Assessment of the Language and Mathematics Skills of Grade 8 Learners in the Western Cape in 2006

K Heugh
GAM Diedericks
CH Prinsloo
DL Herbst
L Winnaar

March 2007

HSRC
Human Sciences Research Council

Table of Contents

Table of Contents 1
CHAPTER 1 - Background 3
1.1 BACKGROUND TO THE STUDY 3
1.2 STRUCTURE OF THE REPORT 3
1.3 TERMS OF REFERENCE 4
1.3.1 Tasks required by WCED and to be undertaken by the HSRC 4
1.3.2 Tasks and Responsibilities of the WCED 5
1.4 PROJECT RISKS 8
1.5 TIMEFRAMES FOR ACTIVITIES 9
CHAPTER 2 – Design and Implementation 12
2.1 INTRODUCTION/BACKGROUND TO THE CHAPTER 12
2.2 THE DESIGN OF THE STUDY 13
2.3 THE FORMAT OF LEARNER ASSESSMENT TASKS 13
2.4 INTERNATIONAL BENCHMARKING AND STANDARD SETTING 14
2.4.1 Benchmarking / Standards Setting of the Language Instrument/s 15
2.4.2 Benchmarking of the Mathematics Instrument 19
2.5 INSTRUMENT DESIGN 22
2.5.1 Instrument design: Language 22
2.5.2 Instrument design: Mathematics 32
2.6 PILOT ASSESSMENT INSTRUMENTS 41
2.7 ADMINISTRATION OF THE FINAL INSTRUMENTS – WCED RESPONSIBILITIES 42
2.7.1 Learner Information Data collection 42
2.7.2 Printing, batching and despatching of instruments 43
2.7.3 Administration of actual instruments 43
2.7.4 Retrieval of instruments and preparation for handover to HSRC 44
2.8 DATA COLLECTION, SCORING, CODING, AND ANALYSIS – HSRC RESPONSIBILITIES 44
2.8.1 Data entry and cleaning 44
2.8.2 Data scoring 46
2.8.3 Data analysis, reporting and interpretation of results 47
2.9 Limitations of the report 49
2.10 CONCLUSION 51
CHAPTER 3 – Learner Performance 52
3.1 INTRODUCTION 52
3.2 RELIABILITY OF THE INSTRUMENTS 52
3.3 PROVINCIAL SCORES 53
3.3.1 Performance per Learning Area 53
3.3.2 Performance per item type and Learning Area 54
3.4 LANGUAGE 56
3.4.1 Performance per language instrument 56
3.4.2 Language by performance levels 58
3.4.3 Language performance by Learning Outcomes 61
3.4.4 Language performance by grade level 62
3.4.5 Language performance by gender 65
3.4.6 Language performance by learning outcome for the different language instruments 67
3.4.7 Language performance by HL and LoLT 69
3.4.8 Language performance by item type 70
3.4.9 Language performance by EMDC 73
3.4.10 Very high incidence of learners exhibiting negligible/minimal evidence of literacy skills at Gr 8 78
3.5 MATHEMATICS 81
3.5.1 Mathematics by performance levels 81
3.5.2 Mathematics performance by Learning Outcomes 83
3.5.3 Mathematics performance by Cognitive Domain 86
3.5.4 Mathematics performance by gender 88
3.5.5 Mathematics performance and language 90
3.5.6 Learner performance related to glossed items 96
3.5.7 Mathematics performance by item type 99

Background

3.5.8 Mathematics performance by the number of books in the home ... 102
3.5.9 Mathematics performance of WCED learners compared to TIMMS 2003 ... 105
3.5.10 Analysis of learner performance / examples of typical learner responses ... 106
3.6 CONCLUSION ... 115
CHAPTER 4 - Conclusion ... 116
4.1 INTRODUCTION ... 116
4.2 KEY FINDINGS ... 117
4.3 FINDINGS WHICH REQUIRE PARTICULAR ATTENTION ... 119
4.3.1 Comparative Language/s Achievement – domestic and international comparisons ... 119
4.3.2 Percentages of learners performing at or above specified levels ... 121
4.3.3 High incidence of students who do not manifest evidence of literacy skills ... 121
4.4 RECOMMENDATIONS ... 122
4.4.1 Language ... 122
4.4.2 Mathematics ... 123
4.5 CONCLUDING REMARKS ... 124
References ... 126
Appendix ... 129

Acknowledgements

The HSRC would like to acknowledge and thank the following people for their assistance during the development, scoring and coding processes of the 2006 Grade 8 Assessment Tasks in the Western Cape:

Mr Matthews Makgamatha for the scoring and coding logistics

Mr Elias Hlongwane, Ms Sophie Strydom and Ms Fredericka Davies for administrative management, especially during the scoring and coding period, and Ms Lindi Basson, Ms Erika Masser and Ms Matselane Tshukudu for general project administration

Ms Lolita Winnaar for data processing, analyses and management

Professor Timothy McNamara (University of Melbourne), international language assessment advisor; and Dr Carol Macdonald (Witwatersrand University), Dr Vuyokazi Nomlomo (University of the Western Cape), and Ms Amanda Siegrühn for assistance with the language instruments' development

Ms Pamella Maseko and Professor Sandile Gxilishe (University of Cape Town), and Dr Monde Mbekwa (University of the Western Cape) for translation quality assurance

The very fine and committed Mathematics and Language moderators and educators (more than 200) from the Western Cape who coded the constructed-response items of 77 936 Grade 8 learners in the Mathematics, English, Afrikaans and Xhosa assessment tasks over a period of two weeks during the July 2006 school holidays

Ms Jenny Rault-Smith and Mr André Clausen, as well as the many other curriculum, examinations, printing, research and data specialists, of the Western Cape Education Department, who were involved in the project as instrument advisors and administrators

The principal and staff of Settlers High School in Bellville
for making available their hall and kitchen facilities; and the school catering team which provided meals and snacks during the wet, windy and cold days of scoring and coding.

CHAPTER 1 - Background

1.1 BACKGROUND TO THE STUDY

The purpose of this study and service was to develop instruments to assess the language and mathematics skills in English and Afrikaans of approximately 75 000 Grade 8 learners in all schools from the Western Cape Education Department (WCED) during 2006, and to analyse and report on the results of the tests. This was in order for the WCED to benchmark learners, prioritise learners with potential and develop and support learning and teaching processes.

The instruments were to be closely articulated with the National Curriculum Statement (NCS) at Grade 8 level.[1] The development and provision of instruments were to measure learner performance in mathematics and language at Grade 8 level as required in terms of Learning Outcomes and Assessment Standards as these are specified in the NCS (DoE 2002).

The WCED further required instruments which were able to benchmark student performance in relation to students in other countries (international benchmarking in terms of developing and developed countries). The Service Provider had to report on performance at provincial, Education Management Development Centre (EMDC), school and individual learner level.

1.2 STRUCTURE OF THE REPORT

Chapter 1 sets out the terms of reference for the study, and the division of responsibilities between the Client (WCED) and the Service Provider (HSRC) in terms of the contractual agreement (9 February 2006) between the two parties; the project risks; and the planned timeframes for the study.

Chapter 2 sets out the design and implementation of the study. In Chapter 3, the findings are presented in relation to the language and mathematics tasks, at provincial and EMDC levels. (Extended sets of tables were also provided separately on 28/29 March 2007 to this effect to report on and document mean and individual learner performance scores at the provincial, EMDC, school and learner levels.) Data will be shown in relation to learning outcomes (LOs), item type (multiple-choice and open-ended/constructed-response items), gender, numbers of books in the home, etc. Chapter 4 presents the conclusion of the study with recommendations for action to be taken by the WCED.

[1] Also known as the Revised National Curriculum Statement Grades R-9 (Schools) Policy (DoE 2002). The WCED confirmed that it was this version and not the earlier C2005 version of the NCS which was to be used as the frame for the assessment instruments.

1.3 TERMS OF REFERENCE

The aim of the WCED was to conduct an assessment and report on the achievement of Grade 8 students in all WCED schools during the second school term of 2006. A further aim of the WCED was to benchmark its learners' achievements against students' achievement in other countries. The objectives were to measure the achievement in mathematics and language of learners. They were also to prioritise learners with potential and inform the WCED's trajectory of developing and supporting effective learning and teaching processes. In order to fulfil the aims and objectives of the study, the WCED identified a set of tasks to be contracted out to an independent provider, and it identified a set of tasks to be undertaken by the Department.

1.3.1 Tasks required by WCED and to be undertaken by the HSRC

In terms of the contractual arrangements, the HSRC had the responsibility to develop the instruments and conduct reporting in a manner which included the following requirements as spelt out in the tender document, B/WCED 375/05 (WCED 2005).

1.3.1.1 The Instrument Development

a. Development of a mathematics instrument.
b. Development of a language instrument (Afrikaans and English).
c. Each instrument needs to be based on the NCS and provide the WCED with international benchmarking (in relation to developed and developing country) information.
d. The development of instruments is to involve close consultation with the WCED via a mutually agreed upon process and schedule.
e. Each instrument should not take more than two hours to administer.
f. Each instrument is a pen-and-paper test, and should be about 15 A4 pages.
g. Instruments need to be delivered to the WCED by 1 April 2006.
h. Answer books/sheets per learner and per school are to be delivered to the Director: Examination Administration, so that these can be sent to the schools with the test instruments.
i. A concise, user-friendly test administration manual is to be developed and delivered to the WCED (electronic and one hard copy) on a mutually agreed upon date.

1.3.1.2 Additional Instrument Development feature agreed upon by WCED and HSRC as per the contract

a. Development of a Xhosa language instrument as a pilot process for 2006.

b. Include language accommodation in the Mathematics instrument for Xhosa-speakers, as a pilot process for 2006.
c. Develop a diagnostic system for reporting learner performance in language and mathematics.

1.3.1.3 Marking, coding, capturing data and analysis

a. Tests need to be marked.
b. Suitable data capturing, cleaning and analysis systems are to be developed.
c. Scores/data need to be captured electronically.
d. Data will require cleaning and checking.
e. The test items data and results of the tests will require sophisticated analysis.

1.3.1.4 Reporting

a. An electronic report and one hard-copy of a report: per learner, per school, per EMDC and for the Province, by the end of October 2006 (WCED 2005: paragraph 3.3).
b. This report must include results of the test and a benchmarking against other countries that wrote the same tests (WCED 2005: paragraph 3.3).
c. Reports, in hard copy, need to be provided for each school and each learner in the school, on the results of that school (WCED 2005: paragraph 3.4).
d. Reporting on the results for the learners in their Home Language (Afrikaans and English) (WCED 2005: paragraph 4.5).
e. Reports should identify learners with potential (WCED 2005: paragraph 1).
f. Reports should provide data and analysis which would advance the WCED's understanding of how to further develop and support learning and teaching processes (WCED 2005: paragraph 1).
g. A certificate, by January 2007, for each learner, in the home language, and which shows the achievement for language and mathematics in terms of the 7- point scale which forms part of the new National Senior Certificate (FET level) from 2008 onwards (WCED 2005: paragraph 3.5). (On 19/20 March 2007, almost 78 000 individually customised learner report cards were submitted to the WCED. They comprised total and selected sub-scale performance scores and levels, as well as related diagnostic comments, for both learning areas.)

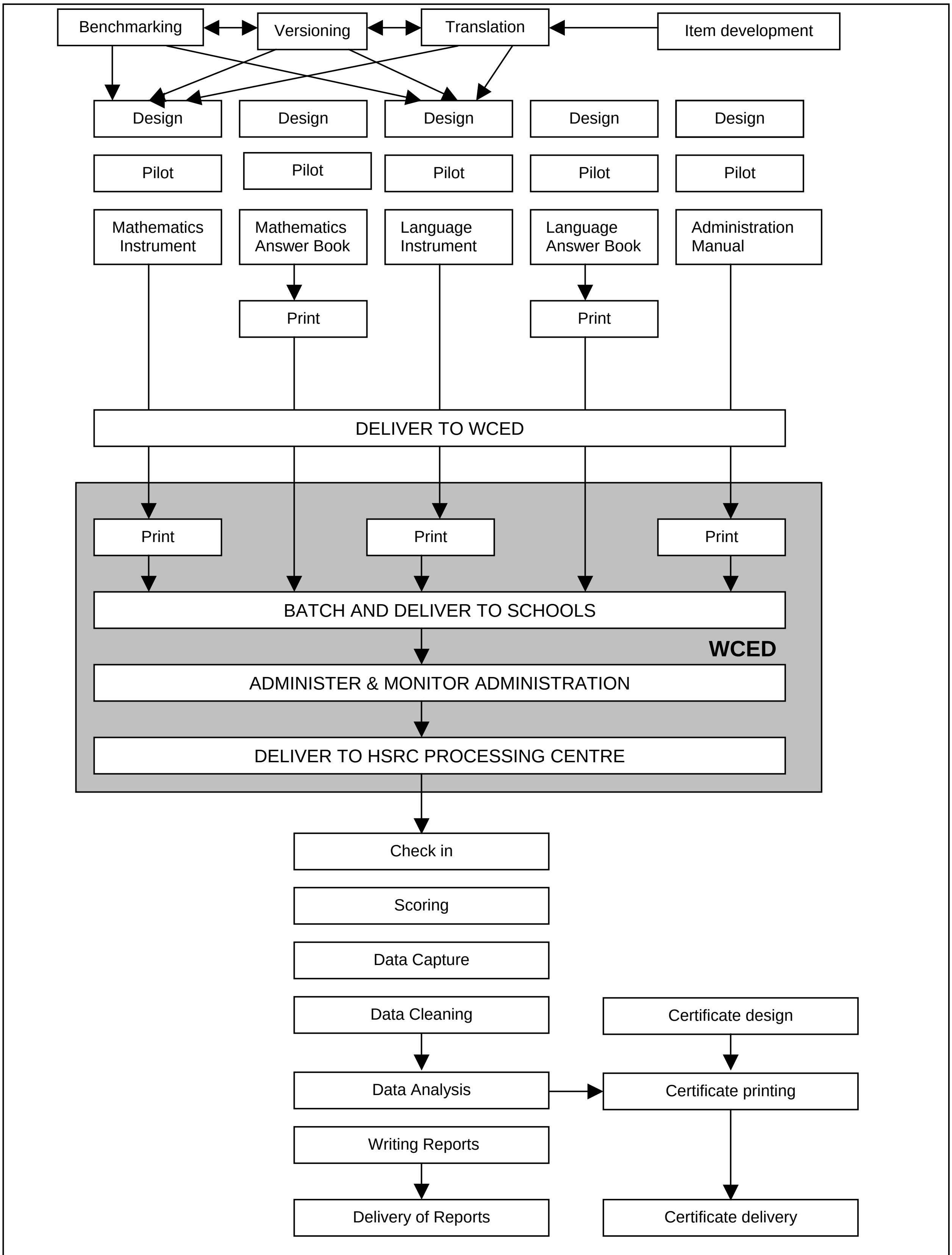
1.3.2 Tasks and Responsibilities of the WCED

a. Printing of tests (WCED 2005: paragraph 5.1).

b. Printing of the administration manual (WCED 2005: paragraph 5.3).
c. Distribution of tests, administration manuals, and answer sheets/booklets to each school (WCED 2005: paragraph 5.1).
d. Administration of tests by 23 June 2006 (WCED 2005: paragraph 5.2).
e. Collection of completed tests (WCED: paragraph 5.1).
f. Distribution to each school, the school reports, and individual learner reports (by implication of WCED 2005: paragraph 3.3 & 3.4).
g. Distribution to the EMDCs, the relevant EMDC report (by implication of WCED 2005: paragraph 3.3).
h. Distribution to each learner, the learner certificates (by implication of WCED 2005: paragraph 3.5).
i. Make available to the Contractor, electronic data per learner, per school, per EMDC, and per teacher if required for further analysis (see Section 8.3 of the HSRC Proposal, HSRC 2005).

Table 1.1 contains a flowchart of activities to express the interlinkages between various elements pertaining to the sets of responsibilities mentioned above.

Table 1.1: Flow of activities pertaining to intended HSRC and WCED responsibilities



MAIN ACTIVITIES IN THE WCED GRADE 8 ASSESSMENT
(Activities in shaded box for WCED; all other activities for contractor)

1.4 PROJECT RISKS

It was important to consider possible project risks, particularly in the case of system-wide or large-scale assessments. This was in order to maximise the opportunity to plan to reduce the potential of such risk. The following project risks were identified by the HSRC and identified in its proposal to the WCED and thence incorporated into the contractual documentation:

> "In projects that involve large-scale data processing, analysis of many records, and where datasets must be linked, the integrity of the master set of data is critical. Therefore the following risks must be ruled out:
>
> - Incorrect or missing learner record data from schools can present problems for linking test data to individual learners. Schools must ensure that the data they supply is correct.
> - Incorrect recording of unique learner numbers on each answer booklet can lead to test data being linked to the wrong learner. Administration Officers must pay attention to ensuring the correct learner number is filled in on each booklet.
>
> To obtain learner data from 520 schools in any environment is a large administrative task, especially where schools may not be able to supply data electronically to the WCED.
>
> - Slow response of schools to requests for learner data can lead to delays in packaging the answer booklets to schools. The collection of data on learners in Grade 8 should begin as soon as possible. It is suggested this can take place soon after the 10th day Survey, or even be linked to this exercise.
>
> In all projects involving fieldwork to collect data, there is risk relating to aspects of fieldwork activities. *(In this project, the service provider did not have control over the fieldwork.)* The following risks have to be ruled out:
>
> - Supply of assessment booklets to the wrong school can lead to delays in the writing of the test and possible leaks of the content of the test. Distribution of tests must be as accurate as possible.
> - Test administration conditions could differ leading to uneven performance between schools and classes. The Administrative Officers must follow the procedure closely.
> - The collection of the completed answer booklets could be slow, leading to delays in data processing and later analysis. The logistics of collecting test booklets must be as efficient as possible.

- The packages from schools could be poorly controlled leading to delays in the data capture and cleaning process. As indicated above a check on the quality of the distribution, administration and returns process must be implemented.
- There could be incomplete or no-returns from schools on account of local conditions or delays. The client and the service provider must decide on a cut-off date beyond which data cannot be included in the full analysis at the provincial and EMDC levels. Reporting of data at the school and learner level may need to be negotiated.
- Numbers of students may be absent on the day(s) of the tests. The service provider and client may decide that the Administration Officer must return the answer booklets the day after the tests are written. No delays of test booklets may be allowed for absent learners to be included. The WCED will be given the opportunity to process late tests on the account of the Department. However, the results of late learner submissions will not be included in the reporting." (HSRC 2005).

1.5 TIMEFRAMES FOR ACTIVITIES

The timeframe of activities were planned in terms of the dates outlined by the WCED during the compulsory briefing meeting of potential bidders for WCED Tender B/WCED 375/05, in early October 2005. The WCED anticipated informing the successful bidder by 15 November 2005 that the contract had been awarded. The HSRC planned to begin work on the project as of 16 November 2005, since the timeframes for development of the instruments and delivery of completed instruments was set at 1 April 2006. An undertaking of this magnitude requires piloting and analysis of instruments and consequent finalization of the instruments in order to meet international assessment criteria for quality, reliability and validity. The success of the tight timeframes would be largely dependent upon the absence of any untimely delays on either side, as well as careful consideration of the risks outlined above. The GANTT chart in Table 1.2 represents the planned and anticipated time schedule of activities in relation to this study.

Table 1.2: Sequence of planned activities for the Grade 8 study implementation

GANTT CHART OF PROJECT TIME SCHEDULE															
Activity	**Nov 05**	**Dec 05**	**Jan 06**	**Feb 06**	**Mar 06**	**Apr 06**	**May 06**	**Jun 06**	**Jul 06**	**Aug 06**	**Sep 06**	**Oct 06**	**Nov 06**	**Dec 06**	**Jan 06**
Instrument Development in Consultation with WCED															
Detailed schedule from micro-level planning and research															
Mathematics instrument design															
Mathematics instrument benchmarking															
Mathematics answer sheet design															
Language instrument design															
Language instrument benchmarking/alternative															
Language answer sheet design															
Administration manual draft															
Selection of sites for pilot															
Permissions for pilot		8													
Formal confirmation notification to pilot schools			15												
Translate instruments															
Pilot instruments, manual and answer sheets/booklets															
Mark, code, and capture pilot data															
Analyse data															
Evaluate pilot instruments, manual, and booklets															
Refine instruments															
Refine answer sheets/booklets															
Refine manual															
Deliver instruments to WCED						1									
Deliver manual to WCED						1									
Print answer sheets/booklets							1-5								
Package answer sheets/booklets							12								
Deliver answer sheets/booklets to WCED							19								
Design certificates															

Data systems development and analysis															
Activity	**Nov 05**	**Dec 05**	**Jan 06**	**Feb 06**	**Mar 06**	**Apr 06**	**May 06**	**Jun 06**	**Jul 06**	**Aug 06**	**Sep 06**	**Oct 06**	**Nov 06**	**Dec 06**	**Jan 06**
Receive digital data(learner, school, EMDC) from WCED															
Develop coding, capturing, analysis systems (formats)								9							
Identify/recruit markers with WCED (constructed-response items)															
Training of markers								26							
Marking / coding															
Data capturing															
Data cleaning										18					
Data analysis (item scoring, sub-scores and marks, other)															
Report drafting												20			

Reporting to WCED (Delivery to Director: Examination Administration)															
Activity	**Nov 05**	**Dec 05**	**Jan 06**	**Feb 06**	**Mar 06**	**Apr 06**	**May 06**	**Jun 06**	**Jul 06**	**Aug 06**	**Sep 06**	**Oct 06**	**Nov 06**	**Dec 06**	**Jan 06**
Report of findings per class, per learner												31			
Report of findings per school												31			
Report of findings per EMDC												31			
Report of overall findings for Province												31			
Report cards per learner												31			
Refined report: Overall findings for Province (if required)															
Certificates per learner															

CHAPTER 2 – Design and Implementation

2.1 INTRODUCTION/BACKGROUND TO THE CHAPTER

This chapter sets out the theoretical considerations for the design and implementation of the assessment instruments. First, an overview of the design of the study is presented; second, the instrument development, piloting, data collection, analysis and interpretation are described; and last, the limitations of the study are noted.

The research methodology addressed the criteria required by the WCED for the assessment of and reporting on achievement of each Grade 8 learner (approximately 75 000 learners) in both language and mathematics tasks in the Western Cape before 23 June 2006, i.e. half-way through the Grade 8 school year. The process followed sequentially the following set of broad activities:

HSRC

- Instrument design
- Benchmarking and / international standards setting
- Piloting of instruments, administration, data capturing, analysis, etc.
- Refinement of the instruments
- Handover of instruments to HSRC (after administration in the schools by WCED)
- Contracting and management of scorers/coders
- Contracting of data capturers, and management and monitoring of data capturing process
- Data cleaning and analysis
- Reporting

WCED

- Collection and processing of learner information in agreed electronic format, including biographic information, home language, and allocation of individual bar-coded learner ID data
- Supply of this electronic data to HSRC for analysis and reporting purposes
- Printing, packaging and distribution of instruments
- Administration of instruments before 23 June 2006
- Recovery of instruments, and preparation for handover to HSRC in July 2006
- Acceptance of individual learner reports and distribution to learners
- Acceptance of school and EMDC reporting tables, completed provincial report

Design and Implementation

- o Acceptance of certificates and distribution to learners.

The account of the development of each of the language and mathematics instruments is presented as these were conceptualised and developed by two different instrument development teams, and within somewhat different sets of frameworks. Therefore the format and layout of this account differs between language and mathematics below.

2.2 THE DESIGN OF THE STUDY

The key elements of the study are the development of assessment tasks in Languages and Mathematics based on the National Curriculum Statements (DOE 2002a, b), designed to accommodate international benchmarking features, and administered to all Grade 8 learners in the Western Cape. The study included a pilot phase, where instruments were developed and piloted in a purposive, representative sample of schools, following which, data were analysed. This provided an opportunity to modify the instruments in order to ensure quality, validity and reliability. The main study instruments were then administered by the WCED. Responses were subsequently scored and coded, data was captured into a large electronic data set, and this was then subjected to comprehensive analysis. The analysis was designed to establish learners' levels of achievement in language and mathematics and to examine whether there were differential levels of achievement according to gender, item type, district (EMDC), language of test, home language, socio-economic background (only through a broad proxy indicator), etc. Data were also analyzed to determine the percentage of learners who demonstrated mastery levels of performance as set out in terms of different Learning Outcomes and Assessment Standards of the NCS for the Home Language Learning Areas and Mathematics of the Curriculum (DoE 2002a, b). The WCED specifically requested that learner achievement would be reported on in terms of the new 7-point assessment scale, as published in the NCS for Further Education and Training (General) (DoE 2005), for levels of achievement as determined by the national Department of Education (DoE 2005a, b). A further dimension to the study has been to include international benchmarking criteria to the mathematics instrument, and international standards criteria for valid language testing in linguistically diverse school systems. Lastly, the WCED required the instruments to offer an indication of the learners' prospects for further education beyond the General Education and Training (GET) band which terminates at the end of Grade 9. To this end, this report will conclude with a critical discussion of the findings with some policy and planning recommendations for the WCED.

2.3 THE FORMAT OF LEARNER ASSESSMENT TASKS

Instruments were developed in a pencil-and-paper format and were to assess the learners' language and mathematical knowledge and skills using English and Afrikaans as the mediums of assessment for the mathematics task, and Xhosa[2], Afrikaans and English as the mediums of assessment for the language instrument/s. Since only English and Afrikaans had been prioritised in previous years of Grade 8 Assessment, the HSRC recommended that the WCED phase in language accommodation to address the *de facto* issue of linguistic diversity in line with the Western Cape Language Policy (Western Cape Legislature, 2001), the national Constitution (RSA 1996), and the Language in Education Policy (DoE 1997) and address the language issues in relation to the three official languages of the province, Afrikaans, English and Xhosa. It was suggested that accommodation of other language mismatches would be attended to in subsequent years of assessment.

The WCED required diagnostic information in regard to learners' potential for achievement in further education and training (FET). A mathematics instrument has the potential to indicate the possible avenues for further study and potential career paths for learners, however, this meant that the language instrument would have to assess more than is normally required of a 'language as a subject' test. The instrument would have to assess learners' Home Language (as a learning area) as well as learners' academic language skills, in terms of how these facilitate or inhibit learning across the rest of the curriculum. To our knowledge, and that of a wide cross-section of international language assessment experts, no language instrument had been designed to accomplish such a double task elsewhere, and this therefore required significant and new development of the construct of the instrument.

The HSRC consulted closely with the WCED in regard to the fine details of the mathematics and language instruments' development and the inclusion of new design features to address the language mismatch of African language speaking (mainly Xhosa-speaking) students and the language of learning and teaching in mathematics.

2.4 INTERNATIONAL BENCHMARKING AND STANDARD SETTING

One of the requirements for the assessment instruments was that they should benchmark student performance in relation to students in other countries (international benchmarking in terms of developing and developed countries). There is a practice whereby education systems wish to compare the health of their systems with those in other countries by using multi-country assessments which do not have 'international benchmarking' status, and a limited number of

[2] The HSRC recognized that there is at least one secondary school where students are mainly Home Language (HL) speakers of Sesotho, and at least one other where students speak Sesotho or Setswana as HL, and that many speakers of Afrikaans have elected to use English as Language of Learning and Teaching (LoLT) and other instances where Xhosa speakers have elected to have either English or Afrikaans as HL as well as LoLT. There are furthermore additional language mismatches in relation to other minority language communities.

'internationally benchmarked' instruments. In relation to Grade 8, there is an obvious internationally benchmarked instrument, Trends in International Mathematics and Science Study (TIMSS) which lends itself to the WCED requirements. In regard to language, there is currently no instrument which qualifies for this status at this time, therefore an alternative approach needed to be found. One of the key contemporary challenges in international assessment contexts is the accommodation of diversity amongst learners, particularly linguistic diversity and ethical practice in language assessment. The HSRC approach was to investigate appropriate strategies to accommodate the linguistic diversity of the Western Cape in both the mathematics and language assessment tasks, and to ensure the highest ethical standards of assessment practice.

2.4.1 Benchmarking / Standards Setting of the Language Instrument/s

In preparation for the design of the language instrument and explicit requirement of the WCED for international benchmarking of the assessment tasks, the HSRC undertook research into the assessment procedures and design of language instruments, which are used across Europe, North America, Australia and New Zealand, South-East Asia and Africa. This investigation included consultation with a wide range of internationally acknowledged experts in the field of language acquisition, literacy development in school education, and language assessment at home language (L1) and first additional language (L2) levels (See Appendix, HSRC 2005). It involved a broad investigation of the language assessment frameworks and available instruments, including: the *Common European Framework*; the British *Qualifications and Curriculum Authority* (QCA); Cambridge TESOL; *National Foundation for Educational Research* (NFER) assessment procedures and research for Britain, Singapore and Hong Kong; *National Education Monitoring Project* (NEMP) of New Zealand; *South Australian Curriculum, Standards and Accountability Framework*; Ontario Achievement Levels (language); and several US state assessments, e.g., the *Tennessee Comprehensive Assessment Programme* (TCAP), and New Jersey Grade 8 Proficiency Assessment.

Owing to the complex nature of language assessment to be benchmarked at Grade 8 level and a requirement to include 'developing' and 'developed' contexts, a brief comparative analysis is therefore presented here. Essentially, a language instrument which would be useful to the WCED would need to have similarities in relation to the language background of learners. In the Western Cape approximately 76%-78% of learners use their Home Language (first language / L1) as LoLT; 22%-24% use their First Additional Language (second language / L2). Therefore, the use of instruments in other countries which cater for the needs of a majority of students who study through their Home Language and a minority who do not, are the ones which need to be evaluated for closest proximity in learner profile and also in terms of curriculum design.

Such assessments are conducted by the education authorities of each state in the USA at Grade 8, in preparation for what is known as the *National Report Card in the USA*, and the results of approximately 660 000 students are collated by the National Center for Education Statistics annually. Whilst the US state curricula and instruments were not specifically influential in the design and 'benchmarking' or 'standards setting' of the NCS language statements, there are features of the instrument design which are similar to those used in Canada and Australia, and which informed the Languages Area Statement in the NCS (DoE 2002a). The UK National Curriculum *Qualifications and Curriculum Authority (QCA)* similarly conducts assessments at 11+ and age 14 (Year 9), which is fairly close in age to Grade 8 students in South Africa, but pegged probably a bit higher than would be realistic or valid for our students. The influential National Union of Teachers (NUT) in the UK has been strongly critical of the QCA language assessments for being culturally biased and inappropriate. The degree of resistance to this instrument reached a point in 2005 that since so few teachers were prepared to grade these assessments, they had to be sent to India and Singapore for grading (Lindsay 2005). A further, significant, consideration is that teachers teach to the specific marking schedule which is used for the QCA instrument. Thus students are specifically prepared for this instrument. Students in the Western Cape would therefore be at a significant disadvantage if we were to use this particular instrument, because they have not been prepared for this framework by their teachers.

The London Education Authority (LEA) is increasingly in favour of using *CAT (Computer Adaptive Testing)* which is particularly suitable for second language learners. The use of on-line CAT is increasing across the US and Britain at the moment. These instruments are useful for predictive purposes as they appear to be able to predict the achievement outcomes of students in the GCSE or secondary school exit examinations. They are sensitive and responsive to different levels of competence and achievement, and would be very useful and offer fairly accurate benchmarking possibilities against L1 and L2 learners in North America and Britain. Whilst these offer significant opportunities and possibilities for South African conditions in the future, the use of computers for testing across the system is unfortunately premature in the Western Cape at present.

The *Common European Framework (CEF)* is a generic system designed to capture the language competencies in each of the languages which students have or are developing, and it is designed for portability across the European Union. The features of this framework are ideally suited to South

African education in principle, since they accommodate multilingualism, allow for the development of competencies over time, and are cumulative. The framework, however, was not designed per grade level, and is entirely criterion based. The instrument is not directly portable to the WCED for the Grade 8 assessment if the requirement is benchmarking at a particular grade level and if the achievement in only one language is being measured. It does, however, have partial portability if the developers of an instrument were to extrapolate and estimate levels of competency for Grade 8. CEF is also based on the notion of a language portfolio which includes tests and assessments which are both teacher- and self-assessed. This limits its use for a pen-and-paper-only instrument.

Of the available language assessment frameworks currently being used by other countries, the Canadian (Ontario), Australian (South Australian) and New Zealand (National Education Monitoring Project) approaches to language assessment are the closest to the South African NCS. This is largely because the NCS have been loosely guided by systems used in these contexts. Assessments conducted within these frameworks show the differences between L1 and L2 learners. English speakers are in the majority, and ‘minority’ or ‘immigrant’ language speakers are in the minority, with a close approximation of the 80% L1 speakers of English.

There is a fairly considerable degree of similarity across North America, the UK, Australia and New Zealand in terms of the technical design of language instruments, and emphasis towards reading, writing, comprehension (interpretation and inference) of different types/genres of texts required across the curriculum. There is also an increasing international tendency towards systemic education monitoring and evaluation of trends which is smaller scale in terms of volume/numbers of schools and students, but more intensive and comprehensive approaches to gathering data. Such data provides a nuanced longitudinal analysis of student achievement in combination with systematic training or preparation of teachers (e.g. the NEMP approach in New Zealand). It is currently argued that large-scale assessments are regarded as expensive and less reliable than more comprehensive approaches.

Since there is no equivalent of the TIMSS for language at Grade 8 level, the HSRC suggested an alternative route to language instrument development which would satisfy WCED requirements to measure learner achievement against students in other countries. The methodology would include a detailed comparison of instruments used at approximate grade / age level from selected international sites: Canada, Australia, New Zealand, the UK, the USA, and the European Union. This would be followed by a careful selection of test items which are relevant to the Assessment Standards for Grade 8 in the Home Language Learning Area Statement (DoE 2002a). Since the

Design and Implementation

NCS is itself designed to meet international standards, a combination of utilizing the selected assessment instruments mentioned above in conjunction with the NCS, was purposively employed to meet international assessment standards at Grade 8 level.

An international language assessment expert was commissioned to monitor and evaluate the language instrument methodology in order to meet international standards in language assessment at Grade 8 level.

2.4.1.1 Framework for the design of the language instrument/s

The language instrument framework developed for the WCED, in compliance with

- international criteria for language assessment in linguistically diverse settings,
- international standards of construct validity,
- the Learning Outcomes and Assessment Standards as set out in the NCS (DoE 2002a) which is accepted as national education policy based on international curriculum and assessment development standards, and
- including specific further requirements of the WCED that the instruments provide some evidence of academic potential of learners beyond Grade 8,

may be presented schematically as in Table 2.1.

Table 2.1: Language instrument/s organising framework

HL Section	Common Writing Task – in HL	LoLT section
Based on literary text originated in each of 3 language versions [Afrikaans, Xhosa, English] Set in a familiar socio-cultural and geographic, context. Familiar literary style for learners.	Two common writing tasks in each of 3 Home Language versions (translations) – for direct comparison across 3 different HL groupings.	Passage selected from an international text originated in English, allowing for comparison of textual reading material expected of secondary school students internationally. Set in an African context – which offered some geographic proximal familiarity, yet also within an international setting o Adapted to match international reading level standards at Grade 8. o Translated into Afrikaans. o Further, minimal adaptation in the version for Xhosa HL learners who study through English FAL.
To test reading, comprehension, writing, and knowledge of language structure. Would provide evidence of achievement in Home Language as a subject and offer evidence of literacy skills in this domain.	To assess productive writing skill development.	To assess ability to read text across the curriculum, understand this text, and to produce written evidence of having engaged with this text, i.e. academic literacy skill development. Would provide some diagnostic evidence of potential achievement in further education.

Of the six Learning Outcomes (LOs) in the NCS for Languages, only four are appropriate for a one-off pen-and-paper test conducted in a limited timeframe. These are identified below together with the planned weighting of items per LO:

- Learning Outcome 3: Reading and Viewing: 30%
- Learning Outcome 4: Writing: 30 %
- Learning Outcome 5: Thinking and Reasoning: 30%
- Learning Outcome 6: Language Structure and Use: 10%.

2.4.2 Benchmarking of the Mathematics Instrument

There is only one internationally benchmarked instrument for mathematics at Grade 8 level, and this is from TIMSS. TIMSS is a large-scale study conducted by the IEA (International Association for the Evaluation of Educational Achievement). South Africa participated in the 1995, 1999 and 2003 studies. The HSRC played a leading role in conducting these three rounds of TIMSS in South Africa and is familiar with the research methodology in relation to this instrument. Benchmarking for the mathematics instrument was based on the TIMSS 2003 instrument which is suitable for international benchmarking in relation to both developing and developed countries.

Because the WCED study was based on TIMSS, the test design and methodology used in TIMSS had a direct effect on these aspects in the WCED study. A short overview of the design and background against which the TIMMS project is undertaken is therefore provided.

2.4.2.1 Background to TIMSS

The central aim of TIMSS is to measure achievement in mathematics and science in order to learn more about the nature, content and context in which learning occur. The assessment frameworks of the TIMSS studies focused on the curriculum as a broad explanatory factor underlying student achievement. It envisaged three “levels” of curriculum, namely what society would like to see taught (the intended curriculum), what is actually taught (the implemented curriculum), and what the learners learn (the attained curriculum). Factors from the educational environment that influence educational decisions are investigated from the perspective of these three curriculum levels.

TIMSS uses written tests of mathematics and science to measure students’ achievement as well as a series of questionnaires (Curriculum, School, Student, and Science and Mathematics Teacher Questionnaires) to collect data from a representative sample from participating countries. The TIMSS study tested learners from a large number of countries over an extensive content domain. Each student is asked to complete one booklet, comprising a subset of items from the total item pool. After each TIMSS assessment cycle, some items are released for public use while the others are kept secure to measure trends over time. Replacement items that closely match the content of the original items are developed. Comparable scores for students that took the test in different years are obtained by making use of Item Response Theory (IRT).

