacy on the list of important language skills. According to Arzipe (2001) the overall importance of developing visual literacy in children is noteworthy. It is particularly important in mathematics as it develops learner understanding of complex concepts and promotes critical and logical thinking and enhances communication skills (Botha et al., 2019).

2.2.5 Studies on textbook analysis

In 2011, the DBE developed a set of workbooks for learners from Grade 1 to 6 for literacy/language and numeracy/mathematics. By 2015, workbooks for English first additional language and life skills in the FP had also been developed. Currently workbooks are available in all 11 languages from Grade R to 6 in the learning area of language. They are available in all 11 languages for mathematics from Grade 1 to 3, and are only available in English and Afrikaans from Grades 4 to 9. According to Mathews et al. (2014), the DBE 2011 intervention strategy included the production and distribution of workbooks as additional support for teachers (DBE, 2011) due to the poor performance of South African learners on national and international literacy and numeracy benchmark tests. Fleisch et al. (2011) suggest that these workbooks have been designed to improve learner performance by supporting the teaching of mathematics.

According to Hoadley and Gallant (2016) in their analysis of Grade 3 Data Handling in the Rainbow workbooks, only five activities across the workbooks cover this content area, with only one of these in the second workbook. These activities are all very similar in form and question types, and even though the content area has a low cognitive demand, the few activities offer few opportunities for practice. A National Education Evaluation and Development Unit [NEEDU] 2012 report concluded that the "workbooks do not provide nearly enough practice exercises in mathematics" (Taylor, 2013). In a further NEEDU (2013) study, as well as in Davis (2012) and Gallant (2013), it was found that in a very small sample of tasks there were subtle misconceptions embedded in some mathematics activities in the workbooks.

According to Valverde et al. (2002), "textbooks are components of opportunities to learn school subjects and have their own characteristic impact on instruction" as a mediator between the planned and enacted curriculum. Textbooks are important components in the learning process and describe the minimal effort that teachers and learners should be making (Pratama & Retnawati, 2018).

Hoadley and Gallant (2016), in their analysis of Grade 3 DBE workbooks, argue that there is weak progression in the topic of data handling. North and Zewotir (2011) suggest that learners need a new set of questioning and justifying skills and that teachers need to instruct learners specifically, being aware of the need to encourage alternative means of interpreting data. Very little time is spent on data handling in the exercises in the Grade 3 DBE Rainbow Workbooks, the recommended CAPS weighting is 5%, yet only 3% of the activities in the book are related to data handling. Hoadley and Gallant (2016) note that key topics and sub-topics are inadequately covered with limited opportunities for practice in some content areas.

According to Hoadley and Gallant (2016) in the same study of Grade 3 English Home Language workbooks, there are many opportunities for learners to read and engage with the text, however this systemic engagement is at a somewhat basic level, with a limited range of texts, especially non-fiction texts. These texts also have a mundane focus on familiar themes, not facilitating key comprehension skills which are shown to be most absent in learner assessments of language (Howie et al., 2008) as the comprehension activities consist mainly of questions at a 'literal' level of understanding. In the mathematics books there were substantial practice opportunities in the content areas of Numbers, operations and relationships and Patterns and functions, but limited opportunities for practicing content and skills for Space and Shape, Measurement and Data Handling. Activities from these content areas are sparsely distributed across the workbooks, providing moderate engagement with these content areas in a term.

There is no "modelling" of answers, solutions or rubrics for assessments provided in the DBE Rainbow English Home Language or Mathematics workbooks, thus no clear examples of what the assessment requirements for a task are. In the mathematics workbook, 'talking heads' of children are used to indicate worked examples. The workbooks offer limited progress tracking opportunities as there are limitations to progression in some areas of learning (Hoadley & Gallant, 2016).

Hoadley and Gallant (2016) suggest that the workbook could be an effective systemic monitoring tool as there is a high level compliance with curriculum coverage. The number of completed pages per content area could be counted to calculate the actual coverage in content areas. This curriculum monitoring does not give an indication of the quality of work, nor learner progress as the NEEDU (2013) report highlights.

In the study done by Hoadley and Gallant (2016) they concur with the DBE suggestion that these workbooks are aimed at being practice tools – for learners to practice their learning. They are limited in providing remediation for weaker learners who may need practice from the previous grade level, because of their undifferentiated nature and curriculum compliance. The possibility of teachers using these workbooks to track learner performance and assessment as an assessment tool is also unlikely because evaluative criteria and conceptual signaling are often implicit (especially in language) and the activities do not model successful learner production of an assessment activity.

A study on the effectiveness of a workbook compared to that of a textbook was conducted by Fleisch et al. (2011), the findings of which concluded that there was no significant difference in learner performance between learners using workbooks and those who used textbooks. Textbooks are used to support and enhance the curriculum either as practice, assessment, monitoring or teaching tools (Hoadley & Gallant, 2016). Primary school teachers report that the textbook is the most important learning resource for mathematics instruction (Mullis et al., 2012) and often it is the only mathematics resource in the classroom (Sibanda, 2020). Hoadley and Gallant (2016) suggest that textbooks can be used as the primary resource in teaching mathematics or languages or as a curriculum coverage monitoring tool. Textbooks do not simply supplement teaching, but often take on a prominent role in South African classrooms in guiding the teaching and learning of subjects. Hoadley and Gallant (2016) argue that using the DBE workbooks as the 'sole' text given to teachers, would not be viable as a teaching or transmission tool as they are limited in the ways they can "structure learning activities for learners" as proposed by the DBE. The workbooks would need further development to ensure that conceptual signaling in the text and assessment expectations are more explicit.

There are several skills and literacies required to effectively navigate mathematical texts like textbooks and workbooks. Most critically, learners need the appropriate level of reading skills and reading comprehension in the language of the text. In addition, they need to be able to interpret visual representations of mathematical ideas, and to work with the vocabulary and symbols that are particularly characteristic of mathematics texts (Hammill, 2010).

Mathematics textbooks are a common resource that is heavily relied on for the learning and teaching of mathematics (Kim & Wang, 2022) and according to Choy et al. (2020) and Rezat et al. (2021) textbooks are one of the main instructional materials. Thus comparative textbook analysis can give new understanding on developing textbooks by recognising the similarities and differences among textbooks (Bayda & Sutliff, 2020; Fan et al., 2013). In a 2022 study by Kim and Pang (2022) on education for sustainable development, a study of the content of textbooks was carried out, to assess the activities related to sustainable development in textbooks from Korea, Singapore and Japan. Since textbooks are regarded as inseparable to teaching and learning, the impact of mathematics textbooks could not be ruled out. In this study it was identified that the analyses of sustainability presented in the textbooks relied on mathematical representations such as tables, graphs, and mathematical equations.

According to Vitta (2021) textbook analysis research on English language teaching textbooks is a quality control check of these products and has become a popular trend in applied linguistics and second language research. It is particularly important because these textbooks have been developed and are seen as supporting language proficiency, while providing everything needed in the lesson, for the teacher. In this study I similarly analyse the Intermediate Phase Mathematics texts as they similarly need to be supportive of learners who are learning English at the same time as learning through the medium of English.

2.2.6 Data handling and statistical literacy

Since data is everywhere and the ability to interpret linguistic and graphical data is imperative, Data Handling is an important topic in the Mathematics curriculum. In FP, learners have their first taste of data handling through collecting basic data about their class and interpreting basic pictographs and bar graphs (DBE, 2011). In IP, the range of data representations that learners are required to generate and interpret expands significantly, and learners are required to start drawing conclusions and making predictions (DBE, 2011). This requires that learners are able to critically analyse data that is presented to them. According to CAPS, learners should work through the full data cycle several times each year (DBE, 2011). It forms part of the specified content for Terms 1 and 3 of each of the grades in the IP (DBE, 2011). This involves collecting, organising, representing, analysing, interpreting and reporting data. CAPS recommends that the contexts for data handling activities should be selected to create an awareness of social, economic and environmental issues. In textbooks and workbooks these contexts are provided through written language as well as symbolic and graphic representations.

The content in data handling builds towards children acquiring statistical literacy. This is an essential literacy due to the constant challenges people face to make decisions that go beyond their knowledge and competences (Engel, 2017). According to Garfield et al. (2010), statistical literacy involves understanding and using the basic language and tools of statistics: statistical terminology, statistical symbols and being able to interpret different representations of data. All of this is gradually built as the curriculum progresses through the school grades.

The ability to discuss personal understandings of data, reactions to data, concerns over conclusions and to communicate about statistical information, is part of being statistically literate according to Gal (2002). Educated citizens should be able to understand basic statistical concepts and interpret and evaluate statistical messages critically so that they are able to identify any abuse of statistical information by the media and individuals in positions of power. Equally important is to be able to communicate the significance and importance of statistics to non-statisticians (Watson & Callingham, 2017). In addition, both Gal (2004) and Watson (2006) agree that fundamental to statistical literacy, is the ability to question claims in social contexts.

The media shares statements and research reports readily, many of which are emotive by nature. Statistics, rather than beliefs and feelings, should be the basis of challenging these and the basis for decision making (Frost, 2013; Ingram, 2015; Tishkovskaya & Lancaster, 2012). Being able to communicate an understanding of statistical data is important so that opinions and implications of the interpreted data can be offered and acceptance of considerations regarding the conclusions provided, be understood by all (Cazorla, Kataoka, & Silva, 2010). According to Monteiro and Teixeira Lima de Carvalho (2020) the above perspective of statistical literacy helps people exercise their citizenship properly as it promotes critical interpretations of statistical information and is an important social need.

Statistical literacy is widely viewed as an important life skill for a functional responsible citizen because of the explosion in information accessibility (Sharma, 2017). Evidence-informed decision making is on the increase and there is concern by governments and producers of large scale data sets, about the capacity of user groups to understand and make intellectual sense from the outputs. Statistical information may have repercussions on people's lives as it refers to many aspects to be found in newspapers, magazines, advertisements and in the work routines of certain professions. Economic indices, sports activities and events, educational rankings, climate forecasts, voting intentions and health are examples of the diversity of aspects statistical literacy. According to Evans (1992), elements of statistical literacy that are related to the types of skills needed by people in everyday life are not limited to the mathematics area. Thus it is an essential requirement for citizens to be statistically literate – it is imperative in defending their rights and vital for democracy. According to English and Watson (2016) people without statistical literacy may not be able to distinguish between probable and improbable information and will have difficulty in understanding, critically evaluating and communicating reactions to such messages. The lack of statistical literacy opens people up to being manipulated because statistics are used against people who are unable to make sense of the statistics. There is thus a social justice aspect to not being statistically literate.

According to Moore (1998) few policy questions have no statistical component and since variation and chance are omnipresent in data, and the media present and analyse data relating to complex issues facing society in a variety of sophisticated ways, our conception of what statistical literacy is, needs to evolve to counterbalance these developments.

Statistical literacy is needed in order to assure active citizen participation in public decision making processes, this need is reinforced by the amount of statistical information on the media, according to Gal (2002) and Engel (2017). Critical statistical literacy would consequently require a basic competence to understand and work with elementary statistical graphs. (Aoyama, 2007; Chick & Pierce, 2012; Ridgway et al., 2019).

Garfield et al. (2010) define statistical literacy as involving understanding and using the basic language and tools of statistics - the knowledge of basic statistical terms, understanding the use of simple statistical symbols, recognising and being able to interpret representations of data. As referred to earlier, data is represented in a variety of ways, both graphical and linguistic.

Wallman (1993, p.1) defines statistical literacy as "the ability to understand and critically evaluate statistical results that permeate our daily lives-coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions". This implies that there is a personal as well as a societal need to develop statistical literacy in learners. Callingham (2007) adds to this by claiming that such a definition requires that students need to develop both the mathematical skills as well as the appreciation of the social context in which the data are set.

The ability to build, read and interpret graphs and to translate these into different representations has been the definition of graphical understanding by different authors. Friel, Curcio and Bright (2001), who defined graphical understanding as the "graph readers' abilities to derive meaning from graphs created by others or by themselves" (p. 132), describe the understanding of graphs to include competences such as recognizing structural elements of a graph and their relationships, assessing the impact of each element on the presentation of the information in the graph, translating the reflected relationships in the graph to the data presented and recognizing the usefulness of a graph, depending on the problem and data being represented, selecting the appropriate graph.

The ability of the child to interpret the representation of the elements and objects in a graph is determined by the graph's semiotic complexity according to Batanero et al. (2010). The author of the graph performs a series of actions and uses concepts and properties that vary in different graphs. Deciding on the use of a particular type of graph and fixing the scale on this graph are examples of how the author establishes semiotic functions which can be more or less difficult for the child to interpret. Examples of these semiotic functions are graphs representing isolated data values like learner's personal data, and representation of a data list, where the data in a list are represented one by one without an attempt to order the data or to combine identical values. When data has been ordered and the frequency for the different values of the variable are obtained in order to interpret the graph, the student needs to understand the idea of distribution. In the case where two or more distributions are represented in the same graph, the interpretation is more complex.

Ben-Zvi and Garfield (2004) suggest that statistical literacy as the lowest of three levels of mental involvement with statistics. This includes basic important skills that may be used in understanding statistical information or research results such as the organization and presentation of data, types of tables and data representation. The understanding of concepts, vocabulary and symbols, as well as an understanding of probability as a measure of uncertainty is part of basic statistical literacy. Statistical literacy is also described by Chick et al. (2005) as 'transnumerative thinking' where students will be able to make sense of and use different representations of data to make sense of the world around them. Since data is all around, social and other media visually display quantitative information, children need to be able to read and recognize data displays.

According to Zewotir and North (2011), statistics is a vital topic at school, since future citizens will be living in a data driven world and will need to possess skills and understanding of issues in data collection, organization, analysis and interpretation, in order to orient themselves. The majority of texts that people are faced with in personal, academic and working lives are based on statistics, e.g. percentages, graphs, rates etc.

In Brazil, in 2000 a working group was established to strengthen statistics education. This was a decisive milestone in bringing interested researchers together (Ribeiro, et al., 2018). This group, GT12, contributed to understanding statistics education as a research area, the epistemology of statistical concepts and the development of materials and teaching methods to promote statistical literacy (Cazorla et al., 2010). The way information is presented in tables, graphs and infographics is linked to the intentions of the developers of the material – who decide what to emphasize, and may omit or mask quantitative and qualitative aspects of the data. The findings of these studies have also indicated that the interpretation of statistical data is not an activity automatically linked to taking in information. Processes relating to the interpretation are established by people, and the context of reading, knowledge of quantity number, proportion, reason, percentage, graphic representation and previous personal experiences are included in the interpretation processes (Monteiro & Teixeira Lima de Carvalho, 2020). Thus, being familiar with the theme of graphs only, does not facilitate interpretation, rather, familiarity is immersed in the relationship between the interpreter and the statistical data presented in the graph. Academic qualifications of participants do not determine their interpretations, rather the person using the information exercising different social roles, bringing beliefs, desires and knowledge linked to their interpretation. When interpreting media graphs, the previous knowledge and experience related to the data being displayed, is mobilized. Thus the process of interpreting data in graphs, is not only made up of formal mathematical knowledge, but immersed relationships with the data in school and out of school contexts (Nunes et al., 1993).

In the South African context the 2003 curriculum paid clear attention to substantial statistics content in direct contrast to previous years when very little or no statistics was included (North et al., 2010). Statistics was introduced into the school curriculum in South Africa for the first time in 2005 (North & Zewotir, 2006). As previously noted, data handling CAPS (2011) is the area of the curriculum where statistical literacy is developed. Children need to understand that information is organized through the use of graphs and charts and that they provide a way of being able to compare data. The interpretation of these graphical representations is reliant on general communication as well as a specific language type. This mathematical language needs well developed CALP.

Watson (2011) describes different tiers of statistical literacy. The first tier focusses on the terminology associated with statistical reasoning and decision making. This includes graphical representations as well as specific phrases. Understanding of this is necessary before being able to make sense of them in context. The second tier focusses on understanding and application of statistical language when it is embedded in the context of wider social discussion. Context forms the central idea in the second tier of statistical literacy. The third tier focus is on critical thinking that enables people to challenge statements made without statistical justification – in order to do this, understanding of both terminology and the context associated with earlier tiers is necessary.

The ability to read and comprehend the mathematical text and interpret the graphical component must be established, before making sense of the information conveyed or solving the problem. When the text is in a second language that must still be translated in context, before even trying to make sense of the problem, these tiers are weakened. The data handling section of mathematics, according to Naidoo and Mkhabela (2017) encompasses real-world situations and assists in developing critical thinking skills in learners and thus it plays an important role in mathematics education.

According to Gal (2002), statistical literacy consists of two broad interrelated components:

1. A knowledge component consisting of five cognitive elements – literacy skills, statistical knowledge, mathematical knowledge, context knowledge and critical questions.
2. A dispositional component, consisting of three distinct concepts – critical stance, beliefs and attitudes.

This is pertinent to this study as the cognitive elements are dependent on the level of language development of the learner. Statistical literacy, the interpretation and representation of data, needs appropriate development of literacy skills in order to 'read' and understand the information presented. Statistical knowledge, mathematical knowledge, context knowledge and critical questions are also dependent on the language skills of the learner, as interpretation of data uses discussion as its means of comprehension. The expectation of the learner in the IP, is to be able to write a paragraph justifying the interpretation of the information presented. This is particularly difficult in a second language, which the majority of learners in SA are learning in.

2.3 Conclusion

In this chapter I presented the theoretical framing of the study. The study is framed by Vygotsky's (1989) sociocultural theory and Cummins's (1981) theory of second language acquisition. Thereafter I presented a review of the literature guided by the key concepts in this study. I reviewed literature on the challenges of learning mathematics in an additional language and the challenge English language learners face when working with mathematical texts in English. I also reviewed literature on reading comprehension, the particular demands of mathematical texts and visual and statistical literacy as well as reviewing studies on textbook analysis.

CHAPTER 3

METHODOLOGY

In this chapter I present the research goal and research questions. I also outline the specific research approach through presenting the ontological and epistemological positioning of the study, and then outlining the methodology. I will also address the issues of validity, reliability and ethics in this chapter.

This study is an interpretivist mixed method case study that analyses the data handling sections of a sample of textbooks used in the South African schooling context.

3.1 Research question

Research questions emerge from identifying questions which are unanswered from previous research. Booth et al. (2003) suggest that contradictions, inconsistencies and incomplete explanations can all provide good starting points for research. The main question in this research is:

> What are the language comprehension demands of English mathematical texts on data handling that are used in South African Intermediate Phase Mathematics?

The sub-questions emerge from the process of descriptive research as described by Mouton and Marais (1996) as the researcher's attempt at discovering patterns or relationships by means of scrutiny of data generated. In the case of this research, the sub-questions expand on the main question in order to capture the linguistic complexity, readability and graphical complexity of the texts. These sub-questions are:

1. What is the readability of mathematical texts on data handling that are used in Intermediate Phase Mathematics?
2. What is the linguistic complexity and what language features contribute to the linguistic complexity of mathematical texts on data handling that are used in Intermediate Phase Mathematics?
3. What is the comprehensibility of the non-textual elements that accompany the mathematical texts on data handling that are used in Intermediate Phase Mathematics?

These questions will be addressed through a close analysis of the data handling sections in four Intermediate Phase book series: three textbook series and one workbook series.

3.2 Ontology and epistemology

A research philosophy outlines the beliefs and values that guide the design of a research study and the collection and analysis of data. It is the researcher's perception to be truth, reality and knowledge (Ryan, 2018). A research paradigm includes the following components: ontology, epistemology, methodology and methods (Scotland, 2012). Since different paradigms have different assumptions and ontological and epistemological views, each research approach is determined by the way of perceived knowledge and reality and reflects these components within its own methodology and methods (Scotland, 2012).

Hudson and Ozanne (1988) define ontology as the nature of reality mainly concerned with the phenomenon in terms of its nature of existence. According to Bryman (2008), the values a researcher holds about what can be known as real and what is believed to be factual is related to ontology. Carson et al. (2001) define epistemology as how reality is being known by the researcher. It is concerned with how the researcher is intending to uncover knowledge to reach reality, it is an internal factor within how the researcher is viewing the world around and it is concerned with distinguishing between right and wrong. It is the belief about how we may come to know the world.

Two different epistemological perspectives of objectivism and subjectivism take different positions of what is real. The position of objectivism is that there is only one version of what is real, irrespective of the researcher's view, and that the only way to find the truth is to observe the world or measure data with as little interference from the researcher or other influences as possible. In comparison, subjectivism takes multiple and varied perspectives of what could be real. Subjectivism states that reality is our own perception, experience or feeling (Alharahsheh & Pius, 2020).

I take the perspective of interpretivism in this research, in which knowledge and the truth are subjective, based on experiences and influenced by cultural and historical factors. The researcher can never be separated from their own values and beliefs and these form the way for collection, interpretation and analysis of data. The interpretivist approach has a relativist ontology: reality is perceived through relationships between facets within the research, for example, the consideration of meanings, understanding of social and experiential aspects (Saunders et al., 2012). Epistemologically, the interpretivist approach is subjective. It provides a clear link between the research subject and the research as it assumes that humans cannot be divided from their knowledge (Saunders et al., 2012).

The origins of interpretivism are in the 18th century with the philosopher Giambattista Vico arguing that there is a difference between the natural and social world and, that social organisation and social experiences form our perceptions of reality and truth (Costelloe, 2016).

Interpretivism considers humans as different from physical phenomena, it has the belief that humans cannot be explored in a similar way to physical phenomena. It is more concerned with in depth variables and aspects related in context, it considers differences of circumstances, cultures and times leading to development which leads to diverse social realities (Alharahsheh & Pius, 2020).

The ontological perspective of interpretivism is of a 'relativist' nature, suggesting that reality is only comprehensible through socially created meanings and that there is no lone shared reality (Ritchie & Lewis, 2003). This perspective reflects the different perceptions individuals have and that there are multiple realities with meanings that are "the categories that make up a participant's view of reality and with which actions are defined… culture, norms, understanding, social reality and definitions of the situation" (Krauss, 2005, p. 758). The interpretivist paradigm accepts that reality can differ in view of different persons, therefore reality is subjective (Alharahsheh & Pius, 2020).

This study is situated in the interpretivist paradigm. According to Lotz-Sisitka et al. (2013) knowledge is constructed in such a study through interaction of researchers and the objects of their study with an epistemological view that knowledge is not objective, but subjective. It is concerned with understanding the world as it is seen from subjective experiences of individuals and perceives reality as existing in the human mind, and conditional on human experiences (Lotz-Sisitka et al., 2013). In the case of this interpretivist study, I will engage with the textbooks and make sense of the characteristics of these texts from my own perspective, with South African learners who are learning in English as an additional language as a particular concern. While the analysis is driven by existing theoretical frameworks, the categorisation of the texts is reliant on my interpretation of the texts.

3.3 Methodology

The general research strategy followed to conduct a research study, would identify and describe the methods to be used. This is known as the methodology which provides clarity on the modes and methods of collecting data (Alharahsheh & Pius, 2020). According to Igwenagu (2016) it is about the research design process, not the methods or instruments involved in the research.

The key influencers of the research methods, procedures and techniques relating to the data handling process, are known as methodological assumptions. These include research strategies, methods and techniques relating to sampling i.e. the size of the selected sample as well as techniques for collection and analysis of data included (Alharahsheh & Pius, 2020).

The nature of the research problem to be explored determines the choice of research method or approach according to Noor (2008). The main research focus of this study is the language features and non-textual elements of the books analysed. The research is descriptive in nature, using both qualitative and quantitative content analysis to develop a comprehensive description of each book with regard to its linguistic complexity, readability and the quality of the non-textual elements on each page.

3.3.1. Mixed method

In this research, quantitative and qualitative data will be gathered from the same document source. Quantitative methods involve quantifying and analysing variables in order to get results and it involves the use of numerical values (Apuke, 2017). In this study, I will be analysing the language features of the texts quantitatively. This will be done through calculations of readability according to established readability formulae. The linguistic complexity will also be calculated as a numerical index using a formula, and the language features contributing to that complexity will be presented according to the frequency of their use in the texts. This is all quantitative in nature.

Qualitative research answers the what, why and how questions (Gay et al., 2009). The visual elements (or non-textual elements) of the data handling texts, will be analysed qualitatively, with their features and functions categorised according to a descriptive framework (Kim, 2012). It is a qualitative evaluation that will result in the categorisation of the non-textual elements. The data will be described qualitatively, but will also be summarised quantitatively according to what proportion of the non-textual elements are assigned to each category.

The selection of a mixed method permits the application of both deductive and inductive reasoning in approaching the analysis of data (Engel & Schutt, 2012). Quantitative research is naturally deductive in nature and according to Engel and Schutt (2012) concepts are theory based and research decisions are made on specific indications of the concept. Inductive reasoning is associated with qualitative research, concepts develop from “the process of thinking about what has been observed” (Engel & Schutt, 2012, p.89).

In this research the dominant method of analysis is quantitative, with qualitative analysis as a supportive method. According to Creswell et al. (2006, p.3) this is called a “nested design” where qualitative data plays a supporting role within the bigger study.

3.3.2 Case study

This research is designed as a case study using Merriam’s (1998) definition of a case study as: “an intensive, holistic description and analysis of a bounded phenomenon such as a program, an institution, a person, a process or a social unit” (p. xiii). Merriam (1998) defines a case as “a thing, a single entity, a unit around which there are boundaries” (p. 27). Noor (2008, p.1602) refers to case study as an “empirical inquiry that investigates a contemporary phenomenon within its real life context using multiple sources of evidence”. The design of a case study is shaped by an overarching methodology and many methods and sources of data can be used (Merriam, 2009; Stake, 2006; Yin, 2014).

When initiating a case study, Stake advises that “for intrinsic case study, case is dominant; the case is of highest importance [and] for instrumental case study, issue is dominant; we start and end with issues dominant (Stake, 1995, p.16). He advises setting one or two focussed research questions to help “structure the observation, interviews and document review” (Stake, 1995, p. 20).

This research will be an instrumental case study, as described by Stake (1998) it places a specific example of the phenomenon being described under close scrutiny to facilitate the “understanding of a particular situation” (Baxter, 2008, p.550). In this research I place examples of mathematics texts that are used in South African Intermediate Phase Mathematics classrooms under close scrutiny in order to provide insight into the complexity of the data handling texts.

A case study should focus on a specific issue. In my research it is the complexity of data handling text in the Grade 4-6 work and textbooks that is in focus. The product of the investigation is an in-depth description of the complexity of the texts, which allows consideration of the suitability of the text for the learners who are using them.

There are criticisms of the case study method, primarily with reliability as it has been argued that there is lack of scientific rigor (Noor, 2004). Gerring (2004) criticises case study for its lack of representativeness in which the question of causal relationships that become apparent for the single unit studied, cannot be presumed to be true of any units not included in the study.

In order to address this critique, I have selected four Intermediate Phase Mathematics book series to analyse. Thus, multiple units are examined in this study. I make no claims that the findings of this study can be generalised to other books not analysed here, but the choice to include multiple cases in the study does allow a more representative picture of the types of texts that learners encounter in the Mathematics classroom than if just one book series was consulted.

There are also advantages to this method as it allows a depth of insight and a "round" (Gerring, 2004, p.346) picture of the phenomenon. It is this 'rounder' quality to the data that is sought in this study. Each book series is discussed in depth according to the language and non-textual components in the texts.

3.4 Method

The method employed in this study is a document analysis of four Intermediate Phase Mathematics book series. The data used in this study is the texts and non-textual elements extracted from the data handling sections of these books. Document analysis refers to the systematic procedure of reviewing or evaluating documents (Bowen, 2009). According to Corbin and Strauss (2008) document analysis requires that data be interpreted in order to elicit meaning, gain understanding and develop empirical knowledge. The process of document analysis entails finding, deciding on, evaluating and synthesising data contained in the documents. The yield of document analysis is excerpts, quotations or complete passages that are organised into themes, categories and case examples specifically through document analysis (Labuschagne, 2003). In this study, the text and images in the documents have been divided into units, each published as an intended whole exercise.

Bowen (2009) suggests that sometimes mixed-methods studies, which include both quantitative and qualitative research methods, include document analysis. According to Stake (1995) and Yin (1994) document analysis is relevant in particular to qualitative case studies, which produce rich descriptions of a single phenomenon, event, organisation or program. Thus, rigorous data collection techniques and the documentation of the research procedure are important.

The advantages of document analysis, according to Bowen (2009) are availability, as most documents are in the public domain and are available without the authors' permission. It is cost effective as the data in the documents has already been collected and the researcher has to evaluate the content and quality of the documents. Documents are not affected by the research process, this is advantageous as they are 'unobtrusive' and 'non-reactive' thus unaffected by the researcher's contribution to the construction of meaning attached to social interactions. Merriam (1998) also suggests that documents are stable due to the investigator's presence not altering what is being studied. According to Yin (1994) the inclusion of exact detail makes document analysis more precise.

There are potential flaws in document analysis, they include insufficient detail as documents are created independent of a research goal – since they have another purpose, they do not always provide enough detail to answer a research question. According to Yin (1994) access to documents may be deliberately blocked, resulting in difficulties in retrieving the documents. Yin (1994) also suggests a 'biased selectivity' or an incomplete collection of documents as a flaw in the process. Bowen (2009) notes that the advantages of document analysis clearly outweigh the limitations, the efficiency and cost effectiveness are the greatest advantages.

3.5 Sample

There are four Intermediate Phase book series that have been selected for inclusion in this study. Three of these are textbooks. Two of the textbook series selected (Platinum series and Premier series) are selected because they are among the most frequently used textbooks in the Eastern Cape, where this research is situated (Mr Schroeder, Publishers Association of South Africa, personal communication). These two book series are listed among those approved by the DBE (DBE, 2021), and need to be purchased by learners or schools. The third textbook series chosen was selected because it is a freely available download. The books were developed by the Sasol Inzalo Institute, in partnership with the DBE. This book series was particularly promoted during the COVID-19 pandemic as a free electronic resource. The fourth book series is the DBE Rainbow workbook series. This is a text that is distributed to all public schools in South Africa. The full references for the books used are listed below. They are referred to as the Platinum books, Premier books, Inzalo books and Rainbow books from this point forward.

All books analysed are compliant with the current Curriculum and Assessment Policy Statement (DBE, 2011).

Platinum Mathematics

Bowie, L., Gleeson-Baird, C., Jones, R., Morgan, H., Morrison, K. & Smallbones, M. (2020).*Platinum Mathematics Grade 4*. Cape Town: Pearson.

Bowie, L., Gleeson-Baird, C., Jones, R., Morgan, H., Morrison, K. & Smallbones, M. (2021).*Platinum Mathematics Grade 5*. Cape Town: Pearson.

Bowie, L., Gleeson-Baird, C., Jones, R., Morgan, H., Morrison, K. & Smallbones, M. (2021).*Platinum Mathematics Grade 6*. Cape Town: Pearson.

Premier Mathematics

Jackson, C. & Raubenheimer, J. (2019). *Shuter's Premier Mathematics Grade 4*. Cape Town: Shuter & Shooter.

Barichievy, M. & Pieterse, K. (2019). *Shuter's Premier Mathematics Grade 5*. Cape Town: Shuter & Shooter.

Tiaden, L., Farrell, N. & John, P. (2020). *Shuter's Premier Mathematics Grade 6*. Cape Town: Shuter & Shooter.

Inzalo books

Human, P., Olivier, A., Lampen, E., le Roux, A., Long, C. & Human, C. (2016). *Mathematics Grade 4 Learner Book*. Rietondale: The Ukuqonda Institute.

Human, P., Olivier, A., Lampen, E., le Roux, A., Long, C. & Human, C. (2016). *Mathematics Grade 5 Learner Book*. Rietondale: The Ukuqonda Institute.

Human, P., Olivier, A., Lampen, E., le Roux, A., Long, C. & Human, C. (2016). *Mathematics Grade 6 Learner Book*. Rietondale: The Ukuqonda Institute.

Rainbow Workbooks

DBE. (2020). *Grade 4 Mathematics in English: Book 1*. Pretoria: DBE.

DBE. (2020). *Grade 4 Mathematics in English: Book 2*. Pretoria: DBE.

DBE. (2020). *Grade 5 Mathematics in English: Book 1*. Pretoria: DBE.

DBE. (2020). *Grade 5 Mathematics in English: Book 2*. Pretoria: DBE.

DBE. (2020). *Grade 6 Mathematics in English: Book 1*. Pretoria: DBE.

DBE. (2020). *Grade 6 Mathematics in English: Book 2*. Pretoria: DBE.

Each data handling section was divided into smaller units for analysis. A unit was defined as a set of text and non-textual elements that were connected either in context or concept. Appendix A provides an example of a unit from the Grade 4 Inzalo book (pp. 92-94). The text and non-textual elements are all connected by the concept of pie charts and the context of Jabu buying fish that together form a coherent and complete unit of work. Appendix B provides an example of a unit from the Grade 4 Rainbow Book 1 (pp. 60-61). These two pages are united by their focus on pictographs in a coherent set of exercises. The units analysed in this research, with their page references, are provided in Appendix C.

3.6 Data Analysis

Each unit was analysed according to its readability and its linguistic complexity.

3.6.1 Readability

Each unit was first assessed for its readability. Readability measures assess a piece of text and compare it to the expected language proficiency of the reader and quantify this as a grade level (Rezaae & Norouzi, 2011). There is debate over their validity, but their use is widespread among writers of educational texts (Rezaae & Norouzi, 2011). Five measures have been selected. The Flesch-Kincaid Grade Level formula measures the number of years of formal schooling that a learner needs in order to interpret a text, as do the Automated Readability Index, the Gunning-Fog formula (Cantos & Almela, 2019), the Coleman Liau Index (Coleman & Liau, 1975) and the SMOG readability formula (McLaughlin, 1969). Each places slightly different weight on characteristics of the texts, and for this reason the average of these readability measures will be calculated for each unit analysed in this research. This will be compared to the grade level of the learners needing to utilise the text. The formulae are:

> Flesch Kincaid Grade Level = (0.39 x average sentence length) + (11.8 x average syllables per word) – 15.59 (Allan et al., 2005, p.5)
>
> Automated Readability Index: Grade Level = 0.50(words per sentence) + 4.71(letters per word) – 21.43 (Thomas et al., 1975)
>
> Gunning Fog Formula: Grade Level = 0.4 x (average sentence length + average number of difficult words) (Allan et al., 2005, p.5)
>
> Coleman Liau Index: Grade Level = 5.89 x (characters/words) – 0.3 x (sentences/words) – 15.89 (https://www.webfx.com/tools/read-able/coleman-liau-index)
>
> SMOG: Grade Level = (square root of the number of polysyllabic words in a sample of 30 sentences) + 3 (McLaughlin, 1969)

In order to assess the readability, the text was copied into the text-checker on the website: https://www.online-utility.org/english/readability_test_and_improve.jsp. This text-checker generates a report on the readability of the text and was used in the studies of Vale, Murray and Brown (2013) and Vale (2018). The readability report contains each of the abovementioned five readability measures. These five results were captured in excel and an average of these readability scores was calculated for each unit. This generated an average approximate grade level reflective of the language proficiency required to comprehend the text. This allowed for a judgment to be made as to the appropriateness of the readability for the grade the particular book was written for.

As an example, the following measures were generated by the online text-checker for the readability of the first unit on data handling in the Grade 4 Inzalo textbook:

Gunning Fog Index:	8.60
Coleman Liau Index:	7.06
Flesch Kincaid Grade Level:	6.80
Automated Readability Index:	5.09
SMOG:	9.21

The average of the five measures is calculated as 7.352. On average, these measures reflect that approximately seven years of formal schooling is required in order to comprehend the text. It is interpreted in this research that this is comparable to the text requiring a South African Grade 7 level of reading.

3.6.2 Linguistic complexity

Shaftel et al. (2006) developed a linguistic complexity checklist to analyse mathematics test items. Shaftel et al. (2006) investigated which language features impacted student performance the most in a general mathematics assessment administered to Grade 4, 7 and 10 students. They define the components that contribute to linguistic complexity of short mathematical texts as including word level, sentence level and paragraph level grammatical features.

At the word level the following were considered to contribute to the complexity of the text: words of 7 letters or more, relative pronouns, slang, homophones, homonyms and the use of abbreviations. Sentence level characteristics that contributed to the complexity of the text included: prepositional phrases, infinitives, complex verbs, complex/compound sentences, conditional constructions and comparative constructions. At the paragraph level the number of cultural- and/or experience-specific references that may be unfamiliar to most learners was considered to add to the complexity of the text, as well as any grammatical and/or spelling errors. The 'Linguistic Complexity Checklist' used is included as Appendix D and appears on the next page. The checklist was applied by counting the frequency of occurrence of the features listed for each unit.

Linguistic Complexity Checklist
Adapted from Shaftel et al. (2006, p. 126)

NUMBER OF SENTENCES: _________

A: BASIC

1. ________ Number of words in item

B: WORD LEVEL CHARACTERISTICS

1. ________ Number of different words with 7 letters or more
2. ________ Number of relative pronouns
3. ________ Number of examples of slang, homophones and homonyms
4. ________ Number of abbreviations

C: SENTENCE LEVEL CHARACTERISTICS

1. ________ Number of prepositional phrases
2. ________ Number of infinitives
3. ________ Number of complex verbs
4. ________ Number of complex / compound sentences
5. ________ Number of conditional constructions
6. ________ Number of comparative constructions

D: PARAGRAPH LEVEL CHARACTERISTICS

1. ________ Number of cultural- and/or experience-specific references
2. ________ Number of grammatical and/or spelling errors

Linguistic Complexity Index:

LCI = (Sum A + Sum B + Sum C + Sum D) ÷ Number of sentences

= __________

In addition to reporting on the frequency of the appearance of the linguistic features listed, a Linguistic Complexity Index was calculated in order to facilitate comparing texts to one another. This was derived by Vale (2018), from Shaftel et al.'s (2006) checklist, to allow comparison of short mathematical texts. These values will allow the comparison of the texts analysed in this research. The formula is as follows:

LCI = (number of words + word level characteristics + sentence level characteristics + paragraph level characteristics) ÷ the number of sentences.

This index doesn't approximate a grade level for the text as the readability measures do, but the values generated can facilitate the comparison of texts to one another in terms of linguistic complexity.

The frequency of use of each of the language features in the linguistic complexity checklist was captured in excel. The formula was applied in excel to calculate the linguistic complexity index. From the spreadsheet generated it was possible to identify which language features were contributing to the index with the highest frequencies.

3.6.3 Non-textual elements

Data handling texts are characterised by many visual elements in the form of charts and graphs and other images. The visual representations and graphics add complexity to text as there is need to have a understanding of content and context both mathematically and graphically. (Lowrie & Diezmann, 2007). Any analysis of the complexity of the text would be incomplete without considering these non-textual elements. The work of MacKinlay (1986), provides a typology of graphical languages that appear in technical texts. These are summarised in the table below, many referring directly to representations of data. Each image in the textbooks or workbooks analysed was classified as belonging to one of these languages. This is a classification and not a measure of complexity or comprehensibility.

Table 1. Graphical languages

Language	**Information encoded by**	**Example**
Single-position	The position of a mark set on one axis	Horizontal axis, vertical axis
Apposed-position	The position of a mark set between two axes	Line chart, bar chart, plot chart
Retinal-list	Retinal properties of marks in a set, independent of position	Colour, shape, size, saturation, texture, orientation
Map	Fixed position with graphical techniques specific to maps	Road map, topographical map
Connection	A connected set of node objects with a set of link languages	Tree diagram
Miscellaneous	A variety of additional graphics	Pie chart, Venn diagram

summarised from MacKinley (1986, pp. 127- 130)

Kim (2012) distinguishes between textual and non-textual elements of mathematical texts. The non-textual elements include the visual representations of data. According to Kim (2012) very little effort has been made to identify systematically the roles and characteristics of non-textual elements in mathematics textbooks, despite the importance of visual representations in learning and teaching. Based on results from a few international pilot studies, Kim (2012) developed a final framework for non-textual elements. This framework's fundamental aspects, 'accuracy' and 'connectivity' are suggested to be the aspects that every non-textual element should possess. The other aspects, ie. 'contextuality, aesthetics and simplicity' are useful aspects that could facilitate the mathematics learning process. Kim (2012) proposes a framework for analysing the quality of non-textual elements of mathematics texts through focusing on their accuracy (whether the image portrays the concept accurately); connectivity (whether the image is fully associated with the mathematical content), conciseness (whether the image is straightforward or contains distractors) and contextuality whether the context and associated image are realistic). The assumption in this study is that a non-textual element that is accurate, connected, concise and contextual would be more comprehensible, and would better support comprehension of the text, than one that is lacking in accuracy, connectivity, conciseness and contextuality.