The mathematics and science achievement tests were developed internationally in a collaborative manner. The tests contain two types of questions. The first type of questions requires learners to select appropriate responses (multiple-choice questions) and the second type requests learners to solve problems and answer questions (constructed-response questions). Each constructed-response question has its own scoring guide, developed to provide data about students’ achievement as well as diagnostic information about misconceptions and common errors. The scoring uses two-digit codes with scoring guides specific to each item. The first digit designates the correctness of the response. The second digit, combined with the first digit, represents a diagnostic code used to identify specific approaches, strategies, or common errors and misconceptions.

The scoring of these booklets is accomplished through the use of a sophisticated and strict set of criteria that is implemented consistently across all nations to assure accuracy and comparability.

The next section provides the reader with an overview of the framework and organization of the TIMSS mathematics test in 2003.

2.4.2.2 Framework for the design of the mathematics instrument/s

The mathematics assessment framework for TIMSS 2003 consisted of two organizing dimensions, a content and cognitive dimension. The five content domains that were identified by TIMSS 2003 for assessment was *Numbers, Algebra, Measurement, Geometry and Data*. The cognitive domains that define the sets of behaviours expected of learners as they engage with the mathematics content were *Knowing facts and procedures, Using concepts, Solving routine problems, and Reasoning*. Tables 2.2 and 2.3 provide an overview of the target percentage of testing time across the content and cognitive domains.

Table 2.2: Mathematical content domains

Content domain	**Target % of testing time**	**Number of items in TIMSS 2003 test**	
		Multiple-choice	**Constructed response**
Whole Numbers	30%	43	14
Algebra	25%	29	18
Measurement	15%	19	12
Geometry	15%	22	9
Data	15%	15	13
Total	100%	128	66

Table 2.3: Mathematical cognitive domains

Content domain	**Target % of testing time**
Knowing Facts and Procedures	15%
Using Concepts	20%
Solving Routine Problems	40%
Reasoning	25%
Total	100%

Table 2.2 indicates that the greatest emphasis in the mathematics instruments was on the Numbers Content Domain. Most testing time and most questions were devoted to this domain. Solving routine problems (see Table 2.3) was the cognitive domain to which most time was devoted.

The NCS (DoE 2002b) categorises the Mathematics Learning Area into five Learning Outcomes: *Numbers, Operations and Relationships; Patterns, Functions and Algebra; Shape and Space (Geometry); Measurement;* and *Data Handling*. The content dimension of the TIMSS framework corresponds with the Learning Outcomes as structured in the NCS (DoE, 2002b) in such a way that comparisons of achievement in the content domains are possible.

The WCED study used data collected on the performance of Grade 8 learners in mathematics during the 2003 TIMSS studies. TIMSS uses a sophisticated process of scaling. For the WCED study the data analysis and reporting were simplified. The results were thus reported on the basis of the average percentage of correct responses. About 50% of the items for were selected from the TIMSS 2003 released items. The performance of the Western Cape learners in these items in this mathematics instrument was compared with the performance of the Western Cape learners on the same items in TIMSS 2003 as well as with the SA and international averages for these items in TIMSS.

2.5 INSTRUMENT DESIGN

2.5.1 Instrument design: Language

2.5.1.1 Construct validity and accommodation of linguistic diversity in the language instrument

A key criterion for international benchmarking or international standards setting is that the instrument should abide with current conventions regarding construct validity. The contemporary considerations in this regard, particularly in diverse linguistic settings, is that the language assessment accommodate diversity in the construction of a valid instrument, in order to eliminate discriminatory assessment practices. The HSRC approach to meeting international standards of language assessment was therefore directed towards ensuring ethical and good assessment practice and this meant in terms of construct validity that the provider was determined that current considerations regarding linguistic diversity were built into the design of the instrument.

To this end the HSRC recognised the historical situation that at Grade 8 level, the majority of students take their HL as a subject, and in the Western Cape, students usually have Afrikaans, Xhosa or English as HL. A minority have Sesotho or Setswana as HL and there are students who have several other languages as HL, which the WCED has not been able to accommodate at this time. Whilst three languages are used by most learners as HL, only two are used as languages of learning and teaching (English and Afrikaans). The language specifications for the assessment task in the WCED tender 2005, however, specified language and mathematics testing in Afrikaans and English only.

The requirements of the NCS are that students cover the same curricular requirements in the languages learning area regardless of Home Language background. The First Additional Language curriculum requirements are substantially different from the HL requirements. Therefore, although the WCED tender specified language assessment in Afrikaans and English only, this would have meant that HL speakers of African languages would have been expected to write an assessment task in English at the same level as HL speakers of English. This would have meant that Xhosa-speaking learners would have been expected to write an assessment task pitched at the wrong language level. Apart from negative discrimination, this would have negated the validity of the instrument, and rendered it ethically unsound. It should be noted at this point that there are very strict ethical considerations which are expected of language assessment internationally, and that the HSRC adheres to these.

It needs to be further noted that the Grade 12 Senior School Certificate administers assessments in both the learners' Home Languages and First Additional Language when assessing language competence. Although African language speaking learners are expected to write their assessments in other learning areas through English, there is no precedent where African language speaking learners write the English HL (as a subject) assessment. They write the English (subject) at FAL level.

Xhosa-speaking students who might have been expected to take a language assessment designed for English HL students[3] may therefore have suffered serious difficulties. These would have been compounded upon receipt of individual reports and certificates of achievement which would compare them with learners who wrote the HL assessments. The HSRC, therefore, suggested a carefully selected use of additional instrument constructs, which would be useful for the WCED in the short, medium and long term and would go some way towards mediating any unintended but negative consequences of an assessment which does not take into account the different circumstances of Xhosa-speaking students.

[3] As was the case in the previous two years of Grade 8 assessments conducted by a foreign assessment agency.

The solution was to test students in both their Home Language proficiency and in their proficiency in LoLT within one instrument. For the majority of students home language and LoLT would be the same language; for about 22-24% of learners this would be different. The details of this would be spelt out below.

2.5.1.2 International standards setting and construct validity

HSRC anticipated a small advisory team including at least one acknowledged expert in language assessment and curriculum (external to the HSRC) to ensure that the language instrument has equivalence with the spread of international instruments currently in use. HSRC subsequently commissioned Professor Timothy McNamara of the Language Testing Research Centre, University of Melbourne, who is an internationally renowned language assessment expert, to advise and monitor the development of the language instrument. Professor McNamara travelled to South Africa during the conceptualization stage of the instrument development and advised the team. Professor McNamara met with the WCED, in late February 2006, in order to advise the Department on the language instrument development. He drew attention to the time constraints under which such a difficult instrument would have to be developed, from scratch, and that these were unrealistic.[4] It normally takes 24 months with 6 full-time language assessment specialists to develop a new instrument to international standards, but in this case there was a period of 3,5 months for development, piloting, revision, finalization and delivery of an instrument in three different language versions. It was agreed therefore with the WCED, that the language instruments would be developed as part of an on-going process of developing new instruments to meet the requirements of the Department, but that the essential features of international good practice would be included and adhered to in the development process. Subsequent refinement and modification would in all likelihood be necessary. Professor McNamara indicated, however, that he was confident that the HSRC would be able to offer a robust instrument and one which, in his opinion, included some features which exceeded international good practice. In essence, this instrument, in three versions, would set new standards for international language testing. It needs to be noted, that both the HSRC and WCED recognized that owing to time and resource constraints, the language instruments in use during 2006 would be considered to be undergoing development and that 2006 would be the first trial of these.

[4] The tender was awarded more than a month later than had been expected [22 December rather than 15 November 2005]. Given the time of year, the first meeting between the WCED and the provider was set for mid- January 2006, and the instruments had to be developed and delivered to the WCED for printing in the first week of May 2006.

2.5.1.3 Instrument development team

Apart from a team of HSRC researchers, an instrument development and advisory team included: Dr Carol Macdonald (University of the Witwatersrand, and a South African expert on language and cognition and the curriculum); and two regional language education experts (Ms Vuyokazi Nomlomo and Ms Amanda Siegrühn, with expertise in language teacher education, text-book writing, curriculum development and translation, and contributing to the Afrikaans and Xhosa Home Language instrument versions. The HSRC project leader co-ordinated the development of the instruments and regular meetings were held with senior language curriculum officials within the WCED to ensure that the instruments were developed according to the criteria mentioned.

The language assessment instrument comprised two sections, namely the Home Language (HL) section and the Language of Learning and Teaching (LoLT) section. The logic for this was that this structure would allow for each learner to be assessed in terms of what s/he is able to achieve in the Home Language, as a subject/learning area, and also in terms of how the learner is able to use her/his language skills in other areas of the curriculum. There were several reasons for separating the language instrument into different components. One of these is that the WCED needed to have an indication of learners' potential for successful engagement with the FET curriculum and beyond. In terms of educational assessment, this would, in part, be indicated by the academic literacy proficiency of learners in areas of the curriculum beyond the language learning area. Therefore an assessment which is limited to Home Language as a learning area on its own would provide insufficient evidence for any predictive assessment of learners beyond this learning area/subject. It is understood in contemporary literacy studies and the genre approach to applied linguistics (which has been extensively researched in Australia) that the language structures for science, mathematics, history and geography, for example, tend to be more cognitively complex than are normally encountered in language as a subject.

The second major reason for dividing the language instrument into two components was that it would accommodate each of the three main linguistic groups best. It would offer Afrikaans, English and Xhosa home language speakers an opportunity to be assessed in their home language and then in the language of learning and teaching which is either Afrikaans or English. Thirdly, this would comply with the international considerations of ensuring language accommodation in the interests of construct validity.

2.5.1.4 Planned Structure of the Instrument/s

The Language Learning Outcomes are as follows:

Design and Implementation

1. Listening
2. Speaking
3. Reading and Viewing
4. Writing
5. Thinking and Reasoning
6. Language Structure and Use

HSRC proposed that the language instrument include only four of these outcomes. For logistical and other reasons, it was not advisable to attempt to include listening and speaking in a two-hour pencil/pen-and-paper test. The WCED specifications made special reference to reading, and it is acknowledged in the research that the crucial area of language development in the senior phase is the development of academic literacy skills, and this implies a focus on reading, writing, thinking and reasoning. By implication this also includes a competent grasp of language structure and use. Thus the instrument would focus on the last four of the outcomes listed above.

Forty per cent of the instrument would be based on open-ended/constructed responses; 60% on multiple choice (MC) items. Although in the two previous years the Australian assessment agency had used MC items only, it is the theoretical and empirical experience of the HSRC that MC-only items do not offer accurate and sufficiently nuanced assessment of learner achievement. This is particularly the case in situations where there is prior evidence of low levels of literacy and reason to have educational concerns about levels of productive language proficiency. Constructed-response items, however, are costly and time-consuming to administer, score and code. Nevertheless, the kind of information which these offer are unlikely to be found in MC-only testing, particularly in settings of low levels of literacy proficiency.

The suggested weighting of the instrument would be as follows:

- Learning Outcome 3: Reading and Viewing: 30%
- Learning Outcome 4: Writing: 30 %
- Learning Outcome 5: Thinking and Reasoning: 30%
- Learning Outcome 6: Language Structure and Use: 10%

The weighting was arranged in this way because the instrument would gather additional information regarding learner competence in language structure and use, *en passant*, through the writing exercises, therefore it was unnecessary to weight this outcome more highly in specific measurement terms.

Multiple-choice questions were graded in order to provide additional information to the WCED in the event that some students did not reach Grade 8 level of achievement. It was envisaged that the items would be weighted according to the following formula.

Questions at Grade 6 level: $\pm$ 10%

Questions at Grade 7 level: $\pm$ 20%

Questions at Grade 8 level: $\pm$ 70%.

It was envisaged that it would be possible to grade the creative writing task according to different grade levels. However, subsequent to a pilot referred to below, and the rubric developed for the constructed response items, it was decided to refine the formula (see below).

2.5.1.5 Accommodation of Language Diversity

It is possible, through a new design construct, for a language instrument to assess learners' language abilities in terms of what they know of their own home language, in the home language; as well as how well they are able to make use of the LoLT across the curriculum. The instrument could be designed in two parts, therefore. Depending upon the administration of the instruments, learners might be presented with a specifically pre-selected language instrument, or under other agreed circumstances, learners might be given the choice of one HL version of the instrument and another LoLT section of the instrument. Whilst the majority of students would be likely to choose the same language for Part 1 and Part 2, Xhosa-speakers and even some Afrikaans-speakers who have English as LoLT might prefer to choose the option of Part 1 in HL rather than LoLT. The design portrayed in Tables 2.4 and 2.5 was therefore planned for the language instrument.

Table 2.4: Language Instrument Part 1: Home Language: Afrikaans, English, Xhosa

Afrikaans HL	English HL	Xhosa HL	Learning Outcomes	% of Test
Reading Extract: literature Comprehension questions including language structure	Reading Extract: literature Comprehension questions including language structure	Reading Extract: literature Comprehension questions including language structure	LO 3, 5, 6	25%
Creative Writing Exercise	Creative Writing Exercise	Creative Writing Exercise	LO 4, 6	25%

Table 2.5: Language Instrument Part 2: LoLT Afrikaans and English only

Afrikaans LoLT	English LoLT	Learning Outcomes	% of Test
Reading & viewing texts from other areas of the curriculum; interpretative Questions – thinking and reasoning	Reading & viewing texts from other areas of the curriculum; interpretative Questions – thinking and reasoning	LO 3, 5	35%
Writing tasks	Writing tasks	LO 4, 5, 6	15%

In consultation, however, with the WCED it was decided that learners would each only receive one consolidated HL and LoLT booklet, in the language combination as appropriate to the learner, and as informed by the CEMIS data collected and processed by the WCED after the 10th-day Survey at the beginning of 2006.

2.5.1.6 Framework for Language Instrument as Developed

The instruments were developed and piloted (see below) and the framework adjusted for finalisation as in Table 2.6. The HL reading extracts were sourced from literary texts which originated in each respective HL in order to eliminate any socio-cultural advantage of one group over another.

Table 2.6: Final language instrument structure for scoring, coding and data capturing purposes

Summary of Learning Outcomes: ALL LANGUAGES	No of items	Score Value
Part 1 Home Language Section A: MC items: DIFFERENT FOR THE 3 HL GROUPS LO 3, 5, 6 Owing to the 3 HL version design of this part of the instrument it was not possible to develop exactly the same number of items per LO at the same grade level, nevertheless there was sufficient overlap and particularly in relation to the rest of the instrument, for any differences to be minimized here.	18	18
Part 1 Section A: OEQs/CR* items: SIMILAR QUESTIONS BUT BASED ON DIFFERENT TEXTS LO 3, 4, 6 Learners have to read and interpret question correctly (LO 3), and then construct an appropriate answer using their writing skills (LO 4) and knowledge of language structure (LO 6). Emphasis on scoring & coding on LO 4&6. Score value reflects: LO 3 & 4. Diagnostic code reflects: L0 6	2	6
Subtotal	20	24
Summary of Learning Outcomes: ALL LANGUAGES	No of items	Score Value
Part 1 Section B: OEQs/CR items: SAME FOR ALL LANGUAGES LO 4, 6 Mark value reflects: LO 4. Diagnostic code reflects: LO 6	2	20 (2 X 5, doubled)
Subtotal	2	20
TOTAL PART 1: HOME LANGUAGE	**22**	**44**
Summary of LO and Grade Level items: SAME FOR ALL LANGUAGES **Part 2 LoLT MCQs**	No of items	Score Value 28 (1 per item)
LO 3 all MCQs test LO3	28	
LO 4	0	
LO 5 in addition, many test LO 5	23	
LO 6 in addition, several test LO 6	5	
Grade 6 level questions	2	
Grade 7 level questions	9	
Grade 8 level questions	17	
Subtotal	28	28
Summary of Learning Outcomes: ALL LANGUAGES **Part 2 OEQs**	No of items	Score Value
LO 3, 4 & 6 Learners have to read and interpret question correctly (LO3), and then construct an appropriate answer using their writing skills (LO 4) and knowledge of language structure (LO 6). Emphasis on scoring and coding is on LO 4 & 6. Score value reflects: LO3 & 4; Diagnostic code reflects: LO 6	5	12
TOTAL PART 2: LOLT SECTION of LANGUAGE INSTRUMENT	**33**	**40**
GRAND TOTAL: WHOLE LANGUAGE INSTRUMENT	**55**	**84**

* Constructed-response (CR) items are also sometimes referred to as free-response (FR/FRQ) or open-ended (OE/OEQ) questions

Summary of LO and Grade Level items: SAME FOR ALL LANGUAGES Part 2 MCQs	No of items	Score Value 28 (1 per item)
LO 3 all MCQs test LO3	28	
LO 4	0	
LO 5 in addition, many test LO 5	23	
LO 6 in addition, several test LO 6	5	
Grade 6 level questions	2	
Grade 7 level questions	9	
Grade 8 level questions	17	

By way of example, the LoLT item characteristics are presented in Table 2.7 in order to indicate the nature of the items included in the LoLT part of the instrument which was common to all students.

Table 2.7: LoLT - Item Characteristics MC and CR

Item	Learning Outcome and specific knowledge required	Grade level difficulty
1	LO 3: low inference	6
2	LO 3: low inference	6
3	LO 3: purpose of text; vocabulary LO 6: use of vocabulary	7
4	LO 3: following contextual clues LO 5: separating cause & effect	7
5	LO 3: following contextual clues; LO 5: inference & deduction	8
6	LO 3: identifying purpose; LO 5: weighs options	8
7	LO 3: makes judgements; LO 5: processes information	8
8	LO 3: draws conclusions; LO 5: processes information	8
9 CR/FR	LO 4: independently generates text; LO 6: critically reflects own writing	8
10	LO 3: identifying text, narrator/voice	7
11	LO 3: demonstrating understanding; ethical issues; LO 5: inference LO 6: vocabulary	8
12	LO 3: theme & character; LO 5: expresses opinions	7
13	LO 3: draws conclusions on basis of evidence; LO 5: inference & deduction	8
14	LO 3: theme & character; writer’s intention; LO 5: inference & deduction	8
15	LO 3: making judgements, drawing conclusions; LO 5: supports arguments with evidence	8
16	LO 3: identifying text, processes information.	8
17	LO 3: character; LO 5: weighs options;	8
18	LO 3/5: making inference; LO 5: analysing cause & effect; LO 6: increasingly complex text.	8
19	LO 3/5: inferring & deducing meaning;	7
20	LO 3: literary devices; LO 5: uses metalanguage/figures of speech	8
21 CR/FR	LO 3: understands idiomatic expression LO 6: knowledge of parts of speech [production of appropriate lexical items only]	6 7

Item	Learning Outcome and specific knowledge required	Grade level difficulty
22 CR/FR	LO 3: inference LO 4: independent writing LO 6: increasingly complex text	8
23	LO 3: finding details LO 5: processes information, weighs options	8
24	LO 3: finding details LO 5: finding cause & effect	7
25	LO 3: reading for detail; LO 5:makes inferences & deductions; weighs options	8
26	LO 3: finding information LO 5: makes inferences & deduces meaning	7
27	LO 3/5: inferring & deducing meaning; LO 6: vocabulary	8
28 CR/FR	LO 3: interprets text; LO 4 & 6: organises ideas coherently, proofreads, corrects drafts, applies knowledge of language at different levels	8
29	LO 3: understands text; LO5: inferring & deducing meaning	8
30	LO 3: reading for detail; LO 5: inference, understanding cause & effect	7
31	LO 3: understanding purpose and reading for detail; LO 5: understanding cause & effect; inference	7
32	LO 3: reading for detail LO 5: inferring meaning LO 6: unpacking language structure/increasingly complex text	8
33 CR/FR	LO 3: applies language knowledge; LO 4: independently generates text; LO 6: works with sentences; punctuation & connectives; critically reflects on own writing	8

2.5.1.7 Diagnostic information

It is important that the data provided by the WCED study should not only be seen as a tool for constructing reporting scales but that the rich source of information constituted in the data should be explored to improve learning and teaching in languages as ‘subjects’ and also in terms of teaching and developing strong language expertise across the curriculum. The HSRC developed a two-digit diagnostic score-code for language, following the example of the TIMSS study and used for the mathematics instrument, discussed below.

By way of example, and in regard to the extended writing / constructed responses (the common items) in the Home Language Part of the instrument, the rubric reported in Table 2.8 was employed.

Although scored out of 5 (and then multiplied by 2, i.e. each was valued at 10 marks for data capturing purposes), the instrument had each item valued at 8 marks on the paper. The discrepancy arose owing to last-minute scoring difficulties and a re-weighting of items. The second digit, the diagnostic code remained the same.

Table 2.8: PART 1, Section B: Extended Writing/Constructed Response: Same for all language versions

First Digit: Value (content)	LO 4: Writing Skills	Second Digit: Diagnostic Code (structure)	LO 6: Knowledge and Use of Language Structure
5: Exceptional	Clear, consistent, logical writing skills	5:	No/minimal errors
4: Above average	Argument generally well presented	4:	Minor errors: spelling/punctuation
3: Average		3:	Minor errors – sentence structure
2: Below average	Lacks coherence	2:	Poor sequence of ideas/logic sentence level
1: Most unsatisfactory	Did not understand the question / completely off the point / completely inappropriate content and register; incoherent	1:	Combination of 2,3,4; or very serious errors In many cases this also indicates learners with extremely low literacy skills. These learners require urgent attention.
99	**Not answered/not attempted**		
98	**Answer attempted and deleted**		
97	**Illegible**		Disturbing signs of inability to write (lack of literacy skills: cannot form recognisable letters or join these together in recognisable words) Learners should be assessed as soon as possible.
96 *	**Rewrote the question or rewrote a question from somewhere else in the text**		Disturbing signs of lack of ability to read and make meaning from text, uses this strategy to hide/disguise lack of literacy.

Whereas, the codes 99, 98 and 97 had been anticipated, what needed to be added was code 96, which had not been anticipated by the provider.

2.5.2 Instrument design: Mathematics

2.5.2.1 Validity

Any assessment instrument must reflect the construct it intends to measure in a valid way. Test scores are valid if they measure what they are supposed to measure. Valid test scores also have small errors of measurement and are of high precision. Validity, therefore, assumes high levels of reliability too. The most important form of validity for scholastic achievement assessment is content validity. To have content validity an instrument should assess all aspects of the curriculum appropriately. The requirements of the NCS and the document of WCED: *Learning Programme Exemplars for Mathematics Senior Phase Grades 7-9 (March 2005)* were used as the basis to determine content validity for the mathematics instrument.

The NCS (DoE 2002b) consolidates the unique features and scope of the Mathematics Learning Area into five Learning Outcomes: Numbers, Operations and Relationships (LO1), Patterns, Functions and Algebra (LO2), Shape and Space (Geometry) (LO3), Measurement (LO4), and Data Handling (LO5). These Learning Outcomes and their Assessment Standards are cognitively dependent on and supportive of each other. The content dimension of the TIMSS framework reflects the Learning Outcomes structured in the NCS (DoE, 2002b) in such a way that comparisons of achievement in the content domains are possible. The NCS was therefore used to develop the framework for the Mathematics instrument in the same format as the TIMSS Framework.

The questions in the mathematics instrument were also classified according to four cognitive domains. The cognitive domains define the behaviours expected from learners as they engage with mathematics content. The following paragraphs define general capabilities expected from learners for the different cognitive domains:

Knowing Facts and Procedures. The ability to perform mathematics depends on the ability to recall learnt mathematical knowledge. The more knowledge a learner is able to recall, the greater the potential to engage in a wide range of problem-solving situations. Facts encompass the factual knowledge that provides the basic language for mathematics, and procedures form a bridge between basic knowledge and the use of mathematics for solving routine problems.

Using Concepts. Using concepts is the process that enables learners to make extensions beyond their existing knowledge to make connections between elements of knowledge that would otherwise be retained as isolated facts.

Solving Routine Problems. Problem solving is a central aim of teaching school mathematics. In items categorized in this domain, the problem setting is more routine than those aligned with the reasoning domain. The routine problems will have been standard in classroom exercises designed to provide practice in particular methods or techniques.

Reasoning. Reasoning mathematically involves the capacity for logical, systematic thinking. It includes intuitive and inductive reasoning based on patterns and regularities that can be used to solve non-routine problems or problems that are unfamiliar to learners.

The framework was developed through an iterative process between the WCED and the HSRC. Target percentages of questions across the content and cognitive domains were established. This was done in close consultation with the WCED with regard to the Assessment Standards to be covered in each of the five Learning Outcomes as well as the percentage coverage of each of the Learning Outcomes.

2.5.2.2 Language accommodation in the instrument

Despite the practice of using only two languages for LoLT, the HSRC was concerned about the validity of instruments which measure achievement in mathematics and language where the majority of students will be assessed in their home language (L1) and a minority in their first additional / second language (L2 – for about 22%-24% learners). Evidence from international research, and domestic research in which the HSRC has been intimately involved (e.g. the Threshold Report, Macdonald 1990; the Grade 6 Systemic Study, DoE 2005; the QLP study, Kanjee and Prinsloo, 2005 and Prinsloo and Kanjee, 2005), shows that students who use their First Additional Language as LoLT, are likely to perform at between 15 to 20 percentage points below those who use their HL/L1 as LoLT.

For this very reason, the national Minister of Education, as early as 2005, initiated research into the viability of translating some examinations into African languages. The risk to both the WCED and to the HSRC in terms of a Grade 8 assessment in the most recent context was that the Department and instruments might face adverse criticism in regard to the validity of assessment instruments if they had not taken this factor into account, or modified the instruments to accommodate students from different linguistic backgrounds. A creative response to such a challenge within the current environment was therefore advisable.

International examples of language accommodation in assessment across the curriculum
A common observation amongst educators and learners alike is that students whose home language is not the same as LoLT, find it difficult to understand the instructions embedded in examination questions.

The Ontario Ministry of Education, for example, insists on ensuring that second-language speakers of English are ‘accommodated’ through one of three strategies: use of dictionaries for translation during the examination; translated directions for test items; or additional time for writing the examination. Similar procedures are offered in South Australia, New Zealand and the UK. A strategy used in the final matriculation examinations in South Africa has been to add a compensatory mark of 5% for African language speaking students in the content subjects. This compensatory measure is not regarded as pedagogically sound. It has been criticised on theoretical grounds for disguising what it is that students do not know and what it is that students have gained from the school system. It was therefore not recommended for this study.

The HSRC recommended a strategy that would limit risk and offer new analysis which may place the WCED at a competitive advantage in relation to other provinces, and even internationally. In a nutshell, the strategy was to make use of explanations of key instructions in Xhosa for a limited number of items (originally 20%) in the mathematics instrument.[5]

In this way the HSRC, in agreement with the WCED, was to use the opportunity to test the assumptions about the possible negative effect of language mismatches in a limited number of test items. The original proposal suggested language adaptation in the mathematics instrument for second language (mainly Xhosa)-speaking learners by including a brief glossary of the instruction in Xhosa for 20% of the items. The idea was further refined in consultation with the WCED. The number of glossed items was increased to approximately 40% and the translation was extended to Xhosa and Afrikaans in the English version of the test and Xhosa and English in the Afrikaans version of the test. These glossed questions were dispersed randomly through the test paper. The results of these questions were compared with a group of similar questions which were not accompanied by the glossed instruction. This allowed a comparative analysis to determine whether or not explanations of instructions have any positive effect on the ability of Xhosa (second-language)-speaking students to answer questions correctly. At the end of the mathematics instrument learners were asked to give their opinion of whether they benefited from the glossed instructions or not. It was hoped that this could provide evidence which either supports or negates the assumption that students, whose home language is not the same as LoLT, find it difficult to understand the instructions embedded in test questions.

2.5.2.3 Framework of the instrument

The test was constructed using the five learning outcomes that are indicated in the NCS for mathematics and the four cognitive domains. The percentage coverage of each area was negotiated with the WCED. Questions were first selected from the TIMSS released question pool. In the areas were the TIMSS questions were not suitable for Western Cape learners or did not ensure adequate curriculum coverage, new questions were developed. The development and editing of the mathematics instrument was marked by extensive engagement with documents relevant to Grade 8 mathematics and TIMSS 2003 as well as ample consultation with officials from the WCED.

Learners’ knowledge and understanding of mathematics were assessed through multiple-choice questions and constructed-response questions. Multiple-choice questions required from them to

[5] The HSRC had monitored an innovative experimental process initiated by the school principal of Zonnebloem Nest College school during 2005, where he had offered Grade 11 learners translated and or glossed terminology in their mathematics examinations in June and November. It was this experiment which was largely responsible for the design of this feature of the mathematics instrument, in conjunction with the literature on language accommodation practices found in Canada and the USA. Should this process be used more widely, it would be important to acknowledge the contribution of Mr Jonty Damsel in this regard.

select their answers from a set of four options included in the instruments. In constructed-response questions learners were required to construct a written response. Constructed-response questions are particularly suited for assessing aspects of knowledge and skills that require learners to explain phenomena or interpret data based on their background knowledge and experience.

The mathematics instrument comprised a test of 54 questions that were scored out of a total of 56 marks. The distribution of mathematics items according to the Learning Outcomes is set out in Table 2.9 and illustrated in Figure 2.1.

Table 2.9: Classification of questions per content area in the mathematics instrument

Learning Outcome (LO)	**No of items**	**Max score**
LO1: Numbers, Operations and Relationships	14	15
LO2: Patterns, Functions and Algebra	12	12
LO3: Space and Shape	10	10
LO4: Measurement	11	11
LO5: Data Handling	7	8
Total	**54**	**56**

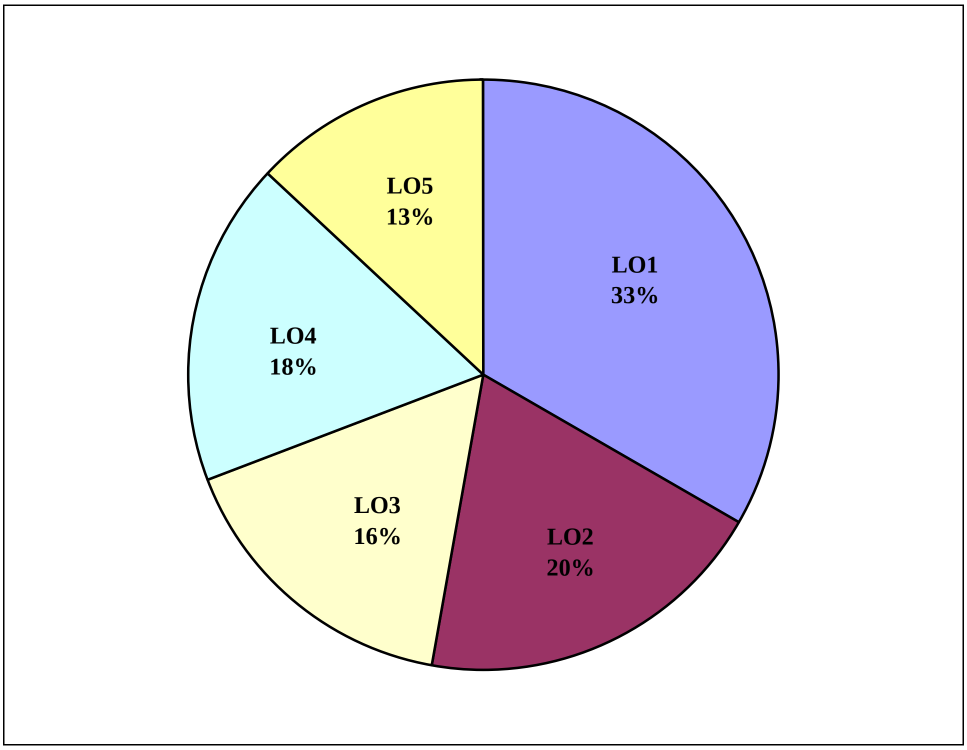


Figure 2.1: Classification of the questions per content in the mathematics instrument

In the mathematics instrument, LO1 comprises 37% of the items, while both LO2 and LO3 comprise 14.5%, and each of LO4 and LO5 comprises 17% of the total number of items. The number of items reflecting each learning outcome is adequate for obtaining relevant information on how learners perform, although additional items for Learning Outcomes 2, 3, 4 and 5 would have made for a more thorough assessment. Despite these apparent constraints, the collection of items per learning area does provide a comprehensive overview of what learners could and could not do in the curriculum.

The distribution of items across learning outcomes for each learning area is reflected in Table 2.10 and Figure 2.2.

Table 2.10: Classification of questions per cognitive domain in the mathematics instrument

Cognitive Domains	No of items	Max score
Knowing Facts and Procedures	15	15
Using Concepts	12	13
Solving Routine Problems	19	19
Reasoning	8	9
Total	**54**	**56**

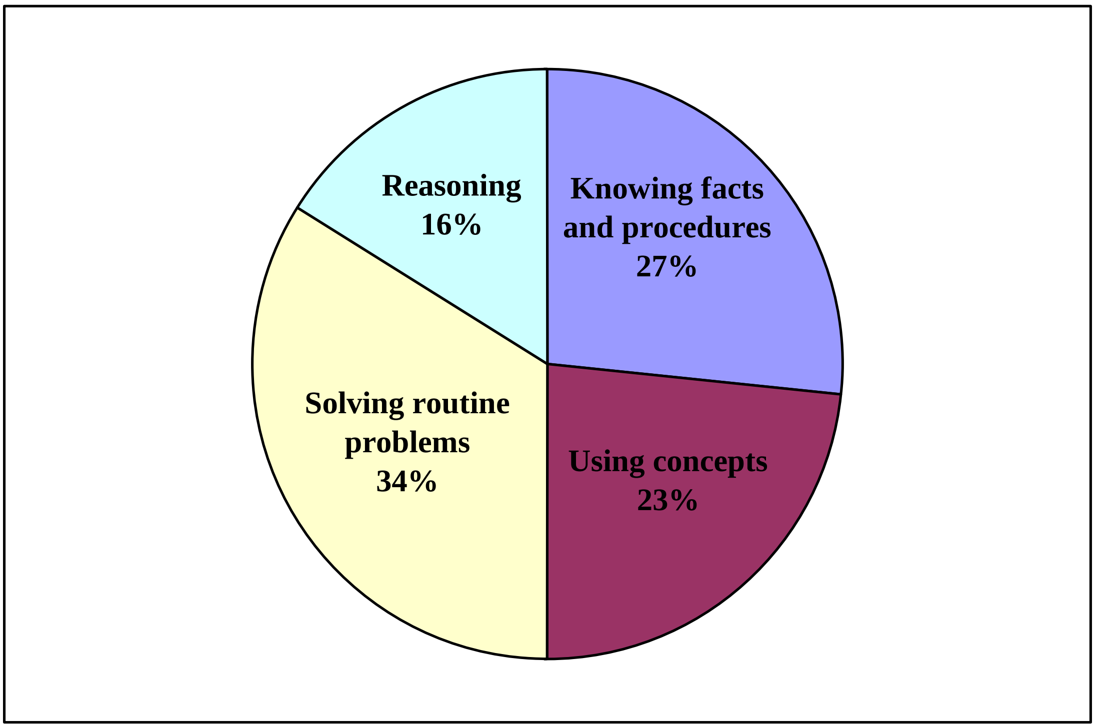


Figure 2.2: Classification of the mathematics questions per cognitive domain

In mathematics the cognitive domain *Knowing Facts and Procedures* comprises 27% of the items, *Using Concepts* comprises 23%, *Solving Routine Problems* comprises 34% and *Reasoning* comprises 16% of the total number of items. The number of items reflecting each cognitive domain is adequate for obtaining relevant information on how learners perform on each cognitive level, although additional items for *Reasoning* would have provided a more thorough assessment.

The questions in the instrument covered both closed (multiple-choice) and open-ended types of questions. Thirty-five of the questions were of multiple-choice type (consisting of a question and four options for an answer; - three distracters and the correct answer) and 19 were constructed-response items (where learners were required to show their working methodology to get to an answer). A further breakdown by item type and learning outcome is provided in Table 2.11.

Table 2.11: Number of multiple-choice and constructed-response questions

Type	Number of items and scores per LO									
	LO1		LO2		LO3		LO4		LO5	
	No of items	Max score	No of items	Max score	No of items	Max score	No of items	Max score	No of items	Max score
Multiple-choice questions	10	10	7	7	7	7	7	7	4	4
Free-response questions	4	5	5	5	3	3	4	4	3	4
Total per LO	14	15	12	12	10	10	11	11	7	8

2.5.2.4 Diagnostic information

It is important that the data provided by the Western Cape study should not only be seen as a tool for constructing reporting scales but that the rich source of information constituted in the data should be explored to improve learning and teaching in mathematics. Data on individual items constitute a rich source of information, not only on the dimension of right/wrong answers provided but also on the diagnostic aspects of which right/wrong responses students actually gave. General trends of what learners can and cannot do could be of value in itself and some of the specific errors identified might be of practical use in teaching. If educators know the proportion of learners making a given error they will know if they are likely to meet the error in their teaching process and how often they can expect to meet it.

Young learners come to the classroom with theories constructed from their everyday and educational experiences. These experiences constitute the foundation for future learning. Similarly, mathematics educators know from experience that simply explaining the same topic over and over again in the same way or assigning additional exercises would not necessarily improve learners' performance in mathematics. Many learners have incorrect conceptions in mathematics that interfere with learning. Simply re-teaching a topic will not help some learners to give up their incorrect conceptions. Since learners actively construct their knowledge, teachers must help them to break down their incorrect conceptions and help them to reconstruct conceptions which facilitate future learning.

By analyzing the common errors made by learners in the answers given to questions in the mathematics test and by identifying misconceptions that may be revealed by these errors, the research will attempt to supply information that can inform the teaching process. The information may help educators to dismantle misconceptions and may result in the prevention of learning problems in mathematics.

Literature related to learning and misconceptions in mathematics (and science) revealed that a great deal of research in mathematics teaching has focused on the investigation of errors made by learners. Many recent research projects on the investigation of errors show that researchers as well as educators in the mathematical field are giving more and more consideration to the fact that learners' errors and misconceptions can be the starting point to effective mathematics teaching (Williams & Ryan 2000). Research done on "error analysis" in different areas of mathematics shows that errors made by learners can be a powerful tool to diagnose learning difficulties and consequently provide valuable contributions to mathematic education. Nesher (1987) believes that a good instructional programme will have to predict types of errors and purposely allow for them in the process of learning. Nesher claims that the road to a state of expertise is paved with errors and misconceptions and that each error has the potential to become a significant milestone in learning.

A number of reports and articles acknowledge the fact that learners benefit from instruction when mistakes, misunderstandings and errors[6] are analysed interactively and faulty reasoning is defined and corrected, or when improper strategies are replaced with more correct ones (Leinwand 1998; Williams & Ryan 2000). However, many educators do not make use of the fact that errors contain valuable information about specific difficulties that learners encounter in mathematics. It would appear that many educators are still unaware of the potential of using diagnostic analysis of errors to guide their teaching methods. Errors may appear to teachers as random mistakes, leading them to overlook the value of reducing the sheer volume of available information into meaningful diagnostics (e.g. the student does not know to borrow, the student is having trouble with zeros).

In the report on the Third International Mathematics and Science Study, Beaton *et al.* (1996) relate the success of Japanese children in mathematics learning to the fact that Japanese mathematics teaching typically makes use of a diagnostic approach. Japanese teachers prepare themselves with a variety of possible responses to a key lead question. They receive guidance as to the thinking these responses indicate and how to teach constructively according to each possible response.

[6] The terms error and mistake and misconception are all used in the literature to indicate an incompleteness or inaccuracy. Errors and mistakes are usually treated as synonyms and indicate sporadic, unconnected and non-systematic errors while misconceptions denotes a line of thinking that causes a series of errors all resulting from incorrect underlying premises, rather than sporadic, unconnected and non-systematic errors.

Design and Implementation

Dossey, Jones and Marin (2002) highlighted the value of analyzing students' responses using two-digit rubrics to provide information on the correctness of responses and methods used by learners, and on misconceptions noted in their work. They make use of several examples to illustrate the enormous amount of information that can be obtained by analyzing the data obtained from the use of the two-digit rubrics scoring system that was developed to score constructed-response items in the TIMSS assessments.

On the exploration of students responses to constructed-response items in TIMSS Angell, Kjaernsli & Lie (2000) described the achievement data in TIMSS and similar projects as significantly more than merely tools for constructing reporting scales. They refer to the constructivist view of learning according to which students learn by constructing their own knowledge and argue that it is of crucial importance that teachers should be aware of students' preconceptions within a topic prior to instruction. The data illustrate that valuable insight into learners' way of thinking can be obtained by analyzing responses to (science) items based on the coding rubrics developed for TIMSS.

This study gives expression to the findings in the existing literature that the analysis of errors provides enriched insight into learners' thinking, their conceptual understanding and the nature of their errors and misconceptions. Errors made in these items were investigated to determine if the items reveal any misconceptions amongst learners. By interpreting the patterns of errors, misconceptions and predominant solution methods of learners we tried to provide some insight into learners' ways of thinking and why they made certain errors.

2.6 PILOT ASSESSMENT INSTRUMENTS

Instrument items were pre-tested in a sample of schools during the last week of March 2006. Before piloting, subject (learning area) specialists in the Western Cape reviewed the instruments and questions for their suitability in terms of difficulty level, content and construct validity. The pilot instruments contained more items than would be required or used in the main study in June. This was in order that weak or problematic items could be discarded after the data analysis of the pilot study had been completed.

The instruments were piloted with Grade 9 learners in seven schools from Western Cape (about 300 learners per learning area) to gather further information on the suitability of the questions. Grade 9 learners were targeted for the pilot because the pilot was done early in the year when very little of the Grade 9 curriculum would have been covered by the Grade 9 learners. Therefore learners were essentially being assessed on Grade 8 material and at that level. During the pilot application subject specialist were requested to comment on the difficulty of the questions for Grade 8 learners in mid-year.

Pre-testing benefited the process in two ways. In the first place the process of instrument administration was evaluated and in the second place item analysis was carried out to provide information on possible weaknesses in the items and to provide additional information on the discrimination value, complexity and suitability of the items.

The administration manuals (for language and mathematics) were drafted alongside the instruments as they were being developed. These were also piloted along with the instrument and redrafted accordingly.

Because the contractual arrangements did not provide for or permit HSRC to monitor the administration process or train the assessment administrators, the administration required a simple, yet complete instruction manual. However, even if a manual is ‘foolproof’ it cannot take into account all the circumstances that can arise during the administration of a test.

Initial intentions to require administrators in the final study to complete a report sheet and so sign off on the administration, recording aspects such as absenteeism, latecoming, completion times, learner queries, or any other incidents related to suitability of venues, or whatever, had to be abandoned for logistical and other reasons.

Apart from the administration manuals, the WCED published a circular to each school in the province to alert school managements and teachers to the dates of the assessment tasks, and the nature and administration requirements of the tasks.

2.7 ADMINISTRATION OF THE FINAL INSTRUMENTS – WCED RESPONSIBILITIES

2.7.1 Learner Information Data collection

In terms of the contract, WCED was required to collect and prepare electronic data sets and make these available to the Contractor. These were to be prepared in terms of: electronic data per learner, per school, per EMDC, and per teacher if required for further analysis. The service provider required the WCED to provide a ‘list of learners’ to the contractor. The ‘list of learners’ refers to an up-to-date, correct and reliable list of learners by class for every school which would be involved in this assessment activity. The reasons for this are:

- The lists of learners would be used to calculate the required number of answer books per school.
- The lists of learners would be printed out by the service provider in the form of control sheets which would account for each learner, in each class in each school.
- The control sheets for each class would include a unique number for each learner which had to be written on the assessment instruments (answer sheets) of each learner for purposes of linking the information from the two assessment tasks automatically for each learner. (This arrangement was replaced at a further stage by WCED pre-printing and affixing bar-coded lables (stickers) to each assessment task booklet.)

The lists of learners from the WCED were to be provided in Excel or Access. These lists were to include the following:

- full names of the learner
- the gender of the learner
- the age/birthdate of the learner
- the home language/s of the learner.

2.7.2 Printing, batching and despatching of instruments

Based on the biographic data collected from the learners, the instrument booklets were to be printed according to the Home Language selected by the learner for study in Grade 8.[7]

2.7.3 Administration of actual instruments

Administration of the instruments included at least one circular sent to schools by the WCED to inform them of the dates, procedures to be followed, and the different versions of the instruments. HSRC-prepared administration or instruction sheets were included with the instrument booklets so that each teacher would receive one and be in a position to follow the instructions carefully. A significant number of instruments did not have bar-coded IDs attached to them. However, a backup arrangement for manual recording of student names and some additional biographical information was in place. (The latter was designated for completion by students in the mathematics instrument,

[7] Operational and administrative circumstances resulted in some degree of mismatch between test language, home language and language of instruction, which may have limited the breadth of reporting for some learners, but also provided an unexpected opportunity to study additional interactions. As a result, the relevant subgroup reports and performance subscales were retained as appropriate.

but eventually collected through the language instrument. The intended broader set of items requesting biographic information was replaced by brief questions on the numbers of books in each learner's home, and their home languages.)

2.7.4 Retrieval of instruments and preparation for handover to HSRC

WCED retrieved the instruments and the bar-coded batches were recorded per school and learner in the WCED offices. The boxes of scripts not bar-coded were separated and marked for manual processing.

2.8 DATA COLLECTION, SCORING, CODING, AND ANALYSIS – HSRC RESPONSIBILITIES

2.8.1 Data entry and cleaning

The HSRC recruited, appointed and trained WCED teachers to code the mathematics and language tests according to coding rubrics designed by the HSRC. The HSRC researchers designed a two-digit coding rubric to assess the constructed-response questions. The coding rubric was modified in the event of unexpected answers, in cases where the test developers had not anticipated these. HSRC researchers managed, moderated and monitored the scoring process which only covered constructed-response (open-ended) items.

After coding the learners' answers to the constructed-response questions, the scripts were batched and sent to Gauteng for data capturing. The two data-capturing agencies (separate ones for mathematics and language to allow for parallel processing and risk management) were briefed in detail about the procedures and standards for capturing, and monitored on a regular basis, especially at the outset. At this stage, all written text, the information on the bar-coded labels, the coded constructed-response items, the item responses on mulitiple-choice questions, and the few responses to biographic information were captured.

All scripts were captured twice (100% verification of all items). Subsequent data cleaning then served to ensure the integrity of the dataset. The data-cleaning process had to confirm that all data fell within pre-specified formats, field lengths and values provided by the HSRC and had to eliminate any inconsistencies, double records, errors and other anomalies picked up through the digital edit process.

Design and Implementation

This extremely large data-cleaning task, but also subsequent scoring, analyses and reporting outputs, was automated wherever possible to avoid any handling of individual records or layouts of outputs and other documents.

In short, the process comprised first digitally matching the language and mathematics booklets / data according to the CEMIS-based bar-codes. Then, unmatched mathematics learners were allocated a unique number through electronic processes. From this point on, these learners' language data / booklets had to be retrieved and numbered manually before capturing.

At the end any further remaining queries had to be solved completely manually and one by one. In just over a 1 000 cases, uncertainties as to the identity of learners resulted in their mathematics and language records and reports being kept separately for schools to solve, rather than assuming identity. In addition, in the case of a small number of about 280 learners, mainly Xhosa-speaking, the assessment tasks were completed partly or in whole in two booklets, being English and Xhosa. These are being solved last, but chances are that these learners should not have completed English HL sections or booklets.

Table 2.12 gives an indication of the complexity and extent of the data management tasks pertaining to large-scale full-population studies of this kind.

Table 2.12: Sub-group sizes for the Grade 8 learner datasets for language and mathematics

Sub-group	Total number of records	Language of the Language / Maths test booklets #	Home language of learner (derived from CEMIS and elsewhere)				
			Afrikaans	English	Xhosa	Other	Missing info
1	35 257	A / A	28 120	300	351	35	6 451
2	18 713	E / E	1 261	10 121	1 757	206	5 368
3	131	A / E	28	26	26	1	50
4	144	E / A	49	39	10	0	46
5	97	X / A	16	3	56	1	21
6	8 110	X / E	48	103	5 684	53	2 222
7	938	A / o	637	12	14	1	274
8	869	E / o	63	449	128	10	219
9	672	X / o	0	6	417	3	246
10	3 727	o / A	0	0	0	0	3 727
11	5 421	o / E	0	0	0	0	5 421
12	2 788	X-HL xE / (E+A) *	0	0	2 706	82	0
13	1 069	?-HL xE / (E+A) *	0	0	0	0	1 069
Totals	**77 936**		**30 222**	**11 059**	**11 149**	**392**	**25 114**

A=Afrikaans; E=English; X=Xhosa; o=not completed/matched; xE=unmatched (English) L1 booklet completion by non-HL learner.

* For sub-group 12, 67 mathematics booklets were completed in Afrikaans, and 2 415 in English. For sub-group 13, 7 mathematics booklets were completed in Afrikaans, and 970 in English.

A range of additional frequencies, not immediately evident from Table 2.12, can be derived from it:

- Total number of language assessment booklets: 68 788 (sum from sub-groups 1 to 9, plus 12-13)
  - 36 326 completed in Afrikaans (from sub-groups 1, 3 and 7)
  - 19 726 completed in English (from sub-groups 2, 4 and 8)
  - 8 879 completed in Xhosa (from sub-groups 5, 6 and 9)
- 40 053 report cards in Afrikaans (from the three language sub-groups mentioned in the first sub-bullet immediately above, plus mathematics sub-group 10)
- 25 147 report cards in English (from the three language sub-groups mentioned in the second sub-bullet immediately above, plus mathematics sub-group 11)
- 8 879 report cards in Xhosa (from the three language sub-groups mentioned in the third sub-bullet immediately above)
- 2 788 and 1 069 semi-complete report cards (only LOLT sections) for the Xhosa and unknown HL groups having completed English HL (L1) booklets (from sub-groups 12 and 13)
- Total number of mathematics assessment booklets: 75 059 (sum from sub-groups 1 to 6, plus those from 10-13, with the numbers for sub-groups 12 and 13 according to the note below Table 2.12)
  - 39 299 completed in Afrikaans (from sub-groups 1, 4, 5, 10, 12 and 13)
  - 35 760 completed in English (from sub-groups 2, 3, 6, 11, 12 and 13).

2.8.2 Data scoring

Data processing and analyses were conducted using a range of software according to need, and included SPSS, SAS, MS Excel and MS Access.

The first step comprised converting item responses into item scores for the multiple-choice questions according to the position of the correct response and the options selected by learners. This can be considered as an electronic process for marking every remaining item. The various item scores for both multiple-choice and constructed-response items got added together next in iterative procedures to provide the many sub-totals and totals required by item type, item content, learning area, and the various other features discussed up to this point. This step underpinned all the subsequent data analysis, generation of report cards, certificates and tables, and the reporting of the findings in Chapter 3.

The number of different responses for each question would also be obtained by means of frequency tables to provide more detailed interpretations of the learners' responses. This information underpins important aspects of the rest of this report, and the feedback to the educators on the type of mistakes made by learners, which will help them in their interventions in the classrooms. The first digit of the scores for the constructed-response questions indicated the correctness of the learner's response and the second digit diagnosed the error made by the learner.

Specific incorrect responses selected by learners for (some of) the multiple-choice items would also provide information on the types of mistakes made by learners. Misconceptions of patterns of learners' understanding were later identified and interpreted by the researchers to inform the teaching and learning process.

2.8.3 Data analysis, reporting and interpretation of results

Without going into any detail at this point, the two main tasks were to describe the performance outcomes of the learners for the language and mathematics assessment tasks in global terms, and to assess if any achievement differences had been associated with belonging to different sub-groups, for example in terms of sex, EMDC, or any biographical variables. The latter aspect is dealt with in Chapter 3 at an overall level for the whole province, and otherwise where meaningful.

The first task above, in addition to Chapter 3, is also covered in quite some detail in the additional tables provided separately to the Department. The mean performance scores of learners for between 11 and 15 variables were provided per learning area at the provincial level and broken down for each separate EMDC. Next, these same mean scores were provided at the EMDC level and broken down into the 50 to 80 schools in each EMDC. These tables, read jointly with the main report, especially Chapter 3, provide EMDCs with detailed information about their relative performance, and that of the schools under their jurisdiction. A third and final set of tables comprises a direct summary of selected performance scores for each learner for every school.

An individual learner report card was also prepared and provided to the Department for four selected fields per learning area (language and mathematics). Each field comprised the learner's performance score (%) and level (on the 7-point achievement scale), with an appropriate diagnostic and forward-looking career- or study-prospect comment, against the relevant sub-group and/or provincial averages as relevant. Because of the sheer volume of this task, it also had to be fully automated to avoid individual handling. A range of transformations from SPSS datasets through

Excel speadsheets to Access databases had to be achieved for this. A similar process, albeit with fewer fields and comments generated, was followed for the preparation and issuing of individual learner certificates.

2.8.3.1 Interpretation of learner-assessment results

For the learner-assessment component, mean score percentages as well as standard errors and standard deviations were used to report the results for each of the assessment instruments. Scores were reported using either tables or graphs and were aggregated by province. As mentioned above, additional detail is provided in the tables for EMDCs and schools. In reporting results, percentages were usually rounded to whole numbers. In some instances, therefore, totals may not necessarily add up to 100%.

In interpreting the performance of learners on different learning outcomes and at an overall level, the following DoE (2005) seven-point scale was used:

Rating code	Rating	Marks in %
7	Outstanding achievement	80 - 100
6	Meritorious achievement	70 - 79
5	Substantial achievement	60 - 69
4	Adequate achievement	50 - 59
3	Moderate achievement	40 - 49
2	Elementary achievement	30 - 39
1	Not achieved	0 - 29

The emphasis in reporting the performance of learners on the assessment tasks has been on what learners know and can do. Learners' mathematical and language achievement has been reported on the instrument as a whole as well as on the separate Learning Outcomes. Language achievement has been reported in terms of HL and LoLT and in terms of:

- EMDC,
- item type (MC or CR items),
- learning outcomes,
- differentiated grade level achievement in MC items,
- gender, and
- number of books in the home.

Mathematical achievement has been reported in terms of:

- EMDC,
- item type (MC or CR items),

Design and Implementation

- learning outcomes,
- cognitive domains,
- gender,
- number of books in the home, and
- home language (HL) and language of learning and teaching (LoLT).

Preliminary analyses (not always reported in all detail, for lack of space) mainly comprised:

- Descriptive statistics (frequencies, means, standard deviations) of items and variables, overall and by sub-group;
- Item analysis (item-total correlations, and the rank-orders and distributions of item-difficulty values) to collect reliability and validity evidence;
- Equivalency comparisons by home language, gender, and other relevant biographical variables, by means of chi-square analyses of contingency tables, and comparisons of means through analysis of variance (ANOVA);
- The calculation and comparison of item-difficulty values and cumulative percentages, to enable comparisons with the benchmarked instruments;
- The calculation of individual scores per learner, and mean scores per school and EMDC, and overall for the province, by item type, learning outcome and other relevant sub-scales, in terms of overall marks and/or percentage conversions, as relevant.

At the learner level, the conversion of relevant information to learner reports ensured that not only pass / fail information is given in crude and unnuanced ways, but a qualitative understanding is given of the types of mistakes made by learners, and their relative areas of strength and weakness in terms of learning outcomes, and type of performance required.

2.9 LIMITATIONS OF THE REPORT

This report only presents an overview of general trends within the education system of Grade 8 learners in the Western Cape. It only provides broad indications of "areas of concern" or areas "in need of further attention". Further research and investigation will be required for more comprehensive or detailed information, especially pertaining to the factors associated with enhanced or hindered performance. By definition, one-off cross-sectional studies cannot address such concerns satisfactorily.

Design and Implementation

The use and interpretation of the information in this report should be understood in the context of the following limitations:

- The information was gathered using a pen/pencil-and-paper mode of assessment. The report is therefore limited to competencies (assessment standards) that lend themselves to be assessed in written form. It is generally considered more appropriate to conduct purposive sample assessments at a deeper level in order to arrive at the most accurate and informative data.

- The limited time that learners had to complete the assessment tasks (80 minutes, with an additional 10 minutes where learners were attempting to answer the tasks in a language other than the home language) dictated that a limited number of questions could be included in the instruments. Some competencies that require direct observation and interaction with learners can only be assessed at site level as part of continuous assessment. The report can therefore neither be comprehensive nor conclusive in its coverage of all the competencies that learners are expected to achieve in either language as a learning area or mathematics at Grade 8 level. The report only provides a "snapshot" of the competencies that learners could demonstrate in a written form.

- The role of item type should be noted when scores are interpreted in a criterion-referenced context. In an item with four answer options, the percentage of correct answers does not represent the same degree of achievement as would be implied if the same percentage of learners correctly answered a constructed-response item. It is usually easier for a learner to select the correct answer from a selection of options in a multiple-choice item than to produce an answer for a constructed-response question. Furthermore, learners may guess the answer or use guessing strategies to select either A, B, C or D options. In some studies learners are believed to be able to guess between 20% and 25% of items as correct. Therefore where learners exhibit very low levels of proficiency in constructed-response items but achieve 25% or above in the MC items, one must assume that there is a likelihood of guessing in the MC items. Despite the limitations of multiple-choice questions in an assessment instrument, resource constraints (time, financial and human resources) left only this as a viable option for the purposes of a large-scale survey.

- Educators were not requested to provide detailed information on the Learning Outcomes and Assessment Standards actually covered up to the mid-year. The fact that educators may have concentrated on certain sections of the curriculum up to this stage of the school year may have had a negative affect on the achievement of these learners in mathematics (but this would not hold true for the language instrument/s to the same extent). Consequently, this aspect of the opportunity to learn could not be reported on.

2.10 CONCLUSION

This chapter describes the research design, methodology, administration and analysis used in conducting language and mathematics assessments with Grade 8 learners in the Western Cape.

The chapter also notes the way in which data had been analysed and reported. The learner scores were also reported separately for each learning area at the provincial level, and district/EMDC levels where meaningful. In the next chapter, the results derived from the assessment instruments, used to measure learning and teaching at Grade 8 level in the Western Cape, are presented.

CHAPTER 3 – Learner Performance

3.1 INTRODUCTION

In this chapter, the performance of Grade 8 learners is presented in terms of the instruments that were used to assess Languages (Home Language, or HL, and Language of Learning and Teaching, or LoLT) and Mathematics.

Results for each learning area are reported separately. All scores are expressed as percentages and disaggregated by EMDC (the identities of which are masked here as per contractual agreement). For each of the learning areas, percentage scores are also converted to performance levels (see Chapter 2) according to current Curriculum Assessment Guidelines. Results are also reported by Learning Outcomes, Home Language and Language of Learning and Teaching (LoLT), gender and EMDC.

Diagnostic information is provided in order to inform classroom practitioners about how learners achieved expected competencies.

3.2 RELIABILITY OF THE INSTRUMENTS

Instrument reliability is the degree to which the instrument score is consistent over time or under similar conditions and refers to the consistency with which an assessment instrument measures a specific construct. Reliability indicates the degree to which the total score obtained in an instrument is free of measurement error. The degree of internal homogeneity of the items in an instrument can be used as an indication of the reliability of the instrument and is measured by the Alpha coefficient. The stronger the interrelationship between items, the closer the Alpha coefficient will be to a maximum score of 1. A reliability coefficient in the vicinity of 0,8 is considered appropriate for achievement instruments. The Alpha coefficient for the two learning areas is reported in Table 3.1.

Table 3.1: Reliability coefficient for the two learning areas (with Language divided further)

Learning Area	Alpha Coefficient
Language: Afrikaans	0.846
Language: English	0.899
Language: Xhosa	0.799
Mathematics	0.900

Learner Performance

The Alpha coefficient for each of the learning areas was relatively high, indicating that results obtained for each of the instruments can be considered to be highly reliable. The reliability coefficient for each Learning Outcome individually is likely to be lower, as they are covered by fewer items. In some cases, a Learning Outcome comprised five or six items only.

3.3 PROVINCIAL SCORES

3.3.1 Performance per Learning Area

The average performance in all learning areas is presented in Table 3.2 and Figure 3.1.

Table 3.2 Provincial averages for Languages and Mathematics

	Average	Standard error	Standard deviation
Language: Afrikaans	30.19	0.07	12.64
Language: English	38.94	0.12	16.37
Language: Xhosa	30.59	0.12	11.37
Mathematics	28.56	0.06	17.00

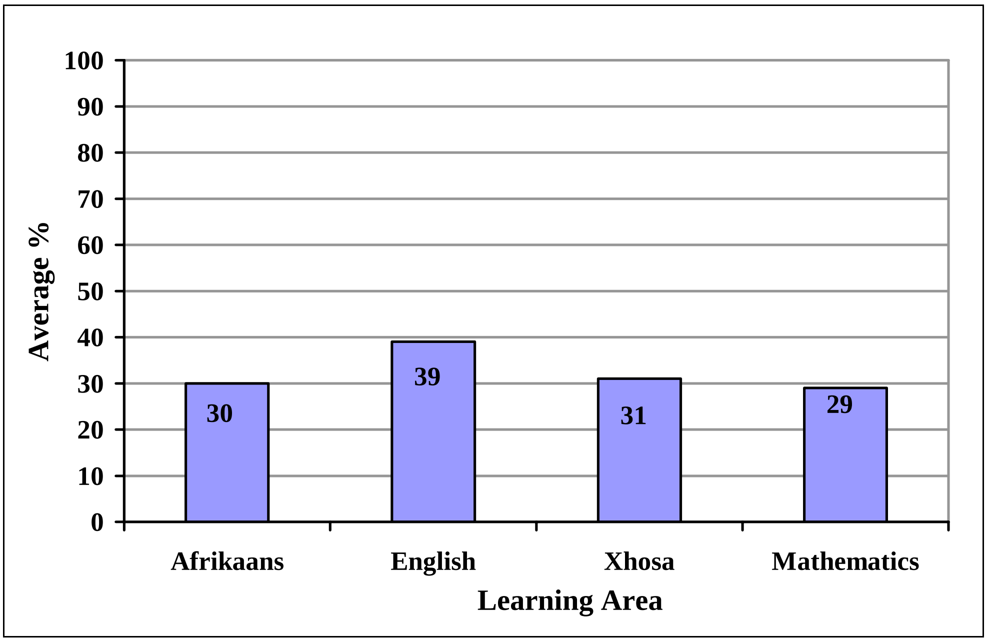


Figure 3.1: Provincial average scores per Learning Area

The performance of learners in all the learning areas was generally low. Learners obtained the highest score in the English Home Language instrument and the lowest score for the Mathematics instrument. Table 3.3 and Figure 3.2 illustrate the performance of learners for the learning areas across the districts (EMDCs).

Table 3.3: Learners' overall scores by learning area and EMDC[8]

	Metro 1	Metro 2	Metro 3	Metro 4	Rural 1	Rural 2	Rural 3
Language: Afrikaans	26	31	31	24	29	31	32
Language: English	38	40	41	36	45	44	47
Language: Xhosa	35	31	32	30	31	31	27
Mathematics	33	25	31	27	25	27	30

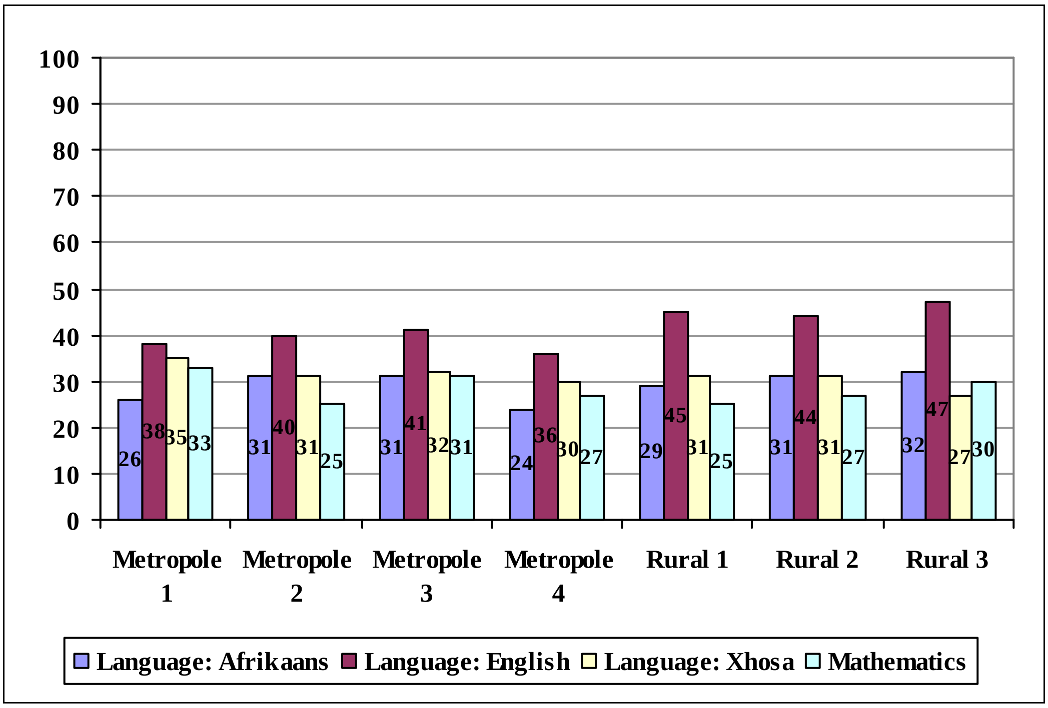


Figure 3.2: Learners' overall scores by learning area and EMDC

An analysis of the data split according to EMDCs shows the patterns reflected in Table 3.3 and the graph in Figure 3.2, with performances across EMDCs ranked mostly in the same order as in the province for the learning areas.

3.3.2 Performance per item type and Learning Area

Figure 3.3 illustrates the performance of learners for the two item types (MC and CR items) across the learning areas and for the two main sub-sections (HL and LoLT) of the language assessment task.

[8] Reporting of performance by EMDC may largely reflect differences between levels of resourcing and facilities, rural or urban contexts, or even the sizes of schools and classrooms, and is no indication of different management or other district qualities.

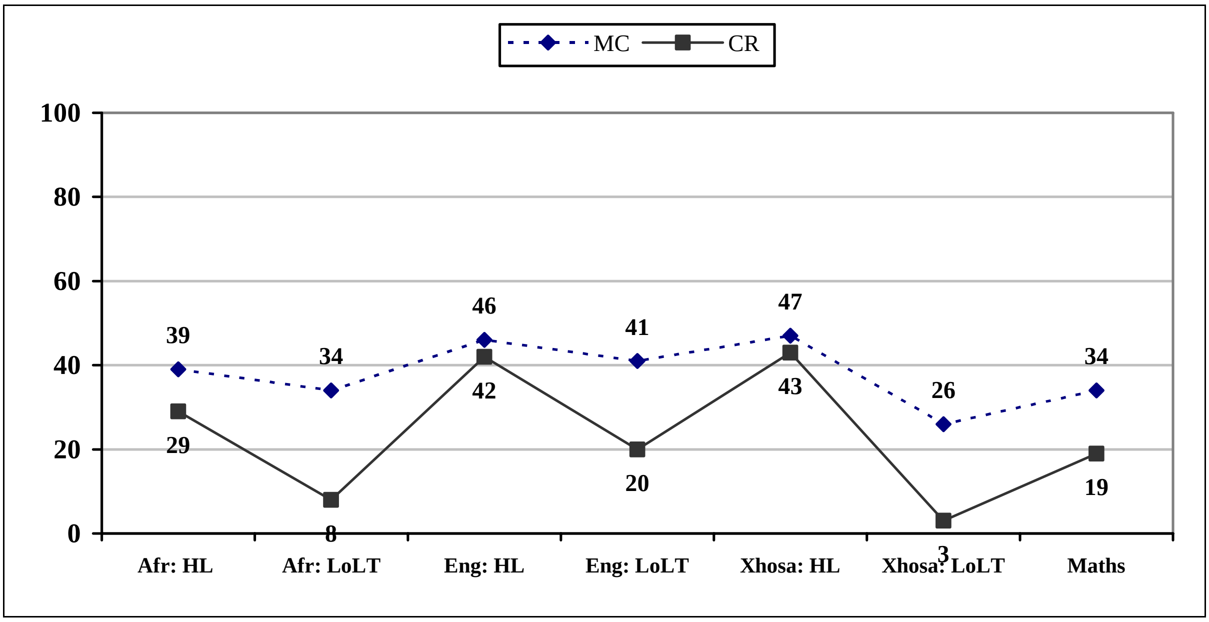


Figure 3.3: Comparison of scores by item type (MC=multiple-choice; CR=constructed-response)

The main trend that became evident is that learners consistently perform better, irrespective of learning area and language assessment sub-section, when answering multiple-choice questions compared to answering constructed-response items. This is consistent with all previous findings from various studies, and signals that it is easier for learners to recall, recognise and select knowledge aspects from options provided within items, than to produce extended and meaningful responses where they have to implement their understanding of the work. The gap for Mathematics remain quite large. This finding suggests that learners continue to have difficulties in applying what they have learned outside the narrow confines of learning area contents and (rote) preparation for familiar assessment strategies or approaches.

This viewpoint or conclusion is supported even more strongly when observing that this MC-CR gap is much wider for LoLT assessment contents compared to HL contents. This suggests that assessment tasks are more familiar to and easier for learners when related to the technical and grammatical contents of a given language as a subject compared to the challenge of applying rules of grammar, vocabulary, and reading and writing skills in general, to materials across the curriculum or elsewhere in real-life situations. This trend is a cuase for even more concern.

A final observation is the relatively large gap for Xhosa learners on both the MC and CR items between the HL and LoLT sections. Not only does it signal the positive outcome of relatively high performance levels by Xhosa-speaking learners on their HL assessment (in comparison to English and especially Afrikaans-speaking learners), but also the extent of the challenge for Xhosa-speaking learners to apply their language abilities to materials not in their mother tongue or to learning materials across the curriculum. (The generally low performance levels of Afrikaans-speaking learners in all this comes out as a serious concern.)

3.4 LANGUAGE

3.4.1 Performance per language instrument

The language instrument was developed with three home-language versions (Xhosa, Afrikaans and English) and two language-of-learning-and-teaching (LoLT) versions (Afrikaans and English). Most students who have English and Afrikaans as home languages wrote the instrument which matched both their home language and LoLT. The majority of African language speaking students in the Western Cape are speakers of Xhosa and it was intended that they should write an instrument with a Xhosa HL section and an English LoLT section.

Within the HL section of each of the three versions, there were two common constructed-response items (two extended pieces of writing). In other words, students were asked to write exactly the same exercises, but in the three different HLs used in these instruments. In Table 3.4 and Figure 3.4, the data refer to those students who wrote:

- Xhosa HL and English LoLT;
- Afrikaans HL and Afrikaans LoLT; and
- English HL and English LoLT.

Where it was clear that some students had written an incorrect instrument, those scores were not included in this data set, in order to keep the data as appropriate as possible.

Table 3.4: Mean scores across language instruments

Xhosa			Afrikaans			English		
HL	**LoLT**	**Common Items**	**HL**	**LoLT**	**Common Items**	**HL**	**LoLT**	**Common Items**
45,1	18,2	23,0	34,7	26,3	29,7	44,2	34,5	36,0

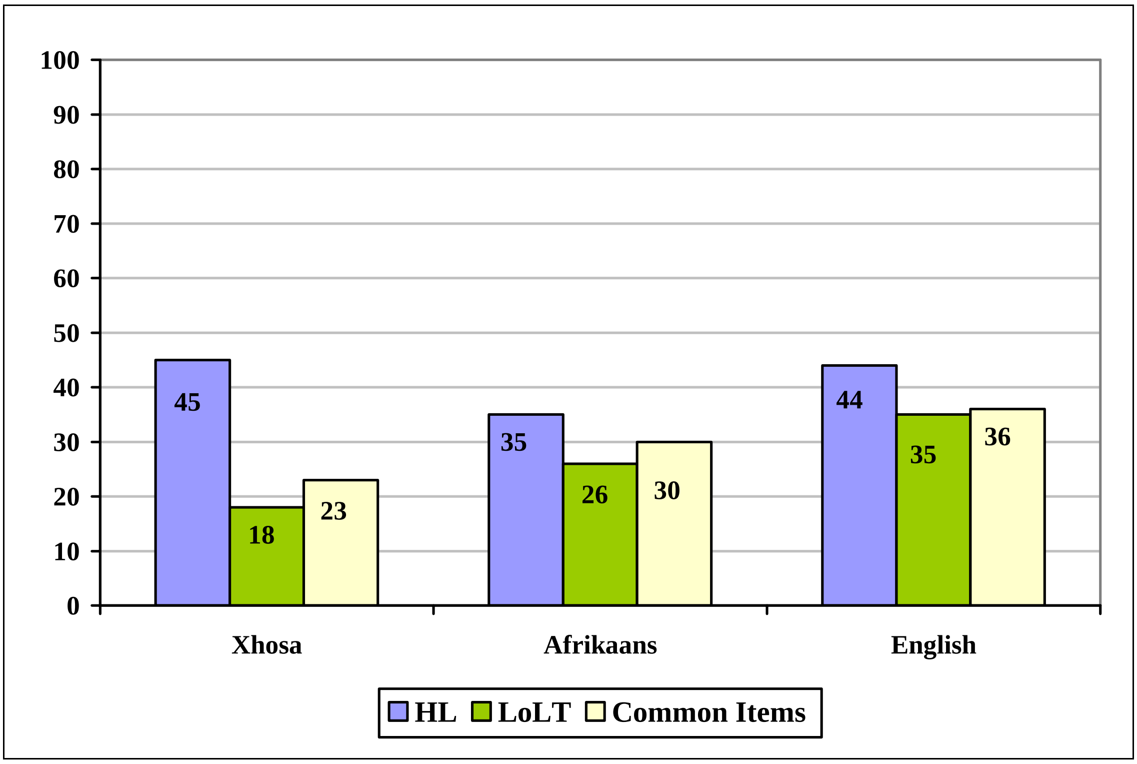


Figure 3.4: Mean scores across language instruments

These data show that in the HL section of the instrument, which was largely based on a literacy extract taken from the reading lists of each HL group, Xhosa-speaking students outperformed both Afrikaans- and English-speaking students. This would suggest that Xhosa HL students are familiar with the genre and are able to read literature at grade level fairly well. However, when only the common extended writing tasks in the HL section were measured, Xhosa-speaking students performed least well. However, no language group demonstrated adequate writing skills on average, and this is a cause for concern since the entire secondary-school curriculum (to the end of GET and through FET) requires competent reading and writing skills. Again, Xhosa-speaking students performed least well in the LoLT section of the paper, and this is not surprising since these students were writing this part of the instrument in their first additional (second) language, rather than the HL as was the case with the Afrikaans-and English-speaking students. However, the low achievement in LoLT for all language groups is of significant concern. The LoLT part of the instruments were designed to measure the extent to which learners might be able to read, understand and construct responses to the kind of text that students in other countries at this grade level would be expected to use. Learners who achieve less than 50% in this section will find it very difficult to continue education in the FET band, since it is unlikely that they would be able to grapple successfully with the academic literacy expectations of learning materials, supplementary reference materials and assessment tasks in Grades 10, 11 and 12.