I use the same coding scheme as proposed by Kim (2012). This is provided as Appendix E and appears on the next page.

Aspect of non-textual elements	Score	Description
Accuracy	2	An NTE correctly shows the definition of a concept, or it is accurate to show a concept based on its definition
	1	An NTE makes sense in terms of the definition or meaning of a concept. But it does not show every required mathematical condition (e.g. some requisite notations are missing or misleading) or some attributes are not appropriate to explain a concept
	0	An NTE is inaccurate in terms of the definition of a concept (e.g. it has an obvious error). Or an inappropriate realistic object is used to present a concept. There is a major error or concern to use the realistic object or context for a concept. Or there is no mathematical concept in the NTE
Connectivity	2	An NTE is explicitly and fully associated with the mathematical content in the text. It directly shows a concept or problem
	1	An NTE is partly related to the mathematical content in the text. There is some missing or irrelevant information in an NTE. It shows the content but it does not explicitly show how it is connected with the content
	0	An NTE has nothing to do with the content mathematically. It can give some clue about contexts in texts (e.g. river when the problem is about length of river)
Conciseness	2	An NTE is straightforward to show a concept or problem without any distracting or other factor
	1	An NTE is straightforward to show a concept or problem with some other factors that might be helpful for the concept
	0	An NTE has distracting or other factors that are useless in addition to factors needed to show a concept or problem
Contextuality	2	A realistic object or context is used in an NTE with mathematical connection
	1	No mathematical ideas or concepts exist, but there is some realistic contextual information (that is used to provide contexts or objects in the problem or to facilitate related activities)
	0	Neither realistic object nor realistic context is included in an NTE

Below is a description of what each code indicates, according to Kim (2012, p. 180):

The *accuracy* of the concept of representation of data explains how accurately the NTE shows the definition of the concept. A coding of 2 means that the concept is accurate to show a concept based on its definition. A coding of 1 means that the NTE makes sense in terms of the definition or meaning of the concept but does not show every required mathematical condition or some attributes are not appropriate to explain a concept. A coding of 0 on the NTE accuracy key, means that the NTE is inaccurate in terms of the definition of a concept, has an obvious error , there has been misuse of a realistic object to present the concept or there is no mathematical concept in the NTE.

The *connectivity* of the NTE describes the association of the NTE with the mathematical concept in the text. The key for a code 2 for connectivity means that the NTE is directly showing a concept or problem. A coding of 1 for connectivity means that the NTE is partly related to the mathematical concept in the text. There may be some missing or irrelevant information in the NTE. It shows the content but not how it is explicitly connected to the concept. A coding of 0 for connectivity means that the NTE has nothing to do with the content mathematically. It may give a clue about the context in the text.

Conciseness of the NTE means the straightforwardness of the representation of the concept. A coding of 2 means that the NTE has no distracting features. A coding of 1 means the NTE is straightforward to show a concept or problem with some other factors which might be helpful for the concept or problem. A coding of 0 means that the NTE has distracting or other features that are useless, in addition to factors needed to show a concept or problem.

Contextuality of the NTE is the appropriate representation of data for the situation/problem for the context. A coding of 2 means that a realistic object or context, with a mathematical connection was used in the NTE. A coding of 1 means that no mathematical ideas or concepts exist, but there is some realistic contextual information, that is used to provide contexts or objects in the problem or to facilitate related activities. A coding of 0 means that neither realistic object nor context is included in an NTE.

3.7 Piloting the instruments

Researchers pilot a study in order to establish the existence of the phenomenon and to test the instruments they plan to use as well as the analysis technique and in order to establish inter-rater reliability. Armstrong et al. (1997) define this as a recognised process in quantitative research where data is independently coded and these codings are compared for agreement. Should disagreement occur, discussion and consensus take place. This is important for the researcher to finalise the criteria which will be used to classify the data.

According to Polit et al. (2001, p. 467) a pilot study is a "small scale version[s], or trial run[s], done in preparation for the major study". A pilot study is one of the most important stages in a research project and is used to identify potential problem areas and deficits in the instruments and practice before implementation during the full study. It is an opportunity to determine whether the project is feasible and to present results.

It is important to determine the sample size and selection, the criteria for the study based on the objectives of the study. It tests the manageability and practicality of the instruments and is a trial run for data analysis. Should components be deemed infeasible or unsatisfactory, they can be modified or removed.

The research question this study wants to answer is to find out what the language demands of mathematical texts on data handling that are used in Intermediate Phase Mathematics classrooms are. The complete study analyses four Intermediate Phase Mathematics (Grades 4-6) book series, to establish the language demands of the text and visual elements used in the exercises on data handling. In the piloting phase I analysed the data handling sections of one Grade 3 book and one Grade 7 book to test the use of the instruments and to establish inter-rater reliability in the coding.

First, each book was divided into manageable, logical units of text and non-textual elements. The rationale behind this was that measuring the complete amount of text and all the non-textual element would not be practical or give a concise indication of the linguistic, readability or non-textual elements. The units naturally separated themselves, as the data handling sections are based on themes and have contextual similarities.

Thereafter, I typed out all the text per unit and put the text through the online readability text checker (https://www.online-utility.org/english/readability_test_and_improve.jsp) to generate the readability scores. Five measures were generated. As expected, the readability scores for the Grade 7 text were much higher than the readability scores for the Grade 3 text. There was however a discrepancy in scores across the measures for some units. This is likely because the different formulae emphasise different aspects of the text. For this reason the decision was taken that for the final study the readability scores would be reported as an average of the five measures generated by the online utility.

Next, the text was then analysed using the Shaftel et al. (2006) linguistic complexity checklist. I counted the frequency of each language feature for each unit. This was also done by my supervisor and we compared the resulting frequencies. There was very little difference in the frequencies reported. Where there were differences, this was due to miscounting and not because of disagreement in how the language feature should be categorised. There was no need to adapt the instrument in any way. The credibility of the instrument was confirmed in this pilot study.

Finally, the non-textual elements [NTEs] in each unit were identified. Two broad categories were identified – the mathematical non-textual elements, which consisted of tables, graphs and charts and non-mathematical non-textual elements, which consisted of images that were not mathematical in nature, for example images that were decorative or that provided supporting contextual information. The NTEs were categorised according to their graphical language. During this process it became clear that a new category in addition to those proposed by MacKinley (1986) was needed as no category accounted for the images that were decorative or provided extra contextual information. A category was added that was labelled 'images'. It was also necessary to differentiate between tables of text and numerical information and other apposed position languages. Tables are apposed position languages because of their arrangement into vertical columns and horizontal rows, but it was desirable to distinguish these from the apposed position graphs. An additional category was thus added for 'tables' because of their frequency of use. It thus became possible to distinguish between tables and graphs.

The NTEs were the coded according to the descriptions of accuracy, connectivity, conciseness, and contextuality of Kim (2012). Here there were discrepancies between the two raters that necessitated discussion to ensure that there was a shared understanding of the categories. Shared understanding was established and agreement was reached about the coding of the NTEs.

3.9 Ethics

The *Ethics in Health Research: Principles, Processes and Structures* (Department of Health, 2015) document indicates that when research depends on material that is already in the public domain and no human participants are involved it does not require ethical clearance. However, it remains important to adhere to ethical research practices. In the case of this research, I have no affiliation with the authors or publishers of any of the textbooks or workbooks, there is thus no conflict of interest. In confirmation of this I received ethical clearance from the Rhodes University Education Research Ethics Committee (clearance number 2020-2804-4841).

3.10 Validity and reliability

Creswell (2014) notes that the combination of qualitative and quantitative data allows for a more comprehensive understanding of the problem, which contributes to the validity and reliability of the results of the analysis in this research.

In order to ensure the validity of the measures used in this research, instruments were selected that have been used in prior research in mathematics education. The linguistic complexity instrument has been used with South African mathematics texts (see Vale, Murray & Brown, 2012; Vale, 2018; Sibanda & Graven, 2018), and the frameworks categorising and analysing the visual representations have been applied in mathematics education research (Kim, 2012). The five readability measures are also widely accepted as measures of the difficulty of texts (Cantos & Almela, 2019). The online readability text-checker has been previously used in the South African studies by Vale, Murray and Brown (2013) and Vale (2018) in which mathematics texts were analysed. All of the documentary sources are also authentic texts that are currently used in South African mathematics classrooms, and this authenticity is essential for the validity of the results.

As reported in section 3.7, the instruments were piloted with a Grade 3 and a Grade 7 text. This enabled the establishment of inter-rater reliability for the application of the Linguistic Complexity Checklist, and the categorisation of the non-textual elements according to Kim's (2012) framework. This process has enhanced the validity and reliability of the findings.

Rigour and quality conceptualize validity in qualitative research, and through attending to these the researcher looks to eliminate bias and increase the trustworthiness of data (Denzin, 1978). In this research, critical friends who are experts in language and primary mathematics were consulted to verify coding for a sample of items at the beginning of the analysis process. Where disagreement in coding occurred, these items were discussed and coding mutually decided. Thereafter, I carried out the remainder of the analysis, checking a final sample of the data with these critical friends at the end of the analysis.

3.11 Conclusion

In this chapter I have positioned this research as an interpretivist study. It is a mixed method case study that takes the form of a document analysis of four Intermediate Phase Mathematics book series. I presented the research question and the sub-questions that the study seeks to answer and in the following two chapters I will present the data that addresses these questions. I also presented the instruments that will be used to analyse the data in terms of the text's readability and linguistic complexity and the non-textual elements' comprehensibility. Issues of ethics and validity and reliability have also been addressed in this chapter.

CHAPTER 4

DATA PRESENTATION: LINGUISTIC COMPLEXITY

In this chapter the readability and linguistic complexity of the text and workbooks are presented in order to answer the research question.

The analysis of the books is presented in the following order: the Platinum textbooks, the Premier textbooks, the Rainbow workbooks and finally the Inzalo textbooks. The readability is considered first, followed by a presentation of the linguistic complexity of the text within each book.

As described in Chapter 3, the readability was calculated as the average of the following readability measures: Gunning Fog Index; Coleman Liau; Flesch Kincaid Grade Level; Automated Readability Index and the SMOG. The five readability measures used to calculate that average were generated by the online text-checker. The readability scores are reported as equivalent to a grade level and are summarised in straight line graphs. To facilitate the comparison across the texts within and between the books, the vertical axis of each graph is standardised with a maximum readability score of 12, considered equivalent to South African Grade 12.

The linguistic complexity was analysed by performing a count of the frequency of use of the linguistic features listed by Shaftel et al. (2006) as contributing to the complexity of texts in mathematics learning and support materials. As a further comparative measure, a Linguistic Complexity Index was calculated using the formula created by Vale (2018), based on Shaftel et al.'s (2006) Linguistic Complexity Checklist. This formula is:

LCI = (number of words + word level characteristics + sentence level characteristics + paragraph level characteristics) ÷ the number of sentences.

This index does not indicate a grade level of linguistic complexity but rather allows for internal comparison between the units within the books and between the books. In order to highlight the language features contributing the most to the LCI, the language features with the three highest frequencies are presented in yellow, per unit, in the tables summarising the linguistic features of each text. The LCI per unit is summarised in bar graphs. Similar to the readability graphs, the vertical axes of all the LCI bar graphs are standardised with a maximum of 22, as this was the maximum LCI calculated, to allow for comparison between books.

4.1 The Platinum textbook series

The Platinum textbook series is one of the most widely used textbooks in South African schools and is on the DBE published list of resources. This textbook is produced on high quality paper and has colour used on every page. The layout of the pages is balanced, with minimal overcrowding or distraction from the content focus. The price per book is R117,60 (DBE, 2021).

4.1.1 Average readability of the Platinum textbook

Appendix F provides all the readability scores for the Platinum textbooks.

In the Grade 4 Platinum textbook there are 12 units addressing the topic of data handling. Figure 4.1 displays the average readability score for all 12 units of the textbook, with a trendline showing a gradual increase in readability score as the book progresses. While the trendline shows a gradual increase, the profile of the graph is jagged, with some units at a Grade 3 readability level and one at a Grade 8 level. This book is aimed at Grade 4 learners, but with the trendline indicating a level of between Grade 5 and 6 readability, it is unlikely that the text would be accessible to Grade 4 learners, particularly those who do not have English as their home language.

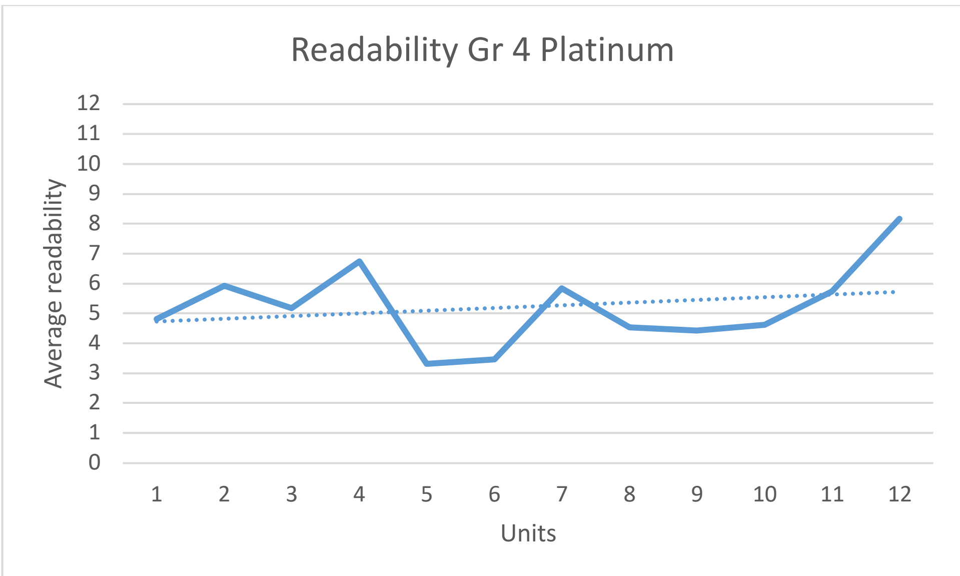


Figure 4.1 Readability scores for Grade 4 Platinum book

In the Grade 5 textbook in the Platinum series, there are 19 units focused on the content area of data handling. Figure 4.2 shows uneven readability levels, as Unit 1 is almost at a Grade 8 readability level, and Unit 4 is lower than a Grade 4 readability level. The trendline shows an increase in the readability levels as the data handling units progress.

The trendline shows that the average readability level of the Grade 5 textbook is between Grade 5 and Grade 6 level, which is appropriate. However, it is notable that five of the twelve units are above Grade 6 level. This likely makes those sections of the book inaccessible to Grade 5 learners if working through the text independently.

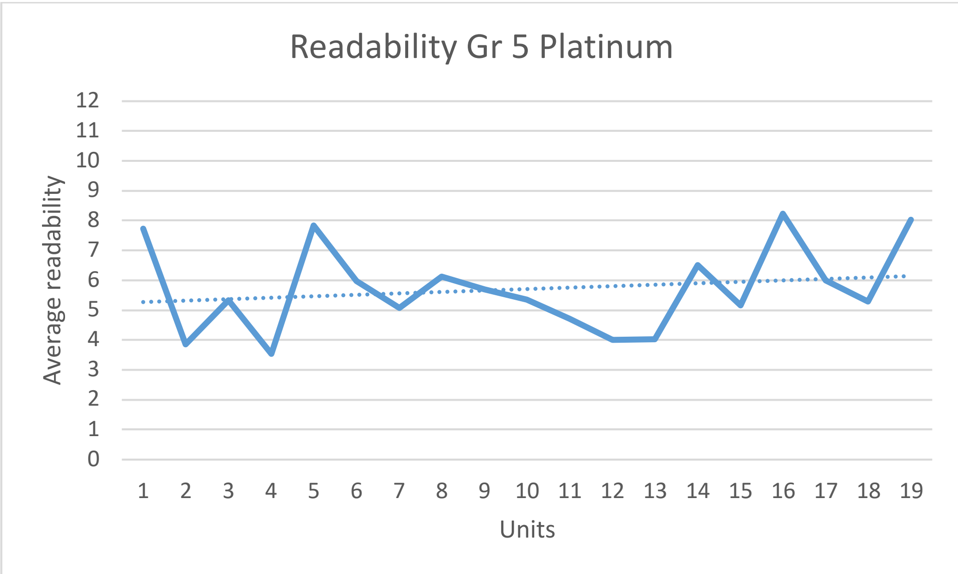


Figure 4.2 Readability scores for Grade 5 Platinum book

In the Grade 6 Platinum textbook, there are 20 units focused on data handling. The readability graph, Figure 4.3, can be described as ‘jagged’ in profile, with the readability scores ranging widely between Grade 3 and Grade 10 levels. The trendline also gives an indication of an increase in the readability score as the book progresses. Eight units (40%) are above Grade 6 level, which indicates that almost half of the text is likely out of reach of learners.

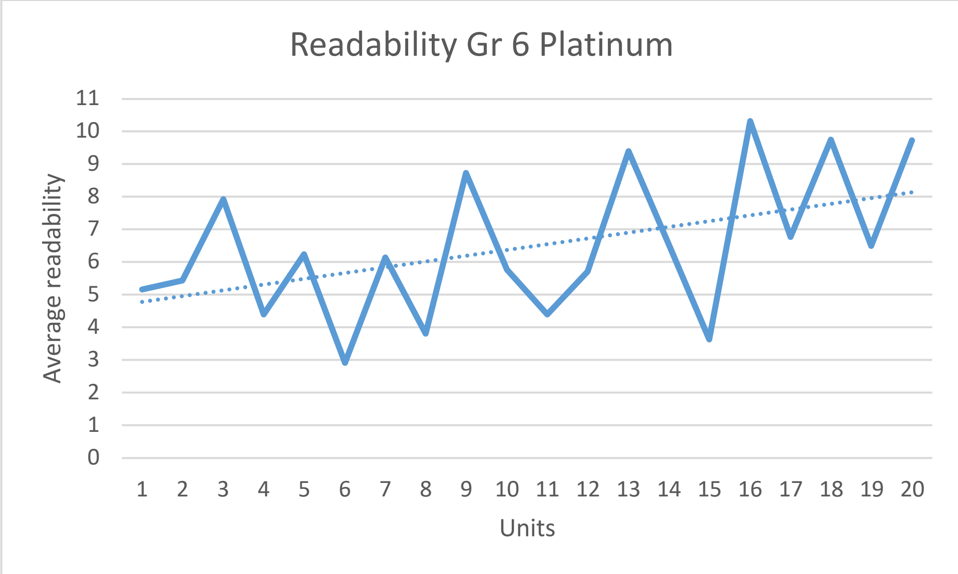


Figure 4.3 Readability scores for Grade 6 Platinum book

The Platinum textbook series has more data handling units, and therefore more text available for analysis, than the other three series selected for this study. The readability scores for each book in the series are generally higher than the grade they have been written for. This is more evident in Grade 6 book, than in the other two grade books. All three of the books in the series contain individual units that utilise text that is potentially out of the reach of learners in the grade concerned, particularly when considering that the majority of the learners may be reading at a level lower than their grade level. This implies that a high level of teacher-mediation would be required for these texts.

4.1.2 Linguistic complexity of the Platinum textbook

The linguistic complexity of each unit was analysed using the linguistic complexity checklist designed by Shaftel et al. (2006) which requires the text to be analysed for the grammatical features at word, sentence and paragraph level. In this section the frequency of the use of these grammatical features per unit is presented in a table. In the table, the language features with the three highest frequencies per unit are highlighted in yellow in order to emphasise which language features are contributing the most to the resulting Linguistic Complexity Index. A bar graph is used to summarise the Linguistic Complexity Index of each unit.

Grade 4 Platinum textbook

Table 4.1 summarises the language features in the Grade 4 Platinum textbook, per unit.

Table 4.1 Linguistic features of the Platinum Grade 4 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural reference	Errors	LCI
1.	44	345	25	3	2	0	4	11	3	0	0	1	0	1	9,02
2.	19	161	13	1	1	0	1	5	0	0	0	0	0	0	9,68
3.	30	258	13	1	4	0	4	2	2	0	0	4	0	0	9,9
4.	18	169	13	1	1	0	1	2	2	0	0	1	0	0	10,6
5.	31	203	13	1	1	0	2	0	0	0	0	0	1	0	7,22
6.	22	138	10	1	0	0	0	0	2	0	0	2	0	0	7,09
7.	15	148	12	1	1	0	1	3	3	0	0	0	0	0	11,13
8.	29	235	12	0	3	0	11	6	2	0	1	0	0	0	9,34
9.	45	416	27	0	2	1	18	10	9	2	0	2	0	0	10,84
10.	26	216	17	0	3	0	16	3	4	0	0	0	0	0	10
11.	48	543	33	5	0	0	24	4	11	4	3	0	0	0	13,14
12.	31	418	35	2	2	0	22	10	12	2	0	0	0	0	16,29

The number of words and sentences per unit is varied. The shortest text is Unit 6 with 138 words. Units 9, 11 and 12 all have more than 400 words, with Unit 11 being 543 words in length. This is a large amount of text for a Grade 4 learner to decode if working independently.

At the word level, words with 7 letters or more contributed the most to the resulting LCI of every unit. All 12 units contained a significant number of words with 7 or more letters. Among these are mathematical terms, e.g. pictographs, vertical, horizontal, fractions, which learners need to become familiar with as part of the outcomes. Also included in the words of 7 or more letters are: popular, dwellings, programme.

At the sentence level, prepositional phrases (8/12 units), infinitives (6/12 units), and complex verbs (6/12 units) contributed with the highest frequencies to the resulting LCI of each unit. These language features add complexity to the text and can make the text less accessible.

A prepositional phrase is a word group that begins with a preposition. A preposition is a joining word that links a noun to another word in a sentence. Prepositional phrases were present in 11 of the 12 units. They can potentially confuse a reader because they mark the existence of an additional phrase in the sentence and hence another concept to be understood (Shaftel et al., 2006). They are, however, necessary when describing how nouns relate to one another. In the

Platinum textbook, the majority of prepositional phrases were ‘in’ and ‘on’. For example, in Unit 4 there is the sentence: “Look at the bar graph about litter **in** a school playground”.

An infinitive refers to the infinitive form of a verb plus any complements and modifiers. Because infinitives begin with the word ‘to’ they are occasionally misidentified as prepositional phrases. In this book, infinitives were present in 10 of the 12 units. For example, in Unit 7 there is the sentence: “This year you will collect data **to answer** questions about different topics”.

Complex verb phrases consist of one or more auxiliary verbs plus a main (lexical) verb. This suggests the use of multiple or difficult verb tenses (Shaftel et al., 2006). Complex verbs were frequently used in this book. They are used in 10 of the 12 units. One example includes the following sentence in Unit 11: “For homework, a grade 4 class was asked to find out what type of TV programme their friends liked best”.

Paragraph level complexity is reflected when there is reference to cultural events or other experiences that learners may not be familiar with. There is 1 instance of added complexity through an experience-specific reference in this book. Unit 5 makes reference to game viewing which might not be an experience that all learners can relate to. This would imply that the vocabulary and descriptions used in the unit may be unfamiliar to learners.

Figure 4.4 summarises the LCI across the units in the Grade 4 Platinum book.

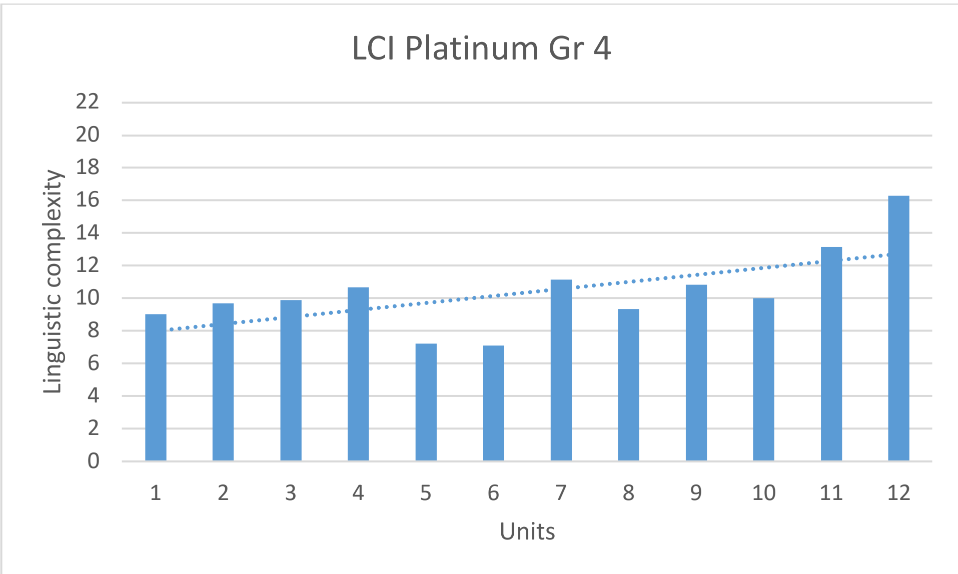


Figure 4.4 Linguistic Complexity Index for Grade 4 Platinum book

The LCI varies between approximately 7 and 16, which is a relatively large variation. The trendline indicates a gradual increase in linguistic complexity over the course of the book.

Grade 5 Platinum textbook

Table 4.2 presents a summary of the linguistic features in the Platinum Grade 5 text.

Table 4.2 Linguistic features of the Platinum Grade 5 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrase	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural reference	Errors	LCI
1.	18	255	23	0	7	1	5	11	11	5	0	0	0	0	17,38
2.	25	231	15	0	3	2	4	14	2	4	0	0	0	0	10,84
3.	10	531	8	5	3	0	0	2	1	2	0	2	0	0	16,2
4.	26	209	17	3	0	0	6	3	3	3	2	2	0	0	9,42
5.	14	165	16	1	4	0	3	6	3	3	0	1	0	0	14,42
6.	19	212	17	0	4	0	11	2	2	1	1	0	0	0	13,10
7.	28	249	18	6	2	0	9	4	0	4	1	2	1	0	10,5
8.	34	296	37	1	1	0	13	1	5	2	2	1	0	0	10,55
9.	29	359	18	4	2	0	5	3	17	9	0	3	0	0	14,78
10.	27	253	19	0	5	0	7	9	2	3	0	2	0	0	11,07
11.	33	324	21	2	1	0	9	6	3	5	0	1	0	0	11,27
12.	14	139	9	3	4	0	5	1	0	1	0	0	0	0	11,57
13.	29	226	19	1	2	0	9	1	4	3	0	0	0	0	9,03
14.	22	258	25	0	3	0	6	7	4	3	0	2	0	0	14
15.	33	268	18	0	6	0	9	3	3	3	0	0	0	0	9,39
16.	19	292	19	2	4	0	7	3	3	3	0	1	0	0	17,47
17.	35	325	31	5	2	0	26	1	5	5	0	0	0	0	11,34
18.	30	283	19	3	5	1	20	2	2	4	0	0	0	0	11,2
19.	38	535	35	4	3	0	20	16	8	7	0	0	0	0	16,36

As with the Grade 4 book, there is some variance in the length of the texts in these units. The shortest text is Unit 12 with 139 words. There are two units (Units 3 and 19) which have more than 500 words. These two units also have among the highest LCIs at 16,2 and 16,36 respectively.

At the word level, the language feature contributing with the highest frequency to the linguistic complexity of every unit in this book is words with seven letters or more. Also contributing with high frequency at the word level are relative pronouns (4 units) and slang, homophones or homonyms (8 units).

The use of slang, homophones and homonyms introduces complexity to a text because not knowing the definition of a particular homonym or homophone can change the meaning of what is read, thus affecting comprehension. For this study, homophones were those words with multiple meanings, like ‘table, key, bar, cell’. These words have more than one meaning, but the implied meaning in the units, is mathematical

Relative pronouns were among the features with the highest frequencies for 4 units. These are pronouns that mark a relative clause, for example, “Nurse Khumalo recorded the names of children **who** visited her clinic on Monday” (Unit 1). Pronouns are important in sentence constructions, even though they may bring difficulty to mathematical text, as they serve to indicate possession and to form questions, among other uses. ‘That’ and ‘which’ pronouns featured in slightly above half the units.

At the sentence level it is prepositional phrases (15/19 units), infinitives (8/19 units), complex verbs (6/19 units) and complex/compound sentences (4/19 units) that contribute with the highest frequencies to the LCI.

Complex verbs were frequently used in this book. Seventeen of the units included complex verbs, for example, “was given” (Unit 10). Unit 9 has the most complex verbs, totalling 17. These included “have to think”, “was collected”, “were sick”, “had been sick”, and “were biased”. Infinitives also appeared as a language feature contributing frequently to the complexity of the text. Unit 2, for example, has 14 infinitives (e.g. “ It is also useful to know the mode of the data”).

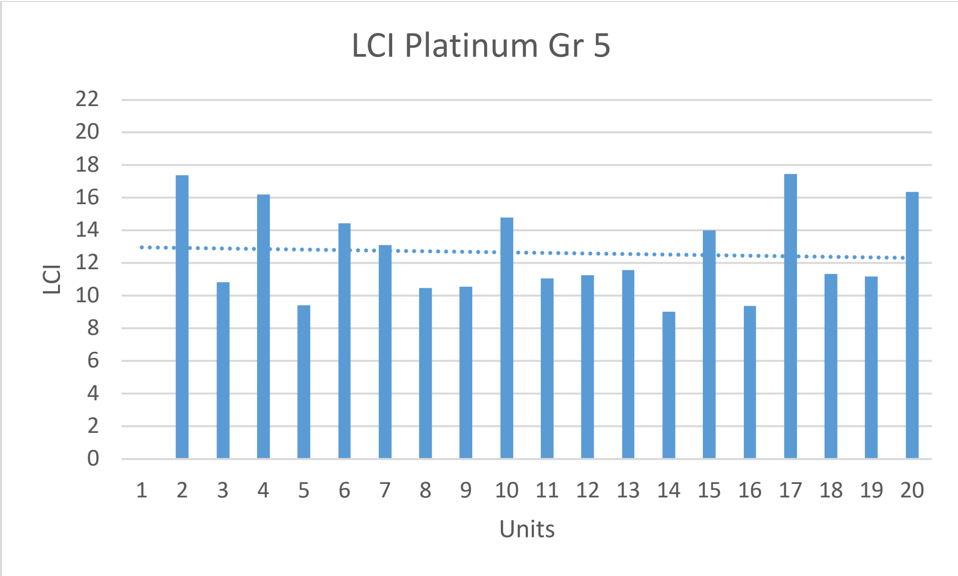


Figure 4.5 Linguistic Complexity Index for Grade 5 Platinum book

As shown in Figure 4.5, the units in this book scored between 9 and 17,47 on the LCI, the majority of the units between 9 and 12. This is a large amount of variation and represents a higher maximum LCI than the Grade 4 Platinum book, which had a maximum of 16,29. In general, the units in the Grade 5 book are more linguistically complex than in the Grade 4 book.

Grade 6 Platinum textbook

Table 4.3 presents a summary of the linguistic features in the Platinum Grade 6 text.

Table 4.3 Linguistic features of the Platinum Grade 6 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural reference	Errors	LCI
1.	36	270	32	2	4	0	10	4	4	2	0	0	0	0	9,11
2.	36	245	28	4	1	0	6	11	7	3	0	0	0	0	8,47
3.	7	85	11	1	2	0	2	2	1	0	0	0	0	0	14,85
4.	19	140	15	1	3	0	9	2	2	1	0	0	0	0	9,10
5.	20	287	15	2	4	0	4	6	4	0	0	1	0	0	16,05
6.	45	280	9	4	3	0	11	3	5	0	0	0	0	0	7

7.	22	378	15	0	3	0	15	0	1	1	0	0	0	0	18,71
8.	39	322	20	3	1	0	6	0	1	1	0	0	0	0	9,07
9.	17	233	33	5	1	0	3	2	3	1	0	0	0	0	16,52
10.	36	261	26	4	3	0	11	1	1	1	0	0	0	0	8,55
11.	11	56	10	0	5	0	1	2	1	0	0	0	0	0	6,81
12.	21	168	20	0	4	0	5	3	7	1	0	0	0	0	9,90
13.	7	98	12	0	0	0	3	2	1	0	0	0	0	0	16,57
14.	28	281	31	0	4	0	7	10	4	0	0	0	0	0	12,03
15.	20	331	11	3	0	0	3	3	3	1	1	0	0	0	17,8
16.	15	238	24	3	3	0	7	0	6	1	0	0	0	0	18,8
17.	47	441	40	9	2	0	7	5	2	0	0	0	0	0	10,76
18.	24	337	29	1	1	0	13	2	3	2	0	0	0	0	16,16
19.	42	461	27	0	0	0	24	4	6	1	1	0	0	0	12,47
20.	22	397	40	1	4	0	12	9	6	2	0	0	0	0	21,22

The length of the texts in the units in the Grade 6 Platinum book vary from 56 words (Unit 11) to 461 words (Unit 19). At the word level, as with the Grade 4 and Grade 5 books, it is words of 7 or more letters that contribute with a high frequency in all units to the linguistic complexity of the texts. Relative pronouns are among the language features with the three highest frequencies for 5 units, and slang, homophones and homonyms contribute among the three highest frequencies for 6 units.

At the sentence level, as with the Grade 4 and Grade 5 books, it is prepositional phrases that contribute among the three highest frequencies for most units (18/20 units). Also contributing with high frequency to the linguistic complexity are infinitives (8 units) and complex verbs (9 units). All units contain complex verbs and 17 of the 20 units contain infinitives.

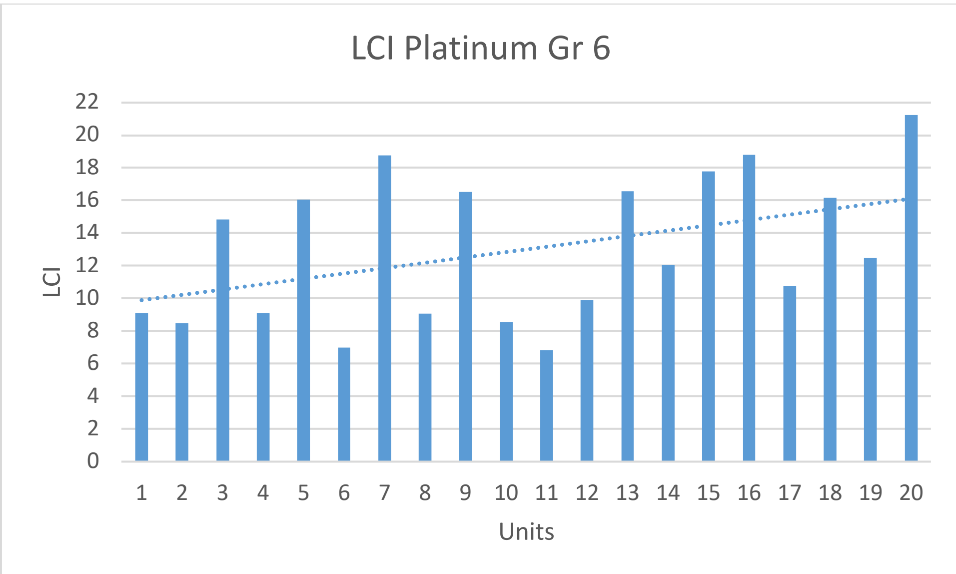


Figure 4.6 Linguistic Complexity Index for Grade 6 Platinum book

The LCI is summarised in Figure 4.6. For this book it is between approximately 7 and 21. This is an even larger variation than in the Grade 5 Platinum book. The trendline indicates a sharp increase in linguistic complexity as the book progresses. The maximum LCI for this book is 21,22 for Unit 20, which is the highest LCI among any of the books in this research.

4.2 The Premier textbook

The second book investigated in the study is the Shuters Premier textbook. This is a textbook that is on the DBE list of approved textbooks and is available for schools and learners to purchase. It is printed in black and white. The price per book is R90,25 (DBE, 2021). A presentation of the data follows.

4.2.1 Average readability of the Premier textbook

Appendix G provides all the readability scores for the Premier textbooks.

The readability of the Grade 4 Premier textbook varies widely across the 12 units in the book. As shown in Figure 4.7, the trendline is upwardly moving and the graph is jagged, with Unit 7 at a Grade 3 level and Unit 8 peaking at Grade 9 level of readability.

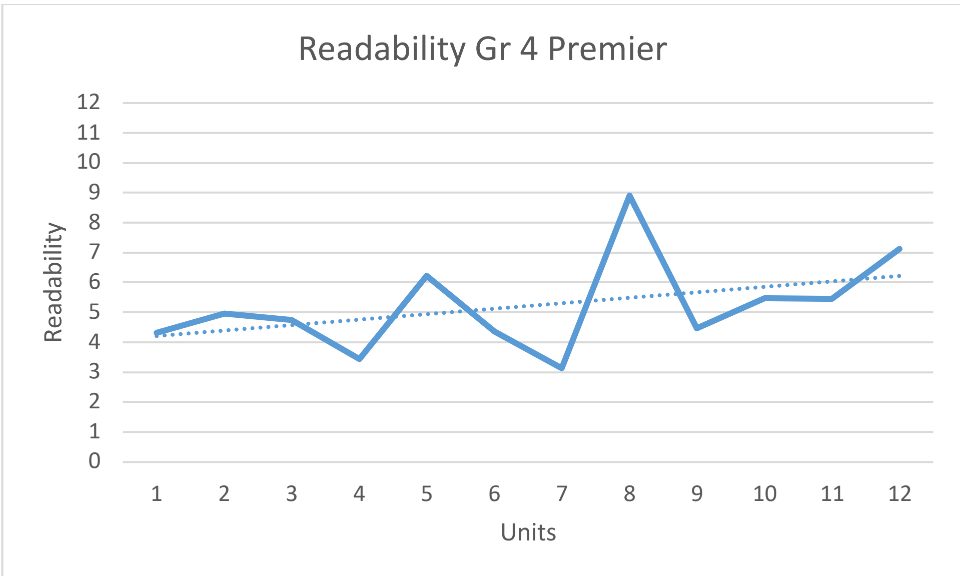


Figure 4.7 Readability scores for Grade 4 Premier book

In the Grade 5 Premier book there are 10 units that are focused on data handling. As shown in Figure 4.8, the first unit is more difficult in terms of readability than all the other units, with a readability score of approximately 7,5, and the last unit is one of the easiest to read with a readability score at approximately a Grade 4 level. Most of the units lie between Grade 4 and Grade 6 on the average readability level. The profile of the graph is jagged at first and then it flattens out, with the trendline indicating a slight decrease in the readability score as the book progresses. Most of the units are at approximately a Grade 5 level of readability, which is appropriate. The following graph shows the readability of each unit in the Grade 5 book.

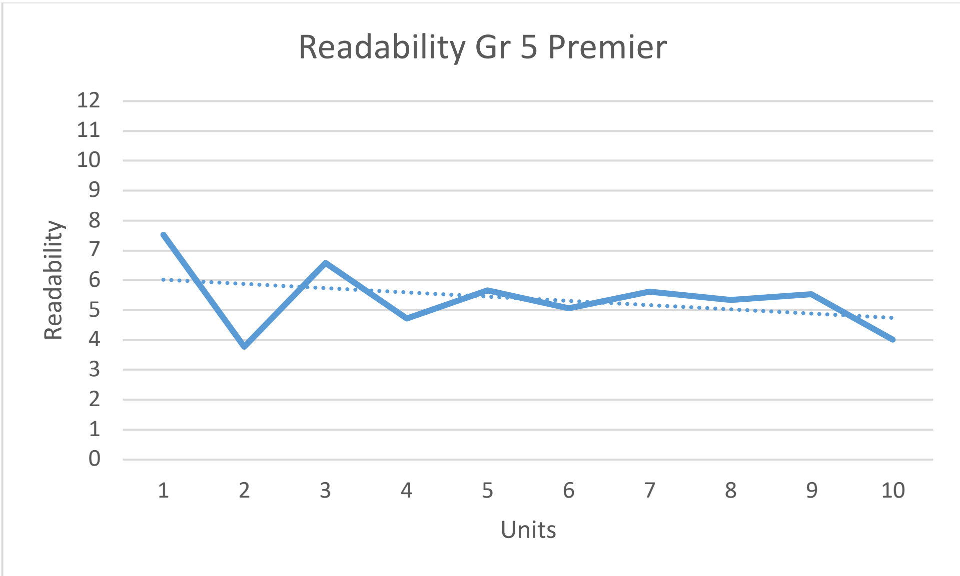


Figure 4.8 Readability scores for Grade 5 Premier book

In the Grade 6 Premier book, shown in Figure 4.9, there are 18 units focused on data handling. As shown in Figure 4.9, the majority of the units are between a readability level of Grade 3 and Grade 7, which is a wide range of variability. Unit 8 is a notable outlier with a readability score indicating a Grade 11 level of readability. The graph below indicates that there are numerous units in this textbook that are easier to read than most of the units in the Grade 4 or 5 Premier textbooks. Fifteen of the 18 units are at a readability level lower than Grade 7 and can thus be considered suitable for the grade level of the textbook. Ten of these units are below a readability level of Grade 4.

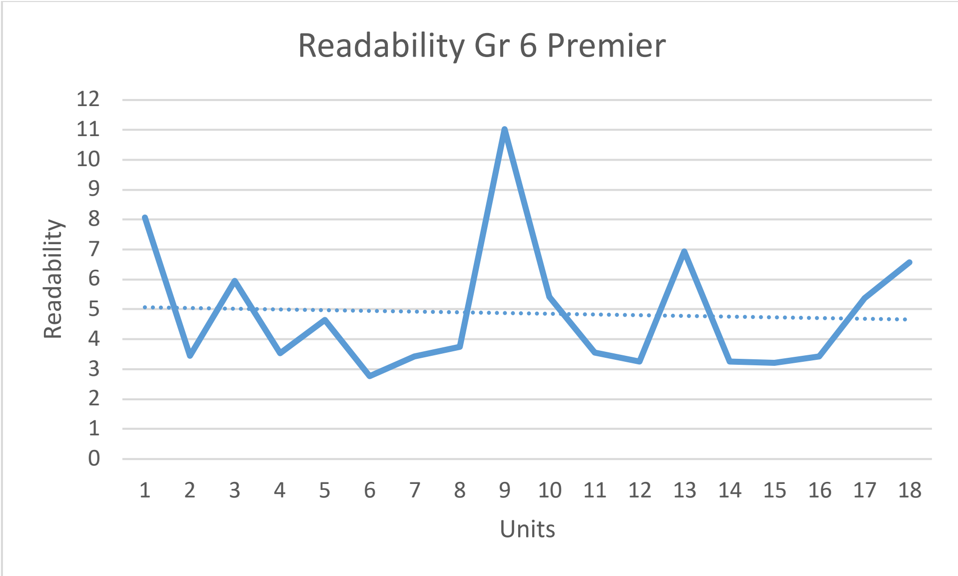


Figure 4.9 Readability scores for Grade 6 Premier book

With the exception of the Grade 4 book, the data indicates that the Premier textbooks have readability levels, on average, of approximately the appropriate grade. There are, however, at least two units per grade that have a much higher readability level than the grade the book is written for. In terms of readability, these particular units would not be accessible to the majority of the learners in that grade, especially if the learner does not have English as their home language. This would particularly be the case if the learner were working with the text independently.

4.2.1 Linguistic complexity of the Premier textbook

The linguistic complexity of each unit was analysed using the linguistic complexity checklist designed by Shaftel et al. (2006) which requires the text to be analysed for the grammatical features at word, sentence and paragraph level. In this section the frequency of the use of these grammatical features per unit is presented in a table. In the table, the language features with the three highest frequencies per unit are highlighted in yellow in order to emphasise which language features are contributing the most to the resulting Linguistic Complexity Index. A bar graph is used to summarise the Linguistic Complexity Index of each unit.

Grade 4 Premier textbook

Table 4.4 represents the data from the grade 4 Premier textbook.

Table 4.4 Linguistic features in the Premier Grade 4 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural reference	Errors	LCI
1.	27	163	20	1	1	0	4	6	6	1	0	0	0	0	7,48
2.	13	67	7	2	0	0	2	1	1	0	0	0	0	0	6,15
3.	12	89	6	1	0	0	3	1	3	1	0	1	0	0	8,75
4.	24	125	10	2	0	0	4	3	0	1	0	1	0	0	6,08
5.	12	129	12	1	0	0	4	4	1	1	0	0	0	0	12,6
6.	11	66	5	0	0	1	1	0	1	1	0	1	0	0	6,90
7.	10	78	5	0	0	1	4	0	1	2	0	1	0	0	9,2
8.	4	47	6	0	0	0	2	3	0	1	0	0	0	0	14,7
9.	25	160	21	0	0	1	3	2	3	2	0	1	0	0	7,72
10.	2	194	21	2	1	1	2	1	3	2	2	0	1	0	8,84
11.	10	86	7	1	0	0	3	2	0	0	0	0	0	0	9,9
12.	14	107	14	0	0	0	2	3	2	1	0	0	0	0	9,21

The texts in the Premier Grade 4 book are shorter than the texts in the Platinum book, ranging from just 47 words to 194 words. At the word level, the main language features that contributed to the complexity of the text were words with 7 or more letters (12/12 units), and relative pronouns (6/12 units). The use of slang, homophones and homonyms and the use of abbreviations appears among the highest three frequencies for just one unit each. Units 9 and 10 had the highest number of words of 7 or more letters, with 21 words. Very few of these words were mathematics-specific vocabulary. Mathematics-specific words included: pictographs, symbols, circular, divided, minimum and fractions. Other words of 7 or more letters included: schedule, experienced, particular, encouraged.

At the sentence level the language features contributing most to the complexity included prepositional phrases (12/12 units), infinitives (9/12 units), complex verbs (8/12 units) and complex/compound sentences (7/12 units). In Grade 4, Unit 1 the sentence "The first four marks are drawn **next to** each other and the fifth mark crosses the first four" provides an example of a sentence which contains a prepositional phrase and is of the complex/compound type. Conditional (1 unit) and comparative constructions (2 units) also contribute among the highest three frequencies for a small number of units.

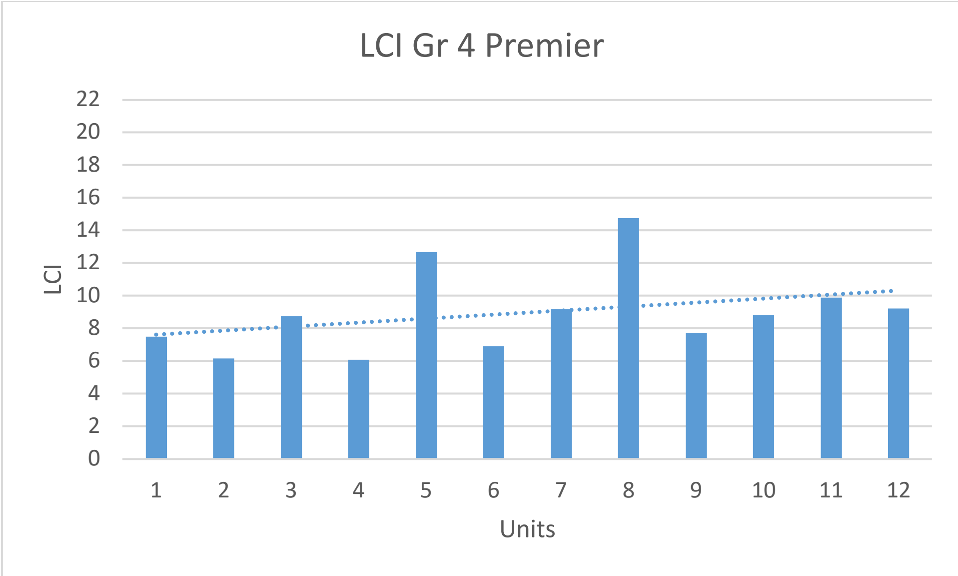


Figure 4.10 Linguistic Complexity Index for Grade 4 Premier book

As shown in Figure 4.10, in the Grade 4 Premier book the LCI varies between 6 and 14, with most of the units scoring between 6 and 10. Only 2 units scored above 10. The trendline indicates a slight increase in linguistic complexity as the book progresses. The linguistic complexity index of the book is similar to that of the Grade 4 Platinum book.

Grade 5 Premier textbook

The linguistic features for the grade 5 Premier book are presented in Table 4.5.

Table 4.5 Linguistic features in the Premier Grade 5 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural reference	Errors	LCI
1	10	135	14	0	4	0	3	1	3	1	0	1	0	0	16,2
2	30	182	19	2	1	0	7	2	0	2	0	2	0	0	7,23
3	33	271	24	4	6	0	11	2	6	1	0	1	0	0	9,87
4	91	572	53	4	0	1	22	11	13	3	1	0	1	0	7,48
5	21	173	16	0	0	1	11	3	0	1	0	4	0	0	9,95

6	41	294	30	2	3	0	8	5	6	1	2	1	0	0	8,58
7	55	339	35	2	0	0	30	2	6	4	0	2	0	0	7,63
8	13	115	15	0	1	0	5	4	2	1	1	0	0	0	11,07
9	23	111	23	0	1	1	10	1	0	1	0	0	0	0	6,43
10	15	118	9	2	2	0	8	5	2	0	0	0	0	0	9,73

The texts in the Grade 5 Premier book are longer than those in the Grade 4 book. They range in length from 111 words (Unit 9) to 572 words (Unit 4). At the word level, words of 7 or more letters contribute with the highest frequency to the linguistic complexity of the text in all units. Included in the words of 7 or more letters are the words pictographs, category, fraction, and numbers. These are all mathematically specific terms and represent words the learners should be gaining familiarity with. Some of the other words of 7 or more letters include: representing, excursion, triathlon, indigenous.

Also at the word level, relative pronouns contribute among the highest three frequencies in one unit, as do abbreviations. The use of slang, homophones and homonyms contributes among the highest three frequencies for 3 units.

At the sentence level, the language features that contribute most frequently to the linguistic complexity of the units include prepositional phrases (10/10 units), infinitives (4/10 units), complex verbs (5/10 units), complex/compound sentences (2/10 units) and comparative constructions (2/10 units).

As with all the books reported on thus far, prepositional phrases are a major contributor to the complexity of the texts. A common example is the use of the preposition 'in', for example "What is being represented **in** the diagram?" (Unit 2). Unit 4 has a large number of prepositional phrases, with 22 examples in its 91 sentences, for example, "Their class teacher compiled a survey **on** the wildlife they saw".

Unit 4 also displays the most infinitives (22) and complex verbs (13) of the units. One example of the use of an infinitive is: "Is it possible **to work out** exactly how many of each animal there are in the botanical gardens?" (Unit 4). Infinitives are used in every unit and complex verbs appear in 7 of the 10 units.

Comparative constructions are also a feature that contributed in this book to the complexity of the text. Comparative constructions refer to sentences in which two or more items are compared or where the comparative form of a word is used, for example, "Which category was most common? Which category was least common?"

Paragraph level complexity is reflected when there is reference to cultural or experience-specific events that learners may not be familiar with. Analysis of the units in the Grade 5 Premier textbook shows 1 instance of added complexity in an experience-specific reference. This is found in Unit 4 where a visit to the botanical gardens and the Kruger Park are used as the context for the problems posed and this may not be a familiar experience for the majority of the learners.

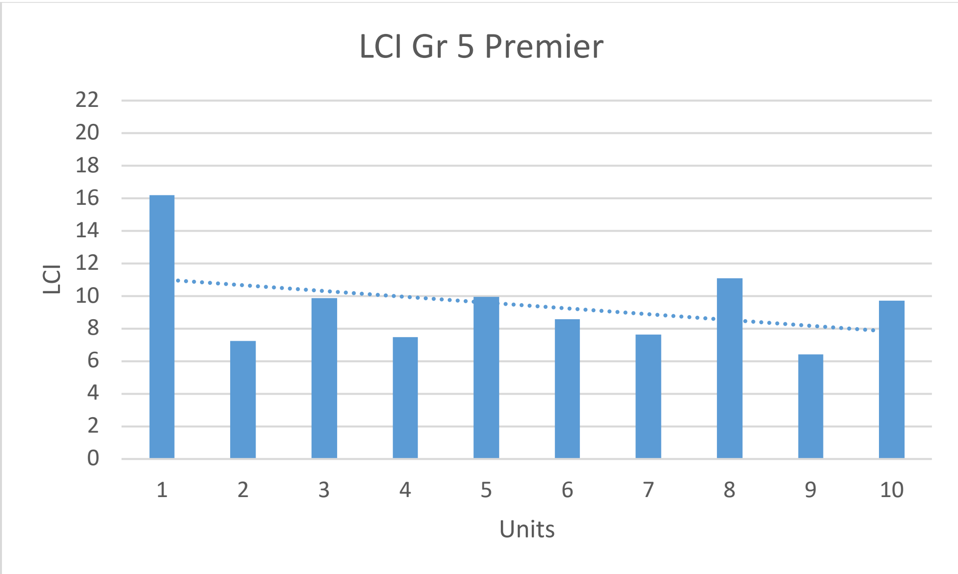


Figure 4.11 Linguistic Complexity Index for Grade 5 Premier book

As shown in Figure 4.11, the units in this book scored between approximately 7 and 16 on the LCI. The trendline indicates a decrease in linguistic complexity as the book progresses. This is likely because the first unit has a LCI that is notably higher than any of the other units (16,2). The units in Grade 5 are more linguistically complex than in the Grade 4 Premier book, but they are generally less linguistically complex than the Grade 5 Platinum book.

Grade 6 Premier textbook

A summary of the language features in the Grade 6 Premier textbook is presented in Table 4.6.

Table 4.6 Linguistic features in the Premier Grade 6 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural reference	Errors	LCI
1.	8	83	10	0	2	0	2	3	2	1	0	0	0	0	12,8
2.	2	161	10	2	2	0	13	0	3	3	0	1	0	0	8,39
3.	17	125	18	0	3	0	7	2	6	2	0	0	0	0	9,64
4.	14	53	0	0	0	2	0	0	0	0	0	0	0	0	3,92
5.	23	169	14	1	1	0	5	1	10	2	1	1	0	0	8,91
6.	19	128	7	1	3	0	3	1	0	2	0	0	0	0	7,57
7.	24	174	15	0	2	0	10	2	4	1	0	0	0	1	8,70
8.	21	158	9	0	3	0	11	0	7	3	0	0	0	0	9
9.	3	102	15	0	0	0	3	2	1	0	0	0	0	0	17,5
10.	14	98	10	0	3	0	5	0	4	1	0	0	0	0	7,85
11.	10	49	5	0	2	0	1	1	1	0	0	1	0	0	6
12.	17	111	2	0	2	2	5	0	0	0	0	0	0	0	7,23
13	17	156	10	0	5	0	7	2	0	1	0	0	0	0	10,5
14	33	126	15	0	1	0	1	2	0	0	0	0	0	0	4,51
15	32	111	8	0	1	0	0	0	0	0	1	0	0	0	3,78
16	72	526	28	0	2	0	24	4	8	4	1	3	0	0	8,54
17	29	159	16	0	2	0	5	3	1	3	0	0	0	0	6,41
18.	24	220	28	1	1	0	11	2	1	4	0	0	0	0	11,0

The Grade 6 Premier book includes a range of text lengths. Several of the units have shorter texts of less than 100 words, while the majority of the texts are between 100 and 200 words. There is one outlier in Unit 16 which is 526 words long.

At the word level, the language features contributing most frequently to the complexity of the texts in this book are: words with 7 or more letters (17/18 units) and slang, homophones and homonyms (7/18 units). At the sentence levels, these features included prepositional phrases (16/18 units), infinitives (5/18 units), and complex verbs (9/18 units). Less frequently, these also included complex/compound sentences (3/18 units), conditional constructions (1/18 units) and comparative constructions (1/18 units). An example of a sentence with a combination of these features can be found in Unit 8: "The **pictograph** below shows the number of bins of **recyclable** waste that **were collected by** a town in one year".

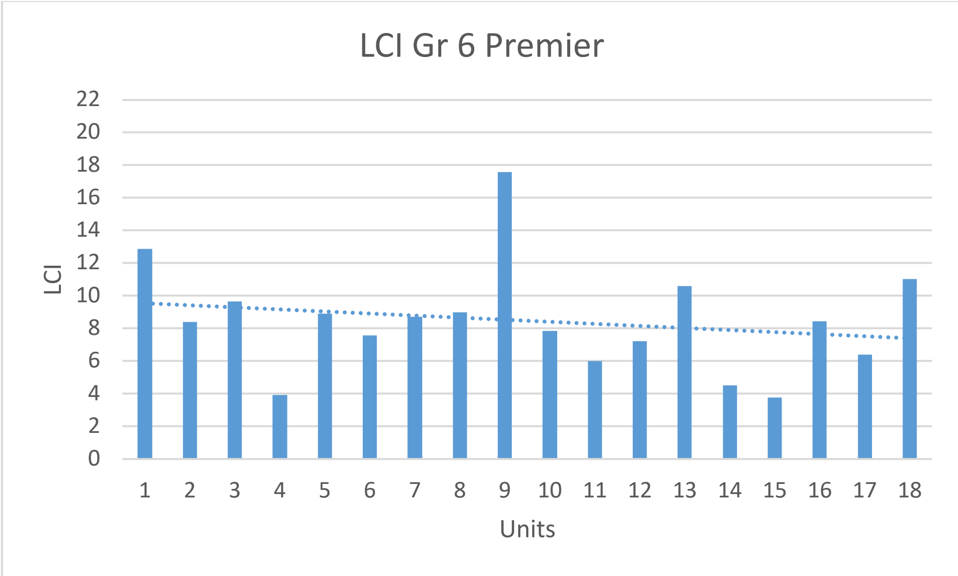


Figure 4.12 Linguistic Complexity Index for Grade 6 Premier book

In Figure 4.12, the LCI in Grade 6 is shown as between approximately 4 and 17,5. Most of the units are between 6 and 10. There is one notably higher LCI. Unit 9 is significantly higher than any of the other units because of the low number of sentences – there are 102 words in only 3 sentences. When compared to the linguistic complexity of the Grade 5 Premier book, we can observe that the language is at approximately the same level of complexity for most of the units. Only four of the 18 units are above 10. When compared to the LCI of the Grade 6 Platinum book, the Premier book is less linguistically complex (see Figure 4.6).

4.3 The Rainbow workbook

The third book series analysed are the Rainbow workbooks, issued by the DBE to every learner from Grade 1 to Grade 6. Here the Grade 4, 5 and 6 books are analysed.

4.3.1 Average readability of the Rainbow workbook

Appendix H provides all the readability scores for the Rainbow workbooks.

Figure 4.13 below displays the average readability scores across the 9 data handling units of the Grade 4 Rainbow workbook.

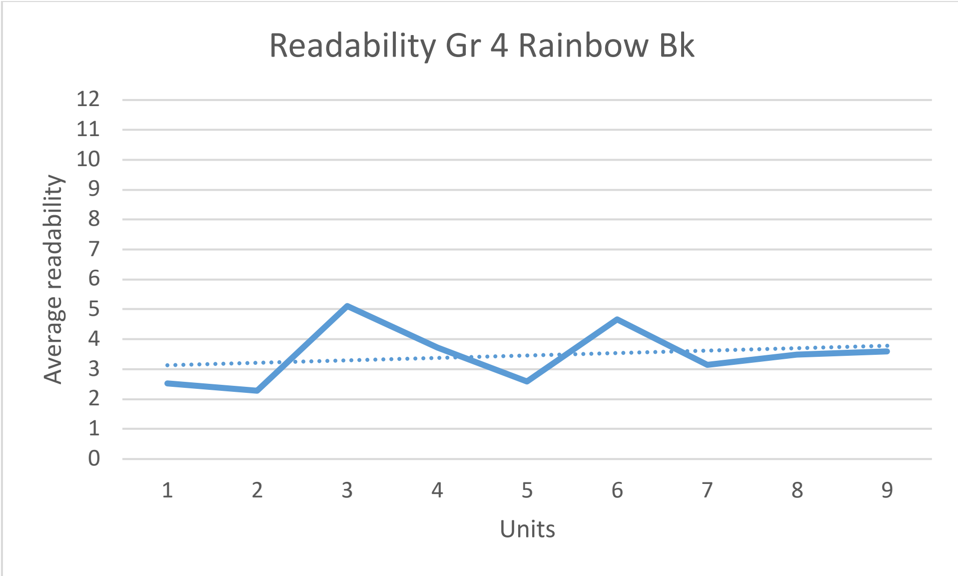


Figure 4.13 Readability scores for Grade 4 Rainbow book

The readability score ranges from just above a Grade 2 level to Grade 5 level. The trendline shows a very slight increase in readability scores as the book progresses, with an average readability score indicating a level between Grade 3 and 4. This is a more appropriate profile of readability scores than seen in the Grade 4 Platinum and Grade 4 Premier books. It is likely that, with the exception of Units 3 and 6, the units may be within the reading ability of learners Grade 4, particularly if the text is mediated by the teacher. Only two units of the nine are above Grade 4 readability.

In the Grade 5 workbook in the Rainbow series there are 7 units that focus on data handling. The readability scores follow a similar pattern as for the Grade 4 book, as shown in Figure 4.14. Only one unit is above Grade 5 readability. As shown by the trendline, the units do not become more difficult to read over the year. Similar to the Grade 4 book, it is likely that a most of the units would be accessible to learners whose reading level is grade appropriate. The readability scores for the Rainbow workbook are more appropriate than that seen in the Grade 5 Platinum and Grade 5 Premier books.

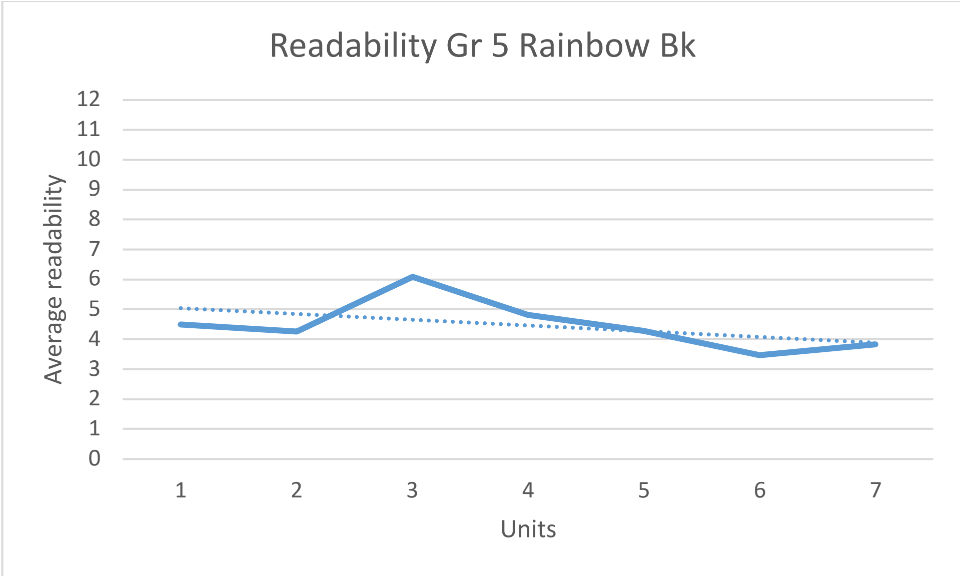


Figure 4.14 Readability scores for Grade 5 Rainbow book

In the Grade 6 Rainbow workbook, there are 10 units focused on data handling. As shown in Figure 4.15 below, the readability of the units ranges from just above Grade 3 level to Grade 8 level, which is a much wider range than for the Grade 4 and Grade 5 Rainbow books.

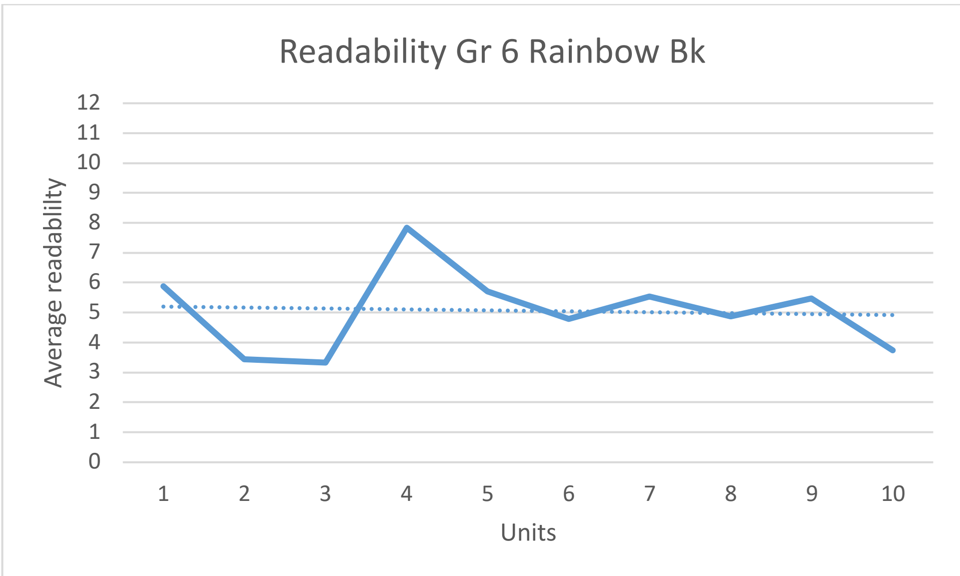


Figure 4.15 Readability scores for Grade 6 Rainbow book

The profile of the readability graph shown above is jagged in contrast to the graphs for Grade 4 and 5. The trendline reveals that on average the readability is at Grade 5 level. Only one unit, Unit 4, is above the intended grade level with a readability score showing a Grade 8 level. The remaining units are grade appropriate or below.

From the data presented and analysed, the Rainbow books largely have an appropriate readability for the grades they have been written for. There are units in each book which are above the grade level of the book, but overall the majority of the units are either at or below the grade level. This means that learners should be able to find the texts accessible, including (for some units) learners who might not yet be at the appropriate grade level in their reading ability.

4.3.2 Linguistic complexity of the Rainbow workbook

The linguistic complexity of each unit was analysed using the linguistic complexity checklist designed by Shaftel et al. (2006) which requires the text to be analysed for the grammatical features at word, sentence and paragraph level. In this section the frequency of the use of these grammatical features per unit is presented in a table. In the table, the language features with the three highest frequencies per unit are highlighted in yellow in order to emphasise which language features are contributing the most to the resulting Linguistic Complexity Index. A bar graph is used to summarise the Linguistic Complexity Index of each unit.

Grade 4 Rainbow workbook

The language features that are most prevalent in the grade 4 text are summarised in Table 4.7.

Table 4.7 Linguistic features in the Rainbow Grade 4 text

Unit	Sentence	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural references	Errors	LCI
1.	25	90	11	0	0	0	1	0	0	0	0	1	0	0	4,16
2.	31	154	11	1	2	5	8	1	0	1	0	0	0	0	5,93
3.	49	404	39	4	4	3	15	5	3	5	1	1	0	0	10
4.	18	127	15	1	0	0	10	2	6	1	0	0	0	0	9,05
5.	25	115	8	0	1	1	2	0	4	0	0	0	0	0	5,24
6.	22	114	15	2	1	0	7	0	1	2	0	0	1	0	6,63
7.	15	66	8	0	1	0	3	1	1	1	0	0	0	0	5,4

8.	16	92	8	0	0	0	4	0	4	0	0	0	0	0	6,75
9.	34	240	20	0	3	0	3	5	11	2	0	1	0	0	8,44

The units in this book vary in length, with 3 units below 100 words and 5 units between 100 and 250 words. Unit 3 is notably longer at 404 words, which is a large amount of text for a Grade 4 learner to decode. At the word level, the number of words with 7 or more letters contributed most to the linguistic complexity. All 9 units contained words with 7 or more letters. Unit 3 had the highest number of words with 7 or more letters, with 39 words. Specific mathematics vocabulary with 7 letters or more was found across in all the units in this book, for example, the word pictograph is used in Unit 2. Also contributing among the highest three frequencies at the word level are relative pronouns (1 unit) and the use of slang, homophones or homonyms (1 unit).

At the sentence level the language features contributing most to the linguistic complexity of the units include: prepositional phrases (8/9 units), infinitives (3/9 units), complex verbs (5/9 units), complex/compound sentences (3/9 units) and comparative constructions (1/9 units). This is consistent with the findings from the Platinum and Premier books reviewed thus far.

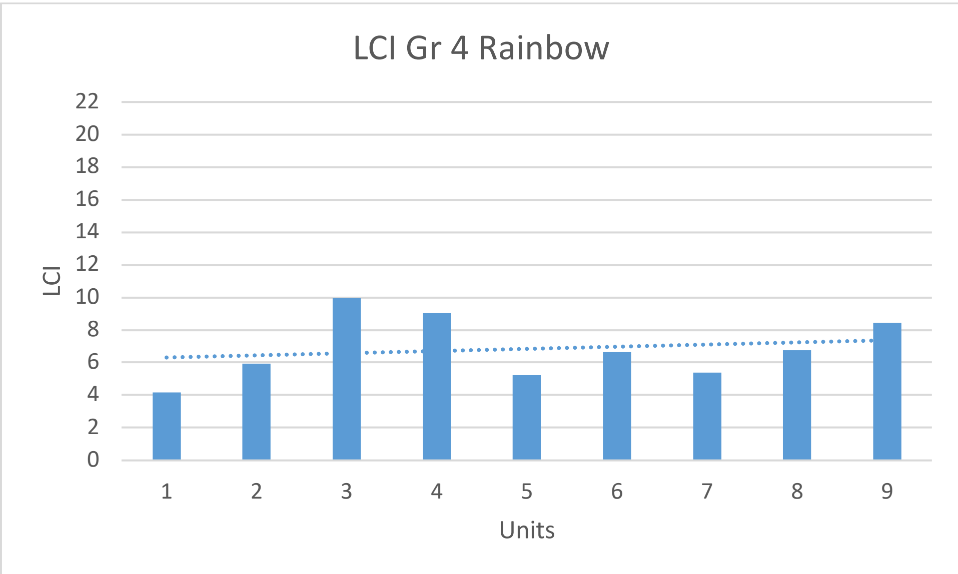


Figure 4.16 Linguistic Complexity Index for Grade 4 Rainbow book

The LCI for the units in the Grade 4 Rainbow book is generally lower than the LCI in the Grade 4 Platinum and Premier books, as shown in Figure 4.16. It varies between 4 and 10. Only 3 units scored above 8. This text is therefore more accessible to learners.

Grade 5 Rainbow workbook

Table 4.8 presents a summary of the language features in the Grade 5 Rainbow book.

Table 4.8 Linguistic features in the Rainbow Grade 5 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural references	Errors	LCI
1.	42	335	28	3	1	0	18	4	8	6	0	2	0	0	9,54
2.	51	329	33	1	2	0	13	3	3	5	0	2	0	0	7,60
3.	19	108	15	2	1	0	7	0	0	3	0	2	0	0	7,10
4.	50	336	36	2	0	0	9	6	5	7	0	2	0	0	8,02
5.	17	563	13	0	3	0	11	2	5	4	0	0	0	0	10,29
6.	35	189	18	2	2	0	2	1	10	3	0	0	0	0	6,45
7.	38	224	16	4	1	1	9	3	3	4	0	0	0	0	6,94

The texts in the units in the Grade 5 Rainbow book are longer than those in the Grade 4 Rainbow book. The shortest text is 108 words. Four of the 7 units are longer than 300 words, with Unit 5 being as long as 563 words in length. This is also the unit with the highest LCI (10,29).

Consistent with the findings in all other books, the language feature at the word level that contributed most to the linguistic complexity of all units is words with 7 or more letters. Included in these words are mathematics specific terms, for example, pictograph and classify appear in Unit 1. Additional words include version, physical, creatively, popular. Also contributing among the three highest frequencies at the word level is the use of relative pronouns (1 unit).

Prepositional phrases were again one of the features to contribute most frequently to the linguistic complexity (6/7 units), for example, "Do the following activity **in** groups of 7" (Unit 5). Complex verbs also contributed frequently to the linguistic complexity (3/7 units), for example, "The pie chart above **was drawn**" (Unit 5). Also contributing to the complexity at the sentence level was the use of complex/compound sentences. This appears as one of the highest three contributing frequencies in 5 of the 7 units. For example, ‘have been sold’, ‘going to collect’ and ‘were collected’.

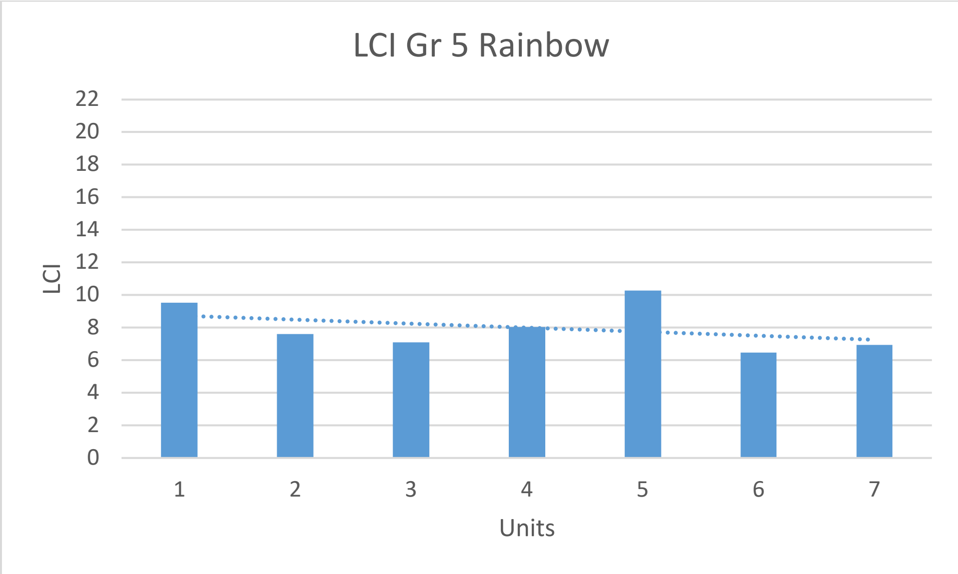


Figure 4.17 Linguistic Complexity Index for Grade 5 Rainbow book

In Grade 5, the units scored between 6,45 and 10,29 on the LCI, as shown in Figure 4.17. Overall the units for the Grade 5 Rainbow book have similar LCI values to the Grade 4 Rainbow book. This is in general also similar to the Premier book. The units in the Platinum book, however, have generally much higher LCI values.*Grade 6 Rainbow workbook*

The language features appearing in the Grade 6 Rainbow book are summarised in Table 4.9.

Table 4.9 Linguistic features in the Rainbow Grade 6 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural references	Errors	LCI
1.	14	107	9	0	3	0	8	2	0	3	0	0	0	0	9,5
2.	60	327	27	3	3	0	14	1	0	4	0	1	0	0	6,45
3.	43	247	21	2	2	1	12	1	4	2	0	3	0	1	6,95
4.	19	194	21	3	1	0	4	3	2	4	0	0	0	0	12,26
5.	28	283	29	1	1	0	6	5	6	2	0	1	0	0	12
6.	15	132	15	2	1	0	6	1	1	2	0	0	0	0	10,93
7.	23	196	11	0	0	0	6	1	8	3	0	1	0	0	9,95
8.	30	204	18	0	0	0	13	2	4	1	0	2	0	0	8,23
9.	22	233	9	0	0	0	5	2	0	1	0	2	0	0	11,5
10.	45	326	18	2	1	1	9	3	3	4	0	1	0	0	8,31

The lengths of the texts in the units in the Grade 6 Rainbow book range from 107 words to 327 words. At the word level, words with 7 or more letters contribute to the linguistic complexity with the highest frequency (all units). There are mathematically specific words of 7 or more letters in every unit, although a small number of them, for example, the words numbers, ascending, temperature, heights and measured appear in Unit 2. Non-mathematical vocabulary of 7 or more letters includes: normally, previous, favourite, collected. The use of relative pronouns and slang, homophones and homonyms appears among the three highest frequencies for 1 unit each.

At the sentence level, it is again prepositional phrases that contribute among the three highest frequencies to the linguistic complexity for all units. An example is “Lilly said: ‘I forgot about the 5 Peter’s living **in** Second Avenue’” (Unit 6). Infinitives (3/10 units), complex verbs (4/10 units), complex/compound sentences (5/10 units) and comparative constructions (1/10 units) also appear among those features highlighted. An example of a complex/compound sentence appears in Unit 10: “Go back to question 1 and work out how many kilograms of each material were collected”.

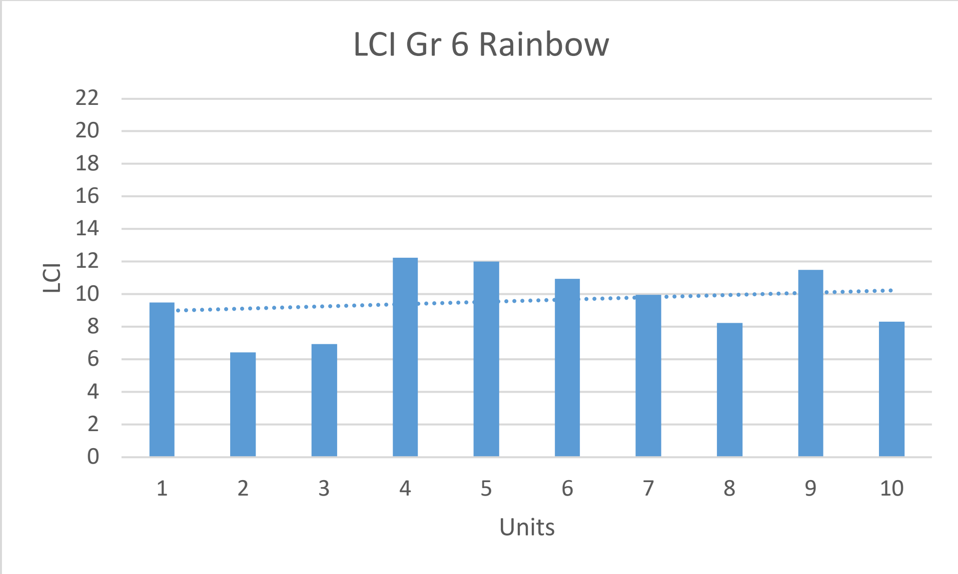


Figure 4.18 Linguistic Complexity Index for Grade 6 Rainbow book

In Figure 4.18 the LCI in the Grade 6 Rainbow book is shown to range between approximately 6 and 12, with most units being above 8. This is higher than for the Grade 4 and Grade 5 Rainbow books. This is comparable to the LCI in the Grade 6 Premier book, although the Premier book had more units, with had several units below 6 and one unit at 17,5 (see Figure 4.12). The Platinum Grade 6 book was more linguistically complex according to the LCI scores for the units (see Figure 4.6).

4.4 The Inzalo textbook

The Inzalo textbook is a free online resource for teachers and learners to download and use. It was developed by the Sasol Inzalo Institute in partnership with the DBE. This book would only have been accessible to learners and schools with internet access and with access to devices suitable for reading an electronic book. The language features and linguistic complexity are summarised in the sections that follow.

4.4.1 Average readability of the Inzalo textbook

Appendix I provides all the readability scores for the Inzalo textbooks.

In the Grade 4 Inzalo textbook there are 8 units that focus on data handling. As shown in Figure 4.19, the units in the textbook start with an average readability score indicating a Grade 7 level and become easier as they progress through the next two units, and then become more difficult to read. The trendline indicates an increase in readability scores as the book progresses, but the profile of the graph is jagged and ranges in readability from Grade 4 to Grade 8 level. There are only 2 units (Units 3 and 4) that have the appropriate readability for Grade 4 learners.

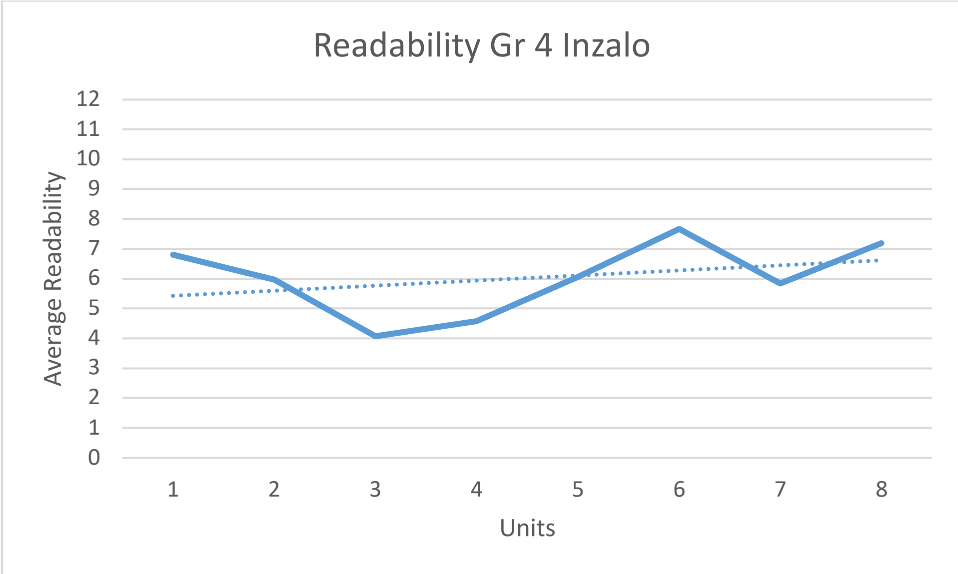


Figure 4.19 Readability scores for Grade 4 Inzalo book

In the Grade 5 Inzalo textbook, there are 6 units focused on data handling. As shown in Figure 4.20, the trendline shows that the readability scores on average decrease as the units progress, however all of the units have readability scores above the grade level of the book. The readability scores range from Grade 6 to Grade 9 level. This is above the appropriate level of readability for Grade 5 learners. In this research, this is the Grade 5 book with the highest readability scores, making it the most difficult Grade 5 book to read. It is likely beyond the reach of Grade 5 learners, and almost certainly inaccessible to those learners who may be reading at a level lower than the grade level.