3.4.2 Language by performance levels

Table 3.5: Linkage between performance levels, descriptions and percentages

Performance level/rating code	Description of rating	Marks in percentages
7	Outstanding performance	80 – 100
6	Meritorious performance	70 – 79
5	Substantial performance	60 – 69
4	Adequate performance	50 – 59
3	Moderate performance	40 – 49
2	Elementary performance	30 – 39
1	Not achieved	0 – 29

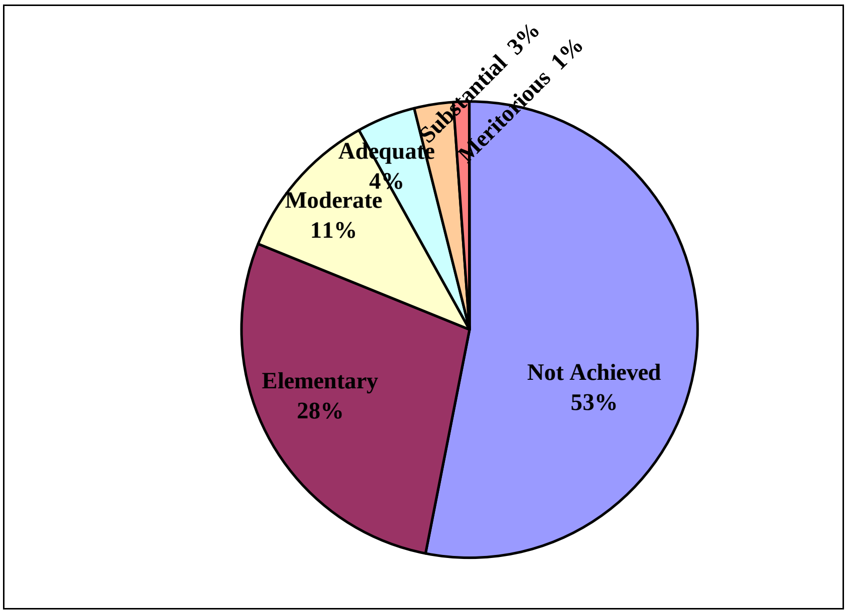


Figure 3.5: Percentage of learners at each performance level in Afrikaans

The pie-graph in Figure 3.5 shows that only 8% of Afrikaans HL speakers achieved more than 50% on this instrument. Only 52 learners achieved 80% or more and only 306 achieved between 70% and 79%. These data are extremely alarming and suggests that urgent attention needs to be focused towards this group of learners. This is particularly the case, because the majority of these learners are receiving ‘mother tongue’ (home language) education throughout and they should be in the best position to develop strong home-language reading and writing skills across the curriculum.

The diagnostic coding of learners’ constructed-response items (those items where learners were expected to produce well-constructed written answers) has shown a high percentage of learners who appear to have almost no basic reading and writing skills (barely literate). This is a cause for concern as it is not clear how so many learners could have reached Grade 8 without having even basic literacy skills. Many of these learners have developed various strategies which create a superficial ability to read and write. Many know how to produce signs and symbols on the page which appear at first sight to resemble individual words and text. However upon scrutiny it becomes obvious that they are using repeated patterns of symbols which may or may not resemble letters of the alphabet, but that they do not form words or sentences.

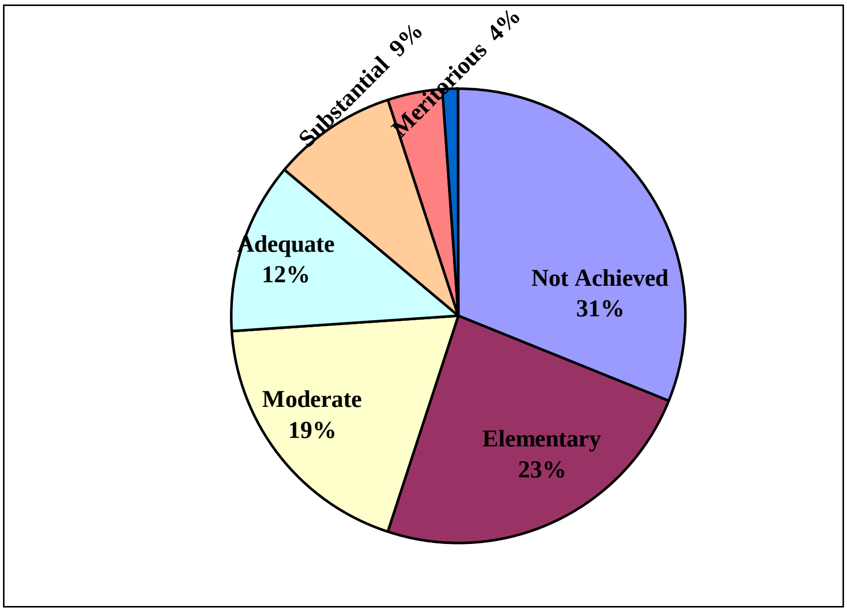


Figure 3.6: Percentage of learners at each performance level in English

The English-HL group of learners appears to have the highest level of achievement, and there is a greater spread of achievement levels in this group (Figure 3.6). Nevertheless, 73% of learners achieved below 50% on this instrument and this is not satisfactory since the majority of these learners receive 'mother tongue' education throughout. A minority are learners who come from other HL backgrounds and have elected to study English at HL level. Of these learners, 750 achieved between 70% and 79%, and 152 achieved 80% or above. The major weakness identified through diagnostic scoring and coding shows that many of these learners are not able to apply the language and literacy skills learnt in English HL to other areas of the curriculum. Alternatively, the learning outcomes and assessment standards in the languages learning area of the curriculum do not adequately prepare learners for the academic literacy skills required in other areas of the curriculum.

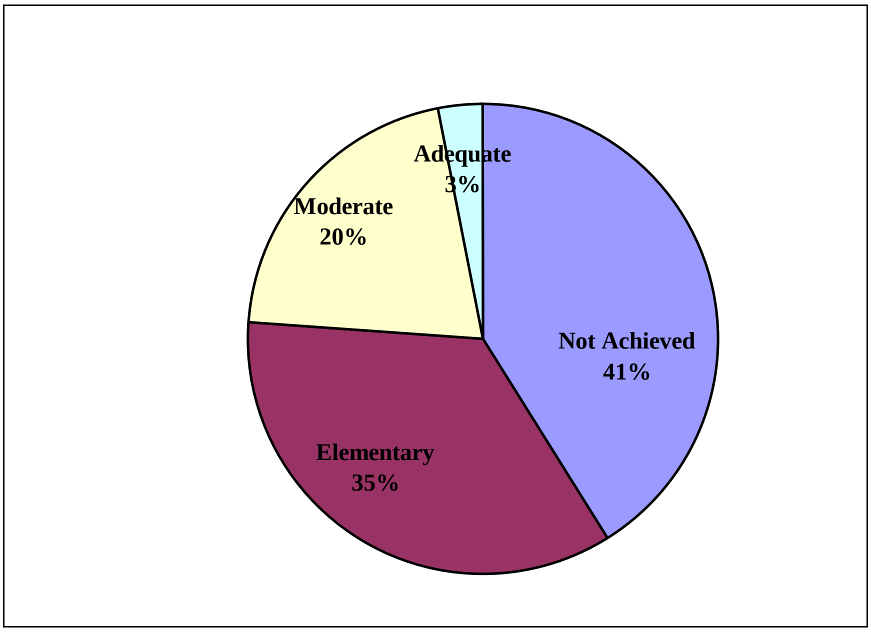


Figure 3.7: Percentage of learners at each performance level in Xhosa

Only 3% of Xhosa-HL learners (Figure 3.7) achieved above 50% on this instrument. No Xhosa-HL learner achieved an average of 70% or more on the instrument. This means that only about 3% of Xhosa-HL learners are likely to have a good chance of completing the GET band and proceeding through FET successfully. The majority of Xhosa HL learners found it particularly difficult to construct text in their home language and almost impossible to do this to grade level expectations of the curriculum at Grade 8 level. The diagnostic scoring and coding of the constructed-response items shows a very high proportion of learners who have extremely low levels of basic literacy, and many demonstrate negligible evidence of literacy. However, as is the case with many speakers of Afrikaans, many of these learners have developed various strategies which create a superficial ability to read and write. Many know how to produce signs and symbols on the page which appear at first sight to resemble individual words and text. However upon scrutiny it becomes obvious that they are using repeated patterns of symbols which may or may not resemble letters of the alphabet, but that they do not form words or sentences. This is a cause for concern as it is not clear how so many learners could have reached Grade 8 without having even basic writing skills (and or literacy skills in general).

A summary is given in Table 3.6 and Figure 3.8 of the percentages of learners in each EMDC who have achieved scores that fall within each of the descriptions according to the 7-point performance-level scale.

Table 3.6: Language performance expressed in terms of achievement level for EMDCs

	Not Achieved	Elementary	Moderate	Adequate	Substantial	Meritorius	Out-standing
Metro 1	40	24	17	9	7	3	1
Metro 2	42	31	18	6	3	1	0
Metro 3	43	25	16	9	6	2	0
Metro 4	47	27	15	6	4	2	1
Rural 1	51	30	12	4	2	1	0
Rural 2	46	30	14	5	3	1	0
Rural 3	45	28	14	6	5	2	0
Provincial average	45	28	15	6	4	2	0

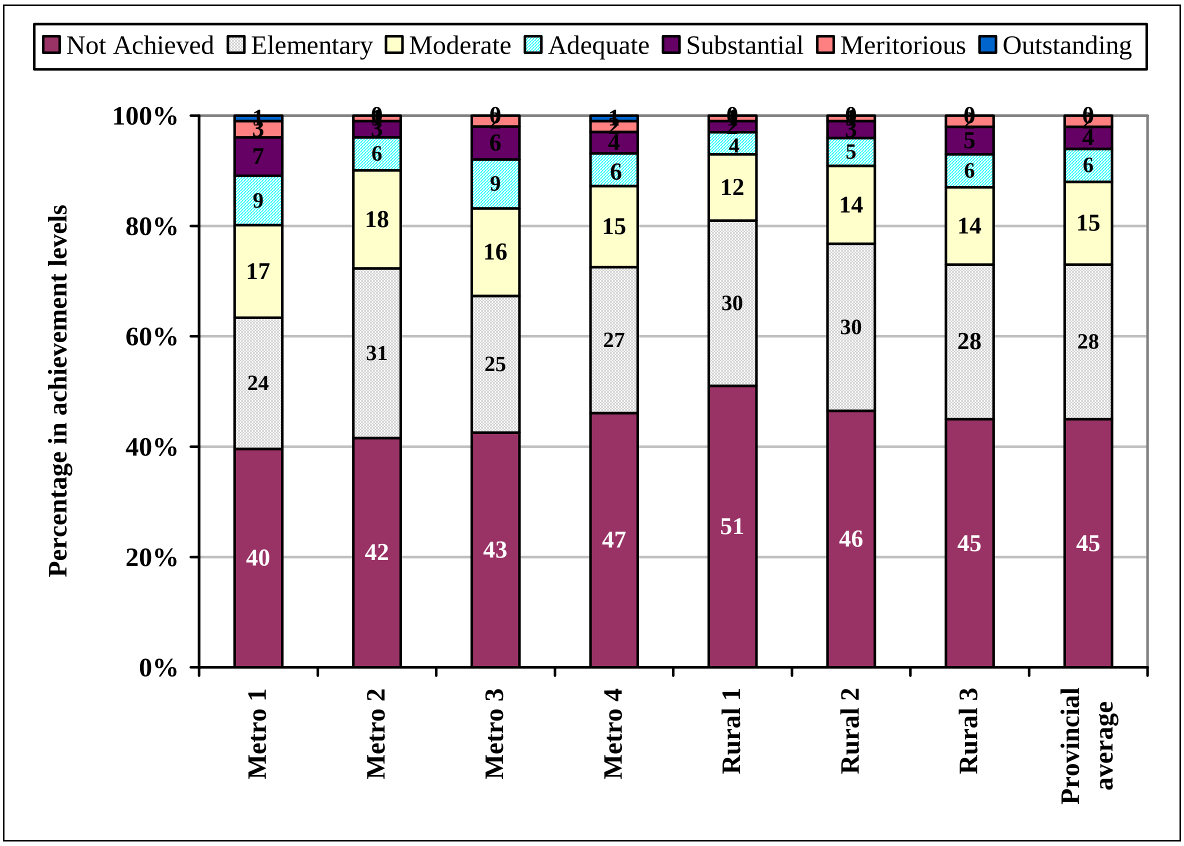


Figure 3.8: Language performance expressed in terms of achievement level for EMDCs

3.4.3 Language performance by Learning Outcomes

Figure 3.9 shows the mean percentages scores of learners on each of the four learning outcomes covered by the assessment tasks. The following key identifies the various learning outcomes:

LO3 = Reading and Viewing

LO4 = Writing

LO5 = Thinking and Reasoning

LO6 = Language Structure and Use

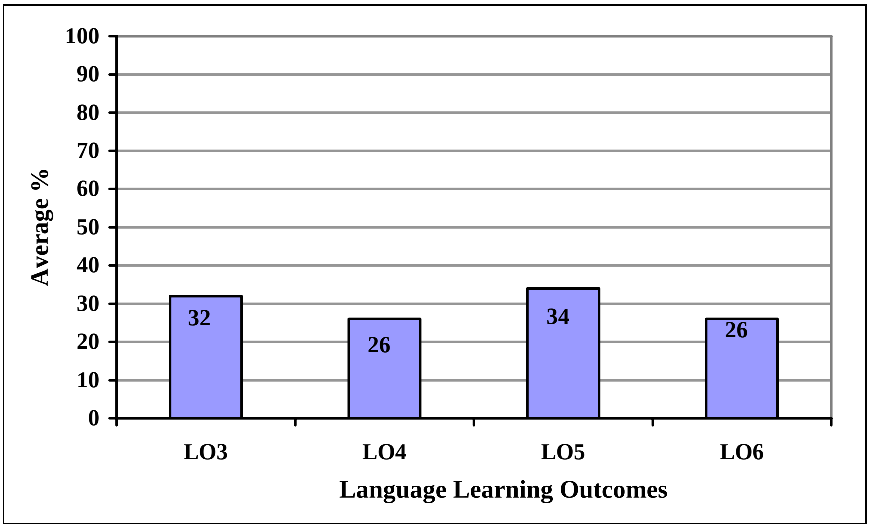


Figure 3.9: Language performance by Learning Outcome

In terms of a ranking of levels of difficulty of language skills, reading (a predominantly, but not exclusively, receptive skill) is considered easiest, followed by knowledge of language structure and use, followed by thinking and reasoning, and followed by writing. Other systemic assessments in South Africa (e.g., the Monitoring Learning Achievement Study, Strauss 1999) show that writing (LO4) appears to be the most difficult task for learners, because this is where they demonstrate their productive competence in language use. Therefore we would expect learners to perform least well in the writing tasks (constructed-response items) overall. This holds true in the WCED assessment tasks, and a similar level of achievement for knowledge and use of language structure (LO6) is consistent with this. However, overall, the achievement levels are disappointing for each of the learning outcomes, and are not positive indicators for successful learning in the system beyond Grade 8.

3.4.4 Language performance by grade level

Figure 3.10 shows the achievement of learners according to the difficulty or grade level that the various tasks were considered to cover.

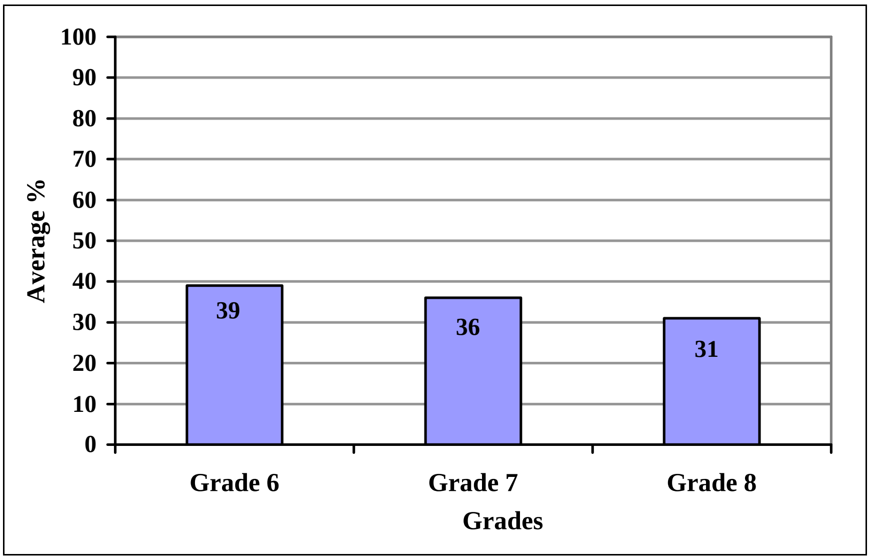


Figure 3.10: Language performance by grade level

Although this was intended to be a Grade 8 assessment, earlier findings from the WCED and national DOE systemic assessments at Grade 3 and Grade 6 levels (e.g., DoE 2005) indicated that many learners were not even reaching optimal literacy levels at earlier grade levels. Therefore it was decided to include and benchmark (against the NCS curriculum statements) the multiple-choice items in each of the HL and LoLT sections of the instruments at Grades 6, 7 and 8 levels. The constructed-response items were diagnostically scored and coded using different criteria and were not included in this subset of the data. It was expected that the average achievement level per grade level would be higher for the earlier grade/s (Grades 6 and 7) and lower for the higher grade (Grade 8). The data correspond with the expectation, but with unacceptably low overall levels of achievement per grade level.

In Table 3.7 and Figure 3.11 learner achievement by the difficulty or grade level of the items is broken down for the various languages in which learners had completed the assessment tasks.

Table 3.7: Performance by difficulty or grade level of items for the language instruments

	Grade 6	Grade 7	Grade 8
Afrikaans	38	32	29
English	41	42	36
Xhosa	39	35	28
Total	39	36	31

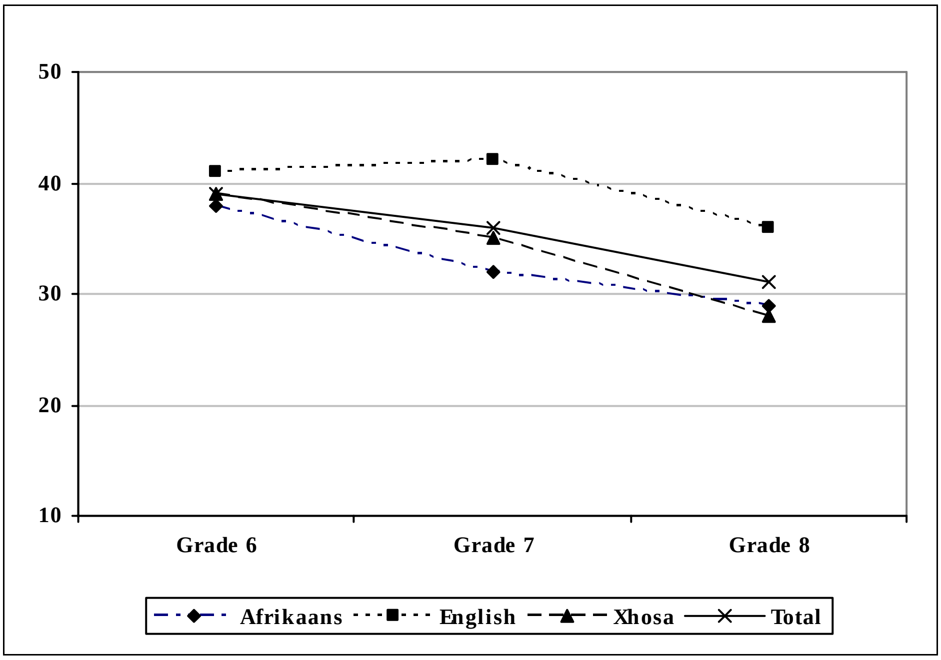


Figure 3.11: Performance by grade level for the language instruments

When the data are split according to home-language group, the English-HL speakers performed best, and the Xhosa-HL speakers performed better than Afrikaans-HL learners on Grades 6 and 7 items, but the Afrikaans-HL learners performed slightly better than Xhosa-HL learners on Grade 8 level items.

In Table 3.8 and Figure 3.12 the relative performance levels of learners in the various EMDCs on items set at the different difficulty or grade levels are summarised.

Table 3.8: Language performance by grade level and EMDC

	Metro 1	Metro 2	Metro 3	Metro 4	Rural 1	Rural 2	Rural 3	Total
Grade 6	42	41	42	38	37	39	41	40
Grade 7	41	36	38	37	32	34	36	36
Grade 8	35	31	33	31	29	30	32	32

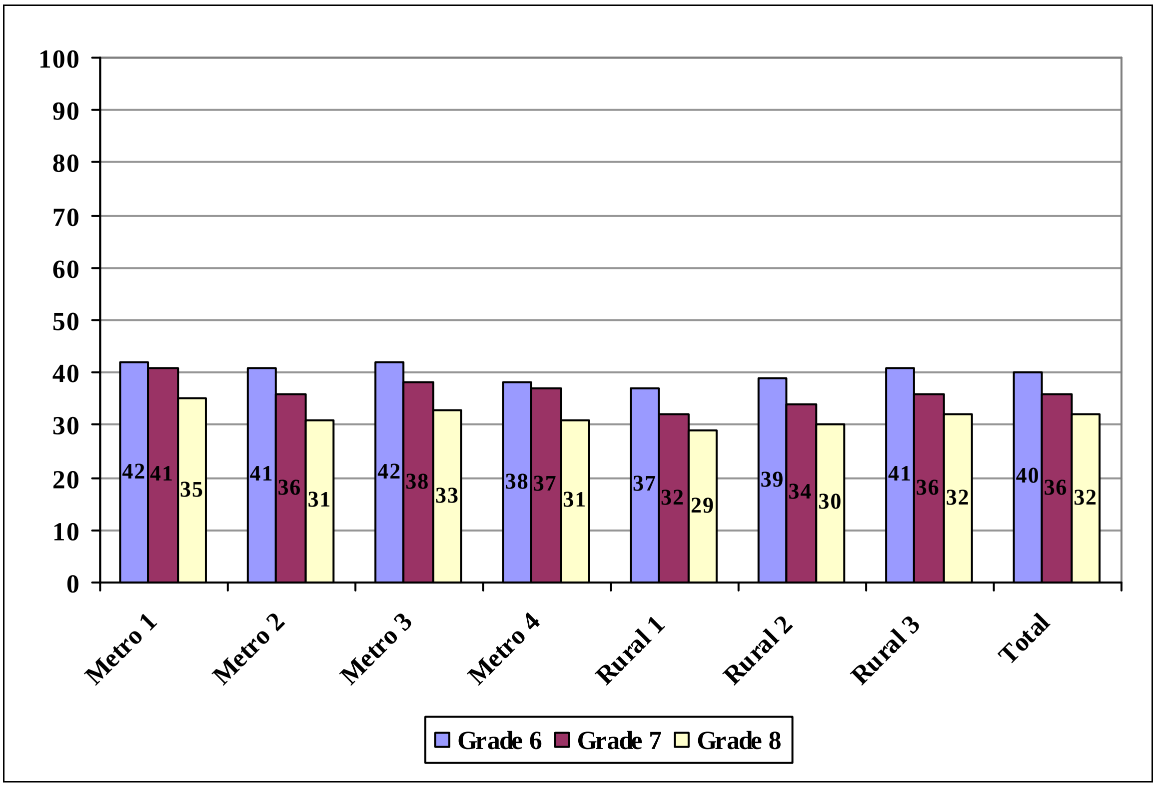


Figure 3.12: Language performance by grade level and EMDC

While there are some differences between the EMDCs, and the Department has been advised to take note of these, the achievement of students across all of these grade levels is not acceptable and does not offer a positive indicator for systemic success in the FET band.

3.4.5 Language performance by gender

The relative performance of boys and girls are summarised in Table 3.9 and Figure 3.13.

Table 3.9: Language performance by gender for the different sections of the instrument

	HL	LoLT	Total
Boys	35	26	30
Girls	42	29	35

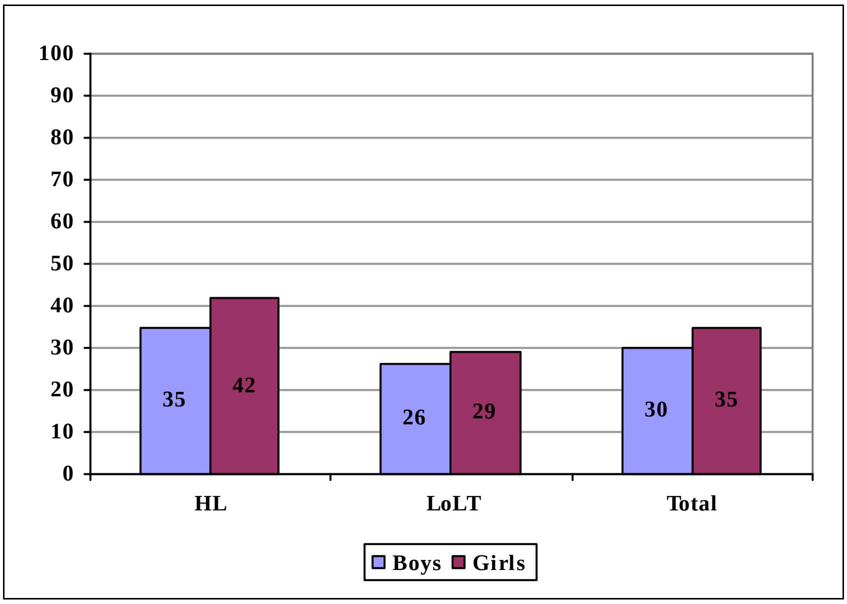


Figure 3.13: Language performance by gender for the different sections of the instrument

The data show that girls performed better than boys in both HL and LoLT, although the gap narrows from 7% in relation to HL, to 3% in relation to LoLT. This may indicate that boys are able to keep up more easily with girls in areas of the curriculum other than the languages learning area/s.

Table 3.10: Language performance by gender and learning outcome (LO)

	LO3	LO4	LO5	LO6
Boys	29	22	32	23
Girls	33	29	35	28
Total	31	25	34	26

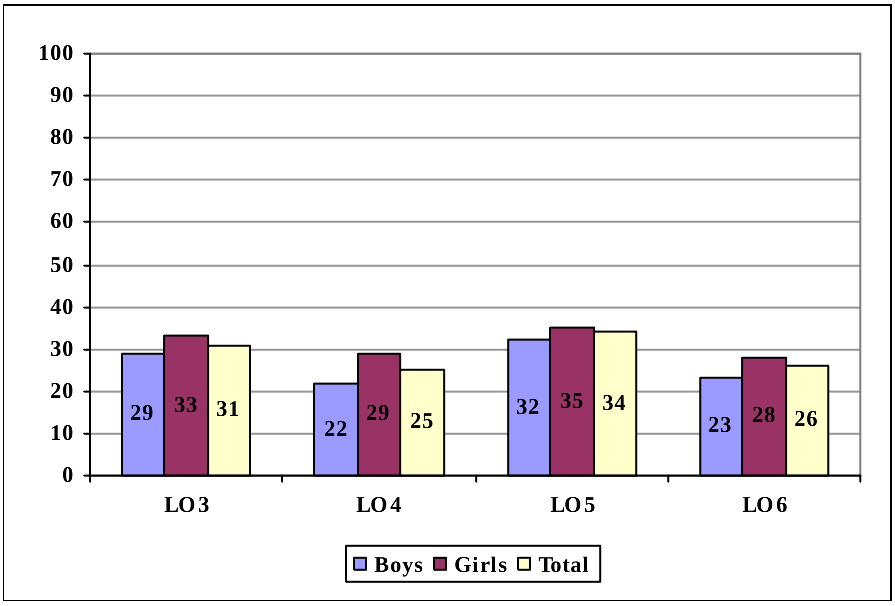


Figure 3.14: Language performance by gender and learning outcome (LO)

When the gender-based data are split amongst the different learning outcomes (Table 3.10 and Figure 3.14), the degree to which girls appear to outperform boys is greatest in LO4 (writing) and LO6 (knowledge of language structure and use). Again this points towards the possibility that girls may be more receptive to language as a learning area and that boys may be more inclined towards other areas of the curriculum.

Table 3.11: Language performance by gender and grade or difficulty level

	Grade 6	Grade 7	Grade 8
Boys	37	34	29
Girls	42	38	34
Total	40	36	31

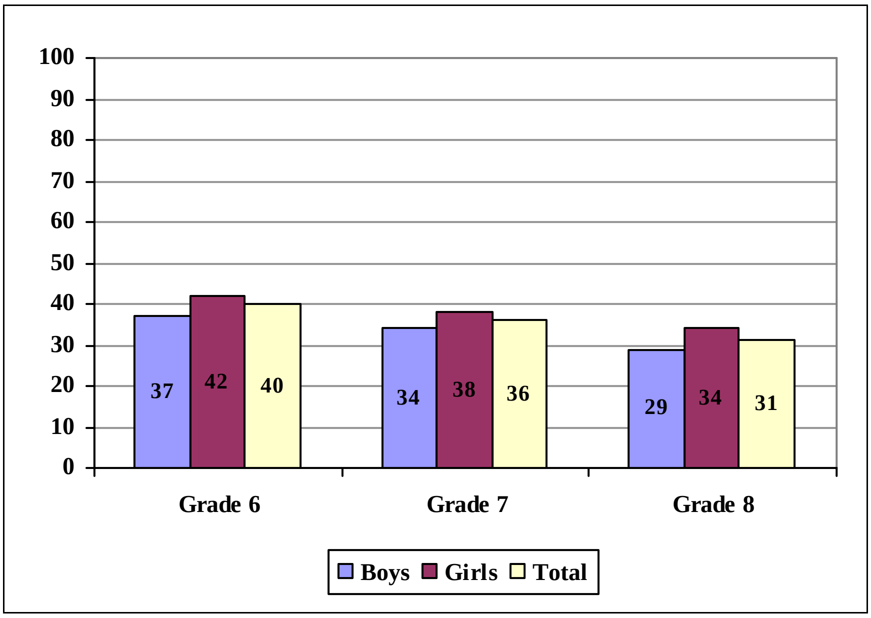


Figure 3.15: Language performance by gender and grade or difficulty level

The gender-based data split according to achievement on items benchmarked against the NCS per grade level (Table 3.11 and Figure 3.15), shows no particularly unexpected findings.

3.4.6 Language performance by learning outcome for the different language instruments

Table 3.12 and Figure 3.16 portray the situation when learner achievement is mapped against the various learning outcomes for each language of assessment.

Table 3.12: Performance by learning outcome for the different language instruments

	LO3	LO4	LO5	LO6
Afrikaans	29	21	31	25
English	35	30	39	24
Xhosa	29	30	30	29
Total	31	25	33	25

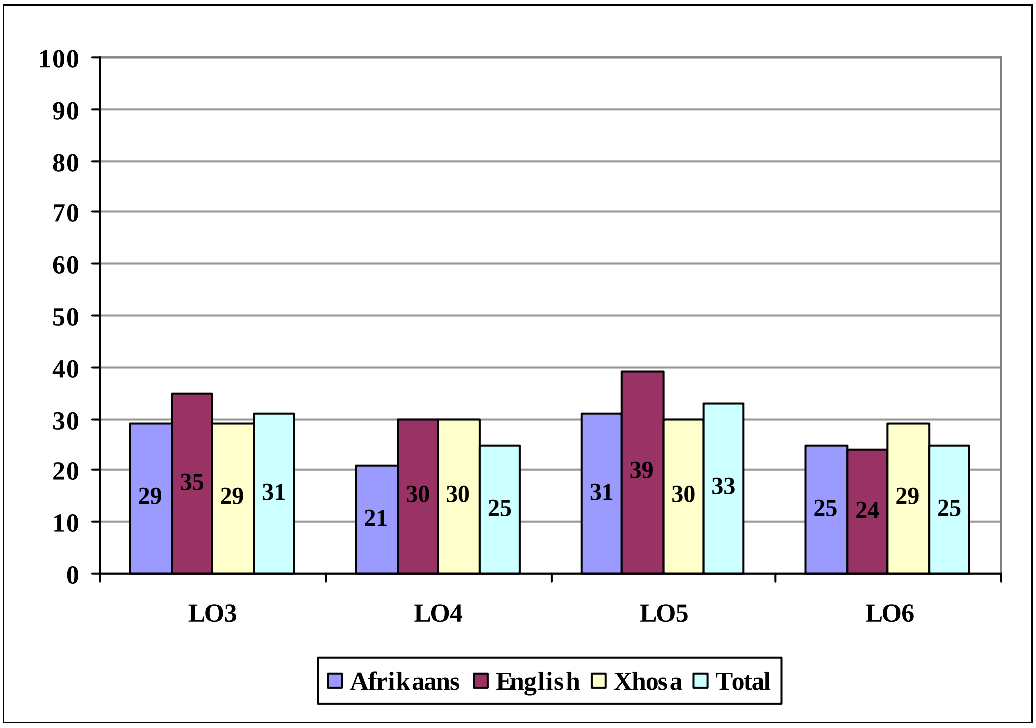


Figure 3.16: Performance by learning outcome for the different language instruments

These data show that English-HL speakers appear to perform significantly better at reading (LO3) and Thinking and Reasoning (LO5) than other learners. However, Xhosa-HL learners appear to have comparatively greater knowledge of language structure and its use (LO6) than do Afrikaans- and English-HL learners. English-HL learners' knowledge of language structure appears to be weakest of all language groups. This indicates that English-HL teachers need to pay more attention to the structure of language (LO6) in mediating the curriculum in their classrooms, and this also applies to Afrikaans-HL teachers. Nevertheless, overall learners achieved poorly across all LOs and educators across the system need to address this.

3.4.7 Language performance by HL and LoLT

Table 3.13: Learner performance for HL and LoLT for the different language groups

	HL	LoLT	Total
Afrikaans	35	26	30
English	44	35	39
Xhosa	45	18	31
Total	39	28	33

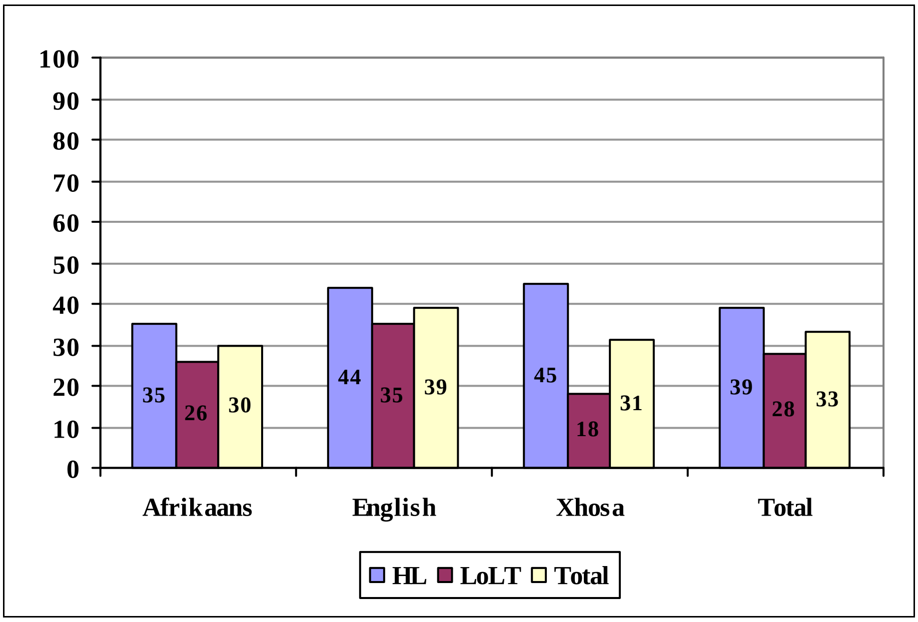


Figure 3.17: Learner performance for HL and LoLT for the different language groups

In general (see Table 3.13 and Figure 3.17), learners achieved better scores with the HL section of the language instruments than the LoLT sections, as discussed earlier. The largest gap between HL and LoLT achievement lies with the Xhosa-HL students who wrote the LoLT section in English, the first additional (or second) language. This is not surprising since English and Afrikaans speakers would have been at an advantage. What this points towards, however, is that those learners who study through the medium of their second language are at a significant disadvantage in comparison with learners who study through the medium of their HL. An achievement of 18% for language across the curriculum in Grade 8 is an extremely serious negative indicator for retention and or success in the FET band. Nevertheless, the achievement of Afrikaans and English HL learners is also of serious concern in this regard.

3.4.8 Language performance by item type

Table 3.14: Language performance by gender and item type

	MC	CR
Boys	37	21
Girls	40	29
Total	39	25

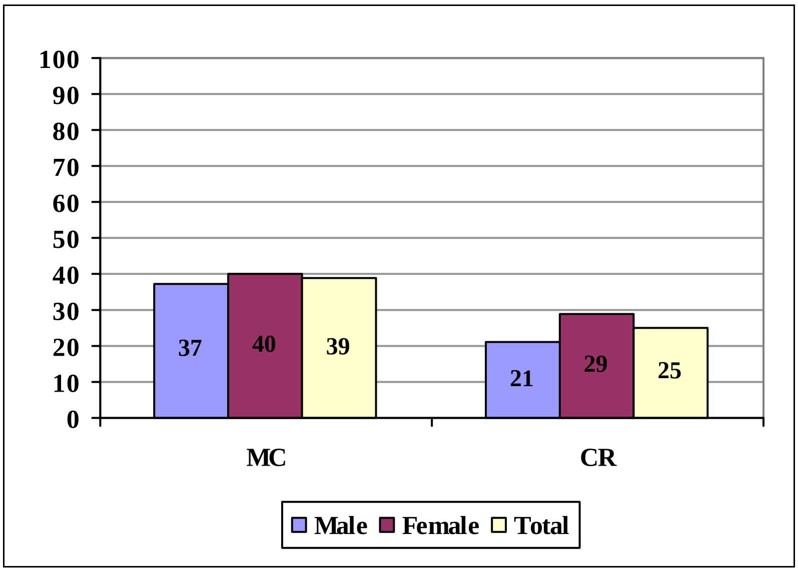


Figure 3.18: Language performance by gender and item type

There is a serious gap in achievement between multiple-choice (MC) item responses and constructed-response (CR) items (Table 3.14 and Figure 3.18). In general, many MC items may permit learners to employ knowledge and skills which are receptive, whereas CR items require learners to demonstrate evidence of having processed the item, and formulated an answer. This is known as productive knowledge. In a language assessment, this points specifically to academic literacy and the level of academic literacy achieved by learners. MC items may also lend themselves to guessing and it is sometimes difficult to know whether or not learners have guessed at answers without reading the items. (It is usually accepted amongst assessment experts that MC items could offer learners a 20% to 25% opportunity to guess correctly across such items within an instrument.) This is particularly likely to be the case in educational contexts where learners exhibit low levels of literacy. Learners' achievement in CR items therefore is an important indicator of the degree to which learners are able to read accurately, and respond appropriately to items. It is these items which are likely to be able to provide the kind of information which an education system requires for predicting future outcomes for the set of learners engaged in the assessment task.