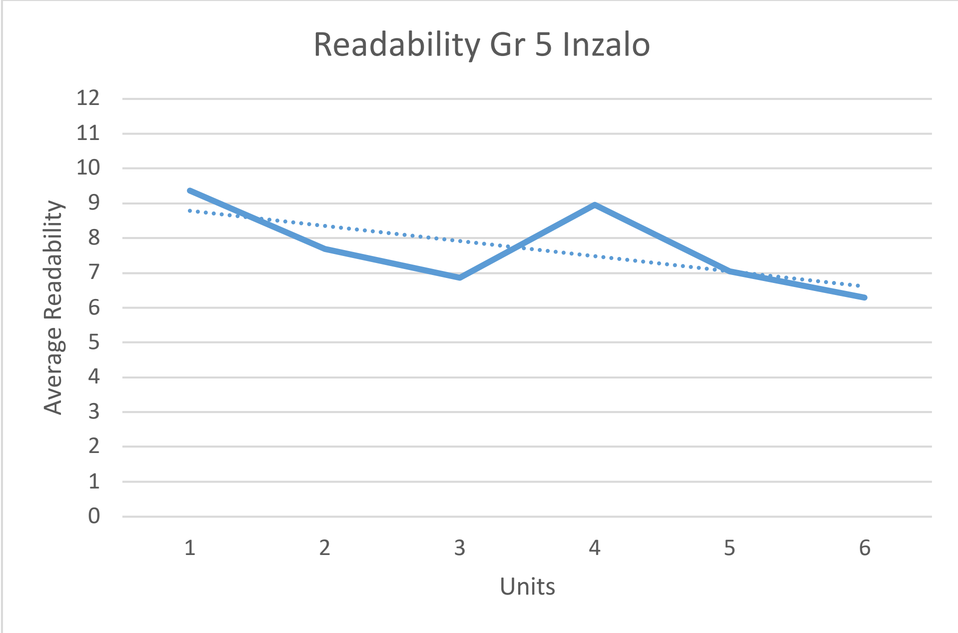


Figure 4.20 Readability scores for Grade 5 Inzalo book

In the Grade 6 Inzalo textbook, there are 8 units focused on data handling. The readability scores range between approximately Grade 5 and Grade 9 level, as shown in Figure 4.21. Only two units are appropriate for the grade of the book (Unit 5 and Unit 6). This book, and the Platinum book, are the two most difficult Grade 6 texts in terms of readability.

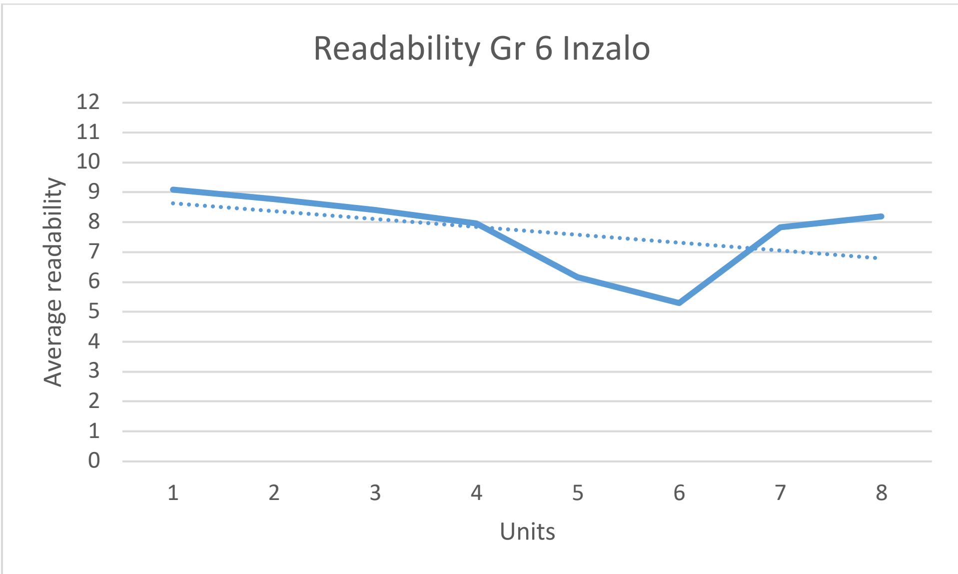


Figure 4.21 Readability scores for Grade 6 Inzalo book

Overall, the Inzalo books have readability scores that are too high for the grade level the books are written for. Only two units out of 8 in the Grade 4 book are grade appropriate in terms of readability, and only two units out of 8 in the Grade 6 book are grade appropriate. No units in the Grade 5 book are grade appropriate.

4.4.2 Linguistic complexity of the Inzalo textbook

The linguistic complexity of each unit was analysed using the linguistic complexity checklist designed by Shaftel et al. (2006) which requires the text to be analysed for the grammatical features at word, sentence and paragraph level. In this section the frequency of the use of these grammatical features per unit is presented in a table. In the table, the language features with the three highest frequencies per unit are highlighted in yellow in order to emphasise which language features are contributing the most to the resulting Linguistic Complexity Index. A bar graph is used to summarise the Linguistic Complexity Index of each unit.

Grade 4 Inzalo textbook

The Grade 4 data is presented in the table below:

Table 4.10 Linguistic features in the Inzalo Grade 4 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/ compound sentences	Conditional constructions	Comparative constructions	Cultural reference	Errors	LCI
1	37	444	22	2	1	0	5	6	6	6	0	1	1	0	13,32
2	15	177	12	1	1	1	4	3	0	2	0	1	1	0	13,46
3	14	141	5	1	2	1	4	2	0	0	0	3	1	0	11,42
4	35	395	11	2	1	1	6	4	0	1	0	1	1	0	12,08
5	13	133	7	0	0	1	2	2	1	0	0	2	1	0	11,46
6	12	159	17	0	1	1	5	3	3	2	0	2	1	0	16,16
7	11	119	7	1	2	0	6	4	4	2	1	1	1	0	13,63
8	19	222	17	1	4	1	11	0	2	1	0	2	1	0	13,84

The lengths of the texts in the 8 units of the Grade 4 Inzalo book range from 119 (Unit 7) words to 444 words (Unit 1). At the word level, the number of words with 7 or more letters contributed the most to the linguistic complexity for all units. Among these words, the mathematics-specific vocabulary included the words: measure, differed, lengths, measurements, pictograph, fractions and estimated. Some of the other vocabulary included: exactly, appropriate, advertisement, statement, poaching. Also contributing among the highest three frequencies for one unit (Unit 8) is the use of slang, homophones and homonyms.

At the sentence level it is again prepositional phrases (all units), infinitives (6/8 units) and complex verbs (4/8 units) that contributed with high frequencies to the linguistic complexity of each unit. Also contributing as among the three highest frequencies are complex/compound sentences (1 unit) and comparative constructions (2/8 units). Comparative constructions appear in every unit. An example from Unit 2 is: “The number of Sybunkins that are 3cm long is the same as the number of Sybunkins that are 6cm long”.

Paragraph level complexity is reflected when there is reference to cultural or experience-specific events that may be outside of the experiences of the majority of the learners. Analysis of the units in the Grade 4 Inzalo textbook, shows instances of added complexity due to experience-specific references. One unit, for example, focuses on rhino poaching in a national park, which may also be outside of the learners’ context. This adds to the complexity as the text may include an unfamiliar set of vocabulary and expressions.

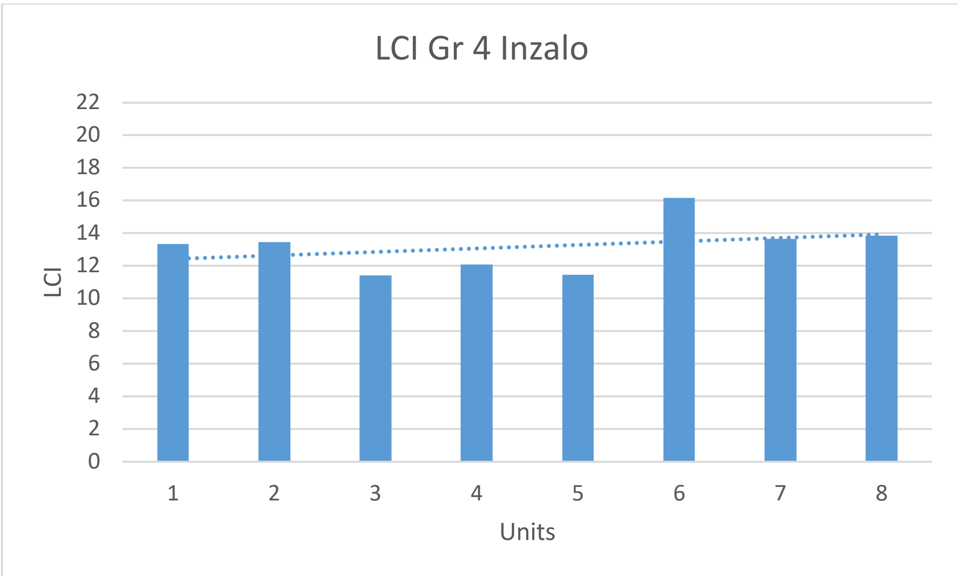


Figure 4.22 Linguistic Complexity Index for Grade 4 Inzalo book

As shown in Figure 4.22, the LCI varies between approximately 11 and 16. Unit 6 is the most linguistically complex with an LCI of 16,16. This is a smaller range, although at a higher level, and is therefore more consistent than the Platinum Grade 4 book (between approximately 7 and 16) and the Premier Grade 4 book (between approximately 6 and 15). The LCI for this book is higher than the Rainbow Grade 4 book, which had a maximum of 10.

Grade 5 Inzalo textbook

The data collected for the Grade 5 Inzalo book is presented in Table 4.11.

Table 4.11 Linguistic features in the Inzalo Grade 5 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/compound sentences	Conditional constructions	Comparative constructions	Cultural reference	Errors	LCI
1	29	390	41	7	1	0	16	6	6	7	0	2	0	0	16,41
2	49	723	57	8	3	0	52	11	7	13	3	4	0	0	17,97
3	39	561	42	8	2	0	19	11	7	8	5	0	0	0	17
4	12	181	30	0	4	0	10	6	2	3	2	0	0	0	19,83
5	29	384	24	6	3	0	13	12	7	4	1	0	0	0	15,65
6	55	696	47	4	3	12	11	4	14	11	6	0	0	0	14,69

There is a wide range of lengths in the texts in the Grade 5 Inzalo units. They range from 181 words (Unit 4) to as much as 696 words (Unit 6) and 723 words (Unit 2). Three units are in excess of 500 words (Units 2, 3 and 6), These are very long texts for a Grade 5 learner to decode. At the word level it is again words of 7 letters or more that contribute the most to the linguistic complexity for all units, with very high frequencies of use of these words (e.g. 57 words in Unit 2 and 30 words in only 12 sentences in Unit 4). Some of these words included: environment, informal, conference and gathering (Unit 1). Some of the mathematics-specific vocabulary included: estimate, seventh, increasing, decreasing and pictograph (Unit 6). Also contributing with a high frequency to the complexity of Unit 6 at the word level was the use of abbreviations (12), for example Jan, Feb, Mar, Apr.

As with all books analysed, it is prepositional phrases that contribute among the most frequently to the linguistic complexity (5/6 units). Also aligned to the findings for the other books, infinitives are another frequent contributor to the linguistic complexity (3/6 units).

Complex/compound sentences also contributed frequently, appearing in all units and highlighted in two units. One example includes: "And there was lots of load shedding, so her account should be less than usual, not more than usual!" (Unit 6).

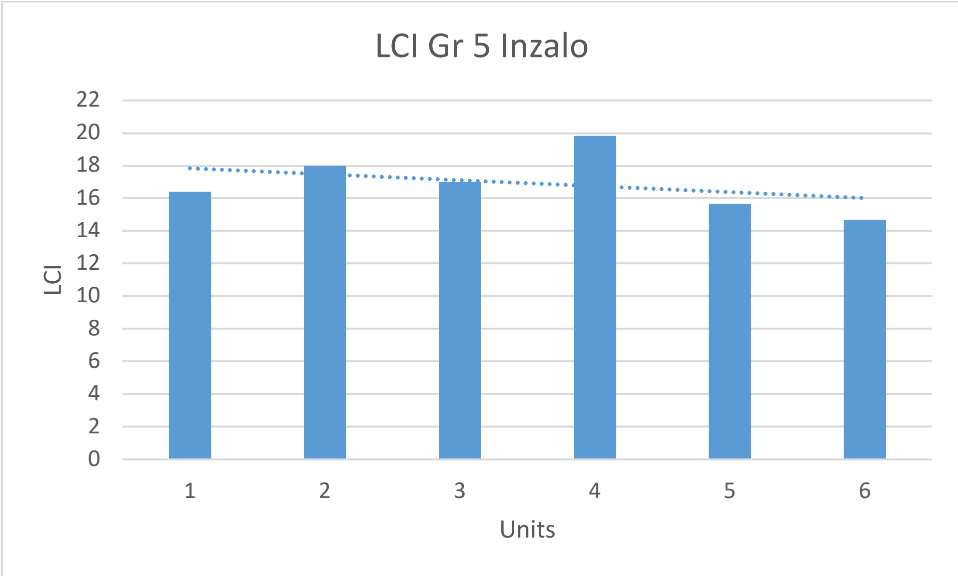


Figure 4.23 Linguistic Complexity Index for Grade 5 Inzalo book

As shown in Figure 4.23, the units in Grade 5 scored between approximately 15 and 20 on the LCI. The units in this book are more linguistically complex than in the Grade 4 Inzalo book. The LCI is approximately double that of the Grade 5 Rainbow book, and is similarly much higher than the Grade 5 Premier book. Most of the Platinum Grade 5 units are below the LCI range for this book. It is the most complex of the Grade 5 texts investigated in this research.

Grade 6 Inzalo textbook

The language features from the text in the Grade 6 Inzalo textbook is presented in Table 4.12.

Table 4.12 Linguistic features in the Inzalo Grade 6 text

Unit	Sentences	Words	7 letters or more	Relative pronouns	Slang, homophones, homonyms	Abbreviations	Prepositional phrases	Infinitives	Complex verbs	Complex/compound sentences	Conditional constructions	Comparative constructions	Cultural reference	Errors	LCI
1	24	297	39	9	2	0	12	1	7	4	0	0	0	0	15,45
2	29	361	30	11	0	0	14	2	6	7	0	5	0	0	15,03
3	45	724	38	8	3	1	20	20	23	19	8	3	0	0	18,82
4	26	366	28	3	2	0	8	11	9	5	1	0	0	0	16,65
5	42	549	44	5	4	3	21	5	15	7	0	0	0	0	15,58
6	22	221	16	2	4	0	2	3	6	4	1	0	0	0	11,77
7	53	679	43	10	3	0	40	6	3	15	0	3	0	0	15,13
8	21	348	20	6	4	0	15	10	6	9	0	0	0	0	19,90

The units in the Grade 6 Inzalo book range in text length from 221 words to 724 words. As with the Grade 5 book, 3 units exceed 500 words. Consistent with all the other texts analysed, it is words of 7 or more letters that contribute most at the word level to the linguistic complexity for all units. Relative pronouns (2/8 units) and slang, homophones and homonyms (1 unit) also contribute among the three highest frequencies at the word level.

At the sentence level, it is prepositional phrases (6/8 units), infinitives (3/8 units), complex verbs (4/8 units) and complex/compound sentences (3/8 units) that contributed most frequently to the linguistic complexity. One example of a sentence with two prepositional phrases, includes the following from Unit 5: "Place the 10 bounce heights **on** a number line exactly halfway **between** the fifth and sixth bounce height". An example of a sentence with two prepositional phrases and a complex verb is the following from Unit 6: "**In** this section, you **will organize** data **about** the size of oranges".

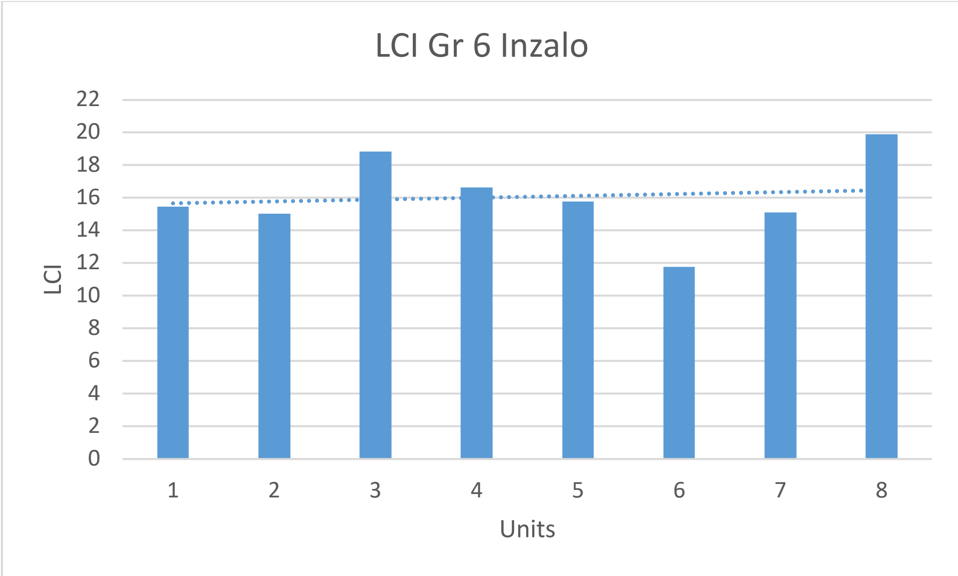


Figure 4.24 Linguistic Complexity Index for Grade 6 Inzalo book

Figure 4.24 summarises the LCI for the Grade 6 Inzalo book. The units in Grade 6 have an LCI between 11,77 and 19,9 and all but one unit have an LCI of over 14. Overall, the LCIs are very similar to the LCIs for the Grade 5 Inzalo book. The Platinum Grade 6 book has 9 of its 20 units with an LCI of over 14, but the book does contain 11 units at much lower LCIs. The LCIs of the Grade 6 Premier book and the Grade 6 Rainbow book have a lower range than the Inzalo book. In terms of LCI, the Grade 5 and Grade 6 Inzalo books are shown to be more complex than the other books analysed.

4.5 Conclusion

In this chapter I have presented the analysis of readability and linguistic complexity for the four book series included in this study. In terms of both readability and linguistic complexity the Rainbow workbooks have been shown to be the least complex, and the Inzalo books have been shown to be the most complex. The Platinum and Premier books all have units that are high in linguistic complexity and are assessed as high in their readability scores. Both of those book series have a wide range of readability and linguistic complexity. In the analysis of the linguistic complexity I have also presented the language features that have contributed the most frequently to the LCI of the units. This has allowed a clearer picture to emerge of why a text might be deemed to be linguistically complex.

CHAPTER 5

DATA PRESENTATION: NON-TEXTUAL ELEMENTS

In this chapter I present the data on the non-textual elements included in the four book series analysed.

First, each non-textual element has been categorised according to its graphical language (MacKinley, 1986). These languages, according to MacKinley (1986) are summarised in Table 5.1.

Table 5.1 MacKinley's graphical languages (from MacKinley, 1986, pp. 127-130)

Language	Information encoded by:	Example
Single-position languages	The position of a mark set on one axis	Horizontal axis, vertical axis
Apposed-position languages	A mark set that is positioned between two axes	Line chart, bar chart, plot chart
Retinal-list languages	One of the six retinal properties of the marks in a mark set independent of position	Colour, shape, size, saturation, texture, orientation
Map languages	Fixed positions with graphical techniques specific to maps	Road map, topographic map
Connection languages	A connected set of node objects with a set of link objects	Tree diagram
Miscellaneous languages	A variety of additional graphical techniques	Pie chart, Venn diagram

In the analysis I add to these languages the 'table'. I distinguish between tables and the remaining graphical languages because of the frequency of their use in the data handling texts analysed. Most tables are examples of an apposed-position language because they are arranged vertically and horizontally into columns and rows. They are, however, slightly different from the apposed-position graphs because they contain text (in the form of numbers and/or words) in their cells. For this reason they have been given a category of their own in this study. Slutsky (2014) differentiates between graphs and tables: graphs are used to visually illustrate relationships in the data and are advantageous in that they summarize a large data set in visual form and are able to clarify trends better than tables. According to Slutsky (2014) tables can make a text more readable by removing numeric data from the text and are used to organize data that is too detailed or complicated to be adequately described. Included in the category of 'tables' are tally tables.

The MacKinley (1986) languages are specifically about "graphical presentations of relational information" (p. 110). There are also, additionally, non-textual elements that are illustrative of concepts or contexts presented in the texts, or are unrelated to the text as decorative elements on the page. They do not present relational information and as such are given a category of their own: ‘images’.

After categorising the non-textual elements into their ‘languages’, I assess the accuracy, connectivity, conciseness and contextuality of these elements. This analysis is guided by Kim’s (2012) framework for the analysis of non-textual elements in mathematics texts, presented in Table 5.2 below. Kim’s (2012) framework assesses non-textual elements and their relationship with the mathematical content of the text. A fully accurate, connected, concise and contextual non-textual element would correctly show the definition of a concept, would be explicitly and fully associated with the mathematical content, would show a concept in a straightforward manner and would make use of a realistic object or context. Each non-textual element in the books analysed is measured against this. It is assumed in this study that non-textual elements that are not accurate, connected, concise or contextual would add complexity to the interpretation of the information provided on the page. Conversely, the inclusion of non-textual elements that are accurate, connected, concise and contextual should facilitate comprehension. It should be added, however, that knowledge of how to interpret some of these non-textual elements are a part of the learning outcomes for Data Handling and so learners may not be familiar with how to interpret them. For example, the interpretation and creation of bar graphs, tally tables, pie charts and pictographs are all part of what is being taught in this section of the curriculum.

Table 5.2 Kim’s (2012) framework for analysis of non-textual elements in mathematics texts

	2	1	0
Accuracy	An NTE correctly shows the definition of a concept, or is accurate to show a concept based on its definition.	An NTE makes sense in terms of the definition or meaning of a concept. But does not show every required mathematical condition (e.g. Some requisite notations are missing or misleading) or some attributes are not appropriate to explain a concept.	An NTE is inaccurate in terms of the definition of a concept (eg. it has an obvious error) Or an inappropriate realistic object is used to present a concept. There is a major error or concern to use a realistic object or context for a concept. Or there is no mathematical concept in the NTE.
Connectivity	An NTE is explicitly and fully associated with the mathematical content	An NTE is partly related to the mathematical content in the text. There is	An NTE has nothing to do with the content mathematically. It can give some clue about the

	in the text. It directly shows a concept or a problem.	some missing or irrelevant information in an NTE. It shows the content but it does not show how it is connected with the content.	contexts in texts (eg. river when the problem is about length of a river.)
Conciseness	An NTE is straightforward to show a concept or problem without any distracting or other factor.	An NTE is straightforward to show a concept or problem with some other factors that might be helpful for the concept.	An NTE has distracting or other factors that are useless in addition to factors needed to show a concept or problem.
Contextuality	A realistic object or context is used in an NTE with mathematical connection.	No mathematical ideas or concepts exist, but there is some realistic contextual information (that is used to provide contexts or objects in the problem or to facilitate related activities.)	Neither realistic object nor realistic context is included in an NTE

5.1 Analysis of the non-textual elements in the Platinum series

In this section I present the analysis of the non-textual elements included in the Platinum series. These elements represented all the graphical languages, except the map languages, and were mostly accurate, connected, concise and contextually supportive of the text.

5.1.1 Graphical languages in the Platinum series

In Table 5.3 a summary is provided of the graphical languages used in the Grade 4, 5 and 6 Platinum textbooks.

Table 5.3. Graphical languages of the Platinum textbook

	Grade 4	Grade 5	Grade 6
Tables	7	15	13
Single position languages	7	5	3
Apposed position languages	16	19	13
Retinal list languages	0	4	0
Map language	0	0	0
Connection languages	1	1	1
Miscellaneous languages (pie chart)	7	4	9
Images	12	7	12

The graphical language used most frequently in all books is the apposed position language, specifically, bar graphs and some pictographs. An example is provided in Figure 5.1.

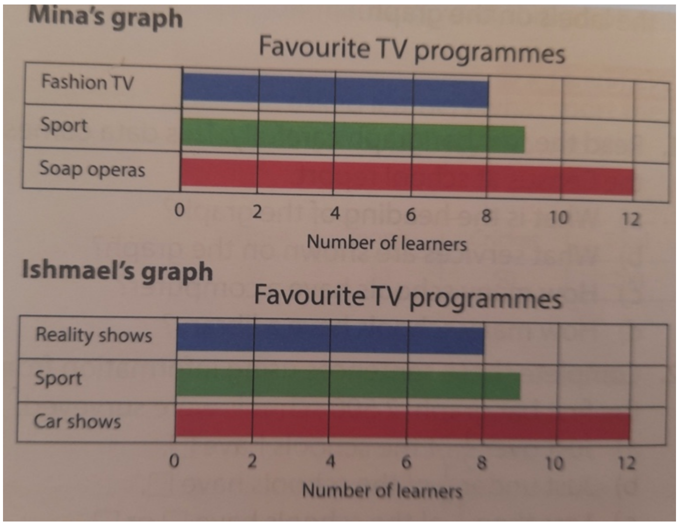


Figure 5.1 Example of an apposed position language (Platinum Grade 6, Unit 11)

Tables are also extensively used in the Platinum textbooks, particularly in Grade 5 and 6. An example appears in Figure 5.2.

Day	Tally	Number of children
Monday	~~IIII~~	5
Wednesday	~~IIII~~ III	8
Friday	IIII	4
	Total	17

Figure 5.2 Example of a table (Platinum Grade 4, Unit 1)

The other graphical language that was most used in the Grade 4, 5 and 6 Platinum textbooks were pie charts, categorised as 'miscellaneous languages'.

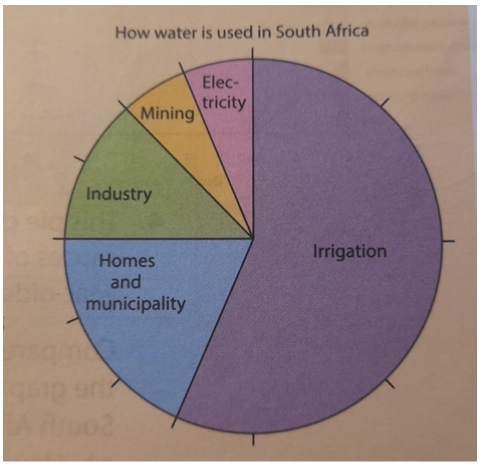


Figure 5.3. Example of a pie chart (Platinum Grade 6, Unit 9)

Single position languages, most in the form of pictographs, were used often in the Platinum textbooks. Also included in the Platinum texts were images that either provided contextual support to the text or were used for decorative purposes.

The Platinum Grade 5 textbook had more retinal list languages than both the other grades. These are marks in a set of data, in which their retinal properties, like colour, size and shape, have meaning. An example is provided in Figure 5.4.

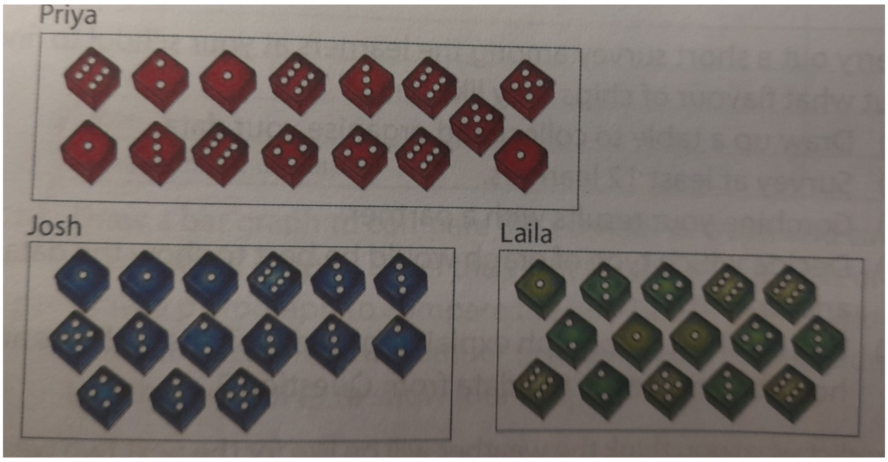


Figure 5.4 Example of retinal list languages (Platinum Grade 5, Unit 11)

5.1.2 Accuracy, connectivity, conciseness and contextuality of non-textual elements

The functions of visual images in texts according to accuracy, connectivity, conciseness and contextuality according to Kim (2012) are discussed below. Each non-textual element has been assigned a category related to its accuracy, connectivity, conciseness and contextuality. A category of 2, 1 or 0 is assigned to each. See Table 5.2 for the full description of each category.

Accuracy

Accuracy of non-textual elements was measured by analysing the visual image against Kim's (2012) framework. The proportion in each category, for each book, is shown in Figure 5.5.

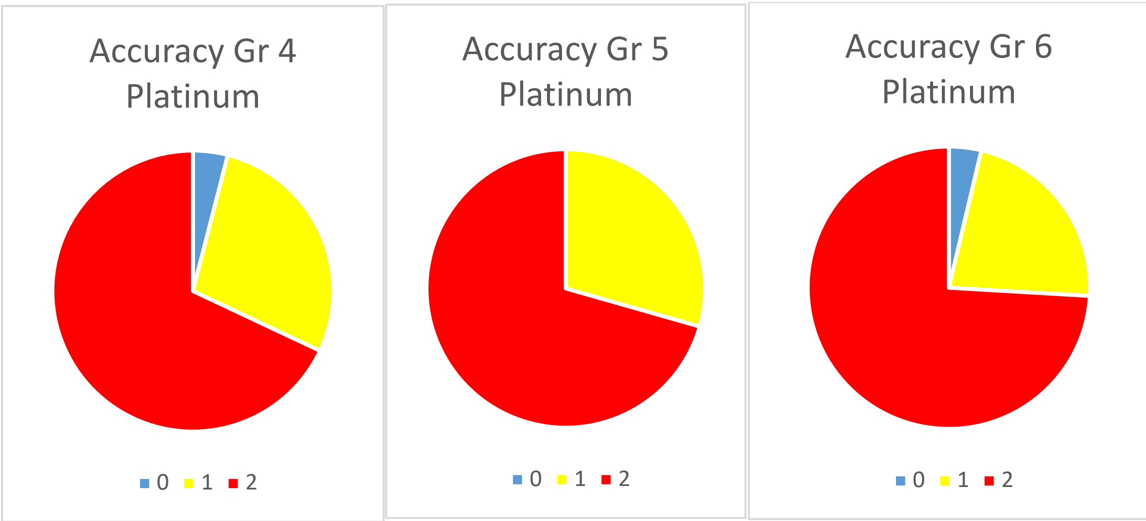
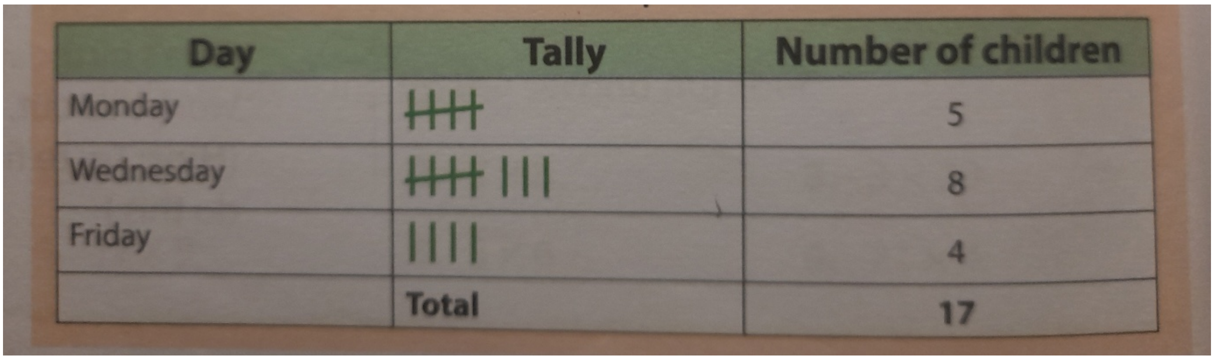


Figure 5.5 Accuracy of non-textual elements (Platinum Gr 4, 5, 6)

In the Platinum textbooks the accuracy of the non-textual elements, shown in Figure 5.5, indicates that the majority of these are accurate to show a concept based on its definition. Some had missing information which categorised them as ‘1’ (Grade 4, 28%; Grade 5, 29%; Grade 6, 22%), but the majority were categorised as ‘2’ (Grade 4, 68%; Grade 5, 71%; Grade 6, 74%).



Day	Tally	Number of children
Monday	卌	5
Wednesday	卌 |||	8
Friday	||||	4
	Total	17

Figure 5.6 Example of accuracy in non-textual elements (Platinum Grade 4, Unit 1)

Figure 5.6 shows a table with the days of the week, the tally and the number of children represented, this is accurate to show the definition of a tally table and was categorised as ‘2’ for accuracy.

Connectivity

In the Platinum textbooks the majority of the non-textual elements are explicitly and fully associated with the mathematical content in the text as shown in Figure 5.7 (Grade 4, 80%; Grade 5, 80%; Grade 6, 74%). This means that they directly show a concept or problem and are fully associated with the mathematical content in the text.

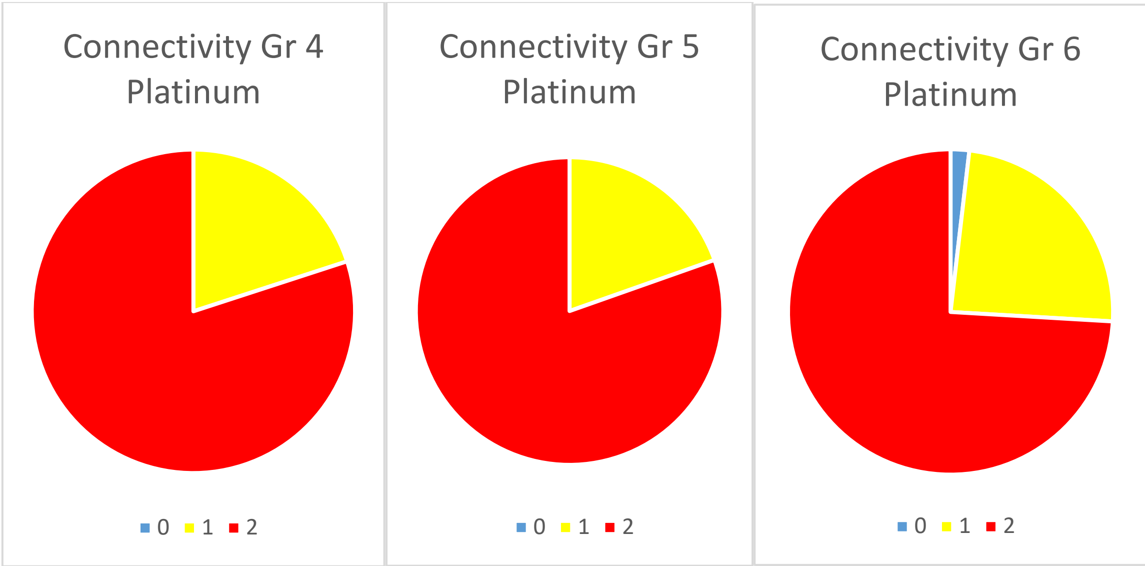


Figure 5.7 Connectivity of non-textual elements (Platinum Grades 4, 5, and 6)

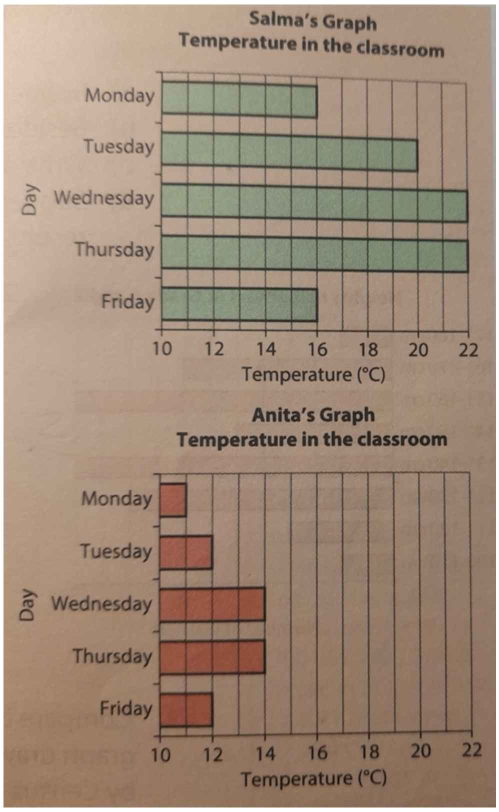


Figure 5.8 Example of connectivity (Platinum Grade 6, Unit 18)

Figure 5.8 shows two bar graphs representing the classroom temperatures measured over a week. It is clearly laid out and and is fully associated with the explanation and questions in the text. It is an example of a non-textual element categorised as '2' according to Kim's (2012) framework.

A smaller percentage of the non-textual elements were only partially related to the mathematical content in the text (Grade 4, 20%; Grade 5, 20%; Grade 6, 24%). There is some missing or irrelevant information in these non-textual elements. It shows content but does not fully align with the mathematical content of the text.

Conciseness

Figure 5.9 shows the relative conciseness of the non-textual elements in the Platinum textbooks.

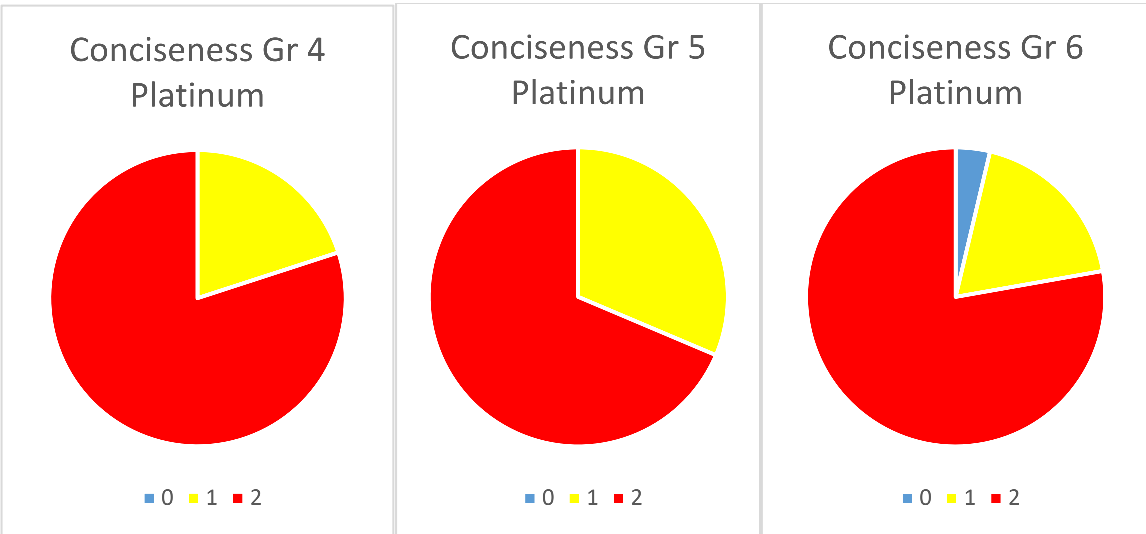
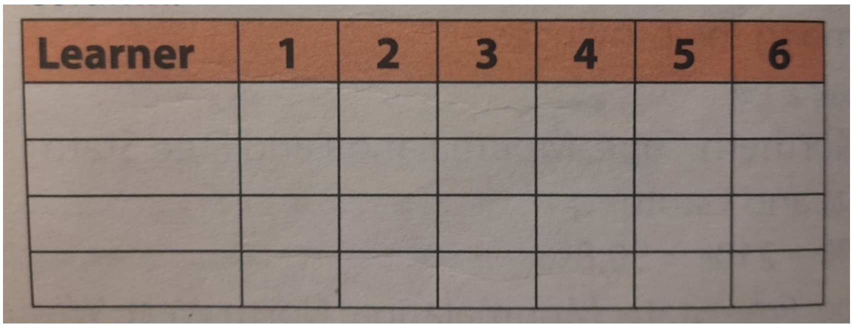


Figure 5.9 Conciseness of non-textual elements (Platinum Grades 4, 5, 6)

The non-textual elements in the Platinum textbooks are mostly concise, being straightforward to show a concept or problem without any distracting factors (Grade 4, 80%; Grade 5, 69%; Grade 6, 77%). There are a smaller proportion of non-textual elements that have some distracting factors (Grade 4, 20%; Grade 5, 31%; Grade 6, 19%). Only 4% of the non-textual elements in the Grade 6 book were evaluated as not concise. Figure 5.10 shows a non-textual element evaluated as ‘2’ for conciseness as it is a table for learners to complete that contains no distracting information and contains no information that is not directly required for the mathematical task.



Learner	1	2	3	4	5	6

Figure 5.10 An example of conciseness (Platinum Grade 6, Unit 5)

Contextuality

Contextuality of the non-textual elements relates to having a realistic object or context used in the non-textual element with a mathematical connection or an appropriate representation of data for the situation or problem.

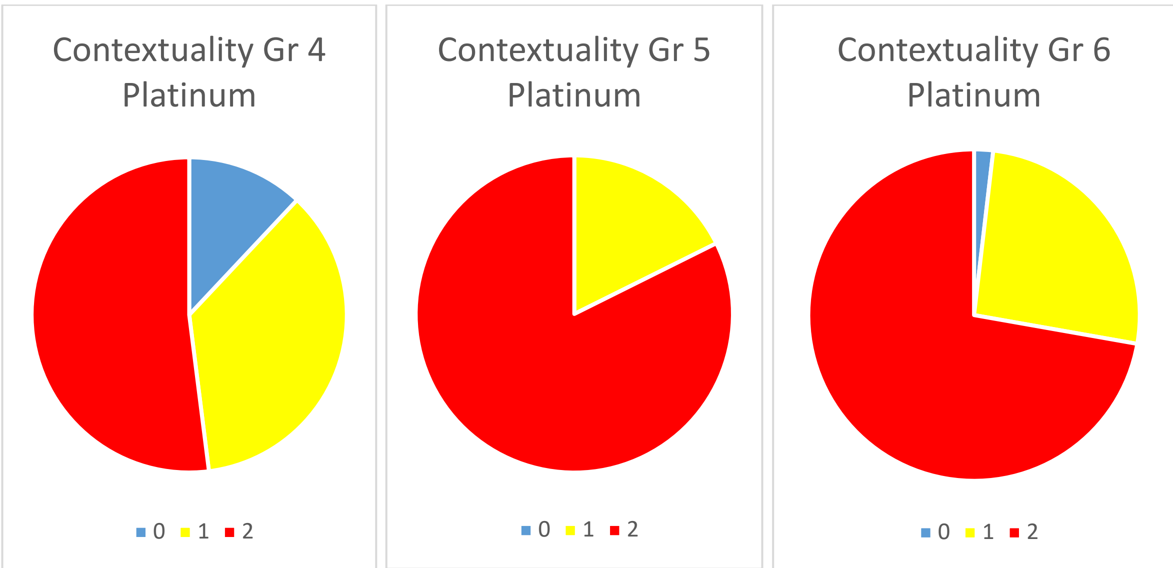


Figure 5.11 Contextuality of non-textual elements (Platinum Grades 4, 5, 6)

While most non-textual elements provide a realistic object or context with mathematical connection (Grade 4, 52%; Grade 5, 82%; Grade 6, 72%), there are some non-textual elements that have not included a realistic object or context in Grade 4 (12%) and Grade 6 (2%).

The Platinum textbooks are colourful and visually appealing, with extra challenges and additional information. This adds to the number of non-textual elements that are included in the units but at times these have little or no mathematical ideas and serve as decoration or additional information that is contextually relevant but only distantly related to the mathematics.

The analysis of the non-textual elements indicated a pattern which generally showed that if the non-textual element is mathematically accurate in representing data, it is generally explicitly connected to the text, straightforward and concise and contextually uses appropriate representation of objects or contexts. An example would be a bar graph that is clearly marked on both axes, with an explanation or questions in the accompanying text (see Figure 5.8). The same can be said for a non-textual element that is completely inaccurate or has obvious errors. It is generally not connected, not concise and contextually unrealistic.

Figure 5.12 shows a pictograph which is categorised as '2' for accuracy, connectivity, conciseness and contextuality. This non-textual element is mathematically accurate, connected to the text, very straightforward and makes use of a realistic context and a contextually-linked icon for the pictograph. Although it would not allow the display of an amount of, e.g. 186 tickets, for the information that it does present it is categorisable at '2' for all aspects of Kim's (2012) framework.

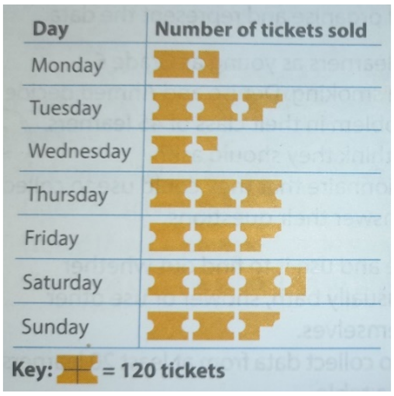


Figure 5.12 An example of a NTE in Platinum Grade 6

Figure 5.13 provides an example of a non-textual element that is inaccurate, partly connected to the text, concise and partly contextual. It is a pictograph representing litter found in a school playground before, during and after an anti-litter project. It has a key indicating that 1 symbol of a cylinder represents 5 pieces of litter. The pictograph uses images of half a cylinder symbol to represent half of 5. This is 'inaccurate' in terms of the concept and context. Half of 5 is 2.5, and it is not possible to have 2.5 items of litter. It is however connected to the text, and straightforward enough to show a concept without any distracting factors. It is partly contextual as some litter may be cylindrical, but not all litter will be.

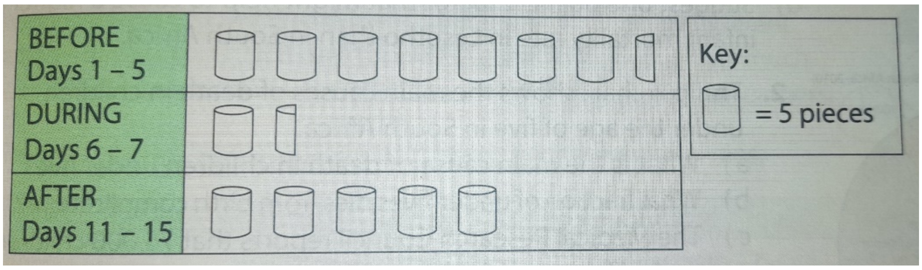


Figure 5.13 An example of a NTE (Platinum Grade 6, Unit 9)

5.2 Analysis of non-textual elements in the Premier series

In this section I present the analysis of the non-textual elements included in the Premier series. The Premier textbook is printed in black and white, there are no colourful images included. There are overall fewer non-textual elements than in the Platinum series.

5.2.1 Graphical languages in the Premier series

The table below summarises the non-textual elements and their graphical languages.

Table 5.4 Graphical languages of the Premier textbooks

	Grade 4	Grade 5	Grade 6
Tables	11	11	9
Single position languages	2	5	1
Apposed position languages	2	14	11
Retinal list languages	0	0	0
Map languages	0	0	0
Connection languages	0	4	0
Miscellaneous languages (pie charts)	3	0	5
Images	7	6	11

Tables were the most frequently used non-textual elements in the Grade 4 Premier textbook (11), apposed position languages were most frequently used in the Grade 5 book (14) and apposed position languages and images were most frequently used in the Grade 6 book (11). Apposed position languages, specifically bar graphs and some pictographs, were used more in Grade 5 than in the other two grades. There was no use of retinal list languages or map languages in any of the books and connection languages only appeared in the Grade 5 book. Miscellaneous languages, specifically pie charts, were used in Grades 4 and 6.

5.2.2 Accuracy, connectivity, conciseness and contextuality of non-textual elements

Accuracy

The accuracy of the non-textual elements in the Premier books is presented in the pie charts below in Figure 5.14. It can be seen that most of the non-textual elements in Grade 4 (72%) and Grade 6 (58%) are categorised as ‘1’, in other words, they make sense in terms of the definition or the meaning of a concept but don’t show every required mathematical condition or some attributes are missing or misleading. In Grade 5 (65%), the majority of the non-textual elements are accurate to show the definition of a concept or are accurate to show a concept based on its definition.

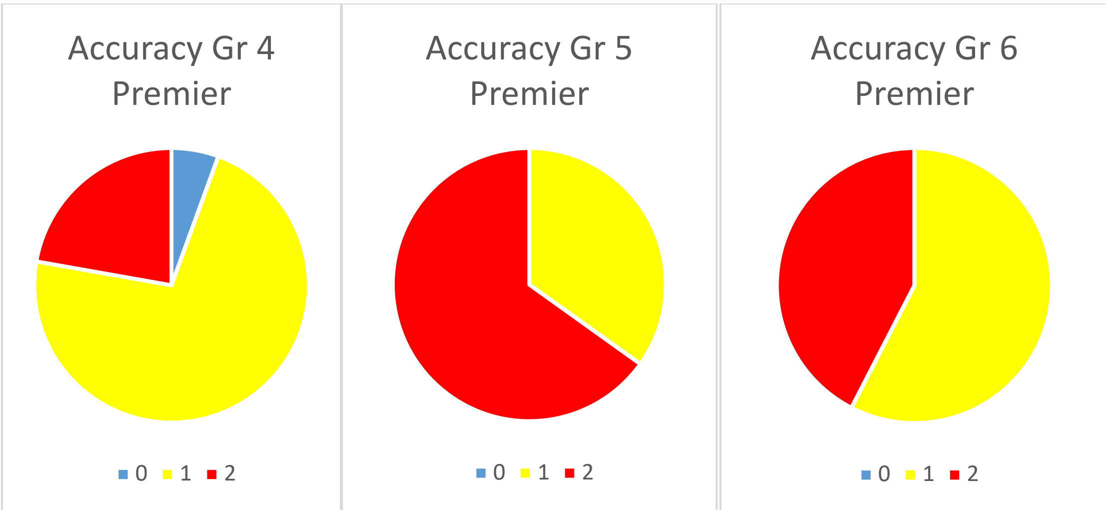


Figure 5.14 Accuracy of NTEs in the Premier textbooks Grades 4, 5, 6

In Figure 5.15, an example of a non-textual element is given that is categorised as ‘1’ in relation to accuracy as it does not show every required mathematical condition. There is no title given, which means that the reader is unable to interpret what the pie chart is in reference to. The key is included, but because of the greyscale used, this too is confusing.

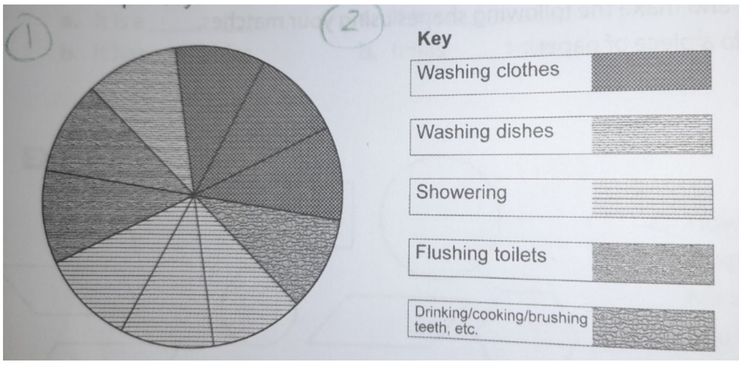


Figure 5.15 Example of a NTE categorised as ‘1’ for accuracy (Premier Grade 4, Unit 8)

Below is another example of a non-textual element that does not show every required mathematical condition.

Colour	Tally	Total number
Red		
Yellow		
Blue		
Orange		
Purple		
Green		
Pink		

Figure 5.16 An example of a NTE categorised as ‘1’ for accuracy (Premier Grade 4, Unit 4)

This is a tally table which is partially filled in, but makes no mathematical sense without the text to guide it. No title is provided, and this missing information means that the table is impossible to interpret without reading the accompanying text. Similarly, the pie chart shown in Figure 5.17 has no key and thus cannot be interpreted without reading and comprehending the accompanying text.

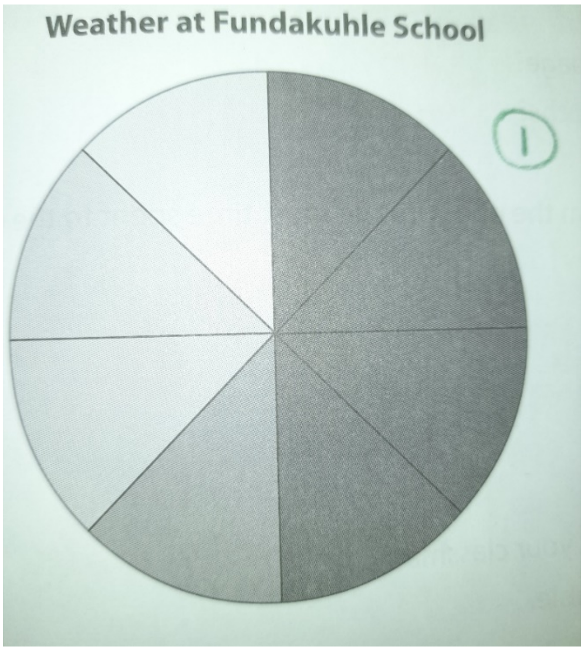


Figure 5.17 An example of a NTE categorised as '1' for accuracy (Premier Grade 4, Unit 6)

Figure 5.18 provides another example of 'inaccuracy' (categorised as '1' on accuracy) in a non-textual element. It is an empty chart for learners to use to draw a bar graph. However, no scale is indicated on the horizontal axis and there is no title included. In addition, no textual guidelines are given for the development of this graph.

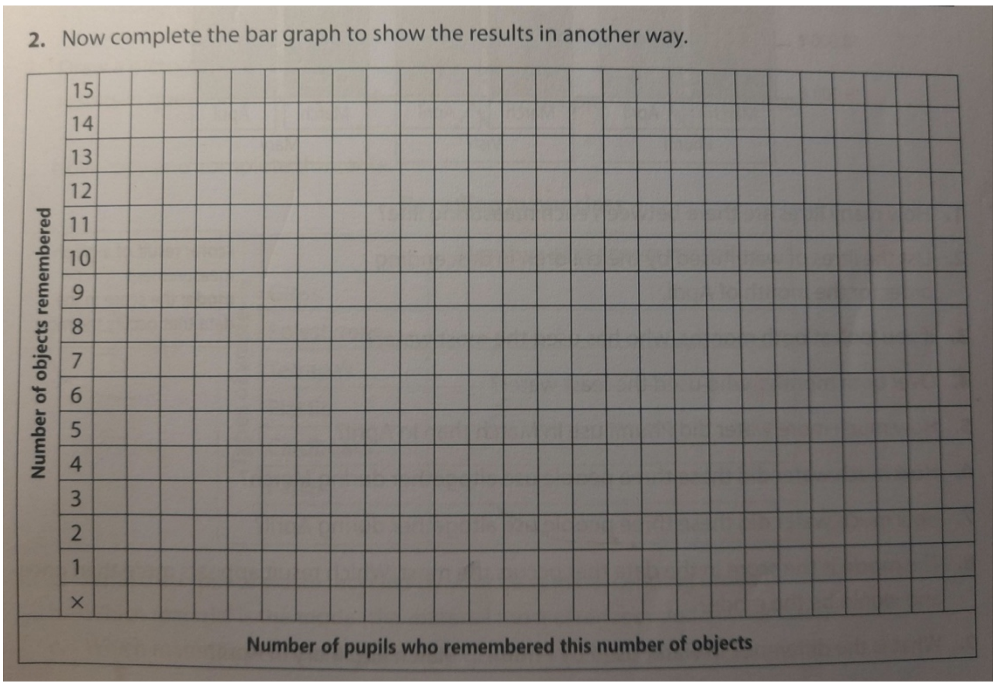


Figure 5.18 An example of 'inaccuracy' due to missing information in a NTE (Premier Grade 6, Unit 1)

The limited accuracy in many of the non-textual elements in the Premier books will make comprehension of these elements more difficult.

Connectivity

The connnectivity of the non-textual elements in the Premier textbooks varied between the books. The Grade 4 book had the smallest proportion (39%) of non-textual elements that were fully connected to the text. The Grade 5 book had the majority (84%) of the non-textual elements as fully connected to the text and in the Grade 6 book there were approximately half of the non-textual elements fully connected to the text (52%). This is shown in Figure 5.19.

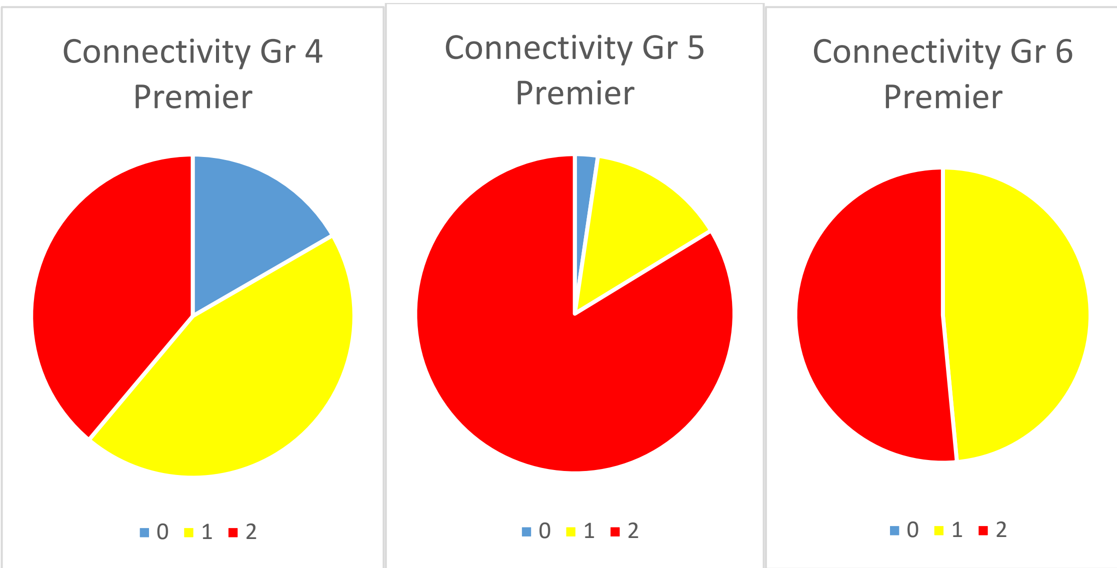
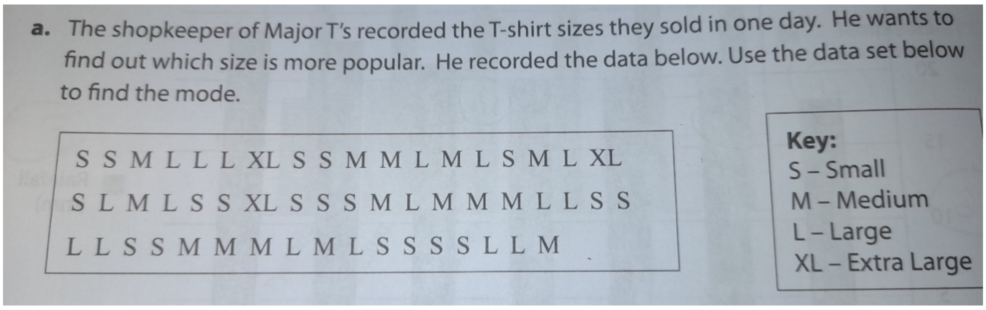


Figure 5.19 Connectivity of NTEs in Premier textbooks Grades 4, 5, 6

In Grade 4 the majority (44%) of the non-textual elements are partly related to the mathematical content in the text (categorised as ‘1’). In Grade 5 there are fewer of these (14%) and in Grade 6, almost half the non-textual elements are in this category (48%). In Grade 5 most (84%) of the non-textual elements have a realistic object or context used in the non-textual element with a mathematical connection or an appropriate representation of data for the situation or problem (categorised as ‘2’). Figure 5.20 shows an example of a non-textual element that is connected to the text. It directly shows the sizes of t-shirts sold when used with the key provided.

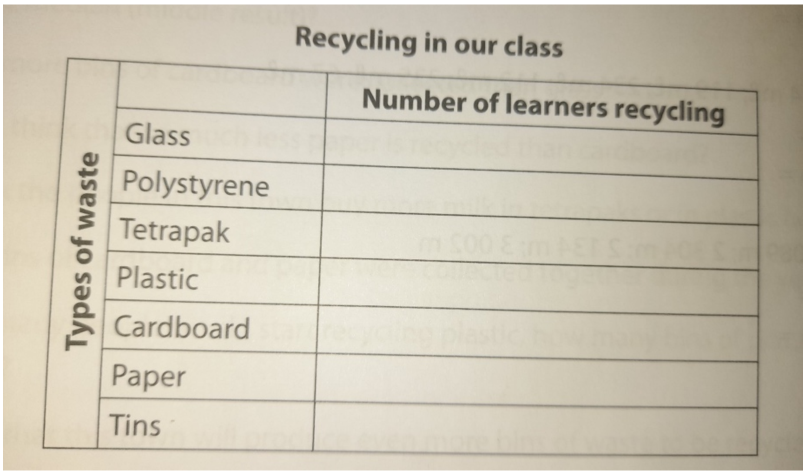

a. The shopkeeper of Major T's recorded the T-shirt sizes they sold in one day. He wants to find out which size is more popular. He recorded the data below. Use the data set below to find the mode.

S S M L L L XL S S M M L M L S M L XL
S L M L S S XL S S S M L M M M L L S S
L L S S M M M L M L S S S S L L M

Key:
S – Small
M – Medium
L – Large
XL – Extra Large

Figure 5.20 A NTE with mathematical connectivity (Premier Gr 5, Unit 9)

In Grade 6, 16 non-textual elements (48%) were only partially related to the mathematical content in the text. There is some missing or irrelevant information in these non-textual elements. Figure 5.21 is an example in which there is some missing information in the non-textual element. The number of learners is not filled in as this is an activity for the learners to complete.


Recycling in our class

Types of waste		Number of learners recycling
	Glass	
	Polystyrene	
	Tetrapak	
	Plastic	
	Cardboard	
	Paper	
	Tins	

Figure 5.21 A NTE that is not fully connected (Premier Grade 5, Unit 3)

Compared to the Platinum textbooks, the connectivity of the non-textual elements in the Premier textbooks shows many more being only partly related to the mathematical content in the text, with some missing or irrelevant information.

Conciseness

The pie charts in Figure 5.22 represent the proportion of non-textual elements in each grade in relation to how concise they are in showing the mathematical concept or problem. As with the connectivity of non-textual elements, this varies between the books. In Grade 5 the majority of them are considered concise (84%), whereas in Grades 4 (50%) and 6 (48%) this proportion is much lower.

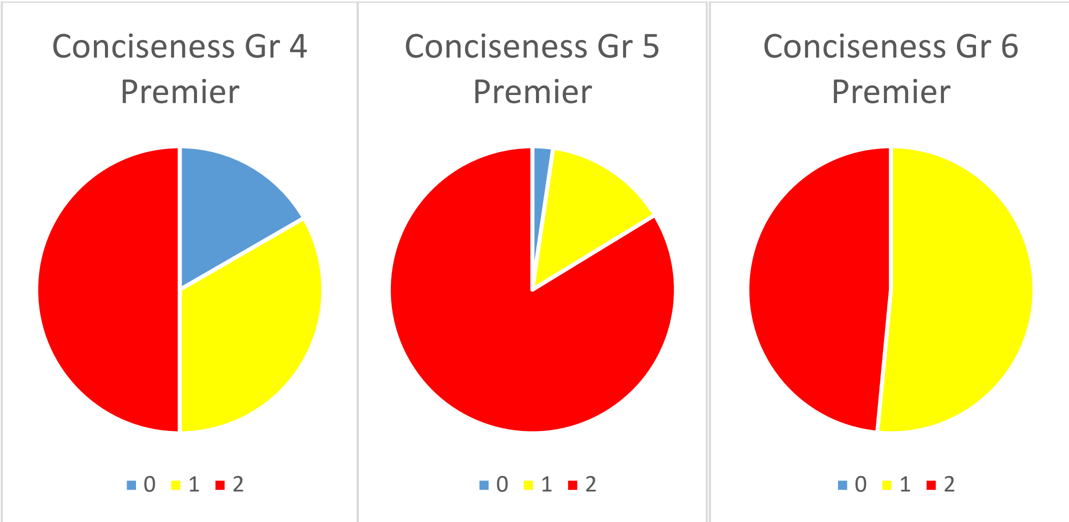
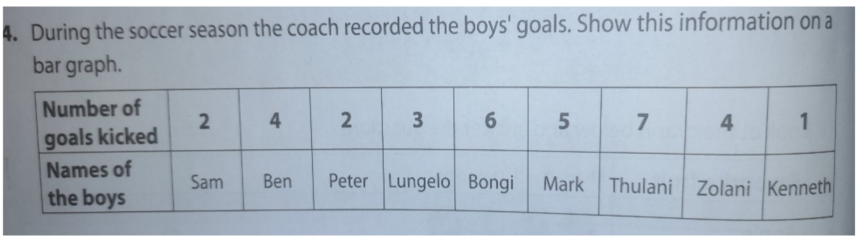


Figure 5.22 Conciseness of NTEs in the Premier textbooks Gr 4, 5, 6.

The non-textual elements in the Platinum textbooks are generally more concise than in the Premier books. There are more straightforward non-textual elements showing the concept without any distractors in the Platinum books(see Figure 5.9). The Premier textbooks had more non-textual elements that had some additional factors. At times these factors were useful in understanding the concept (categorised as ‘1’), but at times they were distractors (categorised as ‘0’). Figure 5.23 provides an example of a non-textual element categorised as ‘2’ for conciseness – there are no additional factors that are either helpful or a distractor.


4. During the soccer season the coach recorded the boys' goals. Show this information on a bar graph.

Number of goals kicked	2	4	2	3	6	5	7	4	1
Names of the boys	Sam	Ben	Peter	Lungelo	Bongi	Mark	Thulani	Zolani	Kenneth

Figure 5.23 A NTE showing conciseness (Premier, Grade 5, Unit 1)

Figure 5.24 provides another example of a non-textual element showing conciseness.

Activity	Bathing and showering	Laundry	Washing up	Watering the garden
Number of litres used per week				

Figure 5.24 A NTE showing conciseness (Premier Grade 6, Unit 7)

Contextuality

Figure 5.25 shows a summary of the proportions of non-textual elements in each category for contextuality in the Premier textbooks. Realistic objects and contexts are used in most of the non-textual elements in the Grade 5 book (79%) and in half of the non-textual elements in the Grade 4 book. In the Grade 6 book the greatest proportion is of non-textual elements categorised as '1' (58%), that is, there is no mathematical idea or concept present, but there is some realistic contextual information to provide contexts or facilitate related activities.

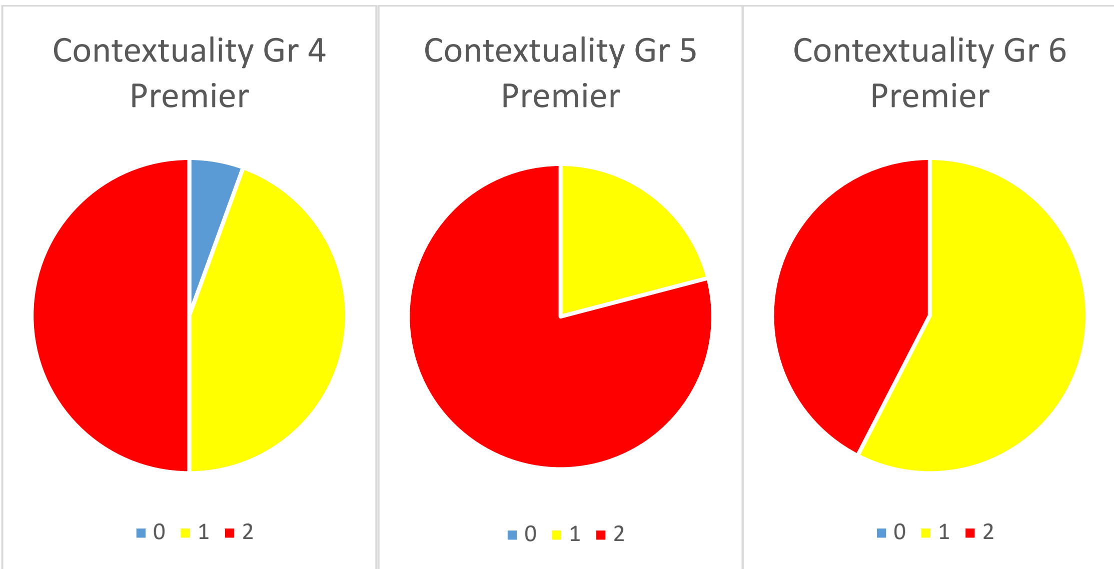


Figure 5.25 Contextuality of NTEs in Premier textbooks Grades 4, 5, 6.

The Premier Grade 4 and Grade 5 textbooks are similar to the Platinum books in terms of contextuality of the non-textual elements. The Premier Grade 6 book, by comparison, has a far larger proportion of non-textual elements that are categorised as '1' than the Platinum books.

5.3 Analysis of the non-textual elements in the Rainbow workbook series

5.3.1 Graphical languages in the Rainbow workbook series

The DBE-issued Rainbow workbooks are printed in colour and contain non-textual elements that include all the graphical languages except map languages and connection languages. A summary of the graphical languages used is provided in Table 5.5. The Rainbow workbooks contain the most non-textual elements of all the books analysed for this research, with particularly high numbers of 'images' used.

Table 5.5 Graphical languages of the Rainbow workbooks

	Grade 4	Grade 5	Grade 6
Tables	7	11	8
Single position languages	7	9	20
Apposed position languages	2	2	5
Retinal list languages	8	2	1
Map languages	0	0	0
Connection languages	0	0	0
Miscellaneous languages (pie charts)	2	2	9
Images	27	22	21

Aside from the non-textual elements classified as 'images', retinal list languages are used the most in Grade 4, tables and single position languages are most prevalent in Grade 5 and the use of single position languages is most prevalent in Grade 6. The majority of these single position languages are pictographs.

This is contrary to the findings in the other books, where apposed position languages are generally used the most frequently. Apposed position languages are not among the most frequently used in the Rainbow books, although there is an increase in Grade 6 and this is the only book in this research that includes a double bar graph.

5.3.2 Accuracy, connectivity, conciseness and contextuality of non-textual elements

Accuracy

The accuracy of the non-textual elements in the Rainbow workbooks is summarised in Figure 5.26. In contrast to the Platinum and Premier books, there are large proportions of non-textual elements that are categorised as '0' for accuracy. This is partly due to the inclusion of many 'images' in which no mathematical concept is present.

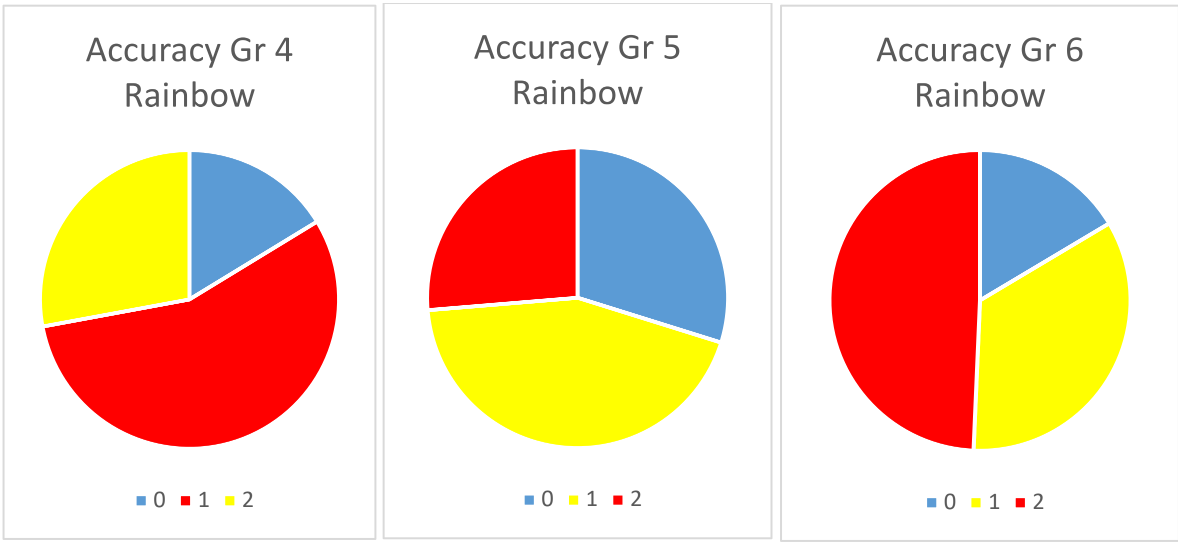


Figure 5.26 Accuracy of NTEs in Rainbow workbooks Grades 4, 5, 6

The accuracy of the non-textual elements in the Rainbow books varies widely. In the Grade 4 and Grade 6 books the largest proportion is of non-textual elements that are accurate to show a concept (Grade 4, 56%; Grade 6, 49%). The largest proportion for the Grade 5 book (44%) is of non-textual elements that make sense in terms of the definition of a concept, but that do not show every required mathematical condition (categorised as '1' for accuracy). Relative to all of the other books, the Rainbow books have the largest proportion of non-textual elements that are classified as '0', that is, they are inaccurate in terms of the definition of a concept, or have no mathematical concept in the non-textual element (Grade 4, 16%; Grade 5, 30%; Grade 6, 16%). An example is provided in Figure 5.27. This element communicates no mathematical concept in the context of the mathematical content on the page.

Figure 5.27 A NTE with no mathematical concept (Rainbow Grade 5, Unit 5)

Connectivity

Regarding connectivity, the Rainbow workbooks again show the most variation of the books analysed in this research. A summary of the connectivity of the non-textual elements is provided in Figure 5.28.

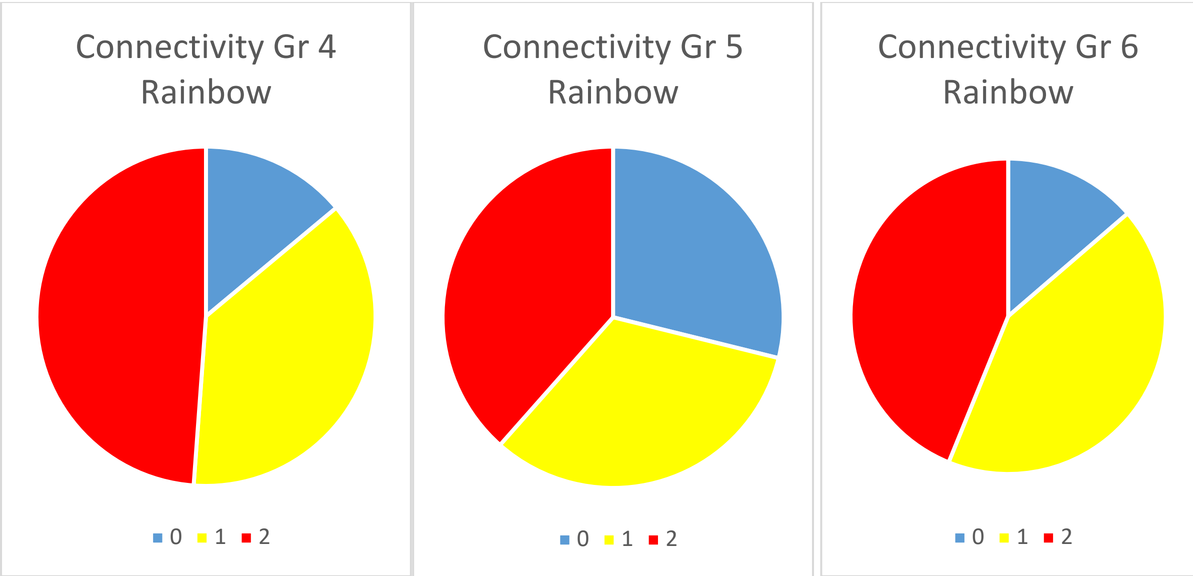


Figure 5.28 Connectivity of NTEs in the Rainbow workbooks Grades 4, 5, 6

The largest proportion of non-textual elements in Grade 4 (49%), Grade 5 (38%) and Grade 6 (44%) are fully and explicitly associated with the mathematical content in the text, directly showing a concept or problem. There are also many non-textual elements that are only partially related to the mathematical content in the text, but do not explicitly show how it is connected with the content (Grade 4, 37%; Grade 5, 33%; Grade 6, 42%). In the Grade 5 book, in particular, there are many non-textual elements that have nothing to do with the content mathematically (29%), only giving clues about contexts or providing extra, unrelated information. This proportion is also relatively high for Grades 4 (14%) and 6 (14%). Figure 5.29 gives one such example.

Figure 5.29 A NTE with no connectivity (Rainbow Grade 4, Unit 1)

Conciseness

When considering the conciseness of the non-textual elements, the Rainbow workbooks again show the most variation of the books included in this research. For Grade 5 (35%) and Grade 6 (49%), the highest proportion of non-textual elements are straightforward in showing a concept without the inclusion of any distracting factors (i.e. coded as '2'). For Grade 4, the highest proportion (40%) is of non-textual elements that are straightforward but that include additional factors that may be helpful to understand the concept (i.e. coded as '1').

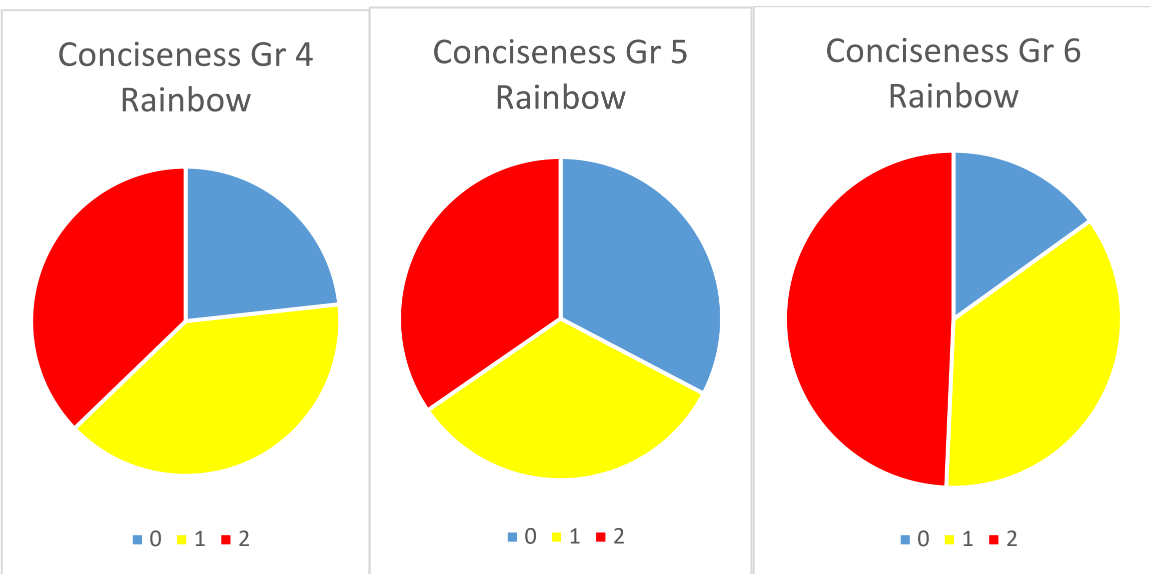


Figure 5.30 Conciseness of NTEs in Rainbow workbooks Grades 4, 5, 6.

Relative to the other books analysed, there is a large proportion of non-textual elements that include distracting or other factors that are not useful to show a concept (Grade 4, 23%; Grade 5, 33%; Grade 6, 15%). An example is shown in Figure 5.31 where the image of the girl is not useful to show a concept and does not provide useful contextual information.