Girls outperform boys on both sets of items, but the largest gap is related to item type rather than gender difference. The educational concern here is with item type, and that learners in general are performing very poorly on CR items. In this case, the academic literacy of learners in Grade 8 in the Western Cape is of serious concern.

Item type by Language group

Table 3.15: Performance by item type for the three HL language assessment groups

	MC	CR
Afrikaans	36	20
English	44	33
Xhosa	34	26
Total	39	25

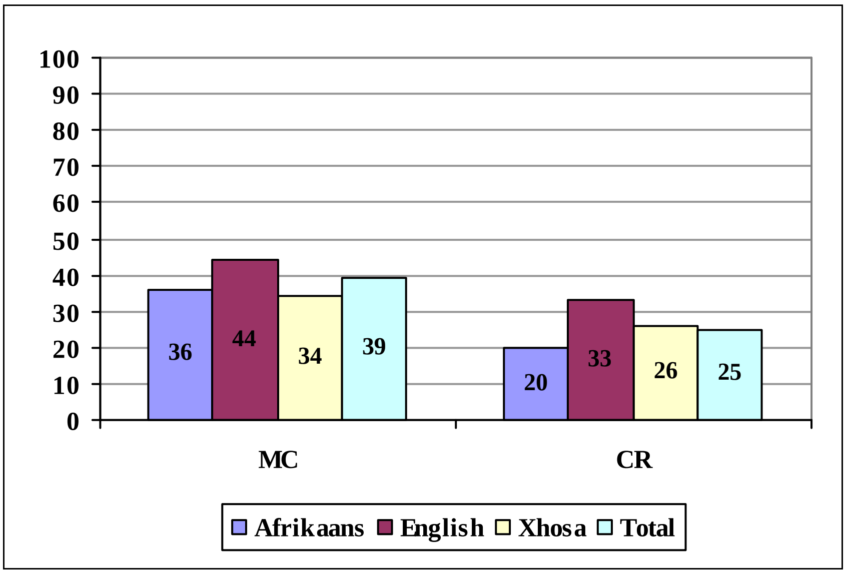


Figure 3.19: Performance by item type for the three HL language assessment groups

The data reflected in Table 3.15 and Figure 3.19 show similar patterns to those which precede these. When split by HL group, one would expect that the English-HL learners would perform best, followed by Afrikaans- and then Xhosa-HL groups. However, in this case, the Xhosa-HL learners outperform Afrikaans-HL learners in the constructed-response items. This is a further indicator of extremely unsatisfactory literacy-level achievement of speakers of Afrikaans, although no group performed well.

In Table 3.16 and Figure 3.20, learner performance by item type is broken down for EMDCs.

Table 3.16: Language performance by EMDC and item type

	Metro 1	Metro 2	Metro 3	Metro 4	Rural 1	Rural 2	Rural 3	Total
MC	41	38	40	37	35	37	39	39
CR	29	25	26	25	21	23	24	25

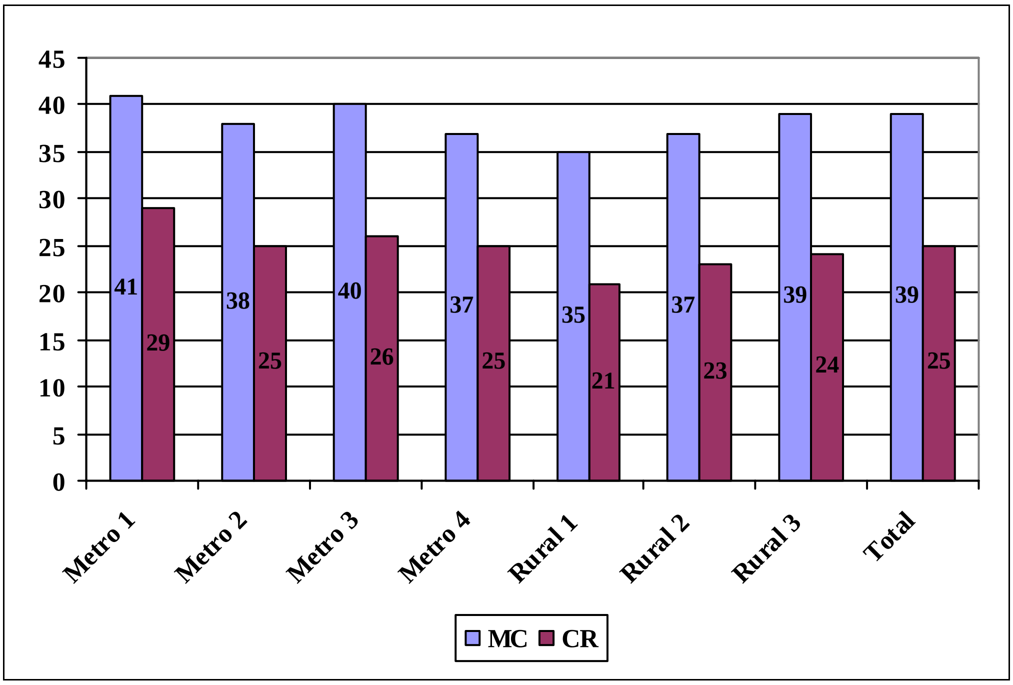


Figure 3.20: Language performance by EMDC and item type

There is a fairly consistent pattern of a gap of achievement between MC and CR items of 12 to 15 percentage points. This is a poor indicator for success across the curriculum in the FET band.

Table 3.17: Language performance of learners by item type and the number of books in the home

	No books	1 to 20 books	21 to 50 books	51 to 100 books	More than 100 books	Total
MC	33	37	42	49	54	39
CR	18	23	29	36	41	25

As may be expected, there is a correlation between the performance of learners in relation to both MC and CR items, and the number of books in the home (Table 3.17 and Figure 3.21). The information which shows numbers of books in the home acts as a proxy also for socio-economic status, in that the greater the number of books, the higher the likelihood of middle-class economic status, and the lower the number of books, the higher the likelihood of low economic status.

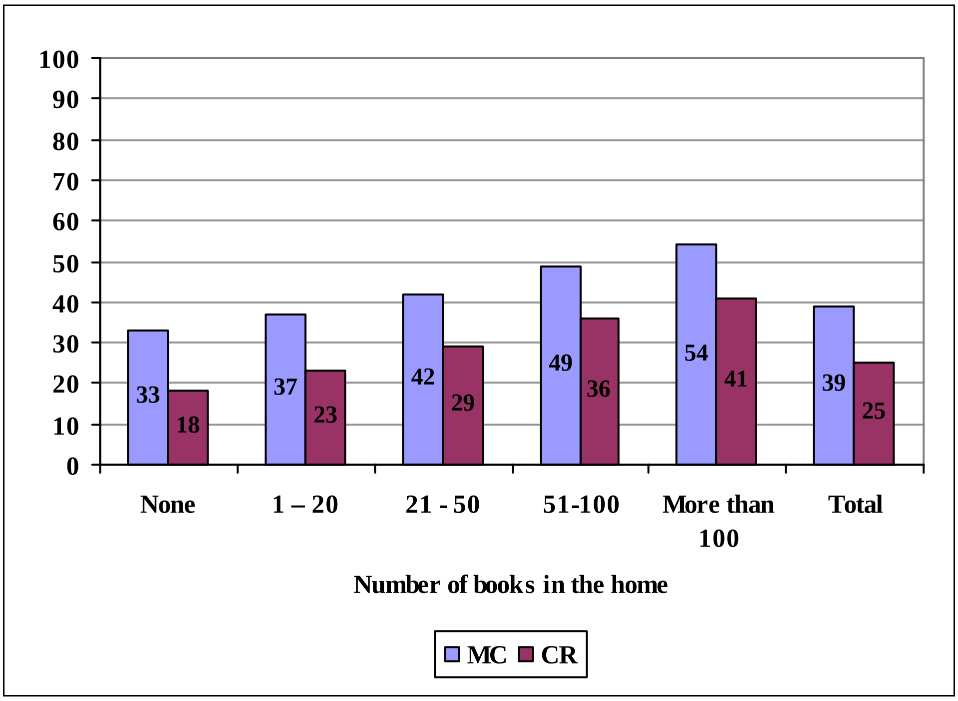


Figure 3.21: Language performance of learners by item type and the number of book in home

3.4.9 Language performance by EMDC

The language performance by EMDC follows a regular and consistent pattern in relation to the gap between achievement in HL and LoLT (Table 3.18 and Figure 3.22). The gap is smallest for one rural EMDC (Rural 3) where probably more learners have Afrikaans as HL and there is a higher incidence of linguistic homogeneity. One would expect that the gap between HL and LoLT would be highest in those EMDCs where LoLT does not match HL.

Table 3.18: Language performance by EMDC

	Metro 1	Metro 2	Metro 3	Metro 4	Rural 1	Rural 2	Rural 3	Total
HL	42	41	40	39	36	38	38	39
LoLT	31	26	30	27	25	26	29	28
Total	36	33	34	32	30	32	33	33

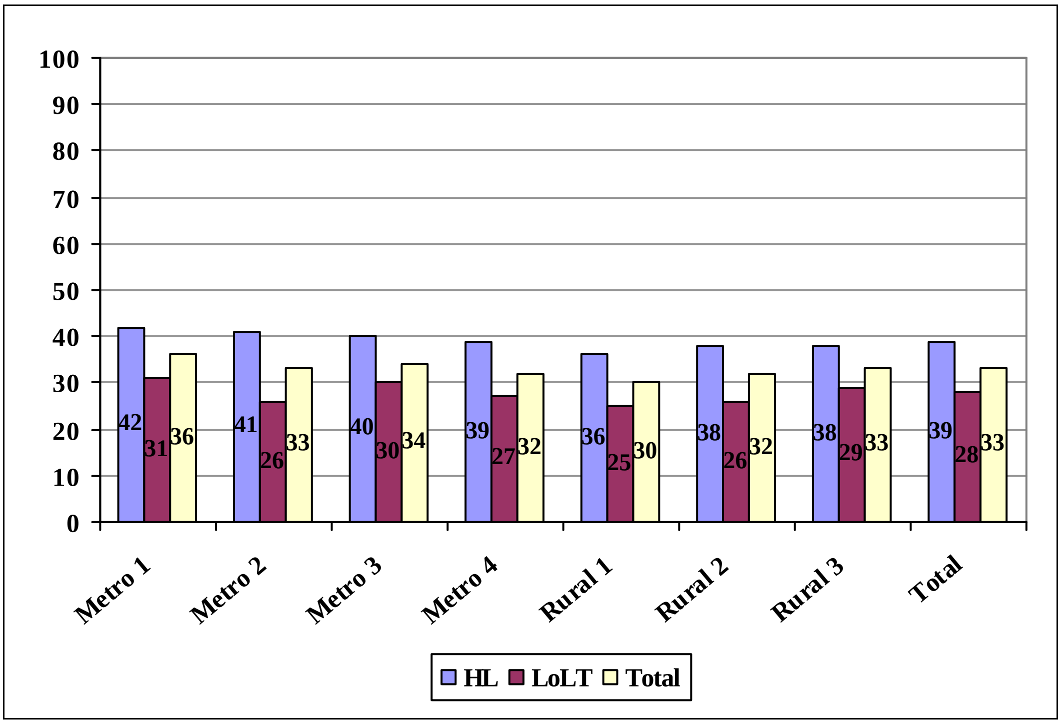


Figure 3.22: Language performance by EMDC

The pattern of learner achievement per learning outcome across the EMDCs follows the same pattern found in the overall data for the province (Table 3.19 and Figure 3.23). In each case, learners performed best in LO5 (thinking and reasoning), followed by reading (LO3). Achievement in writing (LO4), however, was better than that for knowledge of language structure and use (LO6) in two Metropole EMDCs (1 and 4), while achievement in LO6 was higher than that for LO4 in all the "rural" EMDCs. This points towards different emphases in teaching HL. It appears that greater emphasis is given to explicit teaching of language structure in Xhosa and Afrikaans HL lessons, whereas greater emphasis may be directed towards writing in English HL lessons.

Table 3.19: Language performance by EMDC and LO

	Metro 1	Metro 2	Metro 3	Metro 4	Rural 1	Rural 2	Rural 3	Total
LO3	35	31	33	31	28	30	32	32
LO4	30	27	27	26	22	24	25	26
LO5	38	33	36	34	30	32	34	34
LO6	26	27	26	24	26	27	28	26

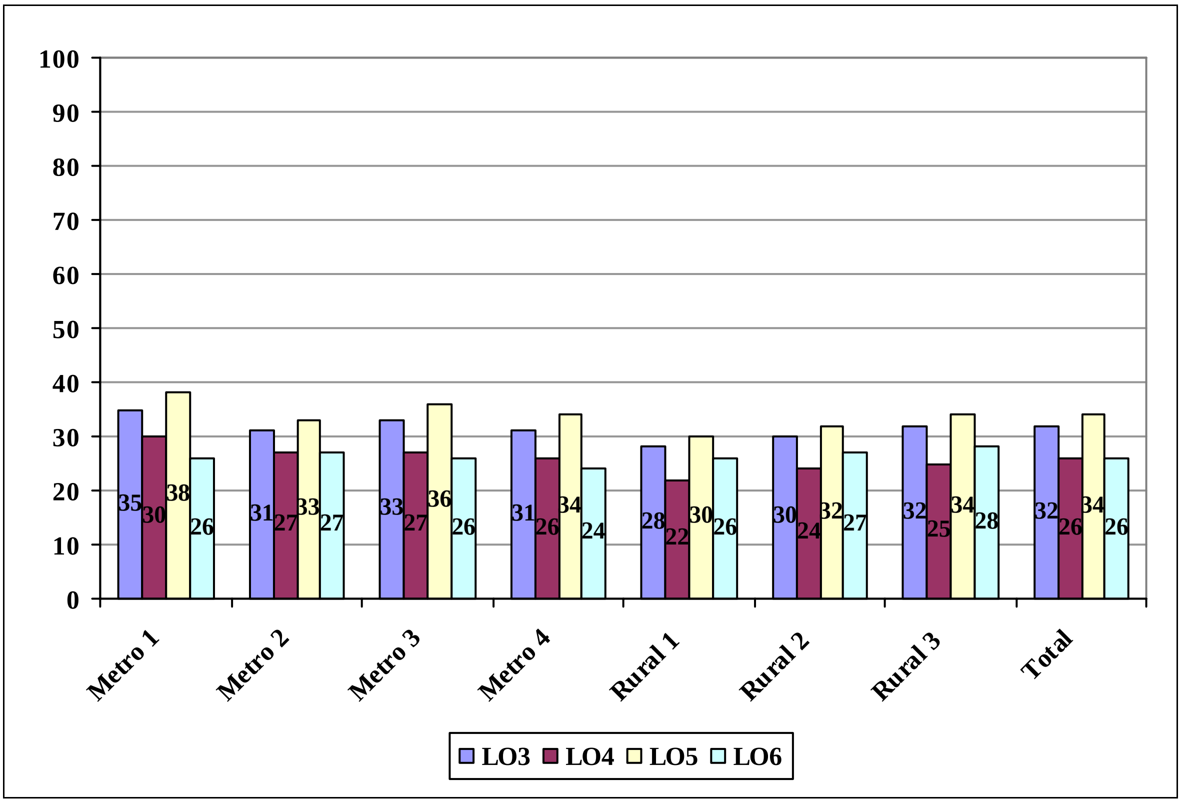


Figure 3.23: Language performance by EMDC and LO

The data reflected in Table 3.20 and Figure 3.24 show that in general the greater the number of books in the home, the higher the level of achievement per learning outcome. However, it is of significance that those learners with the greater number of books in the home tend to do comparatively best in writing tasks (LO4) which are considered to be the most difficult or challenging. Those learners who have fewer books in the home tend to achieve comparatively more highly in LO6 (language structure and use) than writing. This would suggest that learners with more books in the home are likely to achieve higher levels of academic literacy than those who have fewer books. It is the students who can write best who are likely to be able to continue through the FET band and who are more likely to be able to enter and succeed in tertiary education.

Table 3.20: Language performance of learners by learning outcome and the number of books in the home

	No books	1 to 20 books	21 to 50 books	51 to 100 books	More than 100 books	Total
LO3	27	30	35	42	48	33
LO4	19	25	30	37	42	27
LO5	29	33	38	45	50	35
LO6	22	26	29	34	38	27

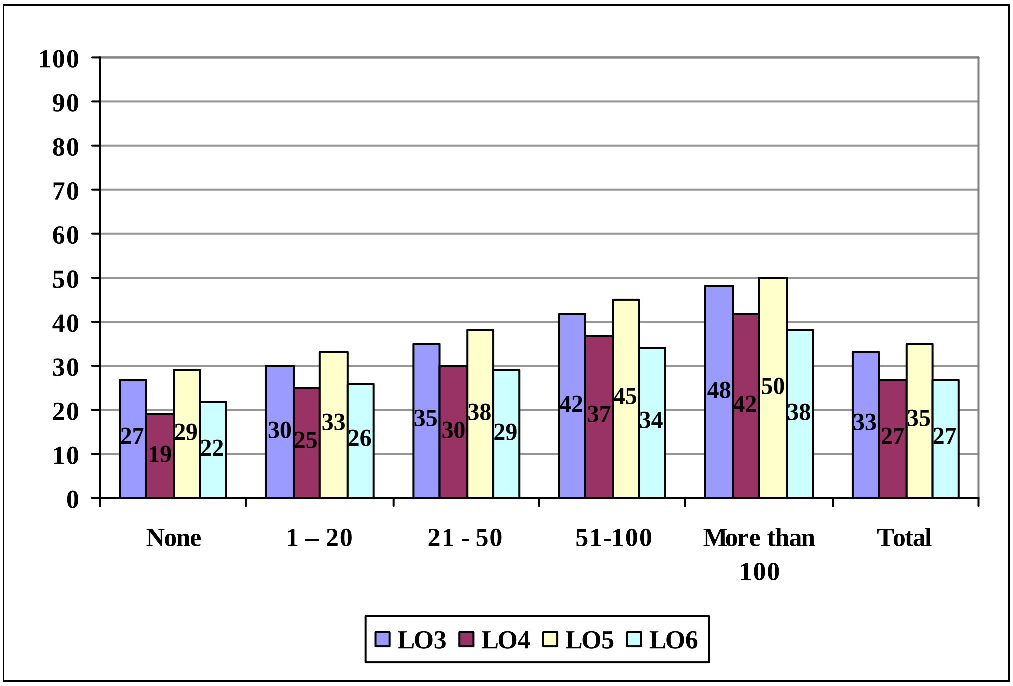


Figure 3.24: Language performance of learners by learning outcome and the number of books in the home

The data reported in Table 3.21 and Figure 3.25 follow the pattern already established in other data from this study. Learners' HL and LoLT achievement is higher amongst those who have more books in the home than those who have fewer books.

Table 3.21: Performance of HL and LoLT by number of books in the home

	No books	1 to 20 books	21 to 50 books	51 to 100 books	More than 100 books	Total
HL	33	39	43	50	54	40
LoLT	23	26	32	38	43	29
Total	28	32	37	44	48	34

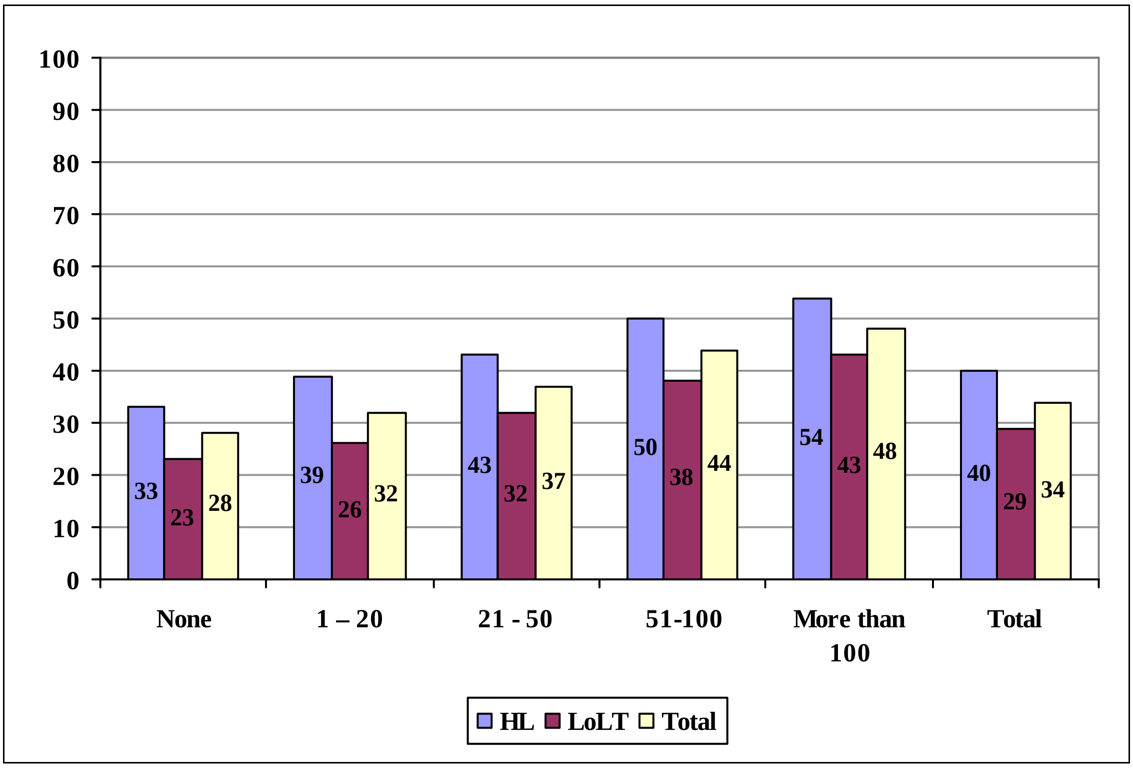


Figure 3.25: Performance of HL and LoLT by number of books in the home

The data reflected in Table 3.22 and Figure 3.26 follow the trends of earlier data in this study. Learners perform better on items covering lower grade or difficulty levels in relation to items from higher grade or difficulty levels, and do consistently more so with a higher number of books in the home, and vice versa.

Table 3.22: Language performance of learners per grade or difficulty level of the items and the number of books in the home

	No books	1 to 20 books	21 to 50 books	51 to 100 books	More than 100 books	Total
Grade 6	34	39	44	52	58	41
Grade7	31	35	41	48	53	38
Grade 8	26	31	36	42	47	33

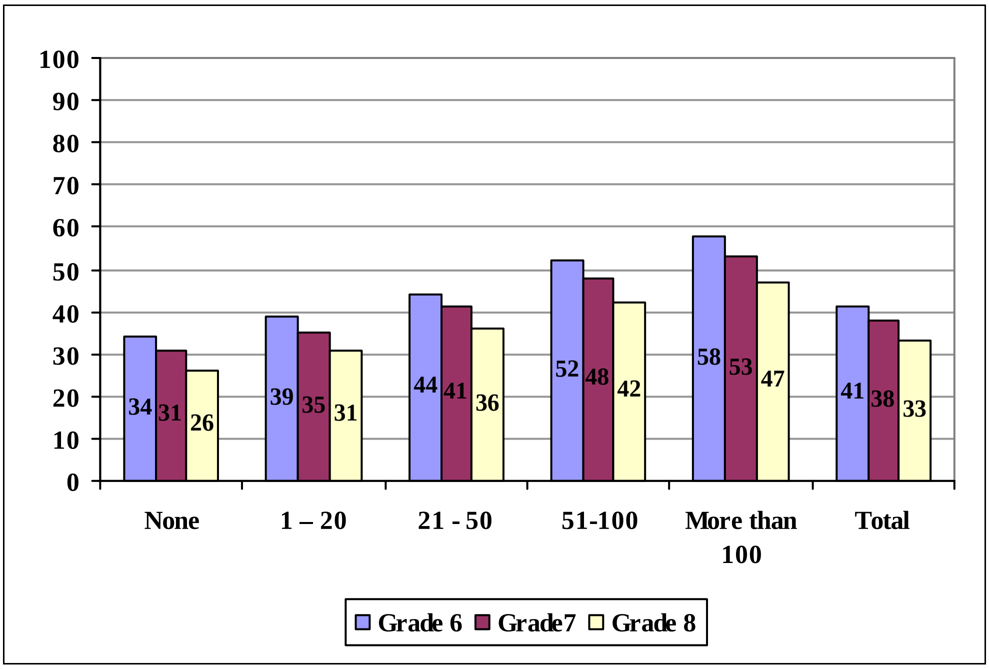


Figure 3.26: Language performance of learners per grade or difficulty level of the items and the number of books in the home

3.4.10 Very high incidence of learners exhibiting negligible/minimal evidence of literacy skills at Gr 8

During the scoring and coding of the open-ended / constructed-response items, it became obvious that an alarming number of students were repeatedly scoring at the bottom end of the spectrum and that the diagnostic codes at this end pointed to either very little or no evidence of literacy. While students appeared to have circled multiple-choice options, and clearly knew the technique for doing this, the fact that they often did not attempt to answer the open-ended questions (left the spaces blank) suggested that they could not read the questions with meaning, and/or did not know how to write an answer. Often learners attempted to answer the questions and produced what would appear at first sight to be text, but this 'text' upon examination frequently fell into the following categories:

- Rewriting of the question
- Offering part or all of a multiple-choice answer/option, selected from immediately above the CR item, or apparently randomly from somewhere else in the assessment task
- Offering an extract randomly selected from a passage in the text and inserting it inappropriately
- Producing repeated patterns of letters which do not form recognizable words in any language (e.g. mnceis nceismnceis …) (see code 97 below)
- Producing shapes which appear to be cursive writing, but are in fact penstrokes which mimic writing, and form non-recognizable words (illegible, see code 97 below).

Once the scorers/coders were alerted to this, it became necessary to add more fine detail to the lower end of the diagnostic coding. Table 3.23 illustrates the diagnostic coding rubric for the two long constructed-response (open-ended) questions at the end of the HL section of each language version as it was adapted during the coding process.

Table 3.23: Part 1 Section B (Common items) extended constructed-response items, diagnostic coding

First Digit: Value (content)	**LO 4: Writing Skills**	**Second Digit: Diagnostic Code (structure)**	**LO 6: Knowledge and Use of Language Structure**
5: Exceptional	Clear, consistent, logical writing skills	5:	No/minimal errors
4: Above average	Argument generally well presented	4:	Minor errors: spelling/punctuation
3: Average		3:	Minor errors – sentence structure
2: Below average	Lacks coherence	2:	Poor sequence of ideas/logic sentence level
1: Most unsatisfactory	Did not understand the question/completely off the point / completely inappropriate content & register; incoherent]	1:	Combination of 2,3,4; or very serious errors In many cases this also indicates learners with extremely low literacy skills. These learners require urgent attention.
99	**Not answered/not attempted**		
98	**Answer attempted and deleted**		
97	**Illegible**		Disturbing signs of inability to write (lack of literacy skills: cannot form recognisable letters or join these together in recognisable words) Learners should be assessed as soon as possible.
96 *	**Rewrote the question or rewrote a question from somewhere else in the text**		Disturbing signs of lack of ability to read and make meaning from text, uses this strategy to hide/disguise lack of literacy.

The rubric for the shorter CR items was similar, but included a possible zero (0) score as a value for the first digit, if the answer was completely wrong. The data analysis has identified the number of instances where learners were allocated any of the scores in the shaded areas, as portrayed in Table 3.24 and Figure 3.27.

Table 3.24: Language assessment task - percentages of learners unable to construct responses to nine selected items

Inability to construct responses	Number of learners	% of learners	Literacy status of learners
Very low (0 or 1 out of 9) hit rate	16 811	24	No need for concern
Low (2 or 3out of 9) hit rate	14 410	21	Insufficient evidence for concern
High (4 or 5 out of 9) hit rate	16 301	24	Disturbing
Very high (6, 7, 8 or 9 out of 9)	21 263	31	Extremely disturbing
Total	68 785	100	

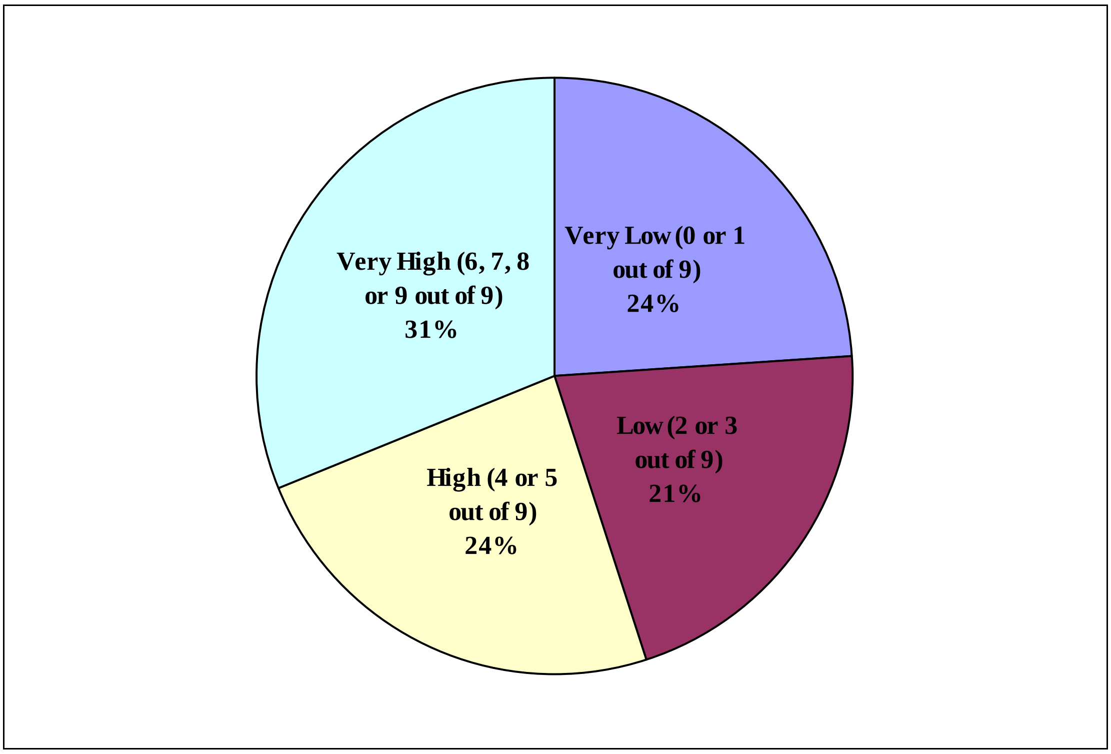


Figure 3.27: Percentages of learners with inability to construct responses to nine selected language items

In the view of the HSRC, those learners who scored in the “High” incidence (or hit rate) bracket here (4-5 out of 9) in terms of this analysis, and as indicated on the tables which are provided separately per school and per EMDC, should be identified by teachers and monitored carefully. Those who scored in the “Very High” incidence bracket here (6-9 out of 9) need to be immediately identified as learners who are at serious educational risk. These learners (who comprise 31% of the 2006 Grade 8 year group) presented evidence which is extremely disturbing. Although the HSRC acknowledges that this assessment task was not designed as a comprehensive diagnostic test for students with learning difficulties, the evidence which was found, suggests that these learners may not be able to read or write at all.

3.5 MATHEMATICS

3.5.1 Mathematics by performance levels

The mathematics instrument was developed in the two language-of-learning-and-teaching (LoLT) versions (Afrikaans and English). Learner performance in Mathematics was analysed according to the 7-point assessment scale, as published in the NCS for Further Education and Training (General) (DoE 2005), for levels of achievement as determined by the national Department of Education (DoE 2005b). The key to these levels are presented in Table 3.25.

Table 3.25: Key to performance levels

Performance level/rating code	Description of rating	Marks in percentages
7	Outstanding performance	80 - 100
6	Meritorious performance	70 - 79
5	Substantial performance	60 - 69
4	Adequate performance	50 - 59
3	Moderate performance	40 - 49
2	Elementary performance	30 - 39
1	Not achieved	0 - 29

Results according to these levels are presented in Figure 3.28.

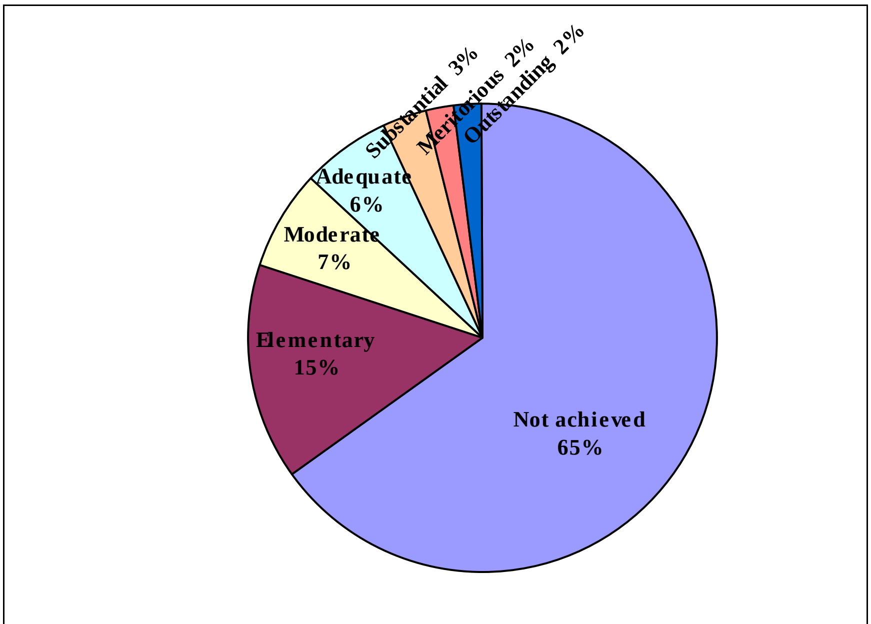


Figure 3.28: Percentage of learners at each performance level in Mathematics

Analysis of Figure 3.28 indicates that very few learners scored at the higher performance levels. Only 7% of the learners performed at the top three levels (above 60%). Of these learners, 1 295 achieved between 69% and 79% and 1 153 achieved 80% or above. Successful advance to tertiary studies in Mathematics, Science and Engineering depends on high performance levels at this stage. These low averages achieved by the best performing mathematics learners is alarming because high performance levels in Grade 8 are considered a good indicator of good performance in the FET band. The results are therefore a strong indication that very few of the learners in this grade will perform well in the FET band.

The majority of the learners (77%) did not achieve the required mathematics knowledge and skills that were assessed and scored below 50% in the instrument. It is expected that these learners will experience serious problems with Mathematics as subject in secondary education. The significance of these poor results is enormous for mathematics performance in the Province and these learners should receive urgent assistance to improve their competency in basic mathematics.

Table 3.26: Mathematics performance by achievement level and EMDC

	Not Achieved	**Elementary**	**Moderate**	**Adequate**	**Substantial**	**Meritorius**	**Out-standing**
Metro 1	54	17	10	8	5	3	3
Metro 2	73	14	6	4	2	1	1
Metro 3	58	16	10	8	4	2	2
Metro 4	67	16	8	5	2	1	1
Rural 1	74	13	5	4	2	1	1
Rural 2	70	14	6	5	2	2	1
Rural 3	63	15	7	6	4	3	2
Provincial average	65	15	8	6	3	2	2

Across the EMDCs similar trends as in the province overall were observed, with the majority of the learners performing at the “Not Achieved” level. However, as shown in Table 3.26 and Figure 3.29, relatively large differences were noted across the EMDCs. Learner achievement across the seven EMDCs range from 4% to 11% in total at the three top performance levels, and from 54% to 74% at the “Not achieved” level across the seven EMDCs.