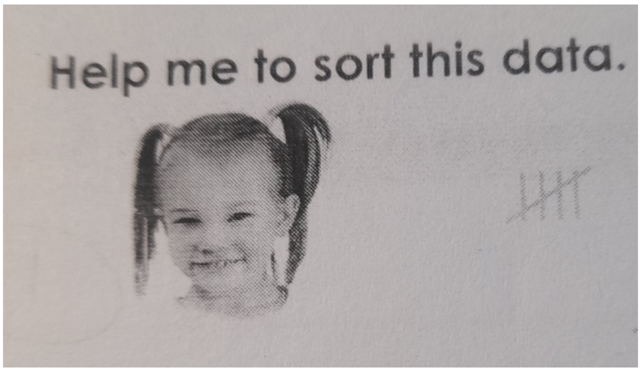


Figure 5.31 A NTE categorised as '0' for conciseness (Rainbow Grade 6, Unit 1)

Contextuality

For contextuality to be coded as '2', according to Kim (2012), a realistic object or context is used with a mathematical connection in the non-textual elements. The Grade 6 workbook has the largest proportion of such non-textual elements (53%). In the Grade 4 (60%) and Grade 5 (58%) books the largest proportion is of non-textual elements coded as '1', indicating that there is some realistic contextual information, but no mathematical ideas are present. As with the other three of Kim's (2012) measures, i.e. accuracy, connectivity and conciseness, the Rainbow workbooks show the most variation of the books included in this research.

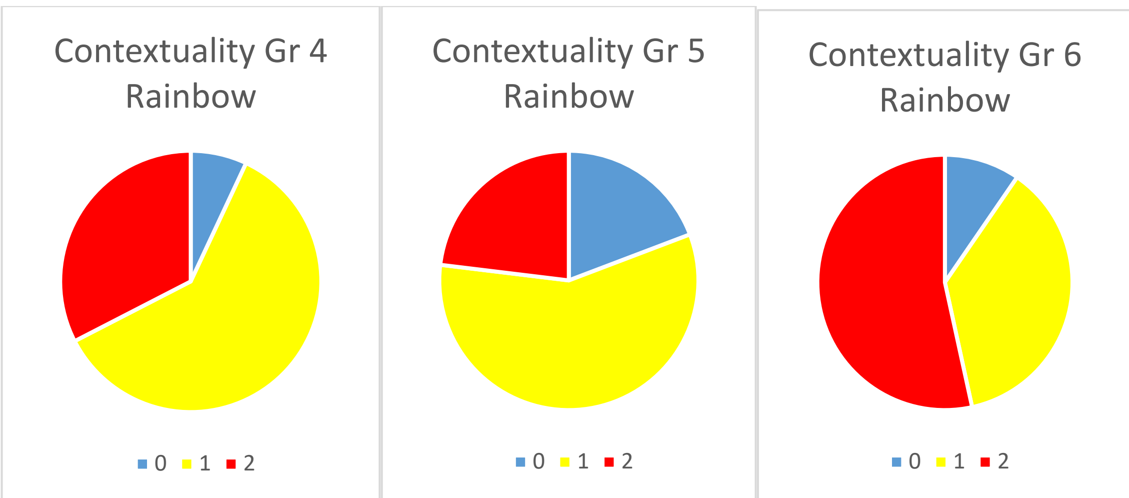


Figure 5.32 Contextuality of NTEs in Rainbow workbooks Grades 4, 5, 6.

The Rainbow workbooks have many non-textual elements which only provide realistic contextual information without the existence of mathematical ideas or concepts. There are also more non-textual elements that have neither realistic context nor realistic objects included, than in any of the other books in the study. This could be due to the number of visual elements on a typical page in the Rainbow workbook. Figure 5.33 gives a typical double page spread of the Rainbow workbook.

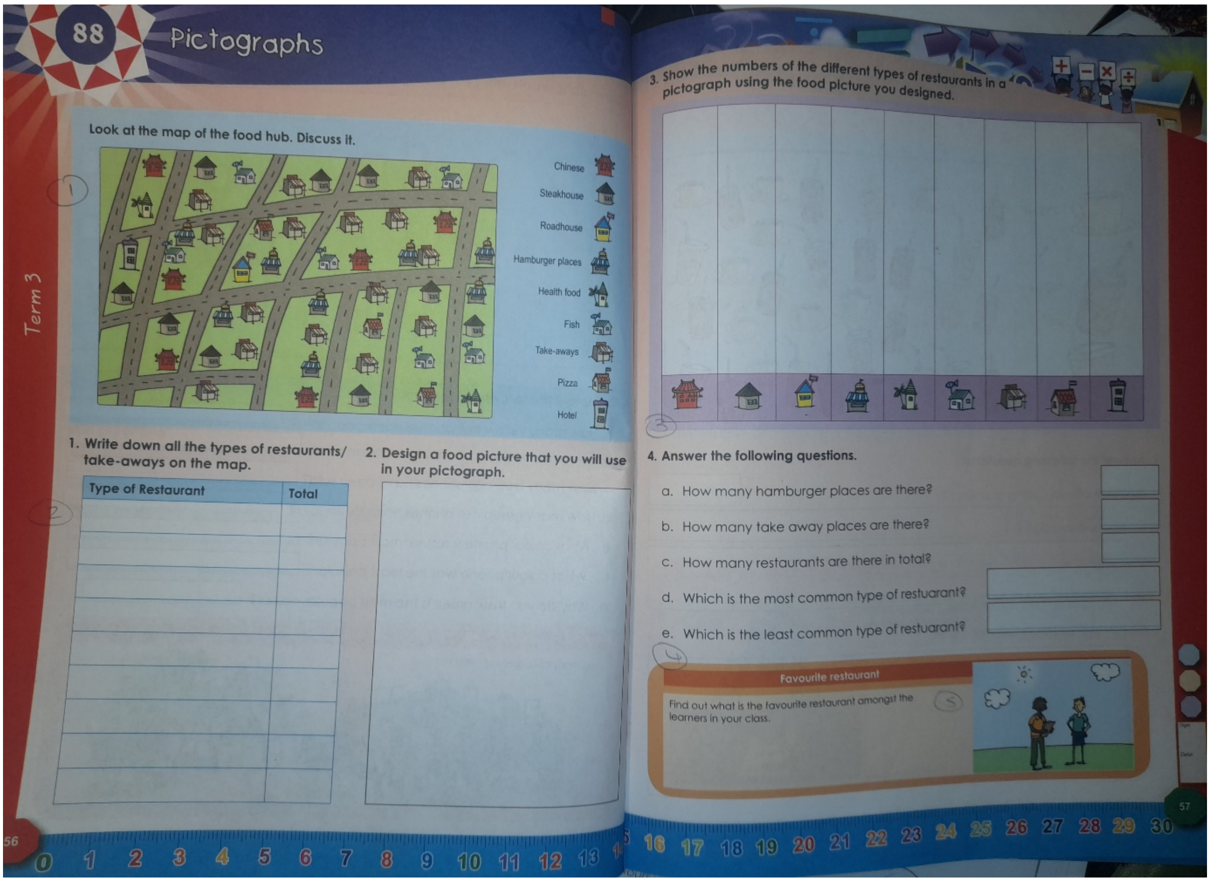


Figure 5.33 An example of a double page in the Rainbow workbook (Grade 4, Unit 6).

The Rainbow books are the books, of all those analysed, with the most non-textual elements due to the high number of images included. They are also the books with the greatest proportion of non-textual elements that are classified as ‘0’ for accuracy, conciseness, connectedness and contextuality. There are many examples of images that are not connected to the mathematical content of the unit or the contextual information in the unit. This may present challenges to learners’ comprehension.

5.4 Analysis of the non-textual elements in the Inzalo series

5.4.1 Graphical languages in the Inzalo series

The graphical languages of the non-textual elements in the Inzalo textbook are presented in Table 5.6. The Inzalo books contained the least number of non-textual elements of the books analysed.

Table 5.6 Graphical languages of the Inzalo textbooks

	Grade 4	Grade 5	Grade 6
Tables	4	9	5
Single position languages	0	2	0
Apposed position languages	5	3	6
Retinal list languages	0	0	2
Map languages	0	0	2
Connection languages	0	0	0
Miscellaneous languages (pie chart)	5	6	3
Images	3	2	14

Apposed-position languages (5) and miscellaneous languages in the form of pie charts (5) are the languages with the highest frequency in the Grade 4 textbook. In Grade 5 it is tables that occur with the highest frequency (9), and in Grade 6 it is images (14) and apposed position languages (6) that occur with the highest frequencies. Single position languages only appear in Grade 5 and map languages and retinal list languages only appear in Grade 6. One example of a pictograph that takes the form of an apposed position language is in Figure 5.34. It is an apposed position language in this case as there is a vertical and a horizontal axis labelled.

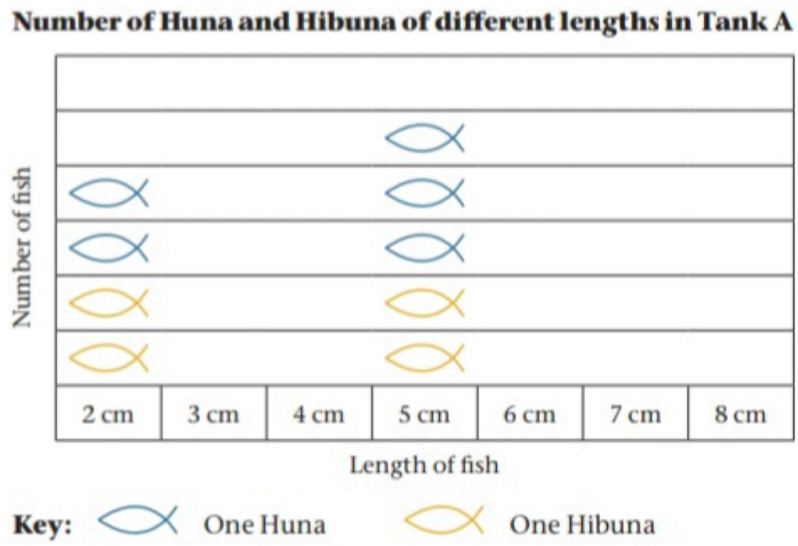


Figure 5.34 An example of a pictograph (Inzalo Grade 4, Unit 3)

5.3.2 Accuracy, connectivity, conciseness and contextuality of non-textual elements

Accuracy

The accuracy of the non-textual elements in the Inzalo textbooks can be seen in the pie charts in Figure 5.35.

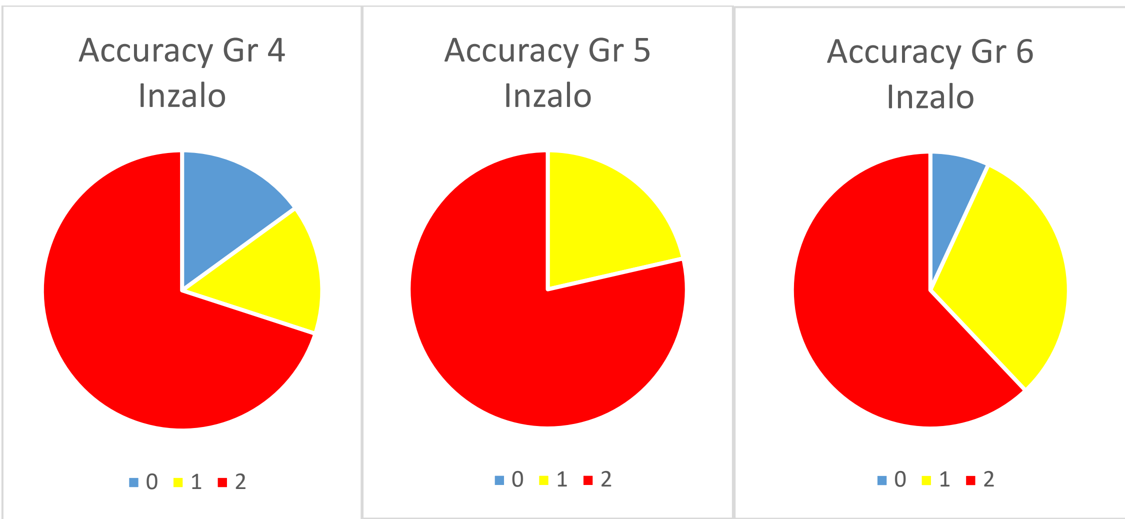


Figure 5.35 Accuracy of Inzalo textbooks Gr 4,5,6.

For all grades the highest proportion is of non-textual elements that correctly show the definition of a concept (Grade 4, 70%; Grade 5, 79%; Grade 6, 62%). The proportions are most similar to the accuracy of the non-textual elements in the Platinum series.

One example of a non-textual element that is accurate is provided in Figure 5.36. The map, with the key, is accurate to show the concept. Without the key, the map and the pie charts would be meaningless. The presence of the key is critical. While it is accurate, this is a complex non-textual element to analyse and understand. It contains map languages and miscellaneous languages in the form of pie charts, as well as containing textual and numerical information to process.

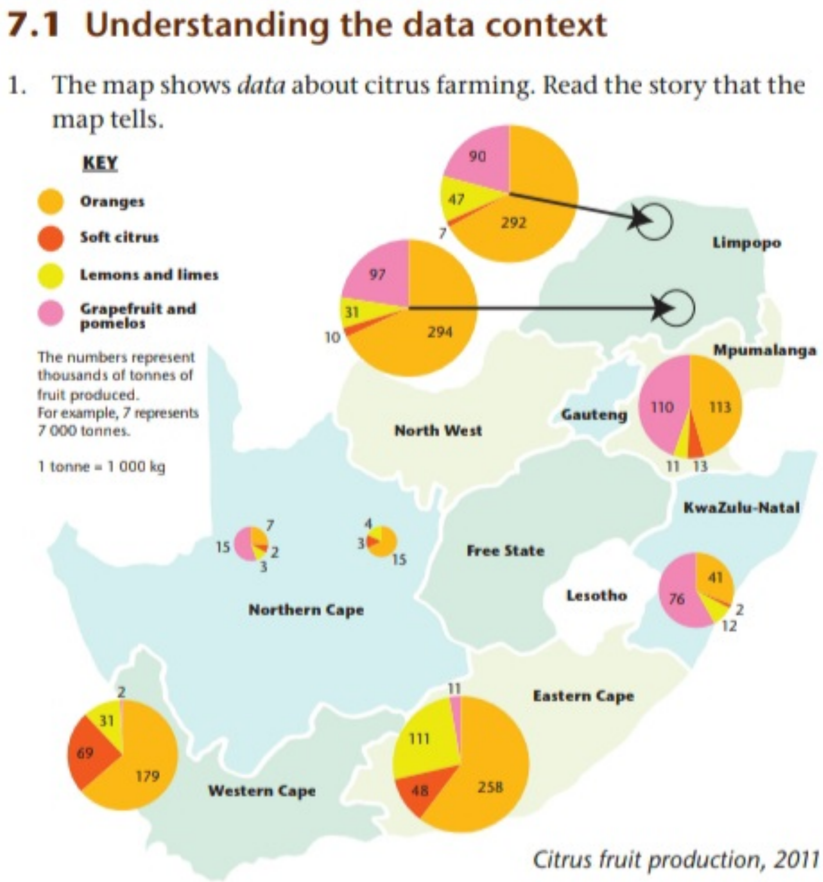


Figure 5.36 An example of an accurate NTE (Inzalo Grade 6, Unit 1)

In the Grade 4 and Grade 6 books there are also several non-textual elements that are classified as ‘0’ for accuracy as they are not related to a mathematical concept. The picture of the three types of fish in Grade 4, Unit 1, is such an example. It provides contextual information, but is not related to the mathematical concepts.

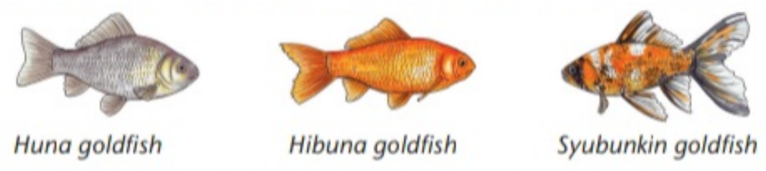


Figure 5.37 A NTE which is classified as '0' for accuracy (Inzalo Grade 4, Unit 1)

Connectivity

Most of the non-textual elements in Grade 4 (80%) and Grade 5 (86%) are fully and explicitly associated with the mathematical content in the text, directly showing a concept or a problem. In Grade 6 this proportion is lower but still represents the majority (52%). The highest proportion of non-textual elements that are only partly related to the mathematical content in the text is in the Grade 6 book (45%). The Grade 6 book is also the only book to have a non-textual element that has nothing to do with the content mathematically (1 NTE, 3%).

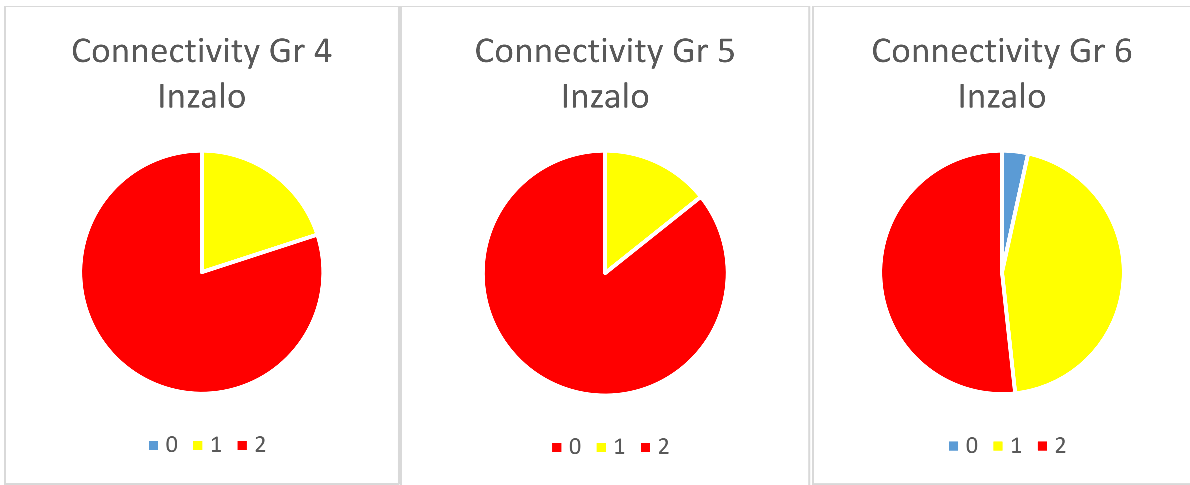


Figure 5.38 Connectivity of NTE in Inzalo textbooks Gr 4, 5, 6

Figure 5.39 shows an example of a non-textual element that is partly related to the mathematical content in the text. Without the accompanying text that names the machine pictured as a sorting machine, the non-textual element has limited connectivity to the mathematical content.

Figure 5.39 An example of an NTE partly related to the mathematical content (Inzalo Grade 6, Unit 3)

Conciseness

The conciseness of the non-textual elements in the Inzalo textbooks is presented in the pie charts in Figure 5.40. The majority of the non-textual elements are straightforward to show a concept or problem, without any distracting factor (Grade 4, 75%; Grade 5, 71%; Grade 6, 55%).

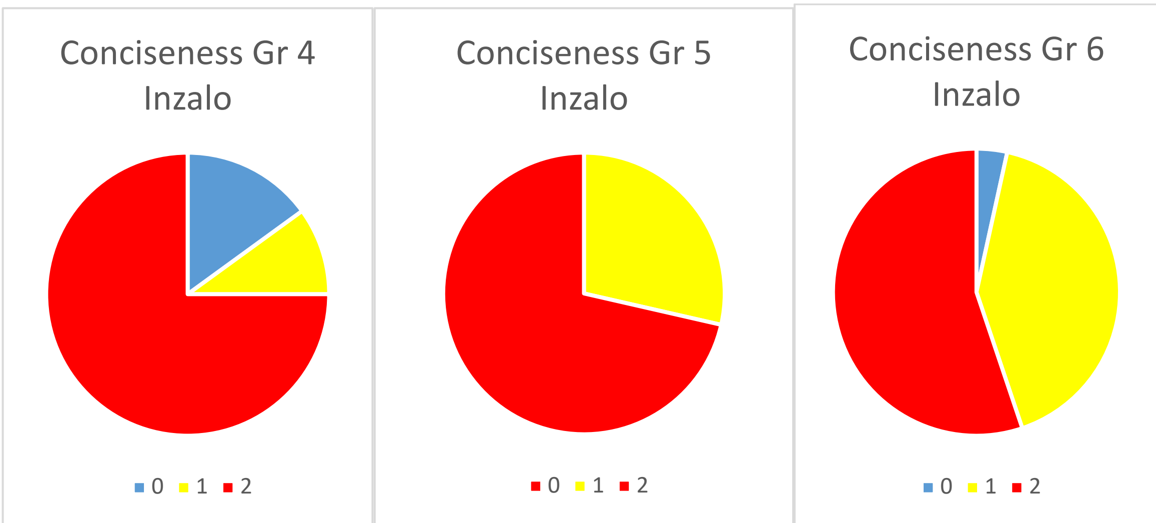


Figure 5.40 Conciseness of NTE in Inzalo textbooks Gr 4,5,6.

Figure 5.41 below shows an image that is straightforward to show a concept. It provides the sample of fish that learners are required to work with in the problems that follow. Learners are to measure each fish from nose to tail (between the dotted lines) and then summarise their lengths. There are no distracting factors in this non-textual element.

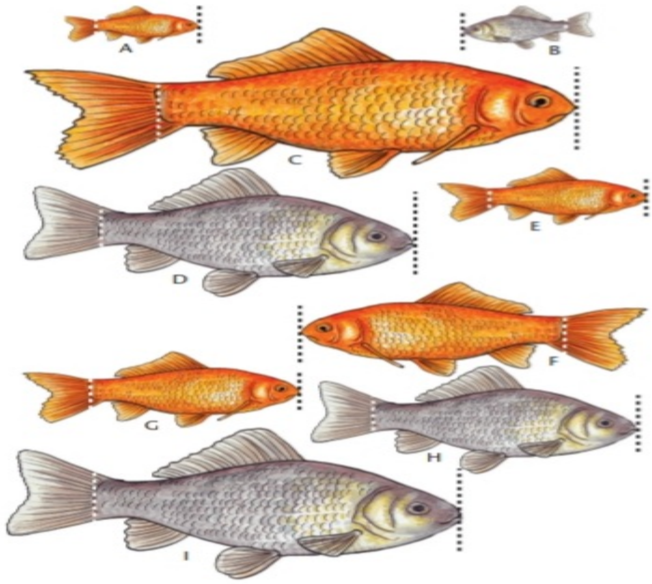


Figure 5.41 An example of a NTE with conciseness (Inzalo Grade 4, Unit 1)

Contextuality

The contextuality of the non-textual elements in the Inzalo textbooks is presented in Figure 5.42. In the Grade 5 (86%) and Grade 6 (72%) books, the majority of the non-textual elements take the form of a realistic object or context with mathematical connection. In the Grade 4 book, the majority are classified as '1' for contextuality (65%), meaning that there is some realistic contextual information used to provide the contexts or objects in the problem. There are no non-textual elements in any of the books that a classified as '0' for contextuality.

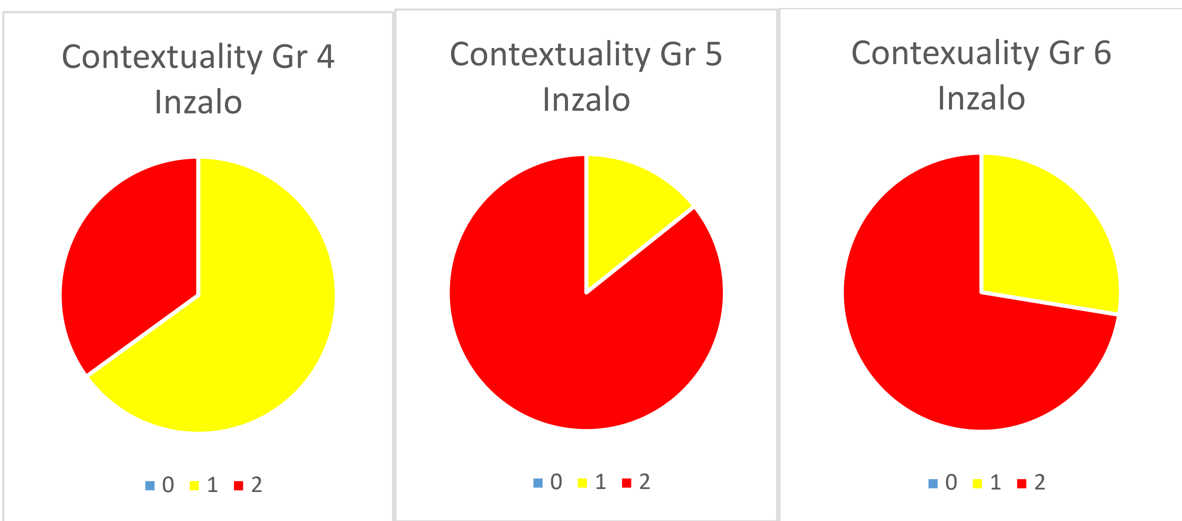


Figure 5.42 Contextuality of NTE in Inzalo textbooks Grade 4, 5, 6.

Figure 5.43 provides an example of a non-textual element classified as '1' for contextuality. It provides some realistic contextual information, but is not connected to the mathematics the learners are engaged in in this unit.

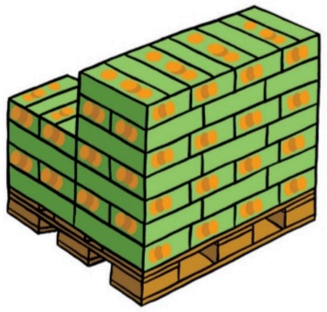


Figure 5.43 A NTE classified as '1' for contextuality (Inzalo Grade 6, Unit 2)

5.5 Conclusion

In this chapter I have presented an analysis of the non-textual elements in the four book series. I have categorised the non-textual elements according to their graphical language (MacKinley, 1986). Certain graphical languages represent the curriculum content of the data handling section: tables, single position languages (e.g. some pictographs), apposed position languages (e.g. bar graphs and some pictographs) and miscellaneous languages (e.g. pie charts). Learners will be learning how to interpret and create these non-textual elements, and thus may need support in comprehending those in particular as they will be becoming familiar with these representations.

In addition I have analysed the non-textual elements according to their accuracy, connectedness, conciseness and contextuality (Kim, 2012). In the three textbook series analysed most non-textual elements were accurate, connected, concise and contextual and thus should support comprehension. In the Rainbow workbooks there were a higher proportion than in the other books of non-textual elements that were not accurate, connected, concise or contextual. These non-textual elements may compromise comprehension as they are not connected to the mathematical content or the context described in the text.

CHAPTER 6

DISCUSSION AND CONCLUSION

According to Allington (2002), the magnitude of a textbooks' usefulness is dependent on how readable and comprehensible they are to the learners. Text that is beyond learners' reading competence results in frustration. Thus, textbooks need to be aimed at the correct reading level, taking comprehension and vocabulary into account. The final chapter of this study aims to provide an interpretation of the analysis of data presented in the previous chapters. The discussion will be structured by the research questions in an effort to address the main research question: What are the language comprehension demands of English mathematical texts on data handling that are used in South African Intermediate Phase Mathematics classrooms? First, the readability of the texts will be considered in order to answer sub-question 1. Thereafter, the linguistic complexity of the texts will be discussed, including a discussion on the particular language features that contribute most frequently to the linguistic complexity in order to address sub-question 2. Finally, the quality and complexity of the non-textual elements will be discussed in order to answer sub-question 3.

Having addressed all the research questions, I will then move on to discuss the practical implications of the findings in this study for teachers and textbook writers. Lastly, I will discuss the limitations of the study and potential avenues for future research.

6.1 What is the readability of mathematical texts on data handling that are used in Intermediate Phase Mathematics classrooms?

Textbooks are the carriers of written language, which needs to be understood/decoded before it can be used to make sense of the concept in the text. The textbook is a key mediating artefact and language is the tool that shapes learning and cognitive development (Vygotsky, 1989). Since learning through a second language is common in the South African context, the second language theory of Jim Cummins (1979) is considered in the discussion of the results.

Dale and Chall (1949) define readability as: The sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it. The success is the extent to which they understand it, read it at an optimal speed, and find it interesting (Dale & Chall 1949, p. 148). Factors influencing the readability of text include vocabulary complexity, sentence and text structure, text length and elaboration, unity and coherence, content familiarity, background knowledge needed, appropriateness for the audience and the quality of the writing (Graves & Graves, 2003).

The use of Cummins (1979) Threshold Hypothesis theory was important in this study, as the majority of the learners in South Africa are learning through a second language, in written form, used in textbooks. Mathematics textbooks, with their content specific academic language, are distinct from other textbooks and the inferred transfer of linguistic skills as suggested by Cummins (1979) needs monitoring and support in the form of images and other textual and non-textual cues.

According to Humphreys and Humphreys (2013) over 200 readability formulas have been developed by different scholars. In this study, the online utility at https://www.online-utility.org/english/readability_test_and_improve.jsp was used. It uses well-established readability formulas, to calculate the readability of the text, reported as an average grade level,. This free online software tool calculates readability using the following measures: Gunning Fog Index, Coleman Liau index, Flesch Kincaid Grade Level, ARI (Automated Readability Index), and SMOG. The measure of readability used here is the indication of number of years of education that a person needs to be able to understand the text easily on the first reading. In general, these tests note the incidence of polysyllabic words and long, complex sentences. Basic text statistics are also displayed, including number of characters, words, sentences, and average number of characters per word, syllables per word, and words per sentence.

The readability of each unit in each book has been presented in Chapter 4. In Table 6.1 I present the range of readability scores per book and the average for the whole book.

Table 6.1 Readability range and average per book

	Grade 4	Grade 5	Grade 6
Platinum	3.3 - 8.2 (ave 5.23)	3.5 - 8.2 (ave 5.71)	2.9 - 10.3 (ave 6.46)
Premier	3.1 - 8.9 (ave 5.21)	3.8 - 7.5 (ave 5.38)	2.8 - 11.0 (ave 4.86)
Rainbow	2.2 - 5.1 (ave 3.46)	3.5 - 6.1 (ave 3.90)	3.3 - 7.8 (ave 5.06)
Inzalo	4.1 – 7.7 (ave 6.02)	6.3 – 9.4 (ave 7.70)	5.3 – 9.1 (ave 7.71)

It can be observed that the readability can vary widely within books. Platinum Grade 6 and Premier Grade 6, for example, range from just below Grade 3 level to Grade 10 and Grade 11 levels respectively. Some books have a large number of units above grade level on terms of readability, for example, Platinum Grade 4 (6/12 units) and Inzalo Grade 5 (all units). It is also important to be reminded that the context in which these books are being used is one in which most learners are not reading at the required grade level. The grade levels reported in the readability measures are for learners who have been learning through the medium of English for that particular number of grades.

The Platinum textbooks have more units than any of the other books analysed, there was more text than in the other textbooks and workbooks. They are slightly too difficult on average for Grade 4 learners, and appropriate on average for Grade 5 and Grade 6 learners. All three books, however, have units in which the readability is well above the Grade level of the learners. There are units in all books where learners will struggle to gain access to the mathematical content presented because of a readability score that is much higher than it should be for the Grade. This would be particularly the case for learners who are struggling with their reading in English. For those units, the textbook as a key mediating artefact fails to function and the object of the activity – the learning of data handling – becomes impossible. Without careful mediation from the teacher, those units become unreachable for the learners.

The Premier textbooks have many units with the appropriate readability for the grades they have been written for. The readability scores indicate that at least 2 units per grade are well below the recommended readability level for the grade intended for and there are just 1 or 2 units per grade that are well above the recommended reading level for the grade (see Figures 4.7, 4.8 and 4.9). This caters for the readability levels of the weaker and stronger learners. On average only the Grade 4 book is above the readability level for the Grade. While the readability for the Premier books is more suited to the grade level than the Platinum books, it remains the case that there are units that will be out of the reach of the majority of the learners. This will prevent the learners from accessing the mathematical learning presented in the units. Again, for those units, the key mediating artefact in the learning of data handling fails to function and becomes an obstacle to learning. Careful mediation from the teacher would be required for learners to make sense of those units.

As shown in Table 6.1, the Rainbow workbooks have on average an appropriate readability level for the grades the books are written for. In fact, the average readability is lower than the grade the book is written for. Of the four book series analysed, the Rainbow workbooks are most appropriate in terms of readability. It is possible that an English home language learner might be able to use the books independently, but likely that English language learners may need some support from the teacher. It is important, however, to note that the Rainbow workbooks have been critiqued on the mathematical content itself. As discussed in Chapter 2, previous research done on the Rainbow workbooks by Hoadley and Gallant (2013) found that in a very small sample of tasks there were subtle misconceptions embedded in some mathematics activities in the workbooks. A NEEDU 2012 report also concluded that "workbooks do not provide nearly enough practice exercises in mathematics" (Taylor, 2013). The positive finding in terms of readability, needs to be balanced by this critique.

The Inzalo textbooks are the most difficult to read of all the books analysed. The average readability level for all three books is above the level of the grade it is written for. The grade 4 text book has a range of readability from Grade 4 to Grade 7, with an average across the whole book of Grade 6. The Grade 5 book ranges from Grade 6 to Grade 9 with no units being at an appropriate readability level for Grade 5s. The Grade 6 book ranges from Grade 5 to Grade 9, with an average of nearly Grade 8 level. These books are not at the reading level of English home language learners. This means that English language learners are particularly going to struggle to use these books independently. The teacher will need to mediate the language of these texts carefully in order for the learners to gain access to the mathematical learning in the units.

According to Zamanian and Heydari (2012) text readability, i.e. the ease with which material can be read and understood, is meant to ensure that the text reaches and affects its audience the way the author intended. The Platinum and Premier books have units which will not reach the intended learners, and the Inzalo books as a whole are not written at a readability level which will allow the intended learners access to the mathematics. It is only the Rainbow workbooks that are written at an appropriate readability level, but other studies have critiqued these workbooks for their mathematical content. As key mediating artefacts in the learning of data handling, it is important for the teacher to be aware of these weaknesses so that they can mediate the language used in order to ensure that the learners are able to access the mathematical learning.

One of the major factors in determining the readability of a text is the number of difficult words per sentence and the length of the sentences. Both these factors link to an increase in readability and linguistic complexity. Table 6.2 provides a summary of the average number of words per sentence for each book.

Table 6.2 Average number of words per sentence per book

Book	Grade 4	Grade 5	Grade 6
Platinum	8,9	10,7	11,1
Premier	7,5	7,7	7,1
Rainbow Books	5,6	6,6	8,1
Inzalo	11,3	13,9	13,4

Table 6.2 shows that the Platinum textbooks and the Inzalo textbooks had the longest sentences, on average of the books. According to Korger (1992) an average sentence for Grade 4 learners learning in English as a second language should have eight to ten words. This puts all the Grade 4 books analysed, except for the Inzalo text book, into this bracket. Sanyal (2006) found that readers find sentences of 8 words or less very easy to read, 11 words, easy, 14 words fairly easy, 17 words standard, 21 words fairly difficult, 25 words difficult and 29 words or more, very difficult. For all the books, the sentence length would be classified as fairly easy, easy or very easy.

Sentence length is not the only determining factor of sentence difficulty, however, longer sentences are generally more complex. For readability, sentence length is measured by the number of words per sentence. The longer the sentence, the higher the readability level of the text. In the case of English language learners this would mean less comprehension of the text. Complex, long sentences affect text readability more than short simpler sentences (Wray & Janan, 2013).

According to Sibanda (2013), long sentences make greater demands on the readers' memory and usually contain complex structures of coordination and subordination, which reduce readability. This is particularly evident in the Inzalo text book which has the longest sentences according to Table 6.2 and which contains many complex verbs, for example, 'would have been'. An observation by Álvarez-Cañizo, et al. (2015) was that reading long sentence makes readers forget the first part, by the time they reach the end of the sentence. Since most South African learners are confirmed to be poor readers (Spaull, 2016), the longer sentences are likely to pose reading challenges. Because of the more complex structure of longer sentences, it is also more likely that these sentences more closely resemble CALP-like language, than BICS language.

MacDonald (1990) conducted a study on textbooks as a source of difficulty when teaching geography through the medium of English in primary schools. The children in the study had poor English language proficiency and MacDonald showed that it was very difficult for them to learn this content subject through English. This failure to perform was partly due to the language used in the textbooks was not accessible to them, and the learners' grade level and texts' reading levels were mismatched. This is likely to similarly be the case for English language learners encountering the units in the books analysed here with high readability levels.

In a Botswana study done by Kasule (2011) textbook readability was measured using an online readability tool and a cloze test. This was administered by student teachers to all Grade 7 English second language speakers. The findings were that the text was too difficult and the majority of the learners read at frustration level. In a 2014 study done by Sibanda, the readability of two Science textbooks was compared. Both of these books were assessed to be above the intended readers' reading levels. In this case the difficult readability was due to vocabulary which was not explained. Technical vocabulary compromised the readability, as Science like mathematics, uses technical jargon. Sibanda (2014) highlights in this study the transitional challenges learners face in Grade 4 which further compromises their ability to comprehend texts.

According to Sibanda (2014), shorter simpler sentences, although more readable, could detract from more from the readability than longer sentences with more comprehensibility support. Graves and Graves (2003) suggest that elaborated text is easier to understand. This is important to note, as simplifying readability is not merely about shortening the sentences of the text.

In conclusion, all the textbooks analysed (Platinum, Premier and Inzalo) had units that measured too high on readability for the grades for which the books were written. Where this is the case, learners will struggle to access the mathematical content for learning if working with the text independently. This is particularly the case for learners who have not yet attained that grade level in English reading. It should be noted that this is not an argument for the reduction in the cognitive demand of the mathematics content in the text. The text can be adjusted to be more 'readable' without reducing the cognitive demand of the mathematics content in the text. The Rainbow workbooks are more appropriate in terms of the readability of the texts, but have other shortcomings as reported in the research by Hoadley and Gallant (2013) and others. Teachers need to be aware of the need to carefully mediate the language in these texts in order to make them accessible to learners.

6.2 What is the linguistic complexity, and what language features contribute to the linguistic complexity, of mathematical texts on data handling that are used in Intermediate Phase Mathematics classrooms?

Readability is just one measure of the complexity of the language used. It does not give detail about exactly what language features are contributing to the complexity. Therefore, I now turn to reporting the findings of applying the Linguistic Complexity Checklist (Appendix D) to the text in an effort to answer sub-question 2.

Data pertaining to linguistic complexity of the text in the workbooks and textbooks analysed was presented in full in Chapter 4. Shaftel et al. (2006) outline the many aspects of language that would make a text more complex at a word level, sentence level and whole paragraph level. These form the items in the Linguistic Complexity Checklist. Linguistic complexity of the texts was determined by counting the frequency of use of these language features and then by applying the formula derived by Vale (2018), based on the work of Shaftel et al. (2006), in order to calculate a Linguistic Complexity Index [LCI] for each unit of text. This LCI allows for comparison of the books.

Table 6.3 presents the range of LCIs in the units in each book, as well as the average LCI for the book as a whole.

Table 6.3 LCI range and average per book

	Grade 4	Grade 5	Grade 6
Platinum	7.1 – 16.3 (ave 10.4)	9.0 – 17.5 (ave 12.6)	7.0 – 21.2 (ave 13)
Premier	6.1 – 17.7 (ave 9.0)	6.4 – 11.1 (ave 9.4)	3.8 – 17.5 (ave 8.5)
Rainbow	4.2 – 10 (ave 6.8)	6.5 -10.3 (ave 8)	6.5 – 12.3 (ave 9.6)
Inzalo	11.4 – 16.16 (ave 13.2)	14.7 – 19.8 (ave 16.9)	11.8 – 19.9 (ave 16.0)

According to Table 6.2 and the average LCI per book, the Rainbow books are the least linguistically complex, and the Inzalo books are the most linguistically complex. This aligns with the findings for readability. The Platinum books are the second most linguistically complex, and the Premier books are third. The difference between the average LCIs for the Rainbow books and that of the Inzalo books is large. The Inzalo books are significantly more linguistically complex. While the LCI doesn't give a grade level comparison, it is clear that the language in the Inzalo books would need mediation from the teacher in order to be comprehensible to learners. It is unlikely that learners would manage to navigate these books successfully independently, particularly English language learners. There are units in the Platinum and Premier books that would similarly require careful mediation from the teacher as they reach high LCIs in their upper ranges.

The application of the Linguistic Complexity Checklist allows for an analysis of the particular features in the text that are contributing most frequently to the resulting LCI. At the word level, the feature that contributed most to the LCI for almost every unit analysed was the use of words of 7 or more letters. At the sentence level it was the use of prepositional phrases that contributed among the highest frequencies to the LCI. Also at the sentence level, the use of infinitives and complex verbs contributed to the LCI frequently, as well as the use of compound/complex sentences. At the paragraph level there were some references in the contexts to the data handling exercises to experiences that learners are possibly unfamiliar with. This contributes to difficulty in comprehending a text, but this was not a major frequent contributor to the LCI.

According to Bergqvist et al. (2012) long words are considered to be a major source of linguistic complexity. According to Muncer et al. (2014) and Schuster and Erickson (2014) word difficulty stems from the number of a word's syllables (poly-syllabic structure): the more the syllables, the more difficult a word is considered to be. According to Álvarez-Cañizo et al. (2015) children often read multiple-syllable words with less accuracy, which affects their comprehension. Dyrvold et al. (2015) noted that a source of linguistic complexity can be word length can be especially when there are so many such long words in text. The fact that all the units have words with 7 plus letters shows the potential for the words to hinder comprehension.

Different sentence types are also factors which can increase linguistic complexity and hinder comprehension. In English there are simple sentences, containing one independent clause; compound sentences, containing two or more independent clauses and complex sentences with one independent and one or more dependent clauses. According to Rubens (2001) for a text to be more readable, it is recommended that simple sentences be used. McIntyre (1996) indicated clearly that reading long complex-compound sentences might be very demanding for readers to comprehend. According to Newman (2012), the complexity of sentences places demands on the language abilities of students, which can result in failure to comprehend. Kim and Wagner (2015) also found that children's sentence comprehension is dependent on accurate decoding. Newman (2012) also observed that complex sentences take longer to read because they contain more information and because there are embedded clauses in these long sentences. An example of a particularly long sentence from the Grade 6 Inzalo book (p. 102) is: “They have to know how many oranges they can expect to harvest, how many boxes to order for packing and how much space to book on ships to transport the oranges to different countries.”

According to Berendes et al. (2018), an acceptable level of linguistic complexity is believed to be of key importance for learning materials. Reading complexity should systematically increase within the grades. Their study used automatic linguistic measures to analyse geography books, to measure readability. Ten language features, considered important for learner understanding, were used in the test. The results indicated that the texts were only partially systematically complex, there were also noticeable differences for these ten language features across grade levels, as well as a distinct difference between publishers. This is congruent with my study. There are certain language features which appear more often in some of the textbooks, there are also marked differences between the different publishers in this study.

In research done by To (2018), four English second language textbooks at four different levels were investigated, in order to conclude findings on the relationship between linguistic complexity and different textbook levels. Findings of this study indicate an increase in linguistic complexity from elementary to upper intermediate, this is also concurrent with the findings of my study: as the grades increase within the IP, the text becomes more complex. To (2018) also found that the upper intermediate, was not the peak of complexity among the levels in the study. This also concurs with my study, in that the book with the highest linguistic complexity is not a Grade 6 book and the average LCI does not systematically increase as the grades increase. This seems to indicate that linguistic complexity might not have been a focus in the preparation of these books.

As the linguistic complexity level increases, the text becomes more CALP-like because of the language features used (Cummins, 1981). Authors need to be mindful of this when designing text. English language learners should have access to text written in simple and clear language. Halliday (2012) noted that difficult texts frustrate learners, and that learners reading at frustration level recognise less than 90% of words in the passage and comprehend only 50% of the text. According to Sibanda (2020), factors making reading complexity at Grade 4 level higher, were the number of words, length of passages, sentence and word complexities and complexity of questions asked. In the same study Sibanda (2020) recommended that the authors of the DBE workbooks needed to aim to develop material that is a match between the textbooks and the reading level of their target readers, as well as ensure that there is a deliberate and systematic gradation of content, starting with the most simple and moving towards more difficult as the books progress.

In conclusion, texts varied in terms of their linguistic complexity. The Inzalo books were particularly complex, and the Rainbow workbooks were the least complex. As evident in the ranges of the LCIs, the Premier and Platinum books had units that were at a comparatively high level of linguistic complexity. The language features contributing particularly frequently included words or 7 or more letters, prepositional phrases, infinitives, complex verbs, and complex/compound sentences. Textbook writers should be aware that these features add complexity when preparing texts and should avoid multiple use of these features within sentences, where possible. Mathematics teachers should be aware that these aspects of texts add complexity and should ensure that learners understand the texts that they are reading through focusing on those words and sentence structures in particular. It is helpful to be able to identify the language features that contribute to the complexity, as this can assist a teacher in identifying what aspects of the text are making comprehension challenging to learners.

6.3 What is the comprehensibility of the non-textual elements that accompany the mathematical texts on data handling that are used in Intermediate Phase Mathematics classrooms?

Children's abilities to identify patterns and generate meaning from data is greatly influenced by the forms in which the data was presented (Leavy, 2015).Visual representations of data can pose many challenges to the learners' comprehension of a mathematical text. Visual images in these texts have specific functions (Kim, 2012), degrees of similitude to reality (Darian, 1997) and present a verbal–visual relationship that determines the thought processes that learners go through while solving data handling problems (Makina & Wessels, 2009). It is therefore important to interrogate the visual representations that support and supplement the written language in describing the data handling problem.

Cucuo and Curcio (2001) write that the importance of representation in learning and teaching mathematics is widely acknowledged, however, there has been little attention given to the interrelationship between numeracy and representation (Pugalee,1999). The ability to decode mathematical information from graphics and to encode mathematical information into graphics (Baker et al., 2001) is involved in this relationship. Learners have to be able to 'break' this code in order to access the information contained in the graphics. As MacKinlay (1999) reiterates, information graphics carry quantitative, ordinal and nominal information through perceptual elements, especially the visual-spatial ability, as graphics are presented in a visual-spatial format (Voyer et al., 1995). Learners have different spatial abilities and may be able to decode NTE's with ease, compared to the learners who may have visual-spatial problems and may find difficulty in processing graphical information (Zangemeister & Steihl, 1995). According to Mayer (1989), graphics should be captioned or supported by text and are only meaningful to the reader if they have this textual support.

The visual and symbolic nature of data handling text, facilitates the need to analyse the intent of meaning of graphs and charts. The work of MacKinlay (1986) provides a typology of graphical languages that appear in technical texts. This typology was used to classify the non-textual elements in the books analysed (see Tables 5.3, 5.4, 5.5 and 5.6). Some of these languages are linked to the curriculum outcomes for data handling. These include the tables, single position languages (e.g. some pictographs), apposed position languages (e.g. bar graphs and some pictographs) and the miscellaneous languages (e.g. pie charts). The interpretation and creation of these non-textual elements forms part of what is to be learned through engaging with these mathematical texts (DBE, 2011). They are thus complex for learners to comprehend as they are still developing the ability to interpret them.

The non-textual elements in the books were also classified according to their accuracy, connectivity, conciseness and contextuality (Kim, 2012). If the NTE is accurate, connected, concise and contextual, it is supportive of comprehension, and in contrast, if the NTE is less accurate, less connected, less concise and less contextual, it is more complex to comprehend. The general observation in this study was that the majority of the non-textual elements in the three textbook series analysed were well suited, accurate, connected, concise and contextually adequate. In contrast, the Rainbow workbooks had a higher proportion of NTE's that were categorised as a '0' for accuracy, connectivity, conciseness and contextuality. While it had the simplest text, the Rainbow workbooks had more NTE's that were distractors, than the other books.

There are few studies that have researched the combined effect of linguistic complexity and nonlinguistic forms of representation. Since mathematics is multimodal, the contextual discourse combines multiple modalities of communication (Martiniello, 2009), and learners make sense of mathematics word problems using both linguistic and non-textual elements which are in the text, according to Lemke (2003), who described the interdependence of verbal language and visual representations in making mathematical meaning in context. Other research has shown that visual representations are of particular importance for English second language learners, when making sense of maths tasks and mathematical concepts (Barton & Neville-Barton, 2003a, 2003b; Turner et al., 2006). These learners also showed greater understanding of the non-textual elements than textual modes of representation, also resorting to non-textual modes to make sense of maths questions when not understanding the text. The other significant outcome of this study is that English second language learners all had the same understanding of mathematical problems when presented with non-textual modes and they also chose to work with the non-textual modes more frequently than the text-only mode (Barton & Neville-Barton, 2003a). The study by Mariniello (2009) posited that as more symbolic/visual representation provides more 'schematic' meaning of an item, the more it will support second language learners, and reduce the effects of linguistically complex text. This is important as the choice of non-textual element could either support the text or prove to be a distractor from the mathematical concept or problem posed in the text. The presence of the non-textual elements in the books analysed could have therefore been supportive of learners' comprehension of the accompanying text, if they were accurate, concise, connected and contextual.

In conclusion, the comprehensibility of the non-textual elements was generally high for the textbooks analysed as a large proportion of their non-textual elements were accurate, concise, connected and contextual according to the framework of Kim (2012). The Rainbow workbooks had the highest proportion of the books analysed of non-textual elements that were not accurate, concise, connected or contextual due to their inclusion of many images that were unrelated to the content or context of the unit. These could be deemed to be less comprehensible as they are distractors from the content and context of the text. It is also important to note that the non-textual elements that were part of the curricular content may have been less comprehensible due to them forming the content of the mathematics to be learned. In other words, the tables, bar graphs, pictographs and pie charts may have been less comprehensible due to learners' relative unfamiliarity with them prior to engaging in the unit.

6.4 Conclusion: What are the language comprehension demands of English mathematical texts on data handling that are used in South African Intermediate Phase Mathematics classrooms?

Taking into account readability and linguistic complexity for the four book series that were analysed, there was a clear text that was written at a more reasonable level of demand – the Rainbow workbooks, and a text that was distinctly more complex – the Inzalo textbooks. It is interesting to note that both of these book series are produced either by the DBE or in partnership with the DBE. The other two textbooks were mixed in their results – some units were written with grade appropriate readability and reasonable linguistic complexity, and some were not. The non-textual elements were largely appropriately comprehensible as judged by Kim's (2012) framework. However, it should be kept in mind that some of the representations formed the curricular content of the sections analysed. Learners are in the process of learning how to interpret tables and graphs through these units and so these non-textual elements may be challenging for learners to comprehend, despite them being accurate, concise, connected and contextual according to Kim's (2012) framework. It is essential that the text and the non-textual elements work well together for comprehensibility.

From the perspective of the South African learner who is likely not reading at a grade appropriate level all of the texts analysed would provide a measure of challenge to comprehend, and this is particularly the case for the textbooks analysed. This presents the mathematics teacher with the challenge of making the texts more comprehensible. The textbook is a key mediating artefact in the mathematics classroom, but will require the mediation of the teacher in order to be comprehensible to the learners.

6.5 Limitations of this study and potential avenues for future research

A major limitation of this study is that, as a document analysis, it did not involve any observation of learners interacting with the texts analysed. I applied established measures of readability and linguistic complexity, as well as a framework for analysing non-textual elements, and from that application I have made conclusions about the comprehensibility of the text for learners, without involving learners in the study. This limits the claims that can be made from this study.

The choice of texts also limits the claims that can be made. This study used a case study methodology through choosing to focus on a defined set of texts. Claims cannot be generalised to all Intermediate Phase Mathematics texts. Similarly, this research has chosen to focus only on the data handling sections of the texts, and so claims cannot be made about the full content of the books analysed.

This research has shown that learners will be confronted by texts that are beyond their current reading comprehension level. A potential avenue for further research would be to investigate how best to mediate these texts in the classroom. This research also only examined the data handling sections of the books, it would also be beneficial to investigate the language demands of other sections of the curriculum.

6.6 Practical implications

It is important to remember that there are no perfect textbooks according to Ansary and Babaii (2002), but that textbook effectiveness improves when accompanied by teacher adaptation. Teachers need to know how to effectively mediate the use of these texts for them to be effective as a key mediating artefact in the mathematics classroom. An awareness of the factors that increase the readability level and that increase the linguistic complexity can help teachers to understand what is contributing to their learners' struggle to comprehend a text. Mathematics teachers need to be equipped with the linguistic knowledge of what contributes to the complexity of mathematical texts in order to be able to produce comprehensible texts themselves, and in order to enable them to effectively assist students in comprehending a challenging mathematical text.

Textbook publishers and authors in the South African market, should be aware of the language used in mathematics texts and the readability of these. Higher readability levels and complex language features add complexity to the text and do not support English language learners' access to the mathematical concepts by using language that is not properly understood. The non-textual elements also need to be carefully evaluated for their accuracy, connectivity, conciseness and contextuality to the text and taking learners' visual-spatial abilities into account, since we are aware that the challenges of decoding and encoding graphical information can also deny learner access to mathematical concepts.

According to Mohammadi and Abdi (2014) textbooks that look appropriate, often lack the standards of a superior book. It is important for the people choosing textbooks to scrutinize all aspects of the text and to have an assessment tool with which to compare it. Whether created or adopted, it will help the person choosing the book to focus on the important factors (Soori, et al., 2011). Teachers need an awareness of readability and linguistic complexity in mathematics texts when selecting suitable textbooks so that they can choose the book that best supports the language needs of learners. It is hoped that this textbook analysis will provide insights to teachers on how to scrutinize textbooks and how to conduct their own analysis during textbook selection.

References

Abedi, J., Courtney, M., Leon, S., Kao, J., & Azzam, T. (2006). English Language Learners and Math Achievement: A Study of Opportunity to Learn and Language Accommodation. Technical Report 702. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.

Abedi, J., & Herman, J. (2010). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record*, *112*(3), 723-746.

Abel, K. L., and Exley, B. E. (2007). Using Halliday's Functional Grammar to Examine Early Years Worded Mathematics Texts. In McNamara, Anne and Rickwood, Janet and van Haren, Rita and Vereoorn, Janette, Eds. Proceedings Critical Capital: Teaching & Learning AATE & ALEA National Conference, pages pp. 1-12, Australian National University, Canberra.

Adams, J. N., & Adams, J. N. (2003). *Bilingualism and the Latin language*. Cambridge University Press.

Alharahsheh, H. H., & Pius, A. (2020). A review of key paradigms: Positivism VS interpretivism. *Global Academic Journal of Humanities and Social Sciences*, *2*(3), 39-43. Allington, R. L. (2011). Reading intervention in the middle grades. *Voices from the Middle*, *19*(2), 10.

Álvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F. (2015). The role of reading fluency in children's text comprehension. *Frontiers in psychology*, *6*, 1810.

Ansary, H. & Babaii, E. (2002). Universal characteristics of EFL/ESL textbooks: A step towards systematic textbook evaluation. *The Internet TESL Journal, 8*(2), available on-line at:http:/iteslj.org/Ansary-textbooks/

Aoyama, K. (2007). Investigating a hierarchy of students' interpretations of graphs. *International Electronic Journal of Mathematics Education*, *2*(3), 298-318.

Apuke, O.D. (2017). Quantitative research methods: A synopsis approach. *Kuwait Chapter of Arabian Journal of Business and Management Review*, *6*(11), 40–47.

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics, 52*, 215–241.

Arizpe, E. (2001). Letting the Story Out: Visual encounters with Anthony Browne's The Tunnel. *Reading Literacy and Language, 35*(3)11.

Armstrong, D., Gosling, A., Weinman, J., & Marteau, T. (1997). The Place of Inter-Rater Reliability in Qualitative Research: An Empirical Study. *Sociology, 31*(3), 597–606.

Arteaga, P., Díaz-Levicoy, D. A., & Batanero, C. (2020). Chilean primary school children's understanding of statistical graphs. *Acta Scientiae, 22*(5), 2-24.

Bailey, A.L. (2007). *The language demands of school: Putting academic English to the test.* New Haven, CT: Yale University Press.

Baker, R. S., Corbett, A. T., Koedinger, K. R. (2001). Toward a model of learning data representations. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society,* (pp. 45-50). Mahwah, NJ: Erlbaum.

Barichievy, M. & Pieterse, K. (2019). *Shuter's Premier Mathematics Grade 5*. Cape Town: Shuter & Shooter.

Bayda, N. I., & Sutliff, G. (2020). Comparing extracted and stipulated definitions in Algebra 1 textbooks and Khan Academy. *International Electronic Journal of Mathematics Education, 15*(2), em0579. https://doi.org/10.29333/iejme/7601

Barwell, R. (2004). What is numeracy? *For the Learning of Mathematics, 24*(1), 20-22.

Barwell, R. (2018). Some thoughts on a mathematics education for environmental sustainability. In P. Ernest (Ed.), *The philosophy of mathematics education today* (pp. 145-160). Springer. https://doi.org/10. 1007/978-3-319-77760-3_9

Baxter, P. & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report, 13*(4), 544-559.

Ben-Zvi, D. & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals,definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3–25). Dordrecht, The Netherlands: Kluwer.

Bergqvist, E., Dyrvold, A. & Österholm, M. (2012). Relating vocabulary in mathematical tasks to aspects of reading and solving. In C. Bergsten (Ed.), *Proceedings of MADIF-8, The Eighth Swedish Mathematics Education Research Seminar*. Umeå, Sweden.

Berendes, K., Vajjala, S., Meurers, D., Bryant, D., Wagner, W., Chinkina, M., & Trautwein, U. (2018). Reading demands in secondary school: Does the linguistic complexity of textbooks increase with grade level and the academic orientation of the school track? *Journal of Educational Psychology, 110*(4), 518–543.

Bloom. B. (1984). *Taxonomy of educational objectives*. Boston, MA: Allyn and Bacon

Booth, W. C., Colomb, G. G. & Williams, J. M. (2003). *The craft of research*. University of Chicago press.

Botha, H., Van Putten, S., & Kundema, I. (2019). Enhancing visual literacy in the mathematics classroom: the case of Dar es Salaam. *Perspectives in Education*, *37*(2), 110-123.

Bowen, G.A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, *9*(2), 27–40.

Bowie, L., Gleeson-Baird, C., Jones, R., Morgan, H., Morrison, K. & Smallbones, M. (2020). *Platinum Mathematics Grade 4*. Cape Town: Pearson.

Bowie, L., Gleeson-Baird, C., Jones, R., Morgan, H., Morrison, K. & Smallbones, M. (2021). *Platinum Mathematics Grade 5.* Cape Town: Pearson.

Bowie, L., Gleeson-Baird, C., Jones, R., Morgan, H., Morrison, K. & Smallbones, M. (2021). *Platinum Mathematics Grade 6.* Cape Town: Pearson.

Brown, E. M. (2012). Helping mainstream teachers learn how to teach English language learners. Masters thesis. Master of Arts in English as a Second Language. Hamline University.

Bryman, A. (2008). Why do researchers integrate/combine/mesh/blend/mix/merge/fuse quantitative and qualitative research. *Advances in mixed methods research*, *21*(8), 87-100.

Callingham, R., & Watson, J. (2005). Measuring statistical literacy. *Journal of Applied Measurement, 6*(1), 1-29.

Callingham, R. (2007). *Assessing Statistical literacy: A question of interpretation*. ICOTS 7

Callingham, R. (2009). Using Rasch measurement to identify cross-cultural aspects of statistical literacy. *Changing Climates: Education for Sustainable Futures*, *1*, 1-10.

Callingham, R., & Watson, J. M. (2017). The development of statistical literacy at school. *Statistics Education Research Journal*, *16*(1), 181-201.

Cantos, P. & Almela, A. (2019). Readability indices for the assessment of textbooks: a feasibility study in the context of EFL. *Vigo International Journal of Applied Linguistics*, *16,* 31-52.

Cartwright, K. B. (2002). Cognitive development and reading: The relation of reading-specific multiple classification skill to reading comprehension in elementary school children. *Journal of Educational Psychology*, *94*(1), 56.

Carson, D., Gilmore, A., Perry, C., & Gronhaug, K. (2001). *Qualitative marketing research*. Thousand Oaks: Sage.

Castanheira, M. L., Crawford, T., Dixon, C. N., & Green, J. L. (2000). Interactional ethnography: An approach to studying the social construction of literate practices. *Linguistics and education*, *11*(4), 353-400.

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Harvard University Press.

Chick, H. L., Pfannkuch, M., & Watson, J. M. (2005). Transnumerative thinking: Finding and telling stories within data. *Curriculum matters*, *1*, 87-109.

Chick, H. L., & Pierce, R. (2012). Teaching for statistical literacy: Utilising affordances in real-world data. *International Journal of Science and Mathematics Education*, *10*(2), 339-362.

Chen, T. (2011). The ability of Taiwanese college freshman to read and interpret graphics in English. *The Asian EFL Quarterly Journal, 13*(4), 300-326.

Choy, B. H., Lee, M. Y., & Mizzi, A. (2020). Insights into the teaching of gradient from an exploratory study of mathematics textbooks from Germany, Singapore, and South Korea. *International Electronic Journal of Mathematics Education, 15*(3), em0592.

Chomsky, N. (1980). *Rules and Representation*. University Press, Columbia, New York.

Coleman, M. & Liau, T.L. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology, 60*(2), 283-284.

Connor, C. M. (2016). A lattice model of the development of reading comprehension. *Child development perspectives*, *10*(4), 269-274.

Constitution of the Republic of South Africa. (1996). Section 29(2).

Cooper, P., & McIntyre, D. (1996). *Effective teaching and learning: Teachers' and students' perspectives*. McGraw-Hill Education (UK).

Corbin, J., & Strauss, A. (2008). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory (3rd ed.). Thousand Oaks, CA: Sage. *Organizational Research Methods*, *12*(3), 614–617. https://doi.org/10.1177/1094428108324514

Creswell, J. (2014). *Research Design Qualitative, Quantitative and Mixed Methods Approaches* (4th ed.). Thousand Oaks: Sage Publishers.

Creswell, J. W., Hanson, W. E., Clark Plano, V. L., & Morales, A. (2007). Qualitative research designs: Selection and implementation. *The counseling psychologist*, *35*(2), 236-264.

Creswell, J. W., Shope, R., Plano Clark, V. L., & Green, D. O. (2006). How interpretive qualitative research extends mixed methods research. *Research in the Schools*, *13*(1), 1-11.

Cummins, J. & Swain, M. (1986). *Bilingualism in education*. London: Longman.

Cummins, J. (1988). The role and use of educational theory in formulating language policy. *TESL Canada Journal*, 11-19.

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. *Encyclopedia of language and education*, *2*(2), 71-83.

Cummins, J. (1979). Cognitive/Academic Language Proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism, 19.*

Cummins, J. (1981). *Bilingualism and minority-language children*. Ontario. OISE press.

Cummins, J. (1994). Knowledge, power, and identity in teaching English as a second language. *Educating second language children: The whole child, the whole curriculum, the whole community*, 33-58.

Cummins, J. (1999). BICS and CALP: Clarifying the distinction. *ERIC,* (1), 1–9.

Cummins, J. (2000). Bics and calp. *Encyclopedia of language teaching and learning*, 76-79.

Cummins, J., Brown, K., & Sayers, D. (2007). *Literacy, technology, and diversity: Teaching for success in changing times* (pp. 91-111). Boston, MA: Pearson.

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. *Encyclopedia of language and education*, *2*(2), 71-83.

Cucuo, A. A., & Curcio, F. (2001). *The role of representation in school mathematics: National Council of Teachers of Mathematics Yearbook*. Reston, VA: NCTM.

Dale, E., & Chall, J. S. (1949). The concept of readability. *Elementary English*, *26*(1), 19-26.

Darian, S. (1997). “The language of experiments in introductory science texts.” *Fachsprache 19*(1/2): 28-42.

Davis, J. D. (2012). An examination of reasoning and proof opportunities in three differently organized secondary mathematics textbook units. *Mathematics Education Research Journal*, *24*(4), 467-491.

de Lange, J. (2003). Mathematics for literacy. In B.L. Madison & L. Steen (Eds.), *Quantitative literacy: Why numeracy matters for schools and colleges* (pp. 75-90). New York: The College Board.

Denzin, N.K. (1978). *The Research Act: A Theoretical Introduction to Sociological Methods* (2nd ed.). New York, NY: McGraw Hill

Department of Basic Education. (2011). *Curriculum and Assessment Policy Statement: Mathematics Grades 4-6.* Pretoria: DBE.

Department of Basic Education. (2013). *Annual National Assessment: 2013 Diagnostic Report and 2014 Framework for Improvement*. Pretoria: DBE.

Department of Basic Education. (2011). *Curriculum and Assessment Policy Statement: Mathematics Grades 4-6.* Pretoria: DBE.

Department of Basic Education. (2020). *Grade 4 Mathematics in English: Book 1*. Pretoria: DBE.

Department of Basic Education. (2020). *Grade 4 Mathematics in English: Book 2*. Pretoria: DBE.

Department of Basic Education. (2020). *Grade 5 Mathematics in English: Book 1*. Pretoria: DBE.

Department of Basic Education. (2020). *Grade 5 Mathematics in English: Book 2*. Pretoria: DBE.

Department of Basic Education. (2020). *Grade 6 Mathematics in English: Book 1*. Pretoria: DBE.

Department of Basic Education. (2020). *Grade 6 Mathematics in English: Book 2*. Pretoria: DBE.

Department of Basic Education. (2021). Grades 4 to 6 catalogue. Pretoria: DBE.

Department of Education. (1997). *Language in Education Policy*. Pretoria: Department of Education.

Department of Health. (2015). *Ethics in Health Research Report 2015*. Pretoria: Department of Health.

Derewianka, B. (1998). *A grammar companion*. NSW, Newton: PETA.

Diezmann, C. M., Lowrie, T. & Kozak, N. (2007). Essential differences between high and low performers' thinking about graphically-oriented numeracy items. *Proceedings 30th Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 226-235). Sydney, Australia: MERGA.

Eldridge, S. M., Costelloe, C. E., Kahan, B. C., Lancaster, G. A., & Kerry, S. M. (2016). How big should the pilot study for my cluster randomised trial be?. *Statistical methods in medical research*, *25*(3), 1039-1056.

Engel, R.J. & Schutt, R.K. (2012). *The practice of research in social work*. Thousand Oaks: SAGE Publications.

Engel, J. (2017). *Learning statistics about society: more than acquiring some technical skills.* In Paper presented at the World Statistics Congress 61*, Marrakech, Morocco*.

English, L. D., & Watson, J. M. (2016). Development of probabilistic understanding in fourth grade. *Journal for Research in Mathematics Education*, *47*(1), 28-62.

Evans, J. (1992). Mathematics for Adults: Community Research and 'Barefoot Statisticians'. In M. Nickison, & S. Lerman (Eds.), *The social context of mathematics education: Theory and practice* (pp. 202-216). London: Southbank Press.

Evans, M. E., Watson, C., & Willows, D. M. (1987). A naturalistic inquiry into illustrations in instructional textbooks. In H. A. Houghton & D. M. Willows (Eds.), *The psychology of illustration* (Vol. 2, Instructional Issues). New York: Springer.

Everhart, V. (2010). Moving beyond social language: How to teach academic language. Slide. CASA Conference. Retrieved from http://www.nmsd.k12.nm.us/publications/documents/MovingBeyondSocialLanguageHowtoTeachAcademicLanguage.pdf [30 November 2015]

Fan, L., Zhu, Y., & Miao, Z. (2013). Textbook research in mathematics education: development status and directions. *ZDM, 45*(5), 633-646. https://doi.org/10.1007/s11858-013-0539-x

Fleisch, B., Taylor, N. Herholdt, R. & Sapire, I. (2011*)*. Evaluation of back to basics mathematics workbooks: A randomized control trial of the primary mathematics research project. *South African Journal of Education*, *31*(4), 488-504.

Flick, W. & Anderson, J.I. (1980). Rhetorical difficulty in scientific English: A study in reading comprehension. *TESOL Quarterly, 14*(3), 345-352.

Friel, S., Curcio, F., and Bright, G. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education, 32* (2), 124-158.

Frost, J. (2013). *Why statistics is important*. The World of Statistics. Retrieved from http://www.worldofstatistics.org/2013/03/04/why-statistics-is-important/

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International statistical review*, *70*(1), 1-25.

Gal, I. (2004). Statistical literacy. In *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47-78). Springer, Dordrecht.

Galant, J. (2013). Selecting and sequencing mathematics tasks: Seeking mathematical knowledge for teaching. *Perspectives in Education*, *31*(3), 34-48.

Gardner, K.M., Rush, A., Crist, M.K., Konister, R.K. & Teegarden, B. (2011). *Cognitive patterns: Problem-solving frameworks for object technology*. Zurich, Switzerland: Global Text Project.

Garfield, J., del Mas, R., & Zieffler, A. (2010). Assessing important learning outcomes in introductory tertiary statistics courses. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 75–86). Chichester, UK: John Wiley & Sons.

Gay, L., Mills, G., & Airasian, P. (2006). *Educational research: Competencies for analysis and applications.* New Jersey: Pearson Education, Inc.

Gerring, J. (2004). What is a case study and what is it good for? *The American Political Science Review, 98*(2), 341-354.

Gibbons, P. (1991). *Learning to learn in a second language*. Portsmouth, NH: Heinemann.

Graves, M.F., & Graves, B.B. (2003). *Scaffolding reading experiences: Designs for student success* (2nd Ed.). Christopher-Gordon.

Gregory, E., Williams, A., Baker, D., & Street, B. (2004). Introducing literacy to four year olds: Creating classroom cultures in three schools. *Journal of Early Childhood Literacy*, *4*(1), 85-107.

Greenfield, T. (1993). Communicating statistics. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *156*(2), 287-297.

Graven, M. H. (2014). Poverty, inequality and mathematics performance: The case of South Africa’s post-apartheid context. *ZDM*, *46*(7), 1039-1049.

Gregory, E., Williams, A., Baker, D., & Street, B. (2004). Introducing literacy to four year olds: Creating classroom cultures in three schools. *Journal of Early Childhood Literacy*, *4*(1), 85-107.

Greene, J., Kreider, H. & Mayer, E. (2004). Combining qualitative and quantitative methods in social inquiry. In B. Somekh & C. Lewin (Eds), *Research Methods in the Social Sciences* (pp.274-282). Thousand Oaks: Sage Publishers.

Hammond, J., Wickert, R., Burns, A., Joyce, H., & Miller, A. (1992). *The pedagogical relations between adult ESL and adult literacy*. Canberra: Commonwealth of Australia.

Hakuta, K. Butler, GY, & Witt, D.(2000). How long does it take English learners to attain proficiency? (University of California Linguistic Minority Research Institute, Policy Report 2000-1). Retrieved November 4, 2002.

Halliday, M. A. K. (1989). Some grammatical problems in scientific English. *Australian Review of Applied Linguistics: Genre and Systemic Functional Studies, 5*(6), 13-37.

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

Halliday, M. A. (1978). Meaning and the construction of reality in early childhood. *Modes of perceiving and processing information*, 67-96.

Halliday, M. A. (1993). Towards a language-based theory of learning. *Linguistics and education*, *5*(2), 93-116.

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

Harries, T. & Spooner, M. (2000). *Mental mathematics for the numeracy hour*. London: David Fulton

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., Chui, A., et al (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. Washington, DC: National Centre for Education Statistics, March.

Hill, S. (2009). Oral Language and Beginning Reading: Exploring Connections and Disconnections. In *Forum on Public Policy Online* (Vol. 2009, No. 2). Oxford Round Table. 406 West Florida Avenue, Urbana, IL 61801.

Hoadley, U. & Galant, J. (2016). An analysis of the Grade 3 Department of Basic Education workbooks as curriculum tools. *South African Journal of Childhood Education, 6*(1), 1-12.

Howie, S., Venter, E., & Van Staden, S. (2008). The effect of multilingual policies on performance and progression in reading literacy in South African primary schools. *Educational Research and Evaluation*, *14*(6), 551-560.

Howell, K.E. (2013). *An introduction to the philosophy of methodology*. London: SAGE Publications.

Hudson, L. A., & Ozanne, J. L. (1988). Alternative ways of seeking knowledge in consumer research. *Journal of consumer research*, *14*(4), 508-521.

Human, P., Olivier, A., Lampen, E., le Roux, A., Long, C. & Human, C. (2016). *Mathematics Grade 4 Learner Book*. Rietondale: The Ukuqonda Institute.

Human, P., Olivier, A., Lampen, E., le Roux, A., Long, C. & Human, C. (2016). *Mathematics Grade 5 Learner Book*. Rietondale: The Ukuqonda Institute.

Human, P., Olivier, A., Lampen, E., le Roux, A., Long, C. & Human, C. (2016). *Mathematics Grade 6 Learner Book*. Rietondale: The Ukuqonda Institute.

Humphreys, A. H., & Humphreys, J. T. (2013). Reading difficulty levels of selected articles in the journal of research in music education and journal of historical research in music education. *Music Education Research Journal, 6*, 15-25

Ingram, N. (2015). Students' relationships with mathematics: Affect and identity. In M. Marshman, V. Geiger, & A. Bennison (Eds.), *Mathematics education in the margins. Proceedings of the 38th annual conference of the Mathematics Education Research Group of Australasia*, (pp. 301–308). Sunshine Coast: MERGA

Igwenagu, C. (2016). *Fundamentals of research methodology and data collection*. LAP Lambert Academic Publishing.

Jackson, C. & Raubenheimer, J. (2019). *Shuter's Premier Mathematics Grade 4*. Cape Town: Shuter & Shooter.

Kaplan, J. J., Fisher, D. G., & Rogness, N. T. (2009). Lexical Ambiguity in Statistics: What do students know about the words association, average, confidence, random and spread?. *Journal of Statistics Education*, *17*(3).

Kim, R. Y. (2009c). Text ? book = textbook? Development of a conceptual framework for non-textual elements in middle school mathematics textbooks. Ph.D. dissertation, Michigan State University, East Lansing, MI.

Kim, R.Y. (2012). The quality of non-textual elements in mathematics textbooks: An exploratory comparison between South Korea and the United States. *ZDM Mathematics Education, 44*, 175-187.

Kim, J., & Pang, J. (2022). An Analysis of Sustainable Activities in Japanese, Korean, and Singaporean Elementary Mathematics Textbooks. *EURASIA Journal of Mathematics, Science and Technology Education*, *18*(2), em2080.

Kim, Y. S. G., & Wagner, R. K. (2015). Text (oral) reading fluency as a construct in reading development: An investigation of its mediating role for children from grades 1 to 4. *Scientific studies of reading*, *19*(3), 224-242.

Klesmer, H. (1994). Assessment and Teacher Perceptions of ESL Student Achievement. *English Quarterly*, *26*(3), 8-11.

Krashen, S., & Brown, C. L. (2007). What is academic language proficiency. *STETS Language & Communication Review, 6*(1), 1-5.

Krauss, S. E. (2005). Research paradigms and meaning making: A primer. *The qualitative report*, *10*(4), 758-770.

Labuschagne, A. (2003). Qualitative Research: Airy Fairy or fundamental? *The Qualitative Report,* 8(1), Article. http://www.nova.edu/sss.QR/QR8-l/labuschagne

Lapp, D., Flood, J., & Martin, D. (1999). Intermediality: How the use of multiple media enhances learning. *The Reading Teacher, 52*(7) 776-780.

Leavy, A. (2015). Looking at practice: Revealing the knowledge demands of teaching data handling in the primary classroom. *Mathematics Education Research Journal*, *27*(3), 283-309.

Lesser, L. M., & Winsor, M. S. (2009). English language learners in introductory statistics: Lessons learned from an exploratory case study of two pre-service teachers. *Statistics Education Research Journal, 8*(2), 5-32.

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communication and Technology, 30*(4), 195–232.

Levin, J. R., & Mayer, R. E. (1993). Understanding illustrations in text. In B. K. Britton, A. Woodward, & M. R. Binkley (Eds.), *Learning from textbooks: Theory and practice* (pp. 95–111). Hillsdale, NJ: Lawrence Erlbaum Associates.

Liesbeck, P. (1984). *How children learn mathematics: A guide for parents and teachers*. London: Penguin.

Lotz-Sisitka, H., Fien, J. & Ketlhoilwe, M. (2013). Traditions and new niches. In R. Stevenson, M. Brody, J. Dillon & A. Wals (Eds.), *International Handbook of Research in Environmental Education* (pp. 194-205). New York: Routledge, New York.

Lowrie, T., & Diezmann, C. (2005). Fourth Grade Students' Performance on Graphical Languages in Mathematics. In *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education (PME 29)* (pp. 265-272). University of Melbourne.

Lowrie, T. & Diezmann, C.M. (2007). Middle school students' interpreting graphical tasks: Difficulties within a graphical language. In C. S. Lim, S. Fatimah, G. Munirah, S. Hajar, M. Y. Hashimah, W. L. Gan & T. Y. Hwa (Eds.), *Proceedings of the 4th East Asia regional conference on Mathematics Education* (pp. 611-617). Penang, Malaysia: Universiti Sains Malaysia.

Lucangeli, D., Tressoldi, P.E. & Cendron, M. (1998). Cognitive and metacognitive abilities involved in the solution of mathematical word problems: Validation of a comprehensive model. *Contemporary Educational Psychology, 23*, 257-275

MacKinlay, J. (1986). Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics, 5*(2), 3-11.

Makina, A. & Wessels, D. (2009). The role Of visualisation in data handling in Grade 9 within a problem-centred context. *Pythagoras, 69*, 56-68.

Martin, M.O, Mullis, I.V.S., Foy, P and Hooper, M. (2016). *TIMSS 2015 International Results in Science*. Chestnut Hill, MA: Boston College

Martiniello, M. (2009) Linguistic Complexity, Schematic Representations, and Differential Item Functioning for English Language Learners in Math Tests. *Educational Assessment, 14*(3-4), 160-179.

Mathews, C., Mdlulu, M., & Ramsingh, V. (2014). The use of workbooks in South African grade 3 mathematics classrooms. *South African Journal of Childhood Education, 4*(1), 80-94.

McLaughlin, G.H. (1969). SMOG Grading – a new readability formula. *Journal of Reading,* 639-646.

Merriam, S.B. (1998). *Qualitative research and case study applications in education.* San Francisco: Jossey-Bass.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation.* San Francisco, CA: Jossey-Bass.

Messaris, P. (1994). Analog, not digital: The roots of visual literacy and visual intelligence. Paper presented at the International Visual Literacy Association Conference, Rochester, NY. NCTE. (2006). Standards for the English language arts. Urbana, IL: NCTE Books.

Mohammadi, M., & Abdi, H. (2014). Textbook evaluation: A case study. *Procedia-Social and Behavioral Sciences*, *98*, 1148-1155.

Morgan, C. R. (1995). *An analysis of the discourse of written reports of investigative work in GCSE mathematics*. Doctoral dissertation, Institute of Education, University of London.

Morgan, C. (2004). *Writing mathematically: The discourse of investigation.* London: Falmer Press.

Monteiro, F., Eduardo.C., Lima de Carvalho. T. (2021) "Statistics education from the perspective of statistical literacy: Reflections taken from studies with teachers," *The Mathematics Enthusiast, 18*(3), Article 10.

Moloi, M., & Strauss, J. (2005). *The SACMEQ II project in South Africa: A study of the conditions of schooling and the quality of education*. Harare: SACMEQ.

Moore, D. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association, 93*, 1253–1259.

Morrow, R. A., & Brown, D. D. (1994). *Critical theory and methodology*. Thousand Oaks: Sage.

Mouton, J. & Marais, H.C. (1996). *Basic concepts in the methodology of the social sciences*. Pretoria: HSRC.

Mullis, I.V.S., Martin, M.O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.

Mullis, I.V.S., Martin, M.O., Kennedy, A.M., Trong, K.L. & Sainsbury, M. (2009). *PIRLS 2011 Assessment Framework*. Boston: TIMMS & PIRLS International Study Centre.

Mullis, I.V.S., Martin, M.O., Foy, P., Kelly, D.L., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. Boston: TIMMS & PIRLS International Study Centre.

Muncer, S. J., Knight, D., & Adams, J. W. (2014). Bigram frequency, number of syllables and morphemes and their effects on lexical decision and word naming. *Journal of psycholinguistic research*, *43*(3), 241-254.

Martin, M. O., Mullis, I. V., Kennedy, A. M., Campbell, J., Foy, P., Gonzalez, E. J., & Schwippert, K. (2007). *PIRLS Technical Report*. Boston: TIMMS & PIRLS International Study Centre.

Martin, M.O, Mullis, I.V.S., Foy, P & Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. Chestnut Hill: Boston College.

National Assessment of Educational Progress. (2015). *The nation's report card: 2015. Mathematics and reading assessments.* Washington, DC: National Center for Education Statistics.

Newman, M.A. (1977). An analysis of sixth-grade pupils' errors on written mathematical tasks. In M.A. Clements & J. Foyster (Eds.), *Research in mathematics education in Australia* (Volume 1, pp.239-258). Melbourne: Swinburne Press.