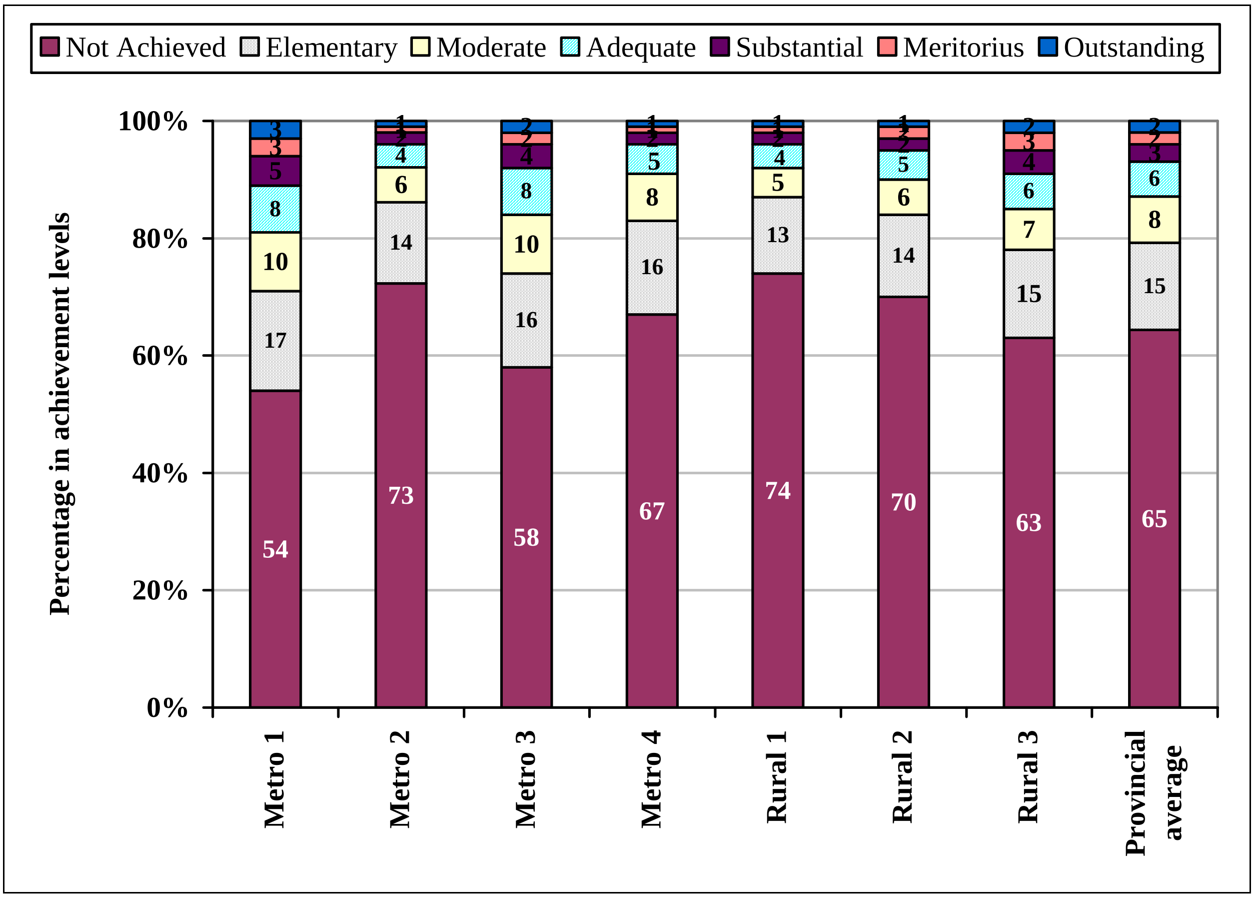


Figure 3.29: Mathematics performance by achievement level and EMDC

Other important features of Figure 3.29 are:

- Metro 1, a "city" district, is the best performing EMDC with 19% of the learners performing at an "Adequate" level and above and the smallest percentage of learners (54%) performing at "Not achieved" level.
- After Metro 1, the next highest performing EMDC is Metro 3, another largely "city" district, where 16% of the learners performed at an "Adequate" level and above and 58% performed at "Not achieved" level.
- The two lowest performing EMDCs is Rural 1 and Metro 2 with 74% and 73% of the learners respectively performing at the "Not achieved" level.

Reasons for the range in performance in different districts (EMDCs) may differ from district to district. It has been suggested to the Department that the low performance level should be investigated at district level so that it can be addressed appropriately.

3.5.2 Mathematics performance by Learning Outcomes

Mathematics performance by Learning Outcomes is illustrated in Figure 3.30.

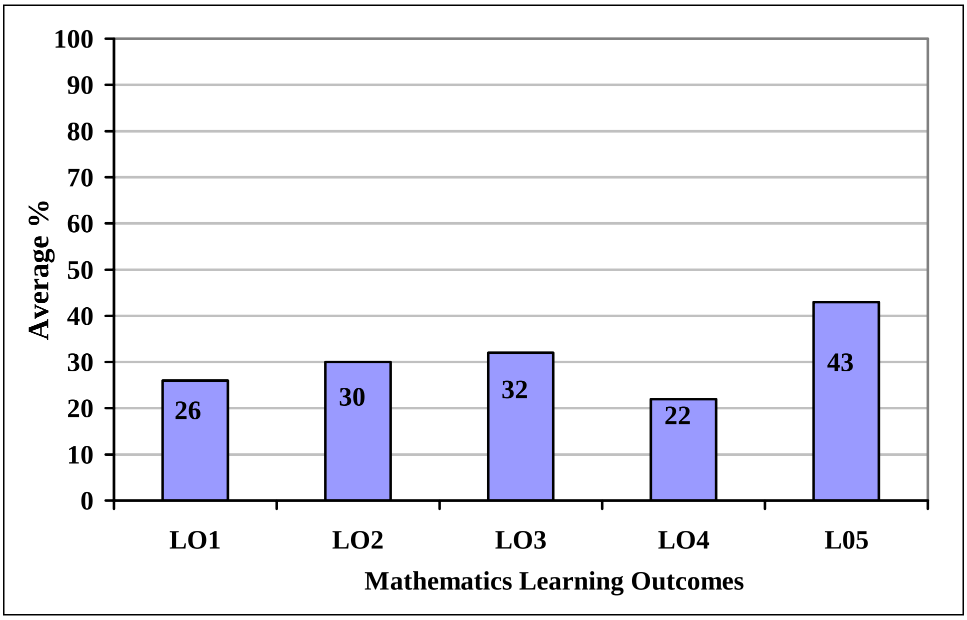


Figure 3.30: Mathematics performance by Learning Outcome

Key to learning outcomes:
LO1 = Numbers, Operations and Relationships
LO2 = Patterns, Functions and Algebra
LO3 = Shape and Space (Geometry)
LO4 = Measurement
LO5 = Data Handling

As can be seen in Figure 3.30, learners performed poorly in all learning outcomes. The highest average score (43%) was in Data Handling (LO5) and the lowest (22%) was in Measurement (LO4). The knowledge and skills required for the performance of Numbers, Operations and Relationships (LO1) are crucial for the attainment of the other learning outcomes; hence the low proficiency in this learning outcome may have contributed to the overall poor performance in Mathematics. The poor performance for LO1 is a cause for concern because numbers, operations and relationships deal with the basic concepts and skills that are used in the other learning outcomes. It was expected that the Western Cape learners would perform poorly in LO1 because analysis of the performance levels of South African learners in other studies (e.g. TIMSS 2003, QLP and Grade 6 Systemic Evaluation) confirms that many of our learners, even in higher grades, experience problems with basic calculations and operations. Educators should therefore pay special attention to the attainment of basic numeracy skills to ensure progress in the other areas of Mathematics.

Learner Performance

The performance of learners per LO across the EMDCs are presented in Table 3.27 and illustrated in Figure 3.31.

Table 3.27: Mathematics performance by Learning Outcomes and EMDC

	Metro 1	Metro 2	Metro 3	Metro 4	Rural 1	Rural 2	Rural 3	Total
LO1	29	23	27	25	22	24	27	26
LO2	34	27	32	29	26	29	32	30
LO3	37	27	36	30	27	30	33	32
LO4	24	19	24	20	20	21	23	22
LO5	51	38	49	42	37	40	47	44

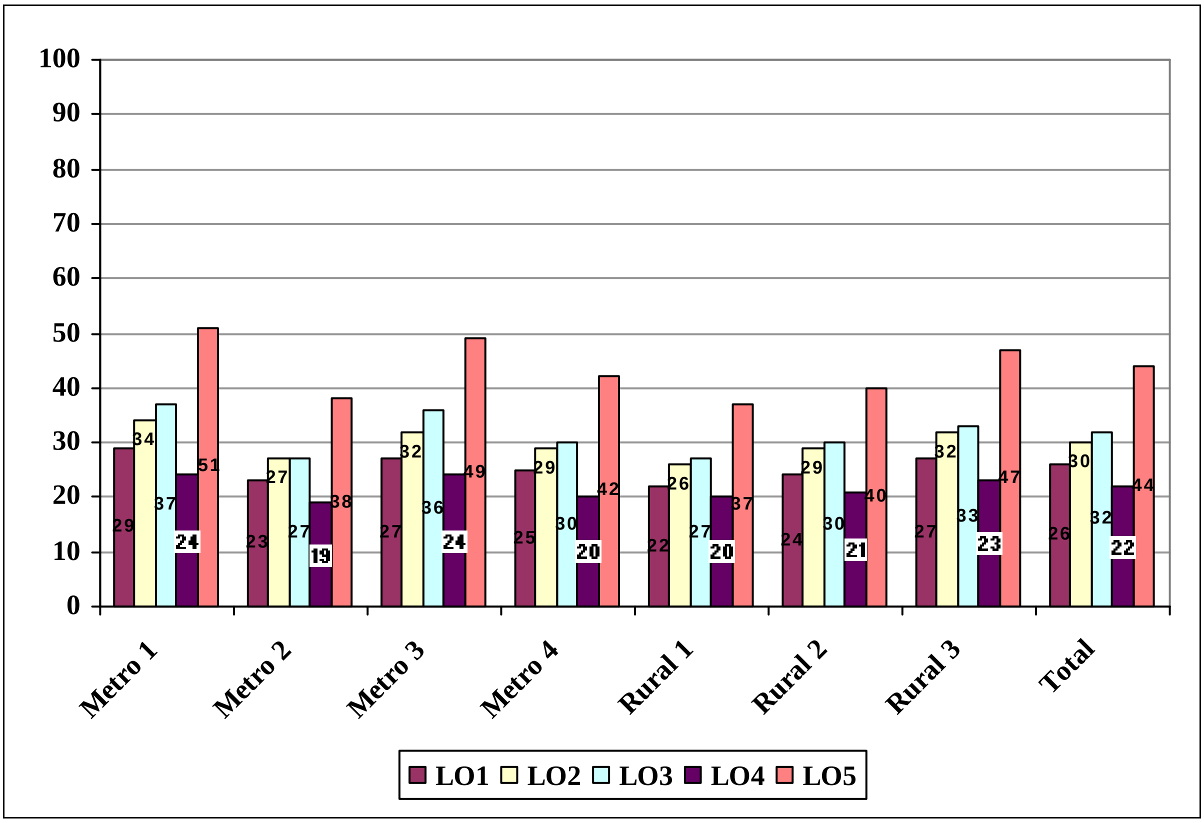


Figure 3.31: Mathematics performance by EMDC and LO

The performance per LO in the EMDCs mirror the performance in the Province with the highest average score achieved in Data Handling (LO5) and the lowest average scores achieved in Measurement (LO4) in all the EMDCs. Learners in two Urban/Metropole EMDCs (1 and 3) obtained higher average scores across all learning outcomes. Among the two lowest performing EMDCs across learning outcomes were one each from an urban and a rural environment, though (Metro 2 and Rural 1).

Learner Performance

The significance of the results per LO across the EMDCs is that in only one of the EMDCs learners were able to achieve an average of above 50% (Metro 1 with 51% in LO1). Learner performance across all the LOs in all the EMDCs is not acceptable and does not provide a positive indicator for success in further education. The differences between EMDCs should also receive attention and the Department has been alerted to the fact that these should be investigated further.

3.5.3 Mathematics performance by Cognitive Domain

For this study, items for the Mathematics instrument were classified according to the four cognitive domains of *Knowing Facts and Procedures, Using Concepts, Solving Routine Problems* and *Reasoning.* The cognitive domains define the behaviours expected from learners as they engage with mathematics content.

Of the fifty-four items included in the instrument, fifteen assessed knowledge of facts and procedures; twelve were on using concepts; eleven covered solving routine problems; and six were on reasoning. The performance of learners according to the cognitive levels is presented in Figure 3.32.

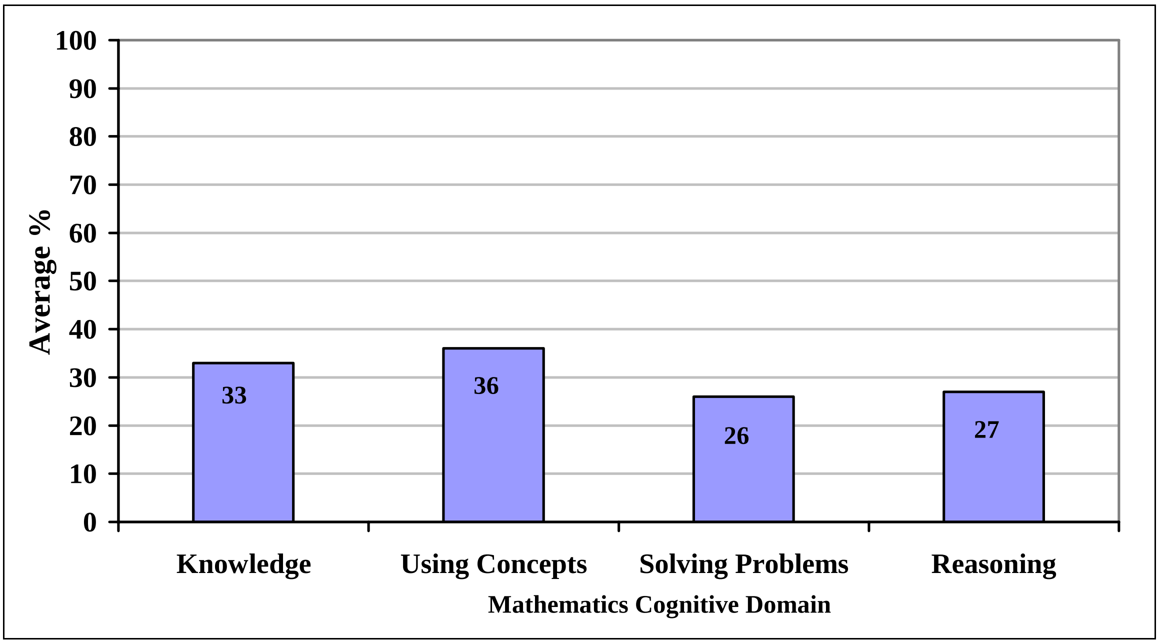


Figure 3.32: Mathematics performance by Cognitive Domain

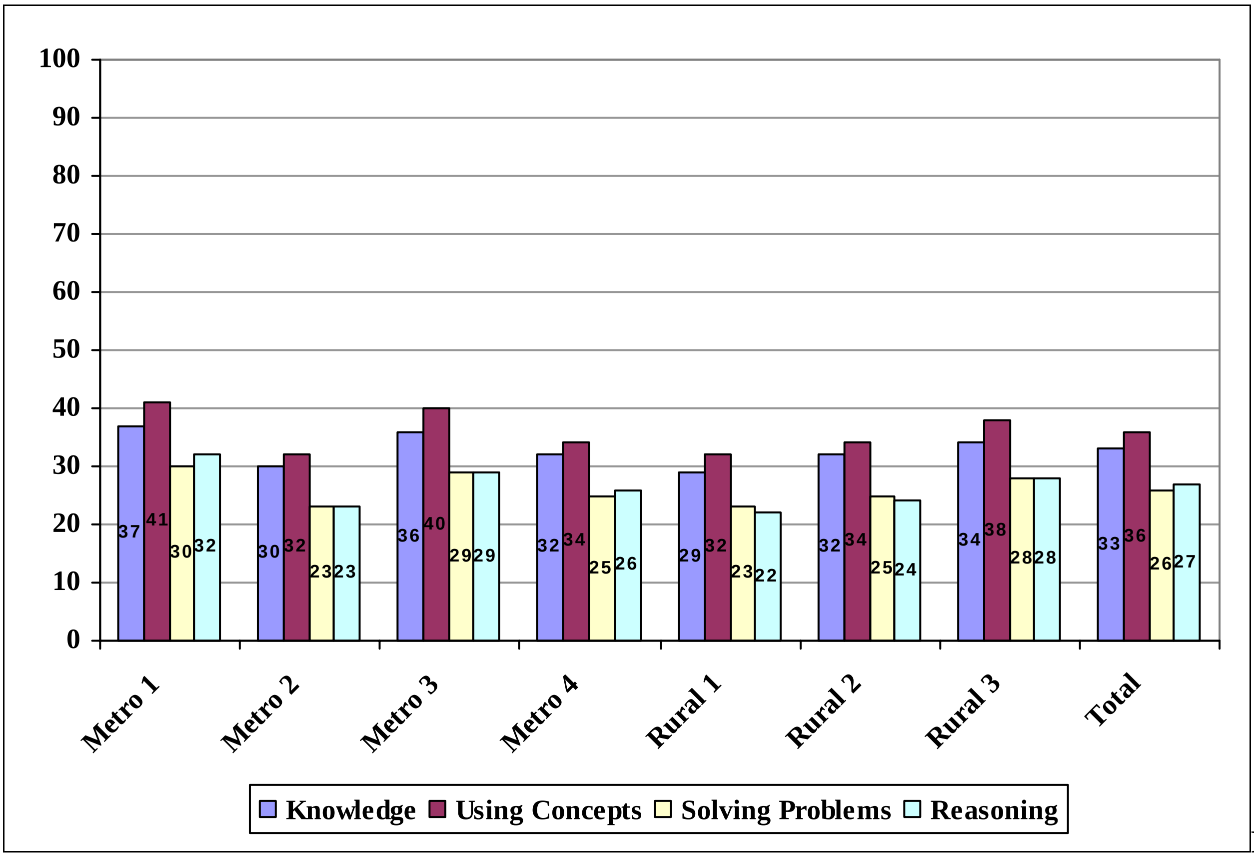
There exists a hierarchical connection between the cognitive domains. *Knowing facts and procedures* domain is considered to be at a lower cognitive level, followed by *Using Concepts and Solving Routine Problems. Reasoning* is considered to be the highest cognitive domain. In an ideal instrument learners generally perform better in the lower cognitive skills than in the complex cognitive procedures of reasoning. The results of this study did not follow the hierarchical pattern of cognitive domains. Learners performed better on *Using concepts* items than on *Knowledge* items. Ideally they should have performed best on *Knowledge* items because that is considered to be the lower cognitive level. These results correspond with the findings of low performance in the learning outcome on numbers, operations and relationships which indicate that learners do not possess basic numerical knowledge and skills. These results are also alarming because the hierarchical connection between the cognitive domains implies that lacking the cognitive skills of *Knowing facts and procedures* may effect the performance of learners at the higher cognitive levels.

Western Cape learners also performed slightly better on *Reasoning* items than on items on *Solving routine problems.* Although it may appear as if learners performed better in the reasoning items compared to their performance in other domains, the difference (anomaly) may be due to the fact that very difficult reasoning items were not included in the instrument, rather than to a real difference in performance. The selection of reasoning items for the instrument was complicated by the fact that the most difficult reasoning items included in the TIMSS study (which was the main source for the development of items) were answered incorrectly by almost all South African learners and therefore the inclusion of these items in the instrument were avoided. It is important to note that the exclusion of the most difficult items from the *Reasoning* domain will not have an effect on the instrument as a whole because a sufficient spread of difficulty levels were attained in the instrument. (Such floor or ceiling effects normally are avoided quite actively because they result in the restriction of the range of scores, which seriously hampers analyses.)

The performance of learners by cognitive domain across the EMDCs is presented in Table 3.28 and in Figure 3.33.

Table 3.28: Mathematics performance by EMDC and Cognitive Domain

	Metro 1	**Metro 2**	**Metro 3**	**Metro 4**	**Rural 1**	**Rural 2**	**Rural 3**	**Total**
Knowledge	37	30	36	32	29	32	34	33
Using Concepts	41	32	40	34	32	34	38	36
Solving Problems	30	23	29	25	23	25	28	26
Reasoning	32	23	29	26	22	24	28	27



Fi

Figure 3.33: Mathematics performance by EMDC and Cognitive Domain

An analysis of the data split according to EMDCs shows similar patterns as those reflected in the preceding graph and table on performance across the cognitive domains for the Province. The performance across EMDCs ranked mostly in the same order as in the Province for all cognitive domains.

3.5.4 Mathematics performance by gender

The performance of boys and girls is compared in terms of EMDC and the averages are reported in Figure 3.34.

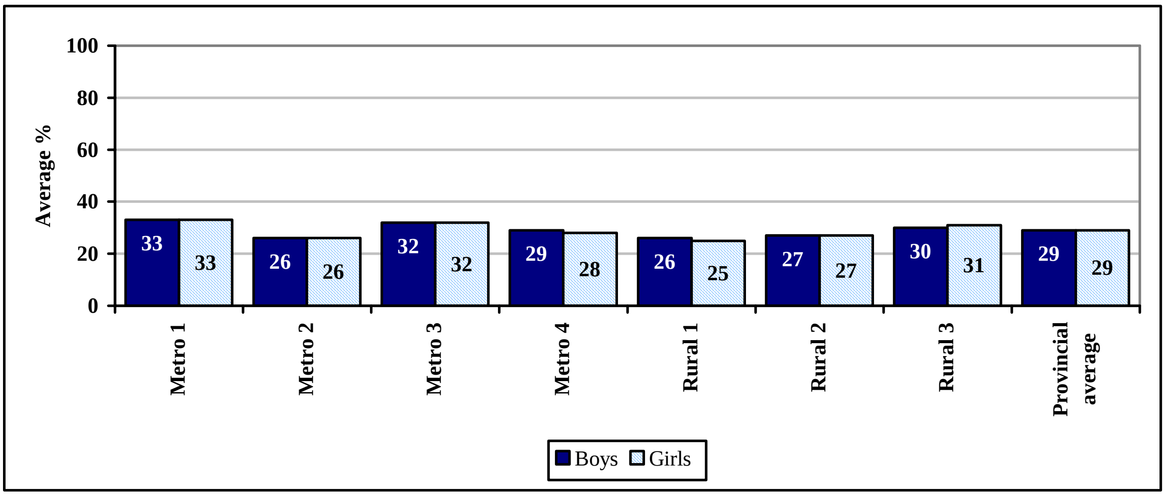


Figure 3.34: Mathematics performance by gender and EMDC

There was no difference in the average score for mathematics for boys and girls (29% for both groups). Across the EMDCs similar trends were observed, with only a one percentage point difference between boys and girls in three of the EMDCs, irrespective of being urban or rural.

Table 3.29: Mathematics performance by gender and learning outcome (LO)

	LO1	LO2	LO3	LO4	LO5
Boys	26	30	32	22	44
Girls	26	31	32	22	44
Total	26	31	32	22	44

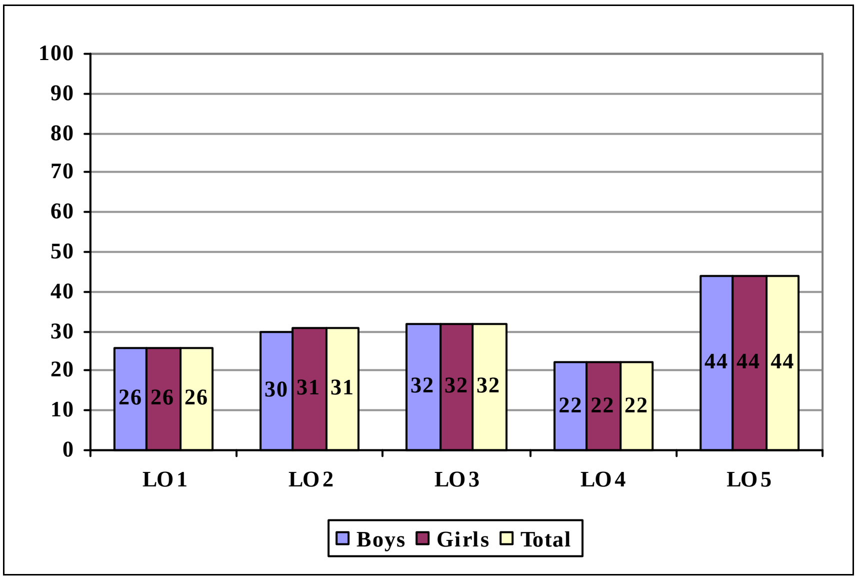


Figure 3.35: Mathematics performance by gender and learning outcome (LO)

Learner Performance

The data reported in Table 3.29 and Figure 3.35 for a gender-based split amongst the different learning outcomes show the same pattern found in the overall data for the Province with no real difference between the average scores for boys and girls.

The performance of boys and girls in relation to the cognitive domains are presented in Table 3.30 and illustrated in Figure 3.36.

Table 3.30: Mathematics performance by gender and Cognitive Domain

	Knowledge	Using Concepts	Solving Problems	Reasoning
Boys	33	37	26	26
Girls	34	36	27	28
Total	33	36	26	27

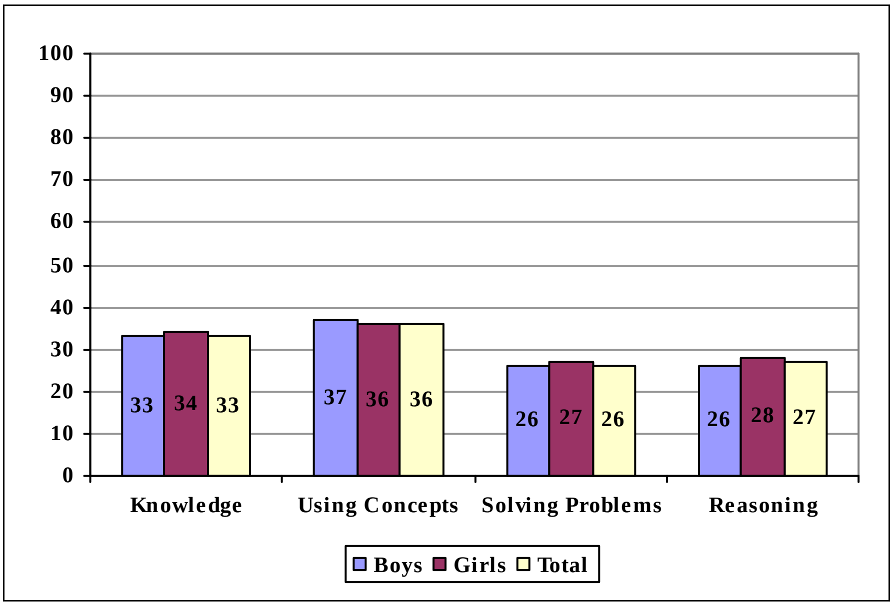


Figure 3.36: Mathematics performance by gender and Cognitive Domain

The gender split across cognitive domains follows the trends of earlier data reported for the Province with very small differences between the average scores of boys and girls.

3.5.5 Mathematics performance and language

It is important to investigate how LoLT relates to the performance of learners in Mathematics. The LoLT section of the language instruments were designed to measure the extent to which learners might be able to read across the curriculum, to understand text and to produce written evidence of having engaged with the text (see the language framework in Chapter 2). The present section on the role of language in performance on the mathematics instrument was included with the intention to provide some diagnostic evidence relating to language achievement, or (academic) literacy, across the curriculum. The performance of learners in mathematics is therefore interpreted in relation to the performance of learners in the LoLT section of the language instrument (Part 2 of the language instrument). It could be expected that the low average scores in the LoLT section of the language test for all three language groups would have a negative impact on mathematics performance.

3.5.5.1 Mathematics performance by languages of instruction (LoLT)

The mathematics instrument was administered in the languages of instruction (LoLT) - English or Afrikaans. The English-speaking learners and the majority of the Xhosa-speaking learners wrote the mathematics test in English and attained a combined average score of 31%. These learners (with average language performance scores of 35% and 19% for the English LoLT and Xhosa LoLT sections respectively) performed slightly better in mathematics than the Afrikaans LoLT group, who averaged with 27% in mathematics, and also had a similar average language score of 26% for the Afrikaans LoLT section. However, the achievement levels are disappointing for both LoLT groups and are of significant concern.

Table 3.31 and Figure 3.37 illustrate mathematics performance by LoLT (also test language) across the learning outcomes.

Table 3.31: Mathematics performance by LoLT and Learning Outcome

	LO1	LO2	LO3	LO4	LO5
Afrikaans	24	28	29	21	40
English	28	32	34	22	47
Total	26	30	32	22	44

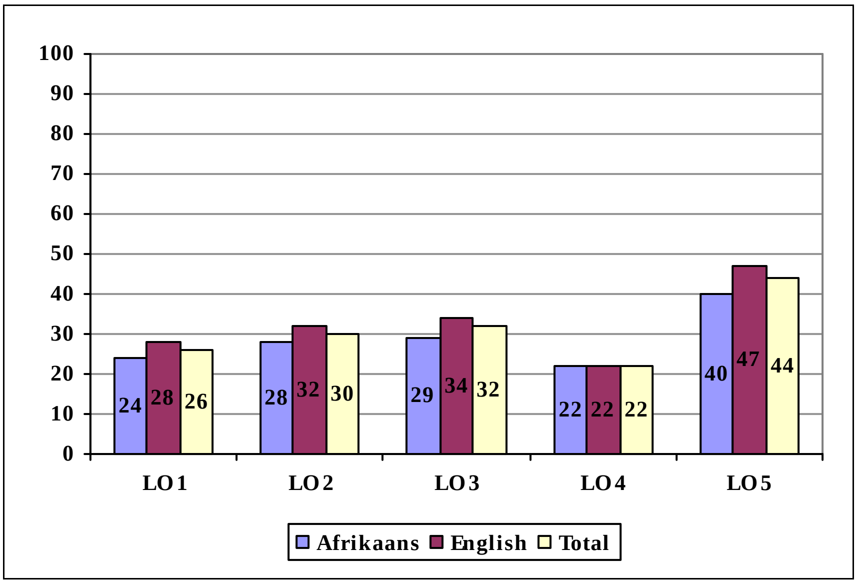


Figure 3.37: Mathematics performance by LoLT and Learning Outcome

The data reported in Figure 3.37 in terms of a LoLT (test language) split across the different learning outcomes show to a great extent the same pattern that was found in the overall data for the Province. The difference between the average scores for the different language groups is small with the largest difference on the items on data handling (LO5). This may indicate that the items in the different learning outcomes are not strongly or systematically biased in favour of either the Afrikaans or English LoLT group. (However, there may also be suggestions that those learners mastering English as LoLT reap the benefit of higher mathematics achievement on completion of a mathematics assessment in English.)

The performance by LoLT (also test language) across the cognitive domains are presented in Table 3.32 and illustrated in Figure 3.38.

Table 3.32: Mathematics performance by LoLT and Cognitive Domain

	Afrikaans	English	Total
Knowledge	31	35	33
Using Concepts	33	38	36
Solving Problems	25	27	26
Reasoning	24	29	26

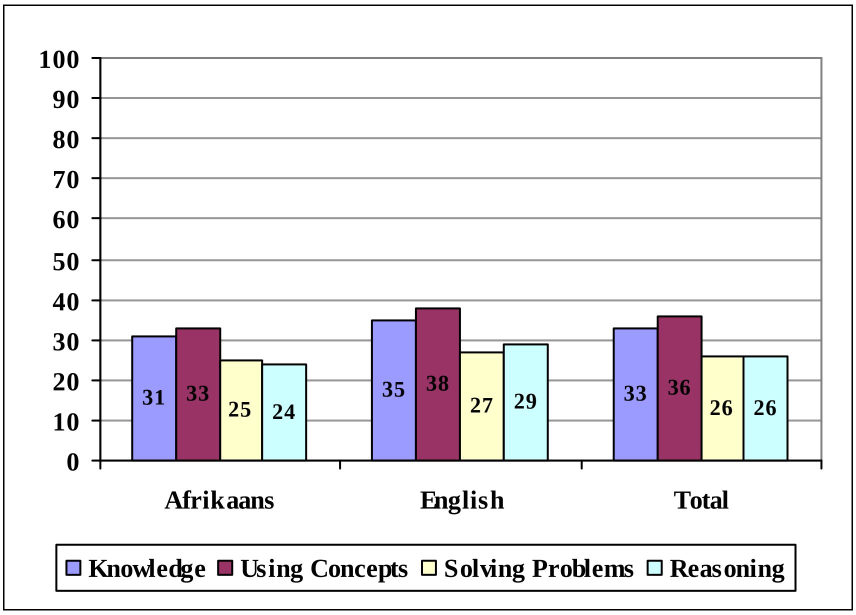


Figure 3.38: Mathematics performance by LoLT and Cognitive Domain

An analysis of the data split according to LoLT and cognitive domain shows similar patterns to those for the Province with the performance across languages ranked mostly in the same order for the cognitive domains as for the Province.

3.5.5.2 Mathematics performance by home language

Figure 3.39 illustrates Mathematics performance as split across the three home-language groups.

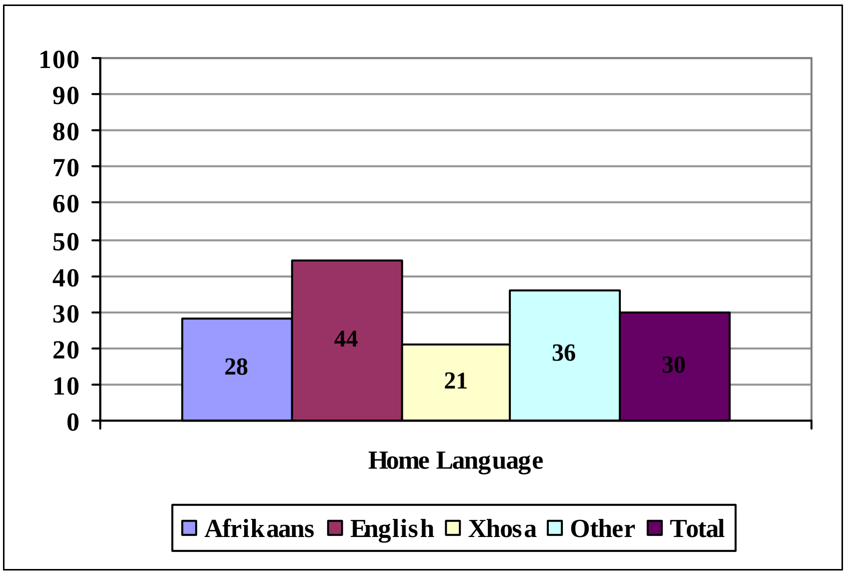


Figure 3.39: Mean mathematics scores across the home languages

An analysis of the data shows that English-speaking learners (44%) outperformed both Afrikaans- (28%) and Xhosa-speaking (21%) learners. This pattern of performance corresponds with the performance predicted in the LoLT sections of the language tests where the English-speaking group performed best with an average score of 32%, followed by the Afrikaans-speaking group with an average score of 27% with the Xhosa-speaking group performing worst with an average score of 18%.

Mathematics performance by home language across the learning outcomes is presented in Table 3.33 and illustrated in Figure 3.40.

Table 3.33: Mathematics performance by home language and LO

	LO1	LO2	LO3	LO4	LO5
Afrikaans	25	30	31	22	43
English	36	45	51	32	71
Xhosa	22	23	22	15	29
Other	35	40	40	26	51
Total	27	32	33	23	46

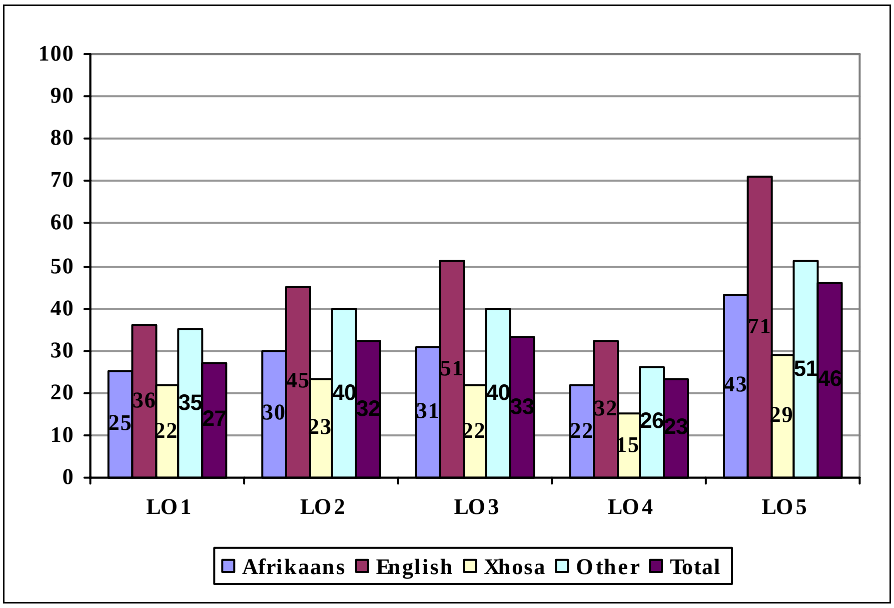


Figure 3.40: Mathematics performance by home language and LO

The data for a home-language split across the different learning outcomes show to a great extent the same pattern found in the overall data for the Province. However, the English-speaking learners perform relatively much better in LO5 compared to the performance for the other home-language groups. Although no evident explanation can be suggested, possible causes, e.g. differences in the standard and suitability of mathematics texts books in English and Afrikaans, should be explored.

Table 3.34 and Figure 3.41 present Mathematics performance by home language and Cognitive Domain.

Table 3.34: Mathematics performance by home language and Cognitive Domain

	Knowledge	Using Concepts	Solving Problems	Reasoning
Afrikaans	32	35	26	25
English	49	54	39	44
Xhosa	26	28	18	18
Other	42	46	32	36
Total	35	38	27	28

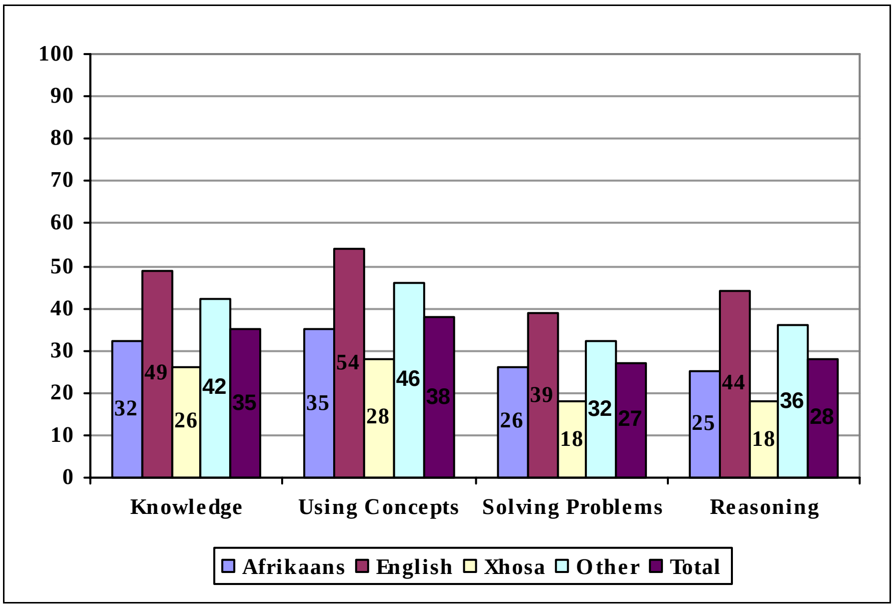


Figure 3.41: Mathematics performance by home language and Cognitive Domain

Mathematics performance across home language and cognitive domain follows a regular and more or less consistent pattern in relation to the differences between English-, Afrikaans- and Xhosa-speaking learners.

3.5.5.3 Comparison of Mathematics performance by LoLT and home language

The differential effect of home language and LoLT on learners' performance in Mathematics was investigated next. Learners were categorised according to whether or not their home language was the same as their LoLT.

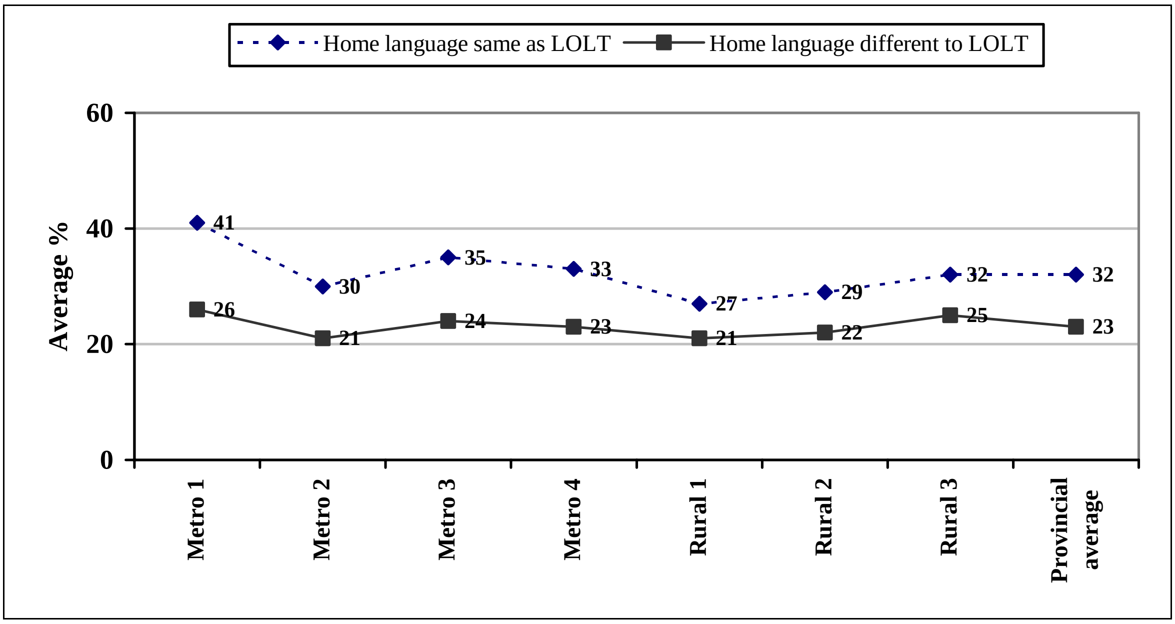


Figure 3.42: Mathematics achievement in terms of correspondence between home language and LoLT

As shown in Figure 3.42, when learners had the same home language as the LoLT in which they completed the instrument, they obtained scores that were nine percentage points (provincial averages) higher than the scores of learners for whom the LoLT was different to their home language. Across the EMDCs, the pattern was consistent with the provincial picture, and the differences between the two sub-groups ranged from 6 to 15 percentage points. The differences revealed by similar previous studies were usually larger (the national Grade 6 systemic evaluation shows a difference of 17 percentage points between the two groups). The strategy of the HSRC of using explanations of key instructions in learners' home languages (glossed items) for some of the Mathematics items may have contributed to the smaller differences between the two groups. It is suggested that the Department explores this strategy further.

3.5.6 Learner performance related to glossed items

To obtain information on whether learners who wrote the instrument in a language other than their home language benefited from the explanation of key instructions in their home language (glossed items), they were asked to indicate whether they made use of these translations and whether the translations helped them in answering the items. Although only 41 978 out of the 75 059 learners answered these questions, valuable information with regard to the need to address language issues in Mathematics assessment could be derived. Table 3.35 presents the numbers and percentages of learners that used the translations and Table 3.36 presents the numbers and percentages of learners that reported benefiting from the translations.

Table 3.35: Numbers and percentages of mathematics learners using translations

			Afrikaans	English	Xhosa	Other	Total
Afrikaans LoLT	Yes	N	13 855	211	272	21	14 359
		%	60	75	81	72	61
	No	N	9 208	69	64	8	9 349
		%	40	25	19	28	39
	Total	N	23 063	280	336	29	23 708
		%	100	100	100	100	100
English LoLT	Yes	N	638	1 988	6 046	116	8 788
		%	57	22	79	40	48
	No	N	486	7 212	1 611	173	9 482
		%	43	78	21	60	52
	Total	N	1 124	9 200	7 657	289	18 270
		%	100	100	100	100	100

Table 3.35 shows that the numbers of learners who made use of the glossed items for the Afrikaans instrument (with instructions translated into English and Xhosa) were 13 855, 211 and 272 respectively for the Afrikaans-home language, English-home language and Xhosa-home language groups. The number of learners who said that they made use of the translations provided for the English instrument (instructions translated in Afrikaans and Xhosa) were 638, 1 988 and 6 046 respectively for the Afrikaans-home language, English-home language and Xhosa-home language groups. In total, 57% of the learners that answered this question reported that they used the translations.

The responses to the follow-up question to the 57% of learners that responded positively to the first question are represented in Table 3.36.

Table 3.36: Numbers of mathematics learners reportedly benefiting from the translations

			Afrikaans	English	Xhosa	Other	Total
Afrikaans LoLT	Yes	N	11507	169	212	18	11906
		%	85	81	80	90	85
	No	N	2006	40	52	2	2100
		%	15	19	20	10	15
	Total	N	13513	209	264	20	14006
		%	100	100	100	100	100
English LoLT	Yes	N	525	1569	5179	100	7373
		%	85	81	88	89	87
	No	N	94	360	682	12	1148
		%	15	19	12	11	13
	Total	N	619	1929	5861	112	8521
		%	100	100	100	100	100

Table 3.36 shows that the number of learners who believes that they benefited from the translations in the glossed items for the Afrikaans instrument (instructions translated in English and Xhosa) were 11 507 (85%), 169 (81%) and 212 (90%) respectively for the Afrikaans, English and Xhosa

home-language groups. The number of learners who said that they benefited from the translations for the English instrument (instructions translated in Afrikaans and Xhosa) were 525 (85%) , 1 569 (81%) and 5 179 (89%) respectively for the Afrikaans, English and Xhosa home-language groups. More than 80% of the learners in all language groups answered positively that they benefited from the translations.

The performance of learners on the glossed and non-glossed items are presented in Table 3.37 and illustrated in Figure 3.43.

Table 3.37: Performance on glossed and un-glossed mathematics items

	MCQ	Total
Glossed items	35	32
Un-glossed items	32	27

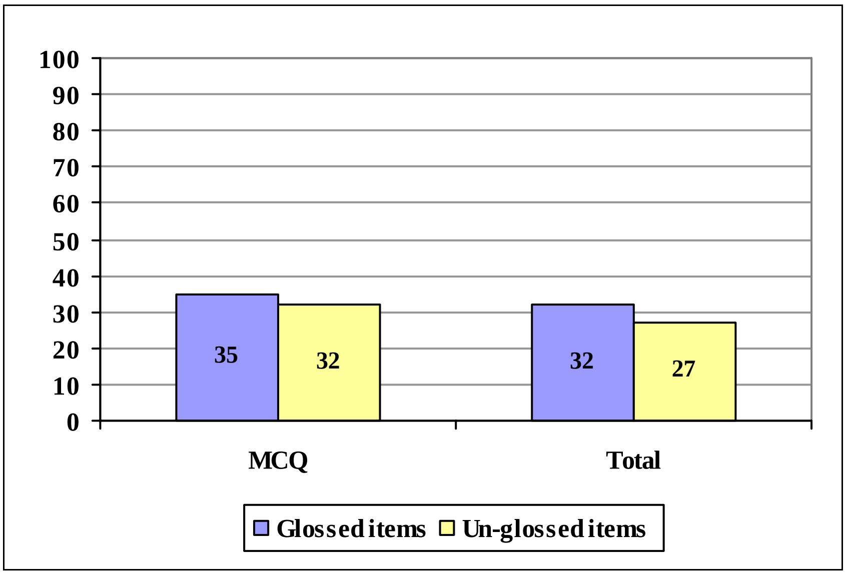


Figure 3.43: Performance in glossed and un-glossed mathematics items

The data show a small positive difference in favour of the average performance of learners on the glossed items compared to the average performance of learners on the non-glossed items. However, the fact that learners believe that they benefited from the translations, even though the difference is perceived as small, suggests that this matter should be investigated further. It has to be remembered that the two sets of items in the present instance have been matched only at face value, without detailed evidence of equivalence. Two specially designed parallel instrument sets have to be admistered. One could either pair or match sets of glossed (translated) and non-glossed (translated) items of equivalent difficulty value and administer such an instrument to the same sample of learners, or administer a glossed and non-glossed version of the test to two samples of learners that are paired or matched precisely on ability level to establish how much such assistance helps to improve learner performance.

3.5.7 Mathematics performance by item type

The mathematics instrument consists of 35 multiple-choice (MC) items and 19 constructed-response (CR) items. The average score obtained for the multiple-choice item sub-total was 34%, which was considerably higher than the average of 19% obtained for the constructed-response item sub-total. The average performance on the individual MC items ranged from 12% to 56% while the average performance for the individual CR items ranged from 3% to 56%. Learners achieved an average of 40% or above on only 11 of the 35 MC items. On the CR items, learners performed very poorly with only two of the 19 items' averages above 40%. Figure 3.44 provides a profile of the percentage of learners who got each item correct.

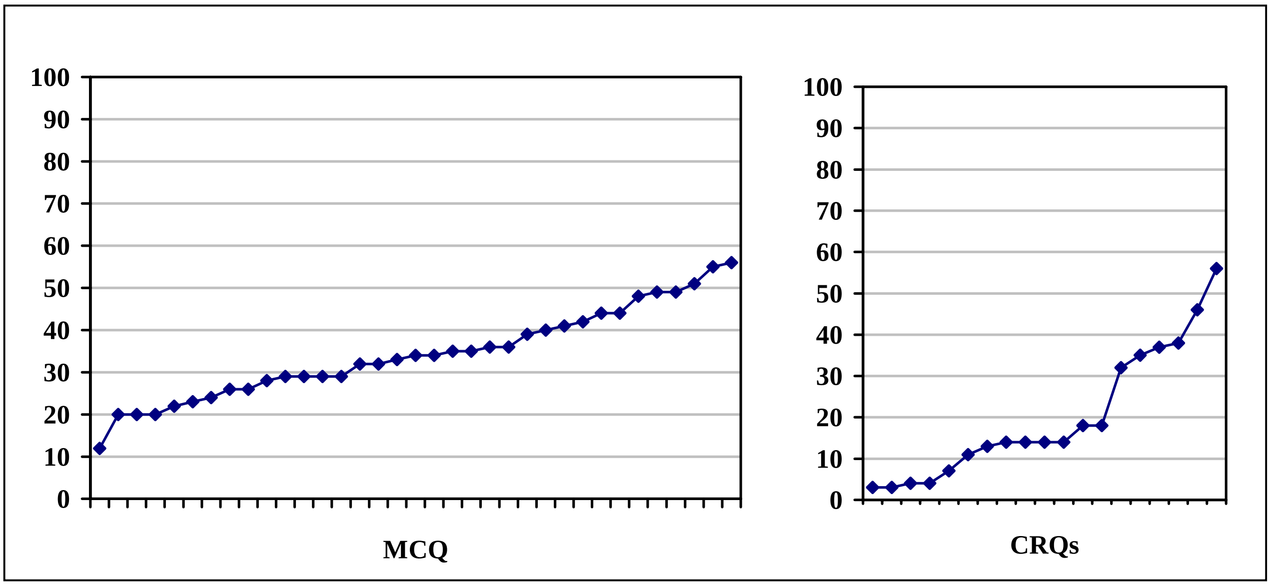


Figure 3.44: Percentages of learners that got each individual mathematics item correct per item type

In Figure 3.45 it can be seen that the same pattern of performance was observed by item type for both boys and girls with almost no difference between the two groups.

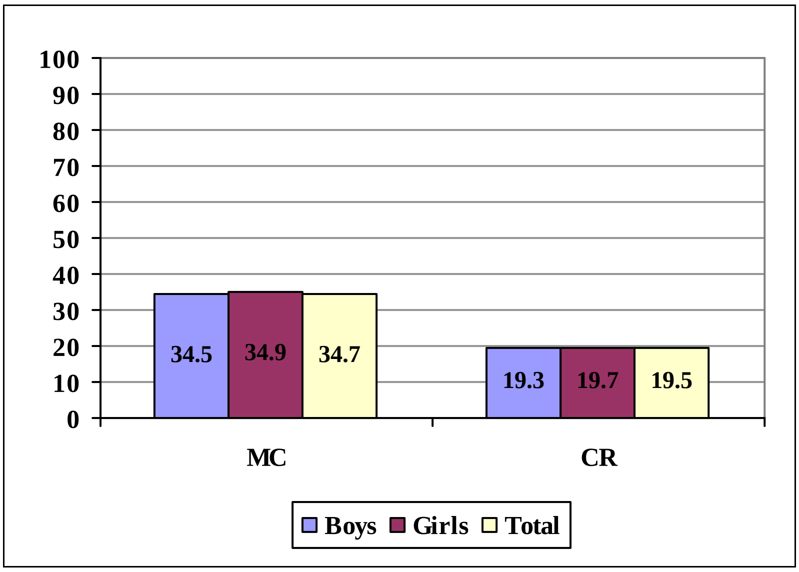


Figure 3.45: Mathematics performance by gender and item type

Mathematics performance by item type across the EMDCs is presented in Table 3.38 and illustrated in Figure 3.46.

Table 3.38: Mathematics performance by EMDC and item type

	Metro 1	Metro 2	Metro 3	Metro 4	Rural 1	Rural 2	Rural 3	Total
MC	38	32	37	33	31	33	36	34
CR	24	15	22	18	16	17	21	19

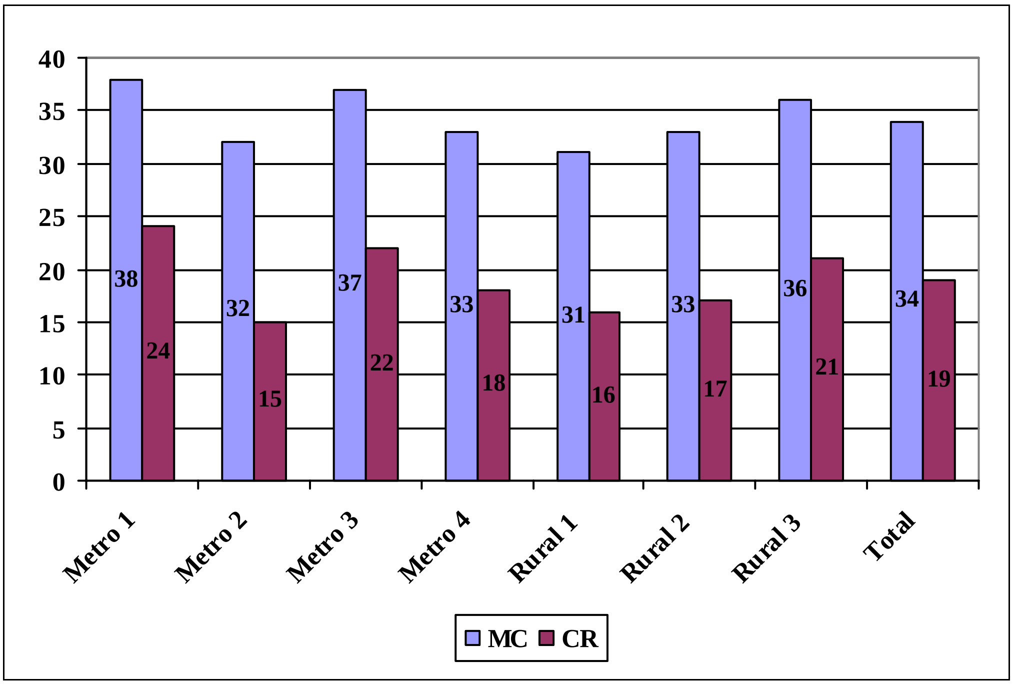


Figure 3.46: Mathematics performance by EMDC and item type

Learner Performance

The item-type split across the EMDCs shows the same pattern found in the overall data for the Province and follows a regular and consistent pattern in relation to the difference between multiple-choice and constructed-response items.

Table 3.39: Mathematics performance by item type and home language

	Afrikaans	English	Xhosa	Other	Total
Multiple-choice	34	49	28	43	36
Constructed-response	19	35	10	26	21

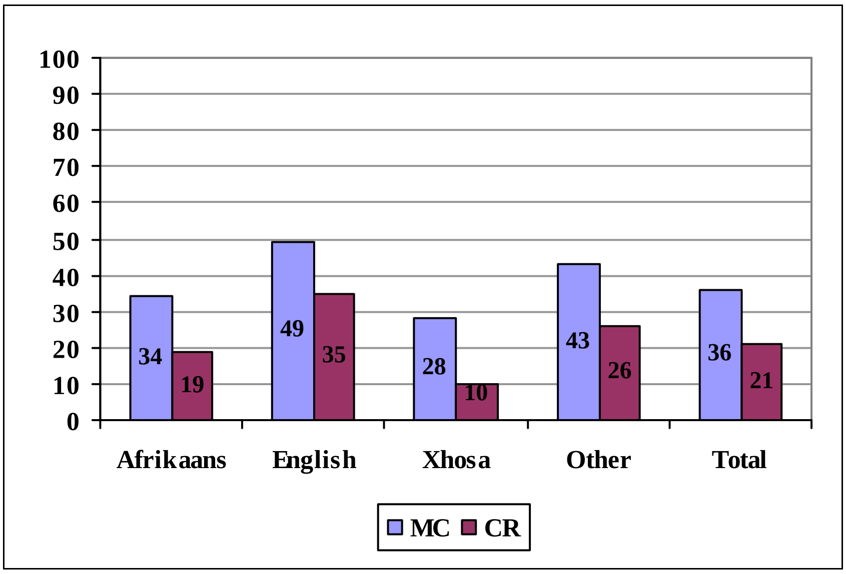


Figure 3.47: Mathematics performance by item type and home language

The data illustrated in Table 3.39 and Figure 3.47 for a home-language split across item type show to a great extent the same pattern found in the overall data for the Province with the average scores obtained for multiple-choice items noticeably higher than the average CR scores for the different home-language groups. However, the gap between the average scores for the MC and the CR items is relatively larger for the Xhosa-speaking learners than for the other two language groups. This could be expected considering the fact that the Xhosa-speaking learners were answering the mathematics instrument in a language different from their home language while the English and Afrikaans-speaking learners answered the test in their home languages.

The data illustrated in Table 3.40 and Figure 3.48 for a LoLT-based split across item type show the same pattern found in the overall data for the Province.

Table 3.40: Mathematics performance by item type and LoLT

	Afrikaans	English	Total
Multiple-choice	32	36	34
Constructed-response	17	21	19

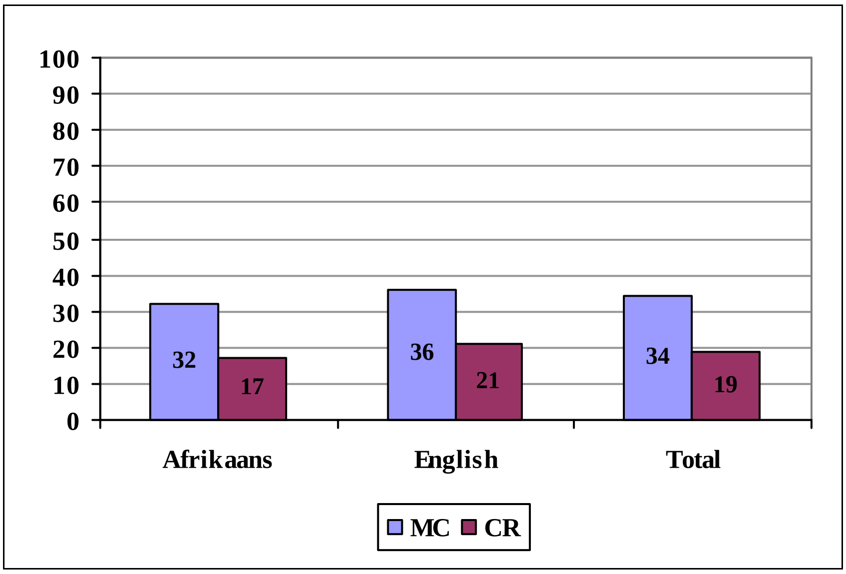


Figure 3.48: Mathematics performance by item type and LoLT

3.5.8 Mathematics performance by the number of books in the home

Figure 3.49 shows a positive correlation between performance on the mathematics instrument and the number of books in the home.

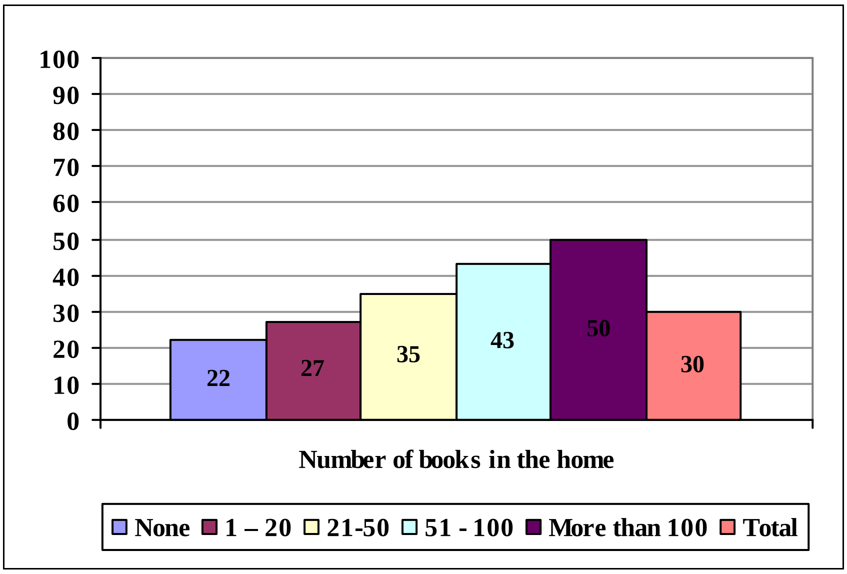


Figure 3.49: Mathematics performance by the number of books in the home

Learner Performance

In Table 3.41 and Figure 3.50 the learners' performance for the number of books in the home across the learning outcomes is presented.

Table 3.41: Mathematics performance by LO and the number of books in the home

	LO1	LO2	LO3	LO4	LO5
No books	21	24	24	18	32
1 to 20 books	24	29	30	20	43
21 to 50 books	30	36	39	26	55
51 to 100 books	37	45	50	34	68
More than 100	44	52	57	41	75
Total	27	32	34	23	47

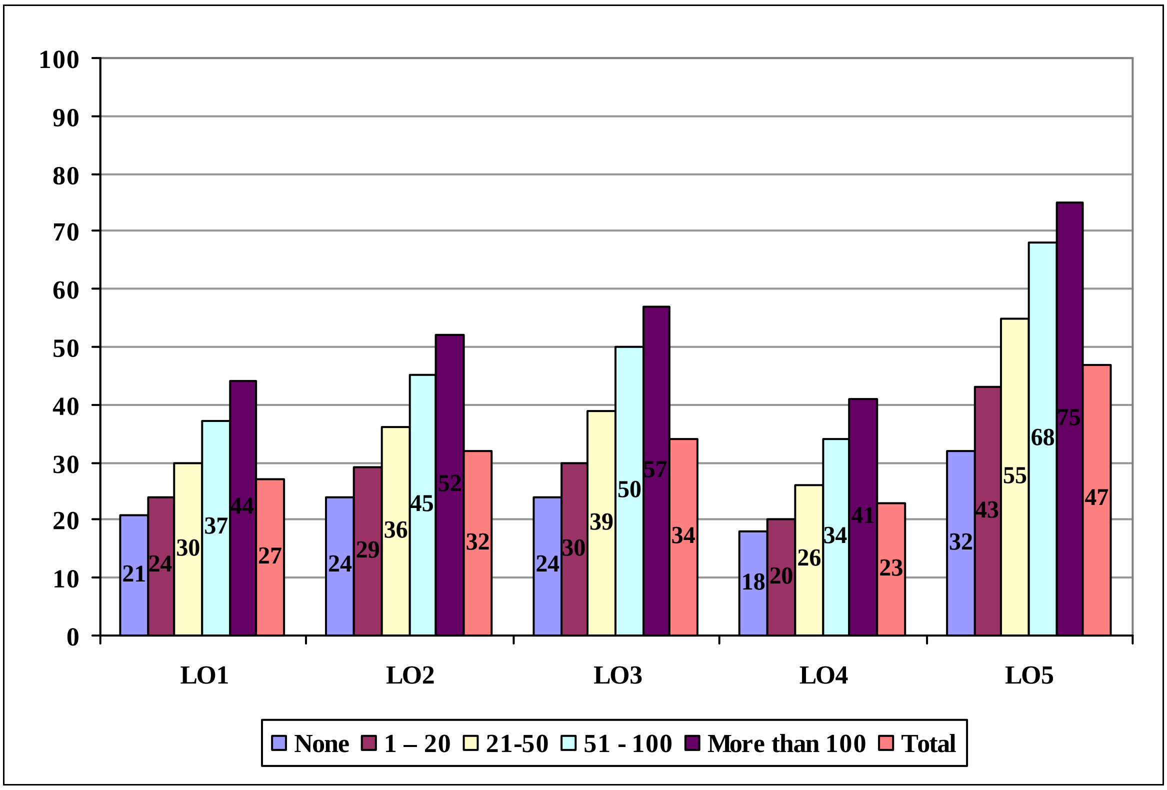


Figure 3.50: Mathematics performance by LO and the number of books in the home

These data shows a consistent increase in performance when the number of books increase, which indicates a strong positive correlation between performance and the number of books in the home across all learning outcomes.

The data illustrated in Table 3.42 and Figure 3.51 for an item-type split across the number of books in the home show to a great extent the same pattern found in the overall data for the Province with the average score obtained for MC items considerably higher than the average obtained for the CR items for each home-language group. However, it is significant that the average score for the MC items increased at a faster rate than the average score for the CR items. The fact that those learners with the greater number of books in the home performed relatively better in the constructed-response items compared to the learners with fewer books in the home is an indication that the number of books is not only a positive indicator of performance but also a strong indicator of the extent to which learners might be able to construct own responses to questions. This would suggest that learners with more books in the home are likely to achieve higher levels of achievement in mathematics (as found in the case of language achievement) compared to those who have fewer books and are more likely to continue through the FET band and enter into and succeed in tertiary education. The Education Department should give attention to this finding because the increase of the number of books available to learners may be a way in which to address poor performance among learners.

Table 3.42: Mathematics performance by the number of books and item type

	MC	CR
No books	29	12
1 to 20 books	33	18
21 to 50 books	40	26
51 to 100 books	48	35
More than 100	55	43
Total	36	21

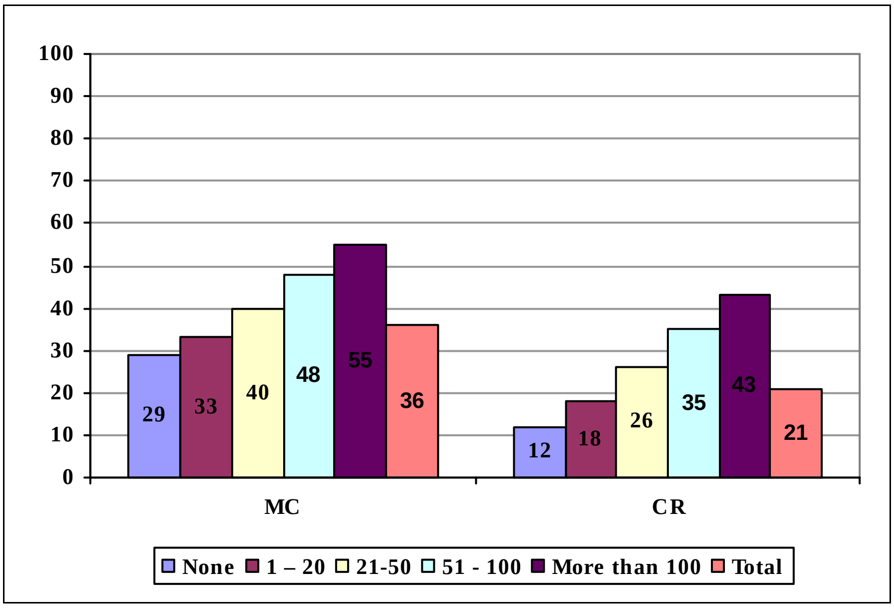


Figure 3.51: Mathematics performance by the number of books and item type

Learner Performance

The learners' performance according to the number of books in the home across the cognitive domains are presented in Table 3.43 and illustrated in Figure 3.52.

Table 3.43: Mathematics performance by the number of books and Cognitive Domain

	Knowledge	Using Concepts	Solving Problems	Reasoning
No books	27	28	20	19
1 to 20 books	32	34	25	26
21 to 50 books	39	43	31	33
51 to 100 books	48	55	40	43
More than 100	55	63	47	50
Total	35	38	28	29

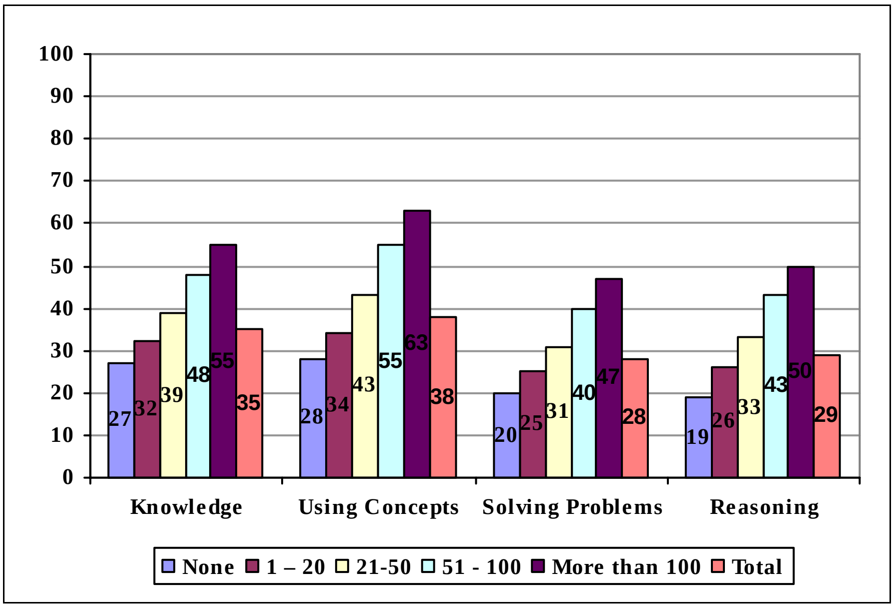


Figure 3.52: Mathematics performance by the number of books and Cognitive Domain

The learner performance trends linked to the number of books in the home across cognitive domains follow the same pattern found in the overall data with a fairly consistent pattern of gaps going along with an increasing number of books across the cognitive domains. The number of books in the home can therefore be considered to be a positive indicator of learner performance across all cognitive domains.

3.5.9 Mathematics performance of WCED learners compared to TIMSS 2003

The average performance of learners on the items from the TIMSS 2003 instrument included in the mathematics test is illustrated in Figure 3.53. The graph compares the average scores achieved by Western Cape learners in the 2006 Grade 8 instrument with the average scores achieved previously on the same items during TIMSS 2003 internationally, in South Africa as a whole, and for the Western Cape learners in that sample.

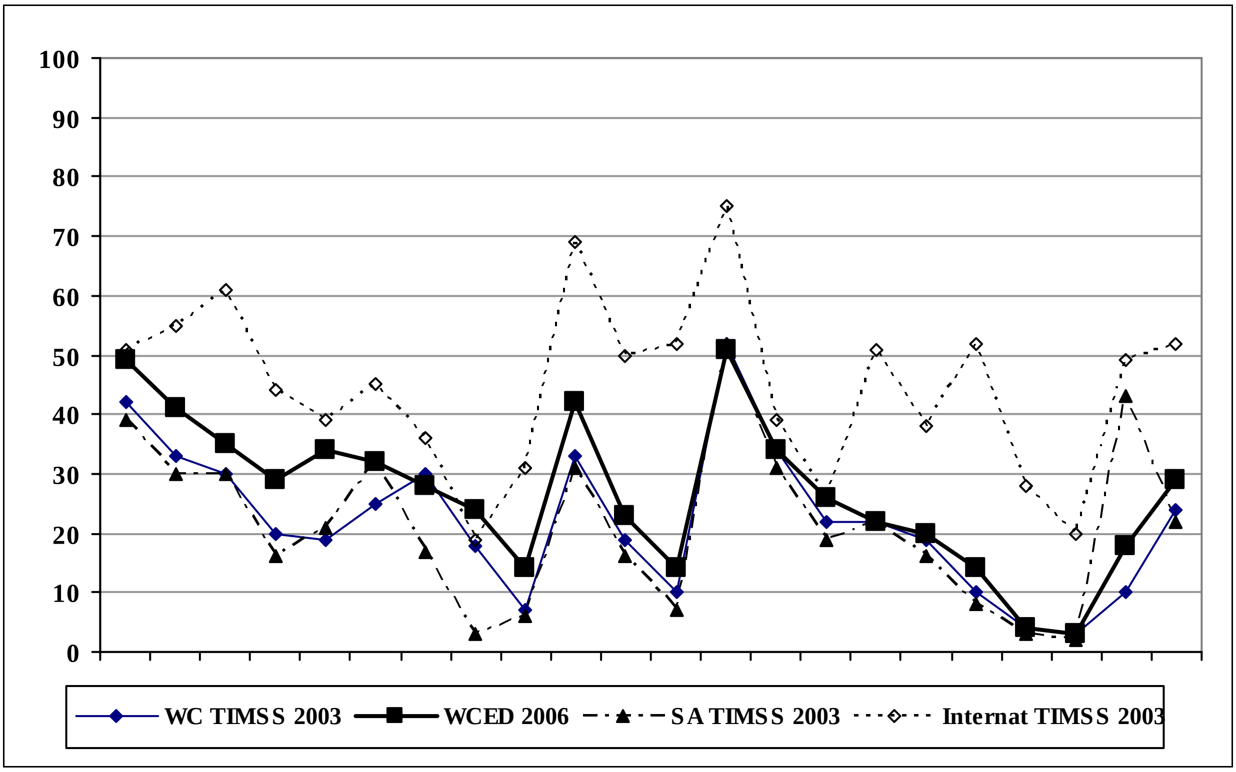


Figure 3.53: Comparison of items included in the TIMSS 2003 and 2006 WCED studies

An analysis of the performances of the learners in Figure 3.53 shows that the Western Cape learners in 2006 on most of the items performed slightly better than the Western Cape and South African learners in TIMSS 2003. However, the average scores for most of the items are still far below the average scores internationally. It is clear that the performance of Western Cape learners compared unfavourably with international standards. To address this matter, teachers should determine and address weaknesses among poorly performing learners. The better performing learners should follow advanced programmes to improve their existing mathematics knowledge and skills further. They should also be stimulated by enrichment programmes to get them interested in the Mathematics and Science study fields so that they can succeed in mathematics-related studies in tertiary education. (Recommendations are systemised in Chapter 4.)

3.5.10 Analysis of learner performance / examples of typical learner responses

Teachers need information that they could use to assist learners and for improving the learning and teaching of mathematics. Since mathematics has a hierarchical structure, more advanced levels will become increasingly out of reach if basic levels are not understood. It is therefore important for teachers to have a system in place that could be used to determine learners' weaknesses and misconceptions regularly. This information could be utilized to address problems that learners encounter before they influence performance at a next level. The example items that are used to illustrate learners' problem areas should also give teachers some idea of the competencies usually assessed at this level internationally, as intended and reflected by the international linkages of the instrument. If teachers integrate the information obtained from this discussion into their lesson planning it could make their classroom practices more effective and should help to improve learner performance.

Responses (the actual correct and incorrect answers) to individual items were analysed in two ways to illustrate the information revealed by the study.

- First, frequencies, using data resulting from the use of the two-digit coding system for the constructed-response items, were analysed. The first digit designates the correctness of the response. The second digit, combined with the first digit, represents a diagnostic code used to identify specific learner approaches, strategies, or common errors and misconceptions.
- Secondly, data resulting from the specific option selected for the multiple-choice items were examined to provide insight into the strengths and weaknesses of the learners in mathematics.