Newman, M. E. (2012). Communities, modules and large-scale structure in networks. *Nature physics*, *8*(1), 25-31.

Naidoo J, Mkhabela N. (2017). Teaching data handling in foundation phase: Teachers' experiences. *Research in Education, 97*(1), 95-111.

Nicol, C. C., & Crespo, S. M. (2006). Learning to teach with mathematics textbooks: How preservice teachers interpret and use curriculum materials. *Educational Studies in Mathematics*, *62*(3), 331–355.

Noonan, J. (1990). Readability problems presented by mathematics text. *Early Child Development and Care, 54*(1), 57-81.

Noor, K.B.M. (2008). Case study: A strategic research methodology. *American Journal of Applied Sciences, 5*(11), 1602-1604.

North, D., Scheiber, J., & Ottaviani, M. G. (2010). Training teachers to teach statistics in South Africa: Realities and attitudes. In *Proceedings of 8th International Conference on Teaching Statistics*.

North, D., & Zewotir, T. (2006). Introducing statistics at school level in South Africa. In *Proceedings of the 7th International Conference on Teaching Statistics*.

Nunes, T., Carraher, T. N., Schliemann, A. D., & Carraher, D. W. (1993). *Street mathematics and school mathematics*. Cambridge University Press.

NEEDU. (2013). *National Report 2012: The State of Literacy Teaching and Learning in the Foundation Phase*. Pretoria: NEEDU.

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

Oller, J. (1979). *Language tests at school: A pragmatic approach*. London: Longman.

Olivier, C. E. (2002). The oral proficiency of ESL teacher trainees in different discourse domains. (Doctoral dissertation). SERP Institute. (2012). Catalysing comprehension through discussion and debate (CCDD). Retrieved from http://ccdd.serpmedia.org/research-academic-language.php

Österholm, M. (2006). Metacognition and reading-criteria for comprehension of mathematics texts. In *Conference of the International Group for the Psychology of Mathematics Education, 16–21 July, Prague, Czech Republic* (Vol. 4, pp. 289-296). The International Group for the Psychology of Mathematics Education.

Patkin, D. (2011). The interplay of language and mathematics. *Pythagoras, 32*(2), 1-7.

Pettersson, R. (1990). Teachers, students and visuals. *Journal of Visual Literacy, 10*(1), 45–62.

Planas, N., & Setati-Phakeng, M. (2014). On the process of gaining language as a resource in mathematics education. *ZDM*, *46*(6), 883-893.

Polit, D. F.; Beck, C. T.; Hungler, B. P. 2001. *Essentials of nursing research: methods, appraisal, and utilization* (5th ed). Philadelphia: Lippincott Williams and Wilkins.

Pratama, G.S., & Retnawati, H. (2018). Urgency of higher order thinking skills content analysis in mathematics textbooks. *Journal of Physics: Conference Series, 1097,* 1-8

Pretorius, E. J., & Spaull, N. (2016). Exploring relationships between oral reading fluency and reading comprehension amongst English second language readers in South Africa. *Reading and Writing*, *29*(7), 1449-1471.

Prediger, S., Erath, K., & Opitz, E. M. (2019). The language dimension of mathematical difficulties. In *International handbook of mathematical learning difficulties* (pp. 437-455). Springer, Cham.

Pugalee; D. K. (1999). Constructing a model of mathematical literacy. *The Clearing House, 73*(1), 19-22.

PISA Governing Board. (2010). *Draft PISA Mathematics Framework*. Copenhagen: OECD.

Rangecroft, M. (2002). The language of statistics. *Teaching Statistics, 24*(2), 34–37.

Rezat, S., Fan, L., & Pepin, B. (2021). Mathematics textbooks and curriculum resources as instruments for change. *ZDM, 53*(6), 1189-1206.

Reddy, V., Zuze, T., Visser, M., Winnaar, L., Juan, A., Prinsloo, C., & Rogers, S. (2015). *Beyond benchmarks. What 20 years of TIMSS data tell us about South African Education*.

Reddy, V., Visser, M., Winnaar, L., Arends, F., Juan, A., Prinsloo, C.H., & Isdale, K. (2016). *TIMSS 2015: Highlights of Mathematics and Science Achievement of Grade 9 South African Learners.* Pretoria: Human Sciences Research Council.

Rezaee, A.A. & Norouzi, M.H. (2011). Readability formulas and cohesive markers in reading comprehension. *Theory and Practice in Language Studies, 1*(8), 1005-1010.

Ribeiro, A., Healy, L., Borba, R. & Fernandes, S. (2018). Mathematics education in Brazil: Panorama of current research. Dordrecht: Springer.

Ritchie, J., Lewis, J. (2003). *Qualitative Research Practice*. Sage Publications, London.

Ridgway, J., Nicholson, J., Sutherland, S. and Hedger, S. (2019). Critical statistical literacy and interactive data visualisations. In J. Evans, S. Ruane & H. Southall (Eds), *Data in Society* (pp. 349-358). Chicago: Policy Press.

Robertson, S.A & Graven, M. (2015) Exploring South African mathematics teachers' experiences of learner migration. *Intercultural Education, 26*(4), 278-295.

Robertson, S.A. (2017). The place of language in supporting children's mathematical development: Two Grade 4 teachers' use of classroom talk. Unpublished doctoral thesis, Rhodes University, Grahamstown.

Robertson, S. A., & Graven, M. (2019). Exploratory mathematics talk in a second language: A sociolinguistic perspective. *Educational Studies in Mathematics*, *101*(2), 215-232.

Robertson, S.-A. & Graven, M. (2020). A mathematics teacher's response to a dilemma: 'I'm supposed to teach them in English but they don't understand'. *South African Journal of Childhood Education*, *10*(1), 1-11.

Roe, J., deForest, R., & Jamshidi, S. (2018). *Mathematics for sustainability*. Springer. https://doi.org/10.1007/ 978-3-319-76660-7

Ross, K., Saito, M., Dolata, S., Ikeda, M., Zuze, L., Murimba, S., et al. (2005). The Conduct of the SACMEQ II Project. In E. Onsomu, J. Nzomo, & C. Obiero, *The SACMEQ II Project in Kenya: A Study of the Conditions of Schooling and the Quality of Education*. Harare: SACMEQ.

Rubens, P. (2001). Science and technical writing, a manual of style. London: Routledge.

Ryan, G. (2018). Introduction to positivism, interpretivism and critical theory. *Nurse Researcher, 25*(4) pp. 41–49.

Sarem, S. N., Hamidi, H., & Mahmoudie, R. (2013). A critical look at textbook evaluation: A case study of evaluating an ESP course-book: English for international tourism. *International Research Journal of Applied and Basic Sciences*, *4*(2), 372-380.

Saunders, M., Lewis, P. & Thornhill, A. (2012). *Research Methods for Business Students.* Harlow: Pearson Education Ltd.

Schuster, J., & Erickson, K. (2014). Text complexity in the Dynamic Learning MapsTM Alternate Assessment System (White Paper No. 14-01). Lawrence, KS: University of Kansas Center for Educational Testing and Evaluation.

Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive and critical research paradigms. *English Language Teaching, 5*(9), 9-16.

Semali, L. M. (2003). Understanding digitization and the visual experience in the age of the internet: Principles, practices and challenges. *Journal of Visual Literacy, 23*, 85-102.

Setati, M. (2002). Researching mathematics education and language in multilingual South Africa. *The Mathematics Educator, 12*(2),6-20.

Setati, M., & Adler, J. (2000). Between languages and discourses: Language practices in primary multilingual mathematics classrooms in South Africa. *Educational studies in mathematics*, *43*(3), 243-269.

Phakeng, M. S. (2016). Mathematics education and language diversity: Mathematics education and language diversity. In *Teaching and learning mathematics in multilingual classrooms* (pp. 11-23). Brill Sense.

Shaftel, J., Belton-Kocher, E., Glasnapp, D. & Poggio, J. (2006). The impact of language characteristics in mathematical test items on the performance of English language learners and students with disabilities. *Educational Assessment, 11*(2), 105-126

Sharma, S. (2017). Definitions and models of statistical literacy: a literature review, Open *Review of Educational Research, 4*(1), 118-133.

Sibanda, L. (2016). Investigating the nature of the linguistic challenges of the Department of Basic Education 2013 Grade 4 Mathematics ANAs and the learners' and teachers' experience of them. Unpublished PhD Thesis, Rhodes University, Grahamstown, South Africa.

Sibanda, J. (2017). Language at the Grade Three and Four interface: The theory-policy-practice nexus. *South African Journal of Education*, *37*(2).

Sibanda, L. & Graven, M. (2018). Can mathematics assessments be considered valid if learners fail to access what is asked of them? *South African Journal of Childhood Education, 8*(1), 77-89.

Sibanda, L. (2017). The linguistic component in Mathematics assessment and attendant tensions and dilemmas the teachers' experience. SAARMSTE (2018)

Sibanda, L. (2017). Do the Annual National Assessments in mathematics unfairly assess English language competence at the expense of mathematical competence? In M. Graven & H. Venkat (Eds.), Improving primary mathematics education, teaching and learning: Research for development in resource-constrained contexts (pp. 147-159). Basingstoke, UK: Palgrave Macmillan.

Sibanda, L. (2020). Impact of linguistic complexity in English language texts on South Africa's primary school Grade 3 to 4 transition. *Issues in Educational Research*, *30*(2), 673-690.

Shohamy, E., Levine, T., Spolsky, B., Kere-Levy, M., Inbar, O., & Shemesh, M. (2002). *The academic achievements of immigrant children from the former USSR and Ethiopia.* Israel: Ministry of Education.

Spaull, N. (2016). Disentangling the language effect in South African schools: Measuring the impact of 'language of assessment' in Grade 3 literacy and numeracy. *South African Journal of Childhood Education, 6*(1), 13-33.

Spaull, N., & Hoadley, U. (2017). Getting reading right: Building firm foundations. In L. Jamieson, L. Berry, L. Lake (Eds.), *South African Child Gauge* (pp. 77-83). Cape Town: Children's Institute, University of Cape Town.

Skinner, B. F. (1957). *Verbal behaviour.* New York: Appleton-Century-Crofts.

Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child development*, 1114-1128.

Soori, A., Kafipour, R., & Soury, M. (2011). ELT textbook evaluation and graphic representation. *European Journal of Social Sciences, 26*(3), 481-493

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks: Sage.

Stake, R.E. (1998). Case studies. In E.K. Denzin & Y.S. Lincoln (Eds.), *Strategies of qualitative inquiry* (pp. 86-109). Thousand Oaks: SAGE Publications.

Taylor, N. (2016). Thinking, language and learning in initial teacher education. *Perspectives in Education, 34*(1), 10-26.

Taylor, M., & Von Fintel, M. (2016). Estimating the impact language of instruction in South African primary schools: A 'fixed effects approach'. *Economics of Education Review, 50* (2016), pp. 75-89

Thomas, W. & Collier, V. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Santa Cruz: Center for Research on Education, Diversity and Excellence.

Tiaden, L., Farrell, N. & John, P. (2020). *Shuter's Premier Mathematics Grade 6*. Cape Town: Shuter & Shooter.

Tishkovskaya, S., & Lancaster, G. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, *20*(2), 1–24.

To, V. (2018). *Linguistic Complexity Analysis: A Case Study of Commonly-Used Textbooks in Vietnam*. Thousand Oaks: SAGE Open.

Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of applied developmental psychology*, *18*(1), 37-54.

UNESCO. (2016). *Global Education Monitoring Report*. Paris: UNESCO.

Unsworth, L. (1997). Scaffolding reading of science explanations: Accessing the grammatical and visual forms of specialized knowledge. *Reading*, *31*(3), 30-42.

Vale, P. & Graven, M. (under review). Strategies implemented by South African teachers to ensure continuing mathematics education during COVID-19.

Vale, P., Murray, S., & Brown, B. (2012). Mathematical literacy examination items and student errors: An analysis of English second language students' responses. *Per Linguam, 28*(2), 65-83.

Vale, P. (2018). Linguistic complexity and related student errors in Mathematical Literacy examinations: Improving inclusive assessment practices for English language learners. In L. Webb (Ed.), *Proceedings of the 26th Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education* (pp. 86 – 100). Gaborone: SAARMSTE.

Valverde, G., Bianchi, L., Wolfe, R., Schmidt, W. & Houang, R. (2002). *According to the Book: Using TIMMS to Investigate the Translation of Policy into Practice through the World of Textbooks.* London: Kluwer Academic Publishers.

Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J. E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology*, *28*(4), 409-426.

Van den Broek, P., & Kendeou, P. (2017). Development of reading comprehension. *Theories of reading development*, 283-306.

Van der Berg, S., Spaull, N., Wills, G., Gustafsson, M.A. & Kotzé, J. (2016) *Identifying Binding Constraints in Education*. Stellenbosch: University of Stellenbosch.

Verhoeven, L., & Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *22*(3), 407-423.

Vitta, J. P. (2021). The functions and features of ELT textbooks and textbook analysis: A concise review. *RELC Journal*, 00336882211035826.

Voyer D, Voyer S, & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*(2), 250-270.

Vygotsky, L.S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and Language*. Cambridge, MA: MIT Press.

Vygotsky, L. S. (1987). *The collected works of L. S. Vygotsky: Problems of general psychology.* (R. Rieber & A. Carton, Eds.). New York: Plenum Press.

Vygotsky, L. S. (1997). *The collected works of L.S. Vygotsky: volume 3: Problems of the theory and history of psychology*. (R. W. Rieber & J. Wollock, Eds.) (1st ed.). New York: Plenum Press

Vygotsky, L. S. (1998). Development of thinking and formation of concepts in the adolescent. In R.W. Rieber (Ed.), *The collected works of L.S. Vygotsky: Volume 5, Child Psychology* (pp29-82). New York: Plenum Press.

Wainer, H. (1980). A test of graphicacy in children. *Applied Psychological Measurement, 4*, 331-340.

Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, *88*(421), 1-8.

Watkins, J. K., Miller, E., & Brubaker, D. (2004). The role of the visual image: What are students really learning from pictorial representations? *Journal of Visual Literacy, 24*(1), 23–40.

Watson, J. M. (2006). Issues for statistical literacy in the middle school. In *ICOTS-7 Conference Proceedings. IASE, Salvador (CD-Rom)* (pp. 1-6).

Watson, J. M. (2011). Foundations for improving statistical literacy. *Statistical Journal of the IAOS*, *27*(3, 4), 197-204.

White, L. (2011). Second language acquisition at the interfaces. *Lingua, 121* (2011), 577-590.

Wyatt-Smith, C. & Cummings, J. (2003). Curriculum Literacies: Expanding Domains of assessment. Assessment in Education. 10(1), 47-49.

Yin, Robert K. (2014). *Case study research: Design and methods.* Los Angeles, CA: Sage

Zamanian, M., & Heydari, P. (2012). Readability of Texts: State of the Art. *Theory & Practice in Language Studies*, *2*(1).

Zangemeister, W. H., & Steihl, H. S. (1995). *Visual attention and cognition*. Amsterdam: Elsevier.

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

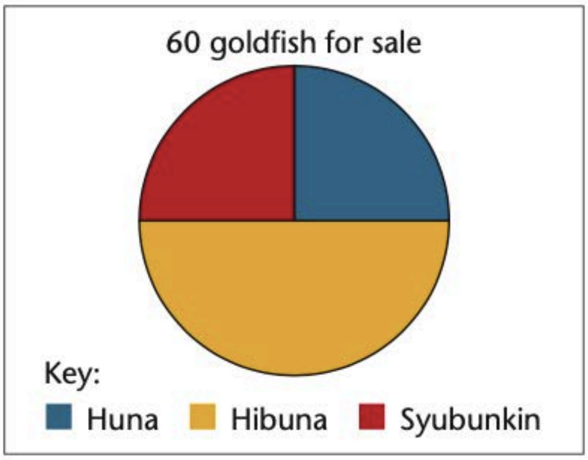
Zevenbergen, R. (2002). The literacy demands of numeracy. *Professional Magazine, 19*, 22-25.

Zewotir, T., & North, D. (2011). Opportunities and challenges for statistics education in South Africa. *Pythagoras, 32*(2), 1-5.

APPENDIX A: Example of a unit – Inzalo Grade 4 Book (pp. 92-94)

7.4 Information in pie charts

Jabu sees an advertisement on the internet of a pet shop owner who wants to sell her tank of fish. She wants R480 for 60 goldfish. The advertisement has only this pie chart.



A pie chart has a **heading** and a **key** to tell the meaning of the pie chart.

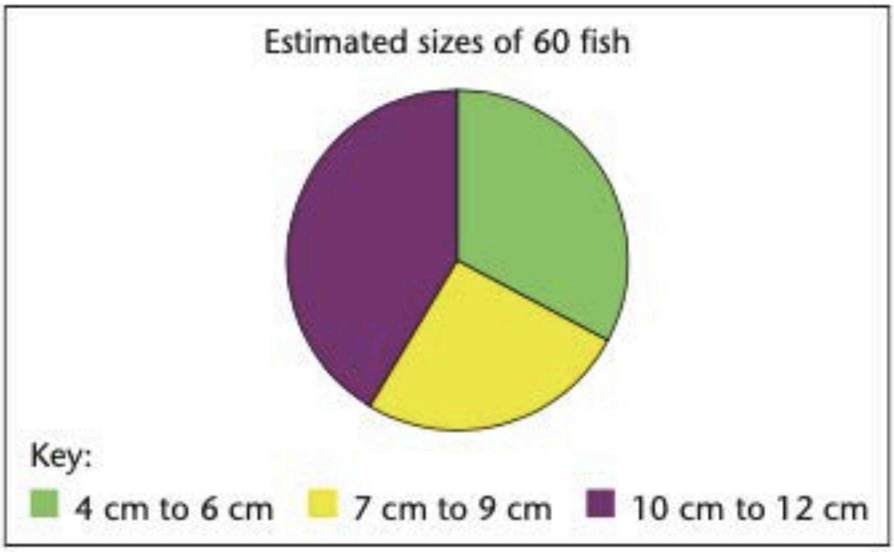
In a pie chart the parts are **fractions** of the pie (circle). For example, we work out *what fraction* of the fish are Huna and then we colour *the same fraction* of the pie.

1. (a) What is the heading of the chart?
   (b) What does the key tell you?
   (c) What does the graph tell you about the fish in this tank?

Jabu says: "A fish tank costs about R240. If the fish tank is part of the sale, then the pet shop owner asks about R4 per fish. That's a bargain!"

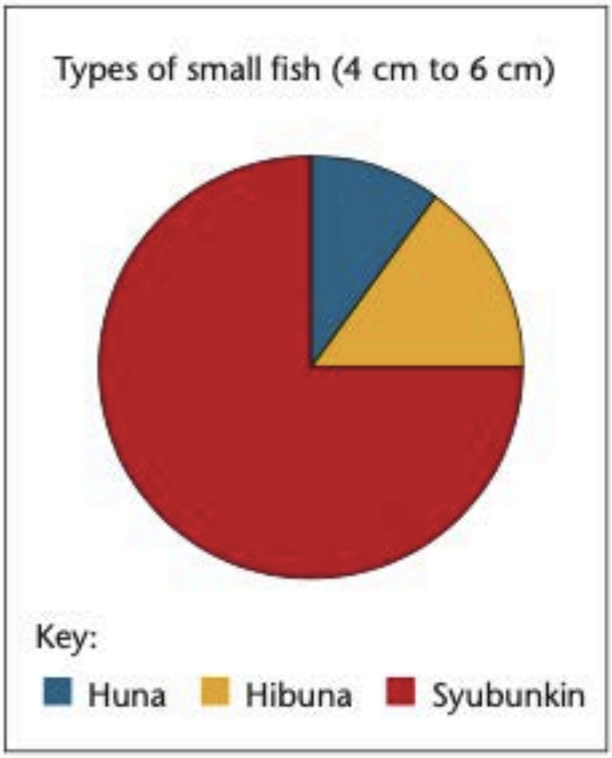
2. Check if Jabu is correct about the price.

Jabu asked the pet shop owner to let him know what size the fish are. She sent this pie chart:



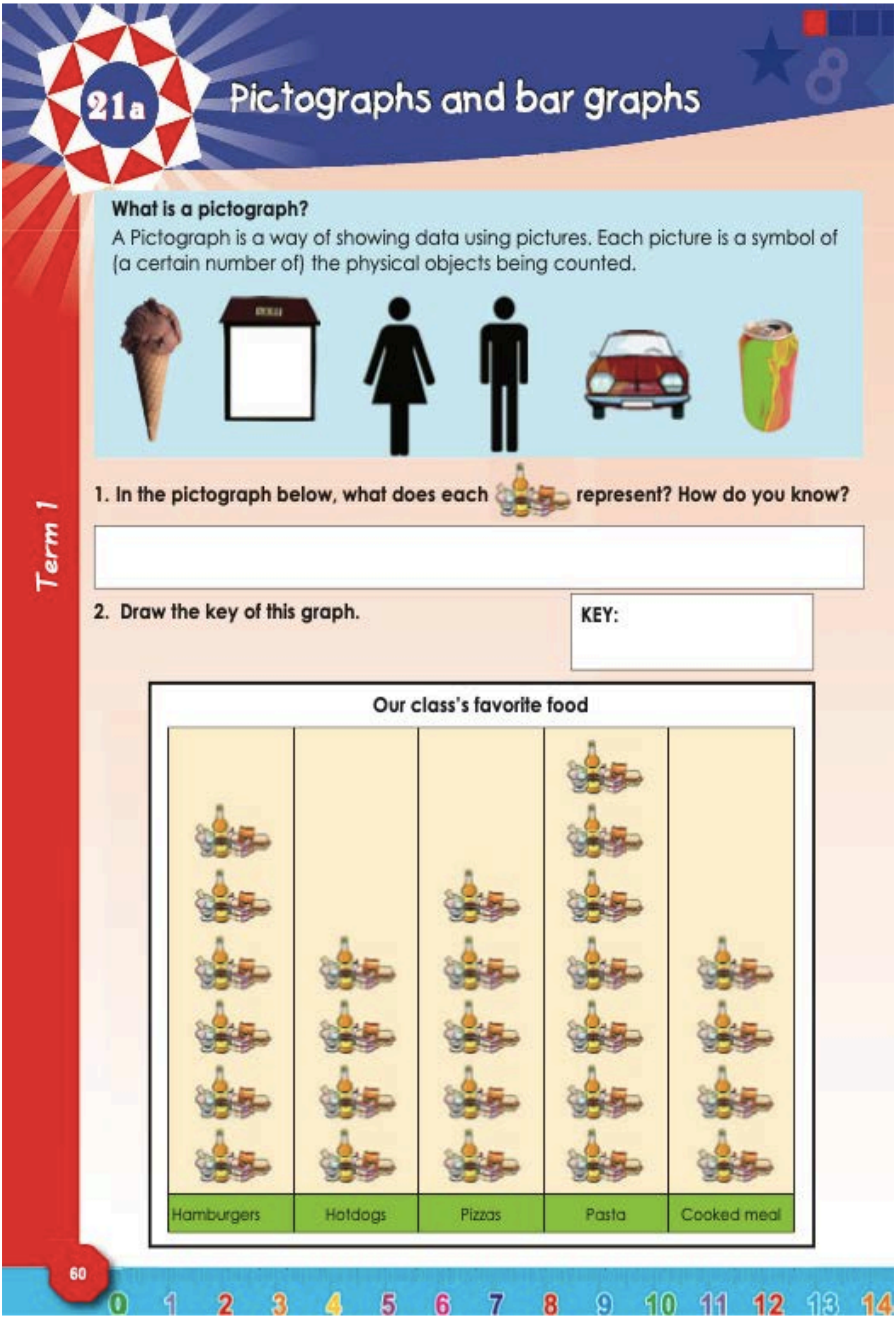
3. What story does the pie chart tell about the sizes of the 60 goldfish? Write the information in your book.
4. Read each statement and say whether you agree or disagree. Explain why.
   (a) More than half of the fish are between 7 cm and 9 cm long.
   (b) About a third of the fish are between 4 cm and 6 cm long.
   (c) About a third of the fish are between 10 cm and 12 cm long.
   (d) The big fish (10 cm to 12 cm) are almost twice as many as the medium fish (7 cm to 9 cm).

Jabu does not want to buy big fish. Big fish need more food and that is expensive. He asked the pet shop owner which fish are between 4 cm and 6 cm long. She sent this pie chart.



5. What story does the pie chart tell about the kinds of small fish? Write the information in your book.
6. Read each statement and say whether you agree or disagree. Explain why.
   (a) About three quarters of the small fish are Syubunkins.
   (b) About one third of the small fish are not Syubunkins.
   (c) About one eighth of the small fish are Hunas.
   (d) The number of small Hunas is the same as the number of small Hibunas.
7. Use the information in all the previous pie charts to work out how many of the fish are small Syubunkins.

APPENDIX B: Example of a unit – Rainbow Grade 4 Book 1 (pp. 60-61)



21a

Pictographs and bar graphs

What is a pictograph?

A Pictograph is a way of showing data using pictures. Each picture is a symbol of (a certain number of) the physical objects being counted.

1. In the pictograph below, what does each [symbol] represent? How do you know?

2. Draw the key of this graph.

KEY:

3. **Draw a pictograph to represent the following information.**
In our science class, our task was to go and search for insects in our gardens in order to see what insects there are at this time of year. I found the following in a section of 2 square metres in my garden: 10 rose beetles, one ladybird, three bees, two flies, nine ants and six caterpillars.

4. **Based on the above graph:**
   a. What time of year do you think it might be? (During which season(s) might certain insects be found generally?)

   b. If I looked in a section of 4 square metres, more or less how many of each kind of insect could I expect to find?

   c. Do you think I was looking at a patch of lawn or a flower bed? Why?

5. **Suggest some data that would be easy and interesting to see/read in a pictograph (rather than in a bar graph).**

6. **Who might be interested in the graph you've suggested above, and why?**

Sign

Date

15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

APPENDIX C: Units for analysis

Platinum Series Grade 4

Unit	Pages
1	34-35
2	36
3	37
4	38
5	39
6	40
7	41
8	128
9	129-130
10	131
11	132-133
12	134-135

Platinum Series Grade 5

Unit	Pages
1	34
2	35
3	36
4	36
5	37
6	38
7	39
8	40

9	41
10	134
11	135
12	136
13	137
14	138
15	139
16	140-141
17	141-142
18	142-143
19	144-145

Platinum Series Grade 6

Unit	Pages
1	44
2	45
3	46
4	46
5	47
6	48
7	49
8	50
9	51
10	52
11	53
12	53
13	148

14	148
15	149
16	150
17	151
18	153
19	154
20	156-157

Premier Series Grade 4

Unit	Pages
1	35
2	35
3	36
4	37
5	38
6	39
7	40
8	40
9	146
10	147-148
11	148
12	148-149

Premier Series Grade 5

Unit	Pages
1	46-47
2	48-50

3	50-53
4	167-168
5	168-169
6	170-171
7	172-173
8	174
9	174-175

Premier Series Grade 6

Unit	Pages
1	37
2	38
3	39
4	41
5	42
6	42-43
7	44-45
8	137
9	138
10	139
11	140
12	141
13	142
14	143

Rainbow Books Grade 4

Unit	Pages
1	Book 1, 34-35
2	58
3	60-63
4	Book 2, 52-53
5	54-55
6	56-57
7	58-59
8	60-61
9	62-63

Rainbow Books Grade 5

Unit	Pages
1	Book 1, xliv
2	Book 2, 66-69
3	72-73
4	74-79
5	80-81
6	82-83
7	84-85

Rainbow Books Grade 6

Unit	Pages
1	64-65
2	66-67
3	68-69

4	70-71
5	76-77
6	78-82
7	82-83
8	84-85
9	86-87
10	88-89

Inzalo Grade 4

Unit	Pages
1	86
2	90
3	91
4	92
5	95
6	258
7	259
8	260

Inzalo Grade 5

Unit	Pages
1	81
2	83
3	86
4	90
5	257

6	260

Inzalo Grade 6

Unit	Pages
1	97
2	99
3	102
4	106
5	268
6	270
7	272
8	277

APPENDIX D: Linguistic Complexity Checklist

Linguistic Complexity Checklist
Adapted from Shaftel et al. (2006, p. 126)

NUMBER OF SENTENCES: _________

A: BASIC

1. ________ Number of words in item

B: WORD LEVEL CHARACTERISTICS

1. ________ Number of different words with 7 letters or more
2. ________ Number of relative pronouns
3. ________ Number of examples of slang, homophones and homonyms
4. ________ Number of abbreviations

C: SENTENCE LEVEL CHARACTERISTICS

1. ________ Number of prepositional phrases
2. ________ Number of infinitives
3. ________ Number of complex verbs
4. ________ Number of complex / compound sentences
5. ________ Number of conditional constructions
6. ________ Number of comparative constructions

D: PARAGRAPH LEVEL CHARACTERISTICS

1. ________ Number of cultural- and/or experience-specific references
2. ________ Number of grammatical and/or spelling errors

Linguistic Complexity Index:

LCI = (Sum A + Sum B + Sum C + Sum D) ÷ Number of sentences

= __________

APPENDIX E: Kim's (2012, p. 180) framework for analysis of non-textual elements (p. 180)

Aspect of non-textual elements	Score	Description
Accuracy	2	An NTE correctly shows the definition of a concept, or it is accurate to show a concept based on its definition
	1	An NTE makes sense in terms of the definition or meaning of a concept. But it does not show every required mathematical condition (e.g. some requisite notations are missing or misleading) or some attributes are not appropriate to explain a concept
	0	An NTE is inaccurate in terms of the definition of a concept (e.g. it has an obvious error). Or an inappropriate realistic object is used to present a concept. There is a major error or concern to use the realistic object or context for a concept. Or there is no mathematical concept in the NTE
Connectivity	2	An NTE is explicitly and fully associated with the mathematical content in the text. It directly shows a concept or problem
	1	An NTE is partly related to the mathematical content in the text. There is some missing or irrelevant information in an NTE. It shows the content but it does not explicitly show how it is connected with the content
	0	An NTE has nothing to do with the content mathematically. It can give some clue about contexts in texts (e.g. river when the problem is about length of river)
Conciseness	2	An NTE is straightforward to show a concept or problem without any distracting or other factor
	1	An NTE is straightforward to show a concept or problem with some other factors that might be helpful for the concept
	0	An NTE has distracting or other factors that are useless in addition to factors needed to show a concept or problem
Contextuality	2	A realistic object or context is used in an NTE with mathematical connection
	1	No mathematical ideas or concepts exist, but there is some realistic contextual information (that is used to provide contexts or objects in the problem or to facilitate related activities)
	0	Neither realistic object nor realistic context is included in an NTE

APPENDIX F: Readability scores for the Platinum books

GRADE 4

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated Readability Index	SMOG	Flesch Reading Ease
1.	5,80	4,41	4,60	1,71	7,52	76,02
2.	6,87	6,34	4,87	3,34	8,18	75,28
3.	5,61	5,89	4,18	3,01	7,24	80,39
4.	7,54	7,60	5,74	4,53	8,32	70,66
5.	3,60	3,88	2,08	1,24	5,78	91,83
6.	4,83	2,85	2,16	0,44	7,05	90,73
7.	6,92	5,59	5,56	3,04	8,10	72,78
8.	5,62	4,16	4,14	1,54	7,19	79,81
9.	5,62	3,96	4,04	1,59	6,92	82,55
10.	5,36	4,91	3,38	2,17	7,30	85,60
11.	7,10	5,44	4,79	3,34	8,00	80,84
12.	9,13	7,80	8,27	5,97	9,67	59,77

GRADE 5

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated Readability Index	SMOG	Flesch Reading Ease
1.	9,43	6,73	7,52	5,36	9,58	66,39
2.	5,50	2,88	3,24	0,73	6,79	88,32

3.	6,46	4,40	5,88	3,43	6,46	77,81
4.	4,17	3,58	3,24	1,07	5,63	86,15
5.	8,05	6,99	6,11	3,95	8,86	67,15
6.	7,86	5,43	4,85	3,28	8,48	80,18
7.	6,61	4,58	4,13	2,01	8,07	81,27
8.	6,86	6,46	5,39	3,48	8,48	71,96
9.	6,51	5,27	4,82	3,55	7,55	82,67
10.	5,96	5,49	5,09	2,84	7,35	75,27
11.	5,90	4,52	3,97	2,17	7,05	84,07
12.	5,41	3,21	3,99	1,15	6,27	84,16
13.	5,06	3,62	3,32	1,08	7,07	85,13
14.	7,64	6,37	6,00	4,20	8,35	72,95
15.	6,83	4,35	4,84	1,70	8,05	74,85
16.	9,30	7,78	7,42	6,67	10,00	69,26
17.	7,16	5,87	4,95	3,13	8,63	76,11
18.	5,75	6,00	4,18	3,26	7,24	81,87
19.	9,44	7,20	8,12	5,71	9,71	61,93

GRADE 6

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated Readability Index	SMOG	Flesch Reading Ease
1.	5,07	6,16	4,29	3,08	7,18	77,65
2.	6,48	5,56	4,93	2,58	7,65	71,82

3.	9,56	7,51	7,19	5,25	10,17	65,12
4.	6,38	3,44	3,89	0,89	7,35	80,31
5.	7,83	5,30	5,71	4,29	8,05	79,67
6.	4,20	2,27	2,01	0,02	6,06	91,75
7.	8,35	3,35	6,94	3,87	8,22	7,92
8.	5,66	0,68	3,20	2,31	7,21	86,79
9.	8,92	9,44	8,14	7,36	9,77	61,12
10.	7,04	6,15	4,73	3,06	7,83	74,03
11.	6,32	3,13	4,52	0,96	7,05	71,75
12.	7,01	5,38	5,44	2,50	8,21	70,30
13.	10,91	9,16	8,65	7,25	11,02	57,96
14.	8,14	6,38	5,50	3,72	8,95	73,51
15.	6,86	0,12	5,27	0,83	5,12	86,78
16.	11,56	9,98	9,39	8,63	12,06	56,01
17.	7,93	7,33	5,43	4,31	8,82	72,81
18.	11,43	9,56	8,72	7,58	11,51	57,52
19.	8,21	5,95	5,66	3,63	9,09	74,02
20.	10,81	8,85	9,90	8,56	10,57	55,88

APPENDIX G: Readability scores for the Premier books

GRADE 4

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated Readability Index	SMOG	Flesch Reading Ease
1.	5,60	3,77	3,70	1,21	7,35	79,26
2.	6,80	4,49	3,86	2,03	7,56	76,60
3.	5,21	4,84	4,54	2,02	7,18	75,73
4.	4,96	2,52	2,77	0,43	6,54	84,46
5.	6,78	6,16	5,98	3,73	8,48	71,32
6.	6,04	3,74	3,73	1,20	7,05	78,97
7.	5,17	2,33	1,67	0,04	6,46	96,96
8.	9,81	9,47	7,82	6,69	10,75	59,91
9.	5,56	3,99	4,46	1,35	6,95	74,50
10.	7,73	4,65	4,66	1,87	8,48	74,98
11.	7,63	4,48	4,78	1,87	8,48	76,12
12.	7,54	8,35	6,25	4,84	8,67	63,88

GRADE 5

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated Readability Index	SMOG	Flesch Reading Ease
1	7,77	7,54	7,33	5,76	9,24	66,55
2	4,84	3,49	3,24	0,99	6,32	82,61

3	8,16	6,87	5,33	3,72	8,80	71,44
4	6,57	3,89	4,13	1,28	7,77	76,66
5	6,30	6,06	4,81	3,08	8,07	75,24
6	6,95	4,80	3,70	1,97	7,91	81,29
7	7,07	5,10	5,12	2,26	8,58	69,31
8	7,02	4,78	4,48	2,16	8,26	78,68
9	7,34	5,10	4,68	2,67	7,85	70,08
10	5,18	3,15	3,58	0,70	7,47	83,42

GRADE 6

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated Readability Index	SMOG	Flesch Reading Ease
1	8,97	8,13	7,79	5,21	10,25	57,68
2	3,79	3,40	2,97	0,85	6,23	86,23
3	5,18	6,55	5,69	3,38	8,94	67,40
4	2,27	5,72	0,09	5,14	4,46	102,43
5	4,60	4,90	3,96	2,06	7,71	79,74
6	4,26	0,59	1,51	1,39	6,08	96,23
7	4,74	2,61	2,70	1,22	6,87	88,62
8	5,54	2,17	3,48	0,11	7,47	83,54
9	19,87	11,90	18,49	18,43	17,49	23,03
10	6,88	5,40	4,24	2,45	8,07	77,15
11	4,41	2,72	3,42	0,73	6,46	79,28

12	2,97	3,89	0,14	4,97	4,33	105,70
13	7,53	7,96	5,76	4,78	8,64	70,08
14	5,65	0,09	3,79	0,53	6,81	74,72
15	4,63	0,78	3,94	0,77	6,21	72,98
16	5,05	2,57	2,72	0,19	6,65	88,60
17	5,97	5,81	5,10	2,97	7,07	68,25
18	8,03	6,31	6,01	3,45	9,02	68,32

APPENDIX H: Readability scores for the Rainbow books

GRADE 4

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated readability index	SMOG	Flesch reading ease
1	4,11	0,47	1,94	0,05	6,10	87,56
2	2,77	0,27	1,90		1,27	90,27
3	6,27	4,79	4,62	2,07	7,82	76,59
4	3,77	3,55	3,70	0,97	6,65	81,10
5	3,58	1,03	2,01	0,45	5,90	88,88
6	6,63	3,83	3,82	1,49	7,52	76,90
7	4,79	0,67	2,93		6,74	81,88
8	4,47	2,67	3,07	0,39	6,87	83,29
9	4,16	3,53	2,95	0,96	6,39	86,52

GRADE 5

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated readability index	SMOG	Flesch reading ease
1.	5,46	4,63	3,44	1,90	7,05	84,59
2.	4,28	4,94	3,85	2,10	6,16	78,92
3.	8,20	6,14	4,76	3,17	8,18	71,03
4.	6,38	4,40	4,10	1,65	7,52	77,65
5.	4,71	4,39	3,81	1,74	6,76	82,29

6.	3,64	3,08	3,12	0,81	6,70	82,29
7.	3,79	4,09	3,25	1,49	6,55	82,26

GRADE 6

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated readability index	SMOG	Flesch reading ease
1.	6,80	6,31	5,04	3,21	8,07	72,57
2.	4,75	2,54	2,92	0,37	6,54	83,85
3.	4,24	2,94	2,42	0,61	6,44	87,88
4.	8,62	8,31	7,19	5,31	9,77	61,72
5.	6,02	6,19	5,24	3,58	7,51	75,51
6.	5,34	4,80	4,38	2,17	7,24	79,33
7.	6,27	6,07	4,59	3,14	7,24	77,33
8.	6,84	4,36	3,72	1,62	7,80	80,50
9.	4,92	6,75	4,70	4,16	6,87	80,26
10.	5,84	2,10	3,05	0,18	7,55	86,08

APPENDIX I: Readability scores for the Inzalo books

GRADE 4

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated Readability Index	SMOG	Flesch Reading Ease
1	8,04	6,53	6,23	4,43	8,84	71,76
2	8,11	4,69	5,15	2,88	9,00	79,19
3	6,30	2,82	3,23	0,88	7,14	88,81
4	6,13	4,04	3,84	2,21	6,70	87,65
5	7,40	6,38	4,81	3,76	8,04	78,77
6	8,57	7,68	7,39	5,78	8,92	65,69
7	6,68	6,03	5,49	3,65	7,37	75,00
8	7,02	8,11	6,61	5,59	8,62	68,46

GRADE 5

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated Readability Index	SMOG	Flesch Reading Ease
1	9,69	10,06	8,44	7,76	10,88	58,48
2	8,28	7,49	7,24	6,19	9,21	69,46
3	8,32	6,05	6,45	4,90	8,62	74,46
4	9,79	9,25	8,29	7,74	9,71	62,52
5	7,59	7,12	7,06	5,33	8,09	68,04
6	6,90	6,32	5,81	4,48	7,95	75,96

GRADE 6

Unit	Gunning Fog	Coleman Liau	Flesch Kincaid Grade level	Automated Readability Index	SMOG	Flesch Reading Ease
1	8,72	10,17	8,27	7,47	10,83	57,83
2	8,64	9,36	7,99	6,84	11,01	59,92
3	9,09	7,51	8,29	6,74	10,44	64,31
4	8,04	8,06	7,76	6,39	9,53	64,49
5	7,20	5,86	5,52	4,26	7,93	78,76
6	5,29	6,43	4,51	3,75	6,50	80,65
7	8,48	8,19	6,85	6,03	9,60	68,74
8	8,70	8,40	7,76	7,66	8,48	68,95