Common errors were identified by analyzing the number of learners that were allocated specific diagnostic codes for the free-response (constructed-response) items and the number of learners that selected different options for the multiple-choice items. These errors were investigated to determine if they reveal any serious or consistent misconceptions amongst learners. By interpreting the patterns of errors, misconceptions and the predominant methods used by learners to solve the mathematics assessment problems, the report attempts to inform the teaching process by providing insight into the ways in which learners process information, apply their reasoning, and end up making certain errors.

Each constructed-response question has its own scoring guide, developed to provide data about students' achievement as well as diagnostic information about misconceptions and common errors.

Examples:

Examples are presented with results from the learner performance assessment of the item next to it. This is followed by a discussion of the results with an analysis of the most frequently made mitakes.

Example 1:

Which of these is closest to $11^2 + 9^2$?

A. 20 + 20
B. 20 + 80
C. 120 + 20
D. 120 + 80

Results for Example 1:

Learner Results for MCQ options

A.	46.2%	✗
B.	9.2	✗
C.	5.7%	✗
D.	29.1%	✓

This multiple-choice item required learners to select the answer closest to $11^2 + 9^2$ from four possible answers. The item was answered correctly by 29.1% of the learners. More than 46% of the learners selected Option A, which was incorrect. These learners may have arrived at the answer by multiplying the approximate value of 11 and 9 by 2, i.e. 10 x 2 + 10 x 2 to get the answer 20 + 20. This type of mistake indicates that learners have not grasped the concept of exponents and may confuse the concept with multiplication.

Example 2:

Which one of the following is equal to
$370 \times 998 + 370 \times 2$?

A. $370 \times 1\,000$
B. 370×998
C. 340×998
D. $370 \times 998 \times 2$

Results for Example 2:

Learner Results for MCQ options

A.	24.0%	✓
B.	8.2%	✗
C.	4.5%	✗
D.	53.5%	✗

This question assesses the ability to determine the order in which calculations with multiplication and addition should be done as well as their ability to do these calculations. The results showed a marked preference towards Option D with 53.5% of the learners selecting this option. Learners who selected Option D as the correct answer may have reached this answer by multiplying the common factor 370 with 998 and then multiplying the answer with 2. The large percentage of learners that selected this answer is an indication that learners experience problems in determining the order in which computations should be done. It may be that learners that could not calculate the correct answer selected Option D because three of the original numbers in the item are included in the answer.

Learners usually find multiplication and addition computations easier than division and subtraction computations. The fact that fewer than a quarter of the learners selected the correct answer to this question that assess multiplication and addition computations is extremely worrying. Therefore it can be speculated that learners did not possess the ability to solve problems that require knowledge on algorithmic procedures and computations.

Example 3:

In which list are the numbers ordered from biggest to smallest?

Translation

Akoluphi uluhlu apho amanani acwangciswe ukusukela kwelikhulu ukuya kwelincinane?
In watter ry is die syfers van die grootste na die kleinste gerangskik?

A. 0,233; 0,3; 0,32; 0,332

B. 0,3; 0,32; 0,332; 0,233

C. 0,32; 0,233; 0,322; 0,3

D. 0,332; 0,32; 0,3; 0,233

Results for Example 3:

Learner Results for MCQ options

A. 7.7% ✕

B. 32.8% ✕

C. 21.1% ✕

D. 28.2% ✓

Double response:s 3.1%

Missing responses: 7%

This item assesses learners' knowledge about place value of decimal numbers. To answer this question learners first have to determine the largest number represented in the options, i.e. 0.332; then they have to find the option/s that starts with the number 0.332 (only Option D starts with 0.332). The final step is to check whether the other three numbers in D is also ordered from large to small. The percentage of learners that selected Option B is higher than the percentage of learners that selected the correct answer (Option D). It is suggested that the majority of learners lack understanding of the place value of decimal numbers because just more than one quarter of the learners selected the correct answer to this question. However, it may also be that some of the learners did not know the terms "biggest" and "smallest" especially considering the findings in the language results that a relatively large percentage of the learners struggle with reading and writing tasks.

Example 4:

What is the value of the following expression:
$1 - 5 \times (-2)$?

A. 11

B. 8

C. -8

D. -9

Results for Example 4:

Learner Results for MCQ options

A. 12.4% ✓

B. 22.6% ✗

C. 45.4% ✗

D. 13.4% ✗

Double response:s 3.1%
Missing responses: 7%

This is a multi-step problem that assesses learners' ability to do calculations with integers. To answer this question, learners should first decide in which order they should do the calculations and then do the calculations in that order correctly.

Although the largest proportion of the learners (45.4%) selected Option C, the percentage that selected Options B (22.6%) and D (13.4%) is also higher than the percentage that selected Option A (12.4%) correctly. Learners who selected Option B appears not to have known the correct order in which the operations should be done because they first subtract and then multiply. Learners who selected Option D could not multiply with negative numbers. The largest proportion of the learners (those selecting Option C) experienced problems with both the order in which the operations should be done as well as with the concept of multiplying with negative numbers.

Example 5:

A computer club used to have 40 members, and 60% of these members were girls. Later 10 boys joined the club. What percentage of the new total of members are girls? Show the calculation that leads to your answer.
Calculations:

Answer: ____________________

Results for Example 5:

Scoring Code	Response	
	2 Correct response	
20	48% with calculations shown	2.5%
	Partial correct response	
10	24 girls	1.4%
11	Correct method but computational error	0.2%
12	48% with no calculations shown	0.4%
	3 Incorrect response	
70	50%	14.2%
71	30% [$\frac{60}{100} \times \frac{50}{1} = 30$]	5.2%
79	Other incorrect (including crossed out/ erased, stray marks, illegible, or off task)	60.7%
	4 Non-response	
99	BLANK (No attempt)	15.6%

Learners had to construct their own responses rather than to select an answer in the MCQ format. The item requires knowledge of percentages and operations with common fractions and involves three steps:

1. Learners first had to find the number of members that were girls, i.e. $\frac{60}{100} \times \frac{40}{1} = 24$ girls.
2. Then they had to find the new total membership for girls, i.e. 40 + 10 = 50 members.
3. Finally they had to find the new percentage of members for girls, i.e. $\frac{24}{50} \times \frac{100}{1}$ to arrive at the new percentage of members for girls that is 48%.

Only 4.3% of the learners answered correctly or partially correctly (Scoring Codes 10, 11, 12 or 20). The Scoring Code 70 represents the number of learners who deducted the 10 boys from the 60% girls originally in the club, and 14.2% of the learners produced this incorrect answer. The largest group of learners (60.7%) was allocated Scoring Code 79, which represents “Other incorrect answers”. More than 15% of the learners did not made any attempt to answer this question and was allocated the Code 99. This suggested that their may be learners in mathematics, as was found in the languages, that could not produce their own answers.

Example 6:

Which of these expressions are equivalent to $k + k + k + k + k$?

A. $k + 5$

B. $5k$

C. k^5

D. $5(k + 1)$

Results for Example 6:

Learner Results for MCQ options

A.	8.1%	✗
B.	39.7%	✓
C.	42.35	✗
D.	3.65%	✗

Double responses: 3.1%

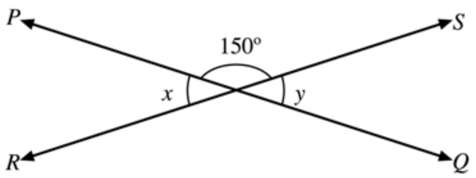
Missing responses: 7%

Example item 5 required from students to identify equivalent algebraic expressions. Only 40% of learners were able to recognise that k + k + k + k + k = 5k. It is very disturbing that distracter C was attractive for more than 40% of the learners. This result corresponds with the result in example one which endorses the previous speculation that learners confuse the concepts of addition and exponents.

Example 7:

In the figure, *PQ* and *RS* are intersecting straight lines.

Translation
Kulo mzobo, u-PQ no-RS yimigca ethe ngqo enqumlanayo.
In die figuur is *PQ* en *RS* twee reguit lyne wat sny.



What is the value of $x + y$?

Translation
Liyintoni ixabiso lika-x + y.
Wat is die waarde van *x + y*.

A. 30

B. 60

C. 180

D. 300

Results for Example 7:

Learner Results for MCQ options

A.	29.9%	✗
B.	23.1%	✓
C.	32.5%	✗
D.	6.9%	✗

Double responses: 1.0%

Missing responses: 6.6%

This item is based on supplementary angles and the solution involves three steps: To answer this question learners first have to:

1. Determine size of either $\angle x$ or $\angle y$ by deducting $150°$ from $180°$ (the sum of adjacent angles on a strait line = $180°$).

Learner Performance

2. They should know that opposite angles formed by strait lines are equal and therefore that $\angle x = \angle y = 30°$.
3. Finally they have to determine the sum of $\angle x$ and $\angle y$, i.e. $\angle x + \angle y = 30° + 30° = 60$.

The incorrect answer selected by the largest proportion of learners was Option C (32.5%). Those who selected Option C probably know the fact that the sum of adjacent angles on a strait line = $180°$. However, they did not know how to apply it in the problem. Option B was also found attractive to learners (selected by 23.1% of the learners). They probably did not read the instructions correctly and only completed the first part of the question where they determined the size of either $\angle x$ or $\angle y$.

Example 8:

A car has a petrol tank that holds 45 litres of fuel. The car uses $8,5\ \ell$ of petrol for each 100 km driven. A trip of 350 km was started with a full tank of petrol. How much petrol **remained** in the tank at the end of the trip?

A. $15,25\ \ell$

B. $16,25\ \ell$

C. $24,75\ \ell$

D. $29,75\ \ell$

Results for Example 8:

Learner Results for MCQ options

A.	26.1%	✓
B.	24.9%	✗
C.	23.0%	✗
D.	17.6%	✗

Double responses: 0.4%

Missing responses: 8.0%

The most obvious explanation for selecting Option B (24.9% of learners) is that they probably did not borrow one unit from 45 when they subtracted the numbers after the decimal point, therefore getting the answer $16.25\ \ell$ instead of $15.25\ \ell$.

It can be speculated that learners who selected Option C calculated the petrol consumed for 250 km to get $21.25\ \ell$. If they subtracted correctly from $45\ \ell$ they would have reached $23,75\ \ell$, but they probably made the same mistake as the learners in Option B by not borrowing when they subtract $21,25\ \ell$ from $45\ \ell$, and therefore got the answer 24.75 as presented in Option C.

Learners who selected Option D did not complete the problem. They only calculated the number of litres used and did not work out the litres that remained in the tank.

Example 9:

The tally chart below shows the scores of learners in a class.

Test Score	Count	Frequency
4	/	1
5	///	3
6	///// /	6
7	//	2
8	////	4
9	///	3
10	/	1

How many learners in the class achieved a score greater than 7?

A. 2

B. 8

C. 12

D. 20

Results for Example 9:

MCQ options		Learner Results %
A.	36.1%	✗
B.	28.6%	✓
C.	9.8%	✗
D.	12.2%	✗
Double responses:		0.4%
Missing responses:		11.8%

This item, based on a test score frequency table, required from learners to be able to interpret the data represented in the table. Learners first had to find the frequencies for a score greater than 7, i.e. the learners that score 8, 9 and 10 marks on the tally chart. To arrive at the answer 8 (Option B) they then had to add up the frequencies: 4+3+1=8. The most frequent error made by learners was selecting Option A as correct. Learners who selected Option A probably did not read the question correctly and may have been under the impression that they should select only the score equal to 7. Once again there is a possibility that the problem may be language related. It may be that learners have not attained the required mathematics vocabulary and do not know what “greater than” implies.

An analysis of the incorrect options selected, or incorrect answers produced by learners for all the items in the instrument revealed the weaknesses/errors illustrated in Table 3.44.

Table 3.44: Weaknesses/errors revealed through analysis of all mathematics items

Weaknesses/errors (or not achieving the set tasks)	Average % of learners who experienced this problem in the instrument
Carry over in operations with whole numbers	22
Select correct operation to solve problem	51
Calculate operations in correct order	67
Read questions correctly	29
Incomplete answering	36
Inability to work with negative numbers	75
Confusion with place value	39

Learner Performance

The high average percentages of learners who experienced problems with basic numeric knowledge and operations such as carrying over in operations with whole numbers (22%), calculating operations in the correct order (67%), and confusion with place value (39) are very disappointing. It is also very alarming that an average 75% of learners made mistakes when they had to calculate with negative numbers.

Educators should be more aware of the type of mistakes that learners make, because learners may be able to improve their performance substantially if they could eliminate the types of mistakes illustrated above. Educators should use a coding system similar to the one used in this study to analyse their learners' mistake on a regular basis. It is therefore suggested that educators should not only be introduced to the diagnostic coding system used in the study, but that they should be trained and supported in its implementation in their teaching.

3.6 CONCLUSION

This chapter reported on the performances of Western Cape learners in languages and mathematics. Their performance for all the learning areas was generally low. Learners obtained the highest score in the English-HL instrument (39%), followed by the Xhosa-HL instrument (31%), then the Afrikaans-HL instrument (30%), and last the Mathematics instrument (29%) with the lowest average score.

The summary of the findings on the study as well as several recommendations which the HSRC would like to make for both language and mathematics in the short to medium term are presented in Chapter 4.

CHAPTER 4 - Conclusion

4.1 INTRODUCTION

The purpose of this report is to present the development of and report on the findings of the WCED system-wide Assessment Task instruments used for measuring the language and mathematics achievement of all school students in the Western Cape during 2006.

The instruments were developed, piloted and refined by the HSRC between the second half of January and April 2006. During the first week of May they were handed over to WCED for printing, dispatching and administering in June 2006. The linguistic diversity of the Western Cape was accommodated in several ways. The language instruments were developed in Afrikaans, English and Xhosa for the Home Language section and in English and Afrikaans in the language of learning and teaching (LoLT) section. The mathematics instrument was developed in the two languages used for teaching mathematics, i.e. Afrikaans and English. In addition, the mathematics instruments were enriched with translations in Xhosa and Afrikaans for some of the items in the English version; and translations in Xhosa and English for the same items in the Afrikaans version. This was a language-accommodation strategy, borrowed and extended from an experimental pilot with Grade 11 mathematics learners during 2005 in the Western Cape.[9] Learners writing tests in their second language were given an additional 10 minutes of time for the instrument/s. The mathematics' instruments were benchmarked against TIMSS 2003 in order to comply with the WCED requirement of international benchmarking. The language instruments were developed within contemporary international language assessment considerations and with specific advice from an international expert.

Completed instruments were retrieved and checked by the WCED and then handed over to the HSRC in July 2006 for coding, analysis and reporting. The HSRC managed the coding, data capturing, cleaning, analysis and reporting process between August 2006 and March 2007.

It needs to be noted that if one conducts a system-wide and one-off assessment of a limited nature, and of learners exactly mid-way through Grade 8, one is effectively assessing the learners' primary school education (Grades 1-7), and a maximum of what learners may have covered of the

[9] Jonty Damsel, principal of Zonnebloem Nest School, initiated and trialled the translation of mathematics questions from English into Xhosa and Afrikaans for the June and November examinations. This is one of the language-accommodation strategies used in countries which are trying to be sensitive towards language diversity and equity issues in their education systems. Language accommodation has been included in the US Grade 4 and 8 annual assessments since 1998.

curriculum by a point half way through the year. In other words, the assessment will not be able to offer a complete assessment of learners at Grade 8 level, as if they had completed the curriculum for Grade 8.

This means that where trends and obvious patterns are noted amongst learners from a particular educational cluster, these reflect conditions in the primary schools rather than the secondary schools at this point. It may be possible to pinpoint particular trends to particular schools and even particular teachers. However, it was within the terms of reference of this study to identify these and at such levels.

Whereas the mathematics assessment task was benchmarked through a use of the TIMSS 2003 study, the language instrument could not be benchmarked against a particular international instrument for reasons advanced in Chapter 2. Chapter 3 included a comparative analysis of the WCED mathematics and TIMSS 2003 findings. In regard to international comparisons regarding language data, this has been reserved for Chapter 4 (see Section 4.3 below).

4.2 KEY FINDINGS

The findings range from very encouraging at the top end of student achievement to very alarming at the bottom end for both language and mathematics. However, as is quite clear from the tables and graphs in Chapter 3, a very small percentage of learners have achieved more than adequately well in either language or mathematics. Rather, the vast majority of learners have performed extremely poorly.

Learners obtained the highest score in the English Home Language instrument (39%), followed by the Xhosa Home Language instrument (31%), then the Afrikaans Home Language instrument (30%) and lastly the Mathematics instrument (29%) with the lowest average score.

An analysis of learner performance according to the 7-point assessment scale (determined by the national Department of Education) shows that the majority of the learners did not attain the knowledge and skills required. Only 8% of Afrikaans-HL speakers, 27% of the English-HL speakers, 3% of the Xhosa-HL speakers and 7% of the Mathematics learners achieved more than 50% on the instruments. Very few learners perform at the higher levels of performance, with only 8% of the Afrikaans-, 14% of the English- and none of the Xhosa-HL speakers performing within the top three levels. Only 7% of the mathematics learners fall within the upper three levels.

Conclusion

The following patterns of achievement across the learning areas were noted:

- Gender differences in learner achievement were observed in the Language instruments, where girls obtained higher average scores than boys. The study found no gender differences in the Mathematics scores of boys and girls;
- This study shows that teaching, learning and assessment in languages other than home language may have negatively affected learner performance in all learning areas. Where the assessment tasks were able to measure teaching and learning activities in a language other than the home language, learners obtained significantly lower scores across all learning areas. The low scores of learners who are taught in a language other than their home language appears to follow difficulty in grasping the meaning of questions and producing answers which are correspondingly meaningful.
- Xhosa-speaking students outperformed both Afrikaans- and English-speaking students in the HL section of the instrument (largely based on a literary extract taken from the reading lists of each HL group). It seems that Xhosa-HL students are familiar with the genre selected for this version of the instrument (folktale) and are able to read literature at grade level fairly well. However, when only the common extended writing tasks in the HL section were measured, Xhosa-speaking students performed least well.
- The average score obtained for MC items was considerably higher than the average score obtained for the CR items for all Learning Areas. Learners found questions for which they had to construct their own answers to be more difficult than questions for which they had to select the answers from given options. A serious cause for concern is that many Afrikaans- and Xhosa-HL learners found it particularly difficult to formulate answers and construct text in their own languages. This highlights the importance of including CR items in assessment instruments to reveal what learners really know and can do.
- There is a consistent increase in performance, particularly in writing skills, when the number of books per household increases. This indicates a strong positive correlation between performance and the number of books in the home across all learning outcomes.
- Learners performed poorly in all learning outcomes in mathematics, but particularly poorly in LO1 (Numbers, Operations and Relationships). Although it was expected that the Western Cape learners should perform poorly in LO1 because analysis of the performance of learners by other South African studies shows this trend, the extremely low performance on LO1 is a cause for concern because LO1 deals with the basic concepts and skills that are used in the other learning outcomes.

Conclusion

- In the language tests, LO4 (writing) appears to be the most difficult task for learners, because this is where they have to demonstrate their productive competence in language use. This is also the area which best indicates where learners are placed in terms of whether or not they would complete the FET band successfully.
- Performance in the EMDCs shows some differences for all Learning Areas and the department has been advised to take note of these differences. Further data analysis might offer more nuanced information and could be undertaken at a later stage.
- In the mathematics instrument, the Western Cape 2006 learners performed slightly better than the Western Cape and South African learners in most of the items included from TIMSS 2003. However, the performance of Western Cape learners still compared unfavourably with regard to international standards.
- Analysis of incorrect answers selected or produced in the mathematics instrument shows that many learners experienced problems with basic numeric knowledge and operations such as carrying over in operations with whole numbers, confusion with place value, doing calculations in the correct order, and calculating with negative numbers.
- More than 80% of the learners in all language groups answered positively that they benefited from the translations (glossed translations of instructions) included in the mathematics instruments. The data, however, show a small positive difference between the average scores for learners who wrote the glossed-items compared with the average performance of learners who wrote the non-glossed items.

4.3 FINDINGS WHICH REQUIRE PARTICULAR ATTENTION

4.3.1 Comparative Language/s Achievement – domestic and international comparisons

English Home Language learners have performed best with a mean score of 39%, followed by Xhosa-speaking students at 31% and lastly by Afrikaans-speaking learners with 30%. One would have expected that HL speakers of English and Afrikaans would have achieved at a similar level and that there would not be a significant difference, since both of these groups have the advantage of ‘mother tongue’ education throughout their school years. One would, however, expect a gap between English and Afrikaans speakers on the one hand and speakers of Xhosa on the other hand, because although the latter take Xhosa at HL level, they usually use English, their second language, as LoLT. This is much more difficult to do than working through one’s HL; therefore we expect a lower mean score for such students. Secondly, in terms of international comparisons with students who study through their second language in other contexts, one expects to find a gap of up to 20 percentage point in the mean achievement score of those students and those who are assessed in

English as HL and those who are second language speakers of English (Thomas & Collier 1997, 2002). Because HL Xhosa-speaking students were tested in English, their second language, for about 50% of the test, we would expect that Xhosa-speakers would, on average score about 10 percentage points less than either English-HL or Afrikaans-HL students.

The gap between English-HL and Xhosa-HL students then on this instrument is consistent with the findings. Xhosa speakers scored 8 percentage points less than English HL students, and therefore they achieved slightly better (up to 2 percentage points better) than one would have expected.

A major concern is that of the Afrikaans-HL group. They have not performed nearly as well as might have been anticipated. Relatively speaking, they have performed at least 9 percentage points less on this instrument than would have been expected in terms of comparisons with other language groups.

Therefore the proximity of mean scores for Xhosa and Afrikaans speakers is disturbing, since it means that Afrikaans-speaking learners, as a group, are performing considerably less well than their English- and Xhosa-speaking peers.

Home-language background is not the only indicator of success in schooling; there are several indicators which are affected by complex issues of diversity. Such indicators include: socio-economic and socio-cultural factors; and geographic location, i.e. urban, peri-urban, rural etc. A significant proportion of Afrikaans-speaking learners, for example, come from rural and relatively poor home backgrounds; whereas a higher proportion of English-HL speakers come from urban and middle-class environments. Most Xhosa-HL learners, however, come from poor urban (township) communities. The WCED learners therefore do come from different contexts and these would affect learner performance.

By way of comparison, in the USA, for example, where about 660 000 Grade 8 students are tested annually, the *National Report Card in the USA,* collated by the National Center for Education Statistics, which falls under the Department of Education annually, shows differences in reading scores between learners who come from three different groups. Although there are no direct comparisons with learners in the Western Cape, it is useful to consider the US Report Card findings for reading against the findings of this study as follows:

Conclusion

4.3.2 Percentages of learners performing at or above specified levels

Group	Proficient (SA Level 4-7; 51%+)	Basic (SA, NCS Level 3-7; 41%+)
W Cape English	23	40
W Cape Afrikaans	9	19
W Cape Xhosa	4	23
USA White	39	82
USA Hispanic	15	56
USA Black	12	52

(US data extrapolated from: Perie et al. 2005)

Grade 8 learners in the Western Cape are simply achieving way below what their counterparts in the USA are achieving. In the USA, 52% of Black learners (considered speakers of African American English) achieve above 41% in English, and they are the lowest achievers in the US on average. Yet their achievement is greater than the achievement of the English-HL learners on the 2006 Grade 8 assessment. Fifty-six percent of Hispanic speakers, who are considered to be second-language speakers of English, score above 41% in English. Again this is significantly higher than English-HL speakers in the Western Cape.

If the WCED wishes to have an idea of how well Western Cape learners are doing in comparison with their counterparts in other parts of the world where learners come from diverse settings, the US National Report Card is a good point for comparison, and it tells us that learners in the WCED are probably performing significantly behind their peers in countries with diverse populations such as the USA.

4.3.3 High incidence of students who do not manifest evidence of literacy skills

Chapter 3 illustrated clearly that there is an unusually high incidence of learners who appear to have either no or marginal literacy skills. More than 30% of learners may not have much more than emergent literacy skills. The question need to be asked: “How is it possible that so many learners have reached Grade 8 without being able to read or write?”

Conclusion

This is a question which the EMDCs will have to take back to the primary schools from whence the learners who manifest this condition were enrolled prior to 2006.

Such a high incidence of this condition raises questions about whether or not the curriculum is sufficiently robust, and / or whether the teachers who have taught these learners sufficiently understand their explicit responsibilities. The bottom line with the low levels of achievement in writing skills means that teachers do not expect their learners to write consolidated text and do not monitor the writing activities of their students. Had they done this with the necessary regularity, there is little chance that so many learners who cannot read , write or perform basic mathematical tasks would have been able to pass through the education system, without the education system noticing that this was the case.

4.4 RECOMMENDATIONS

There are several recommendations which the HSRC would like to make for both language and mathematics in the short to medium term. These are indicated briefly here, however, the implications exceed the terms of reference of this study and should be addressed more systematically through other means.

4.4.1 Language

The EMDCs and WCED at Provincial level should prioritise the need to facilitate and develop academic literacy proficiency for both HL and LoLT by the time learners reach secondary school. To this end it is advisable in the short term:

- To concentrate on building reading skills in the languages learning area throughout the Foundation and Intermediate Phase (Grades 1-6).
- To concentrate on building reading skills across the curriculum in the Intermediate and Senior phases (Grades 4-9) (i.e. this involves all teachers of all learning areas).
- To not assume that children learn to read unless they are exposed to consistent and explicit reading strategies.
- To encourage parents and communities to engage in reading and writing activities, but to not assume that they will do this. It has to be assumed that students have to be **taught** how to read in school.
- To have Foundation Phase teachers ensure that learners develop writing skills and that learners can write independent sentences by the end of Grade 3.
- To have Intermediate Phase teachers (in all learning areas) ensure that learners are writing logical and well-structured paragraphs.

Conclusion

- To have Senior Phase teachers ensure that learners are writing independent text comprising several well-constructed paragraphs which follow on from one another.
- To restrict the use of: gapped/cloze exercises and worksheets.
- To limit the use of MC questioning; extend the use of constructed-response assessment.
- To have teachers build their own expertise and knowledge of language structure, reading and writing.
- To increase capacity in the use of diagnostic scoring processes in order to identify the kinds of difficulties which learners are experiencing with reading and writing.

4.4.2 Mathematics

The EMDCs and WCED at Provincial level should prioritise the need to facilitate and develop mathematics proficiency by the time learners reach secondary school. To this end it is advisable in the short term:

- To concentrate on building mathematics skills in all learning outcomes throughout the Foundation and Intermediate Phases (Grades 1-6).
- To pay special attention to the attainment of basic knowledge and skills in numbers, operations and relationships (LO1) to ensure progress in the other areas of Mathematics.
- To concentrate on teaching at the correct level according to the curriculum in the Intermediate and Senior phases, while ensuring that prior knowledge of the previous phase/s are in place (Grades 4-9).
- To limit the use of MC questioning; extend the use of constructed-response assessment. Insist that learners show the procedures followed and explain their reasoning.
- To increase capacity in the use of diagnostic scoring processes in order to identify the kinds of difficulties which learners are experiencing with reading and writing. Teachers should be introduced to diagnostic coding and trained and supported in the implementation of it in their teaching.
- To have teachers build their own expertise and knowledge of mathematics and the methodology of teaching and learning in mathematics.
- To further explore the strategy of the HSRC to make use of explanations of key instructions in learner’s home language (glossed items) for mathematics items. Other ways in which learners who are taught through languages other than their home language could be assisted, should be included in the study.
- To pay attention to the finding that suggests that learners with more books in the home are likely to achieve higher levels of achievement in mathematics. The increase of the number

of books available to learners (even at school through reading corners) may be a way to address the poor performance of learners in mathematics by addressing their low language ability.

- To address weaknesses of poorly performing learners by special programmes developed at Provincial or EMDC level.
- To develop advanced programmes to improve the existing mathematics knowledge and skills of higher performing learners further. These learners should be stimulated by enrichment programmes to get them interested in mathematics and science study fields in tertiary education.
- To encourage parents and communities to be more positive in general to mathematics as learning area and to learners' ability to perform well in it. .

4.5 CONCLUDING REMARKS

Both the WCED and the HSRC have learnt a great deal about the needs of system-wide assessment at Grade 8 level, particularly in relation to diagnostic assessment which is designed to relate the findings on achievement to biographic detail and other learner information which the WCED collects and prepares electronically. The key lesson here is the allocation of appropriate timeframes for preparation. Future assessments of this scale need to be allocated considerably longer periods of time for instrument development (including piloting), and collection and adequate electronic preparation of learner information (CEMIS data). A provider should be appointed at least 12-15 months ahead of the assessment date, in order to allow for adequate instrument development, piloting, refinement etc. The provider's capacity to deliver adequate diagnostic data analysis is as dependent upon good instrument developing and coding as it is on good, error-free instrument administration (e.g., efficient and correct matching of completed instrument results across learners, and to stable electronic learner ID information from the client).

Be it as it is, the HSRC is in a good position to advise on the flow of information and the design of the underpinning systems to enable the processing of large volumes of learner demographic and assessment data towards producing automated, but individually customised, reporting at learner, school, EMDC and Provincial level by using a range of software applications.

The Grade 8 learners in the Western Cape have performed very poorly across all languages and mathematics. There is an unacceptably high level of underachievement at the lowest possible floor level. Far too few learners are achieving at the highest levels in the Province and these learners appear to come from a handful of schools. If 31% of learners can reach Grade 8 with what appears to be no ability to read or write, and a further 24% can barely read or write, this is a shocking indictment on a system. The performance of learners in mathematics is also disappointing, with 77% of the learners scoring below 50% on the instrument. The gap in performance between answers to MC and CR items in this instrument lead the HSRC to argue that a reliance on MC or cloze procedure exercises on hand-outs provides a veneer of educational assessment which disguises the fact that learners are not developing literacy or numeracy in primary schools in the Province. There is no short-cut to teaching reading, writing and numeracy. It requires consistent effort, with explicit teaching and learning methodologies on the part of each learner and teacher in the system throughout the primary-school years. This study points to a need to look more closely at the NCS, teacher education programmes, learning materials, classroom practice and the relationship amongst them.

Finally, the HSRC advises the WCED to consider the cost-benefit relationship amongst different forms of systemic assessment and the kinds of information which each can offer a system. At this point, it would appear that the kind of assessment which would be of most benefit to the system would be of a purposive sample type that allows for deeper penetration of the failure in literacy and numeracy.

References

Angell, C., Kjaernsli, M. & Lie, S. 2000. Exploring students responses on free-response science items in TIMSS *(*as cited by Shorrocks-Taylor, D and Jenkins, E.W (Eds) Learning from others.

Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.l. & Smith, T.A. 1996. Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study. Chestnut Hill, MA: Boston College.

DoE. 1997. Language in Education Policy. Pretoria: Department of Education.

DoE. 2002a. Revised National Curriculum Statement Grades R-9 (Schools). Languages English - Home Language. Pretoria: Department of Education.

DoE. 2002b. Revised National Curriculum Statement Grades R-9 (Schools) Policy. Mathematics. Pretoria: Department of Education.

DoE. 2005a. Subject Assessment Guidelines. Home Language. National Curriculum Statement Grade 10-12 (General). Pretoria: Department of Education.

DoE. 2005b. Subject Assessment Guidelines. Mathematics. National Curriculum Statement Grade 10-12 (General). Pretoria: Department of Education.

DoE. 2005c. Grade 6 Intermediate Phase Systemic Evaluation Report. Pretoria: Department of Education.

Dossey, J.A., Jones, C.O. and Marin, T.S. 2002. Analysing students' responses in mathematics using two-digit rubrics. As cited byRobitaille, D.F. & Beaton, A.E. (Eds). Secondary analysis of TIMSS data.

Kanjee, A. & Prinsloo, C.H. (2005). Improving learning in South African schools: The Quality Learning Project (QLP) summative evaluation (2000 to 2004). Pretoria: HSRC.

Conclusion

Leinwand, S.J. 1998. Classroom realities we do not often talk about. Mathematics Teaching in the Middle School, **3(5)**, 330-331.

Lindsay, M. 2005. Senior Official in London Education Authority and National Union of Teachers. Personal Communication. Abbreviations?

Macdonald, C. (1990). Crossing the threshold into Standard Three. Main Report of the Threshold Project. Pretoria: Human Sciences Research Council.

Mullis, I., Martin, M., Gonzalez, E., Gregory, K., Garden, R., O'Connor, K., Chrostowski, S. & Smith, T. (2003). TIMSS assessment frameworks and specifications (2nd edn). Chestnut Hill, Ma: Boston College.

Nesher, P. 1987: Towards an instructional theory: The role of students' misconceptions. For the Learning of Mathematics, **7(3)**, 33-39.

Perie, M., Grigg, W.S. & Donahue, P.L. 2005. The Nation's Report Card: Reading 2005 (NCES 2006–451). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, D.C.: U.S. Government Printing Office. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006451

Prinsloo, C.H. & Kanjee, A. (2005). Improving learning in South African schools: The Quality Learning Project (QLP) summative evaluation (2000 to 2004) technical report. Pretoria: HSRC.

Reddy, V. 2006. Mathematics and Science achievement at South African Schools in TIMSS 2003. Pretoria: Human Sciences Research Council.

RSA. 1996. The Constitution of the Republic of South Africa, 1996. Act 108 of 1996. Cape Town: The Constitutional Assembly & the National Parliament, Republic of South Africa.

Strauss, J. (1999). Results of the Monitoring Learning Achievement (MLA) Project. Bloemfontein: University of the Free State and Department of Education.

Thomas, W. & Collier, V. 1997. School effectiveness for language minority students. National Clearinghouse for Bilingual Education (NCBE) Resource Collection Series, No. 9, December, 1997. http://www.ncela.gwu.edu/pubs/resource/effectiveness/

Thomas, W. & Collier, V. (2002). A National Study of School Effectiveness for Language minority students' long-term academic achievement. Santa Cruz, CA and Washington, DC: Center for Research on Education, Diversity & Excellence. http://www.crede.org/research/llaa/l.les.html

Western Cape Legislature. 2001. Western Cape Language Policy. Cape Town: Western Cape Legislature.

Wile, J. (2005). Director, International Development, International Reading Association. Email correspondence.

Williams, J. & Ryan, J. 2000. National testing and the improvement of classroom teaching: Can they coexist? British Educational Research Journal, **26(1)**, 49-73.

Appendix

Language Assessment Instruments Reviewed and Evaluated for WCED Study

Source country	Grade/Age/Competency	Suitability
International		
PIRLS (Progress in Reading Literacy Study) 2001	Grade 4 – literacy	Below grade level required; but useful for instrument framework design at Grade 8 level.
PISA (Programme for International student Assessment)	Age 15 – literacy	Too advanced
IEA (International Association for the Evaluation of Educational Achievement)	Age 13 – language/literacy	Last assessment at this level 1990/1991- Out of date; Planned assessment of 13-year olds in 1996 abandoned.
Council of Europe/EU		
Common European Framework	Generic competency levels, includes language knowledge and skills across all languages known by students (interactive, accumulative building of a comprehensive language portfolio/passport for each learner throughout school)	Very relevant to needs of students in SA; but not for this WCED assignment.
Australia		
South Australia Curriculum, Standards and Accountability Framework	Year 7 literacy test; Year 8 language strand assessment criteria	Could be adapted, used for standards setting.
Canada		
Ontario Achievement Levels	Grades 1-12 Language	Could be adapted, used for standards setting.
Hong Kong		
National Foundation for Educational Research (NFER)	Hong Kong certificate of education examination in English Language, age 16. NFER contracted to explore equivalence between grade descriptors and other internationally recognised standards.	Relevant since instrument set for L1 and L2 speakers of English; but level/grade too high for WCED requirements now. Publication confidential to the Hong Kong Examinations and Assessment Authority (HKEAA) - not available.
New Zealand		
National Education Monitoring Project (NEMP), Otago University, for Ministry of Education	Year 4 (8- to 9-year olds); and Year 8 (12- to 13-year olds). Ongoing systemic evaluation to monitor trends.	Students from 254 schools assessed each year. Four-component instrument design, only one is pen and paper. This can be used as a partial guideline for benchmarking purposes.

Conclusion

Source country	Grade/Age/Competency	Suitability
Singapore		
National Foundation for Educational Research (NFER)	Measurement of English proficiency of students on entry and at various levels of school in Singapore. Tests administered over period ending 2006.	Relevant since students in Singapore are L2 users of English; however the NFER report on the study of assessments is only due 2006 or later – therefore too premature for immediate use by WCED.
UK		
Qualifications and Curriculum Authority (QCA)	Age 11 and 14 (Year 9). Tied to National Curriculum National tests in English, mathematics and science. English language instruments closely linked to test marking schemes for which teachers prepare their students ahead of time.	Appears to be similar to requirements of WCED. However, the instrument is very controversial and has received considerable adverse criticism, including cultural inappropriacy, and and should be used with due caution. WCED teachers will not have prepared students for this marking scheme. The philosophy of the South African NCS is compatible with the NUT critique which increases possibility of mismatch of instrument with needs of WCED. Will nevertheless be reviewed by HSRC for possible partial benchmarking.
Computer Adaptive Testing and Second Language Assessment (CATs)	On-line testing sensitive to user level/grade. Used in UK, and USA.	Instrument very useful and transportable to SA curriculum, but dependent upon computers which are not available per learner.
National Foundation for Educational Research (NFER)	Manage Computer Adaptive Testing in First and Second Language in Britain; and systemic assessment in Singapore and Hong Kong	Not readily transportable to our educational conditions in South Africa at present.
Cambridge	TESOL	Not appropriate for L1 and L2 learners/ too focused on entry to HEIs.
USA		
New Jersey Grade 8 Proficiency Assessment	Language and mathematics assessment provide accommodation strategies for L2 speakers (translation dictionaries, translation of test directions, extended testing time).	Use a procedure of statistical equating to ensure that tests are at the same level.
Tennessee Comprehensive Assessment Programme (TCAP)	Grades 3-8.	Achievement Test Items, Grade 8 can be used for partial benchmarking.