UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

NUMERACY PERFORMANCE OF GRADE 3 LEARNERS IN RURAL AND URBAN PRIMARY SCHOOLS

by

PHILLIP MASIBI TSHABALALA

Submitted in partial fulfilment of the requirements for the degree

Master of Education

in Assessment, Evaluation and Quality Assurance

in the

FACULTY OF EDUCATION

at the

UNIVERSITY OF PRETORIA

Supervisor: Prof. S.J. Howie

December 2008

ABSTRACT

The aim of this study was to undertake a comparative analysis of numeracy performance of Grade 3 learners in urban and rural primary schools. A secondary analysis of the Systemic Evaluation (SE) survey data that was conducted by the Department of Education in South Africa in 2001 at the Foundation Phase (Grade 3) level of the schooling system was done. It focused on numeracy achievement data and explored the relationship between achievement and gender, age, location, home language and a number of identified equity indicators. The equity factors were explored further with a view to predict learner performance in numeracy.

The findings of the descriptive analysis showed that the overall performance in numeracy of Grade 3 learners in rural schools compared to their counterparts in urban schools was not satisfactory. From the exploratory analysis a few of the equity variables (e.g. language, funding, LTSM) were found to have an effect on learner performance in numeracy by location of the school. Only two variables, namely home language (HL) and the number of books in the learner's home, selected as home background equity factors could be used to predict numeracy performance in the final logistic regression model. The variables in the final model that accounted for pedagogical equity factors were availability of computers for teaching, inadequate physical facilities, availability of numeracy reference materials, calculators and others like equipment in the classroom, for example an OHP available for teaching.

List of key words: numeracy/mathematics achievement, equity factors/indicators, educational indicators and national assessments.

TABLE OF CONTENTS **PAGE**

TABLE OF CONTENTS **PAGE**

LIST OF FIGURES **PAGE**

LIST OF TABLES **PAGE**

LIST OF APPENDICES

LIST OF ACRONYMS

ABET	Adult Basic Education and Training
AMI	Assessment Modelling Initiative
CEPD	Centre for Education Policy and Development
C2005	Curriculum 2005
DDSP	District Development Support Programme
DoE	Department of Education
EFA	Education for All
HSRC	Human Sciences Research Council
IEA	International Association for the Evaluation of Educational Achievement
ISRDP	Integrated Sustainable Rural Development Programme
JET	Joint Education Trust
LOLT	Language of Learning and Teaching
LTSM	Learning and Teaching Support Material
MLA	Monitoring Learning Achievement
NAEP	National Association for Educational Progress
NEPA	National Education Policy Act (RSA Act No. 27, 1996)
NCES	National Centre for Education Statistics
OTL	Opportunities to Learn
PISA	Programme in International Student Assessment
QIDSUP	Quality Improvement, Development, Support and Upliftment Programme
RIEP	Research Institute for Education Planning, University of Free State
RSA	Republic of South Africa
SACMEQ	Southern Africa Consortium for Monitoring Educational Quality
SE	Systemic Evaluation
SES	Socio-Economic Status
SO(s)	Specific Outcome(s)
STATS SA	Statistics South Africa

TIMSS	Third International Mathematics and Science Study – now known as Trends in International Mathematics and Science Study
TIMSS-R	Third International Mathematics and Science Study-Repeat
UNESCO	United Nations Education Scientific and Cultural Organisation
UNICEF	United Nations Children's Education Fund
URP	Urban Renewal Project
USAID	United States Agency for International Development

DEDICATION

To my late mother who passed on during the early stages of the conception of this study. The loss was immense to bear but thanks, it lifted me up.

ACKNOWLEDGEMENTS

This study has been possible with the support of the following people and my heartfelt thanks go out to them:

- To the Almighty God, for providing me with the knowledge, courage, patience and strength to complete this study.
- Prof. Sarah Howie, for her expert guidance, and demand for excellence and patience throughout this research.
- Prof. Tjeerd Plomp, for his invaluable time and guidance.
- Mr Sollie Millard and Ms Jacqui Sommerville, of the Statistics Department, University of Pretoria, for their assistance and support during data analysis.
- Ms Elsie Venter, of the Centre for Evaluation and Assessment, University of Pretoria, for the technical support.
- Ms Cilla Nel, of the Centre for Evaluation and Assessment, University of Pretoria, for the academic support and language editing.
- My wife Stella and our lovely daughters, Sibongile and Nandipha for the love, support and presence in time of need without which I do not know how I would have made it to here.
- My father, for who I am, and all my siblings, for reminding me of where I come from and the warm place that I hold in their dear hearts, “Gaabo motho go thebe phatswa”.

DECLARATION

I declare that **NUMERACY PERFORMANCE OF GRADE 3 LEARNERS IN RURAL AND URBAN PRIMARY SCHOOLS** is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

Signature **Date: 15 December 2008**

CHAPTER 1

INTRODUCTION AND BACKGROUND TO STUDY

The aim of this study is to undertake a comparative analysis of Systemic Evaluation data of Grade 3 learners in urban and rural primary schools. In particular, the learners' performance in numeracy will be analysed to examine differences in achievement in numeracy between learners from these schools. In this study, the extent to which there are gaps with regard to equity between urban and rural schools will be explored with a view to predict learner performance in numeracy.

In this chapter, the purpose and aim of the study are given (1.1). This is followed by an overview of the background to the South African education context (1.2). A thorough background is then given on the introduction and implementation of Systemic Evaluation (SE) in South Africa (1.3). Thereafter the problem statement is elaborated (1.4) and this is followed by the rationale for the study (1.5). The main research questions guiding the study are then presented and explained with a brief mention of research premises (1.6). To conclude, the structure of the dissertation (1.7) is outlined.

1.1 STATEMENT OF PURPOSE

This study seeks to describe and analyse the performance of the South African Grade 3 learners in numeracy. A comparative analysis will also be undertaken in this study in terms of how Grade 3 learners from rural schools fared in numeracy as compared to their counterparts in urban schools.

A secondary analysis of the SE data from a survey that was conducted in 2001 at the Foundation Phase (Grade 3) in the mainstream by the Department of Education (DoE) will be conducted. This study will focus on numeracy achievement data of rural and urban primary school learners and explores the relationship between achievement and gender, age, location, home language and a number of identified equity indicators.

The extent to which the transformational goal of equity has been implemented in the primary school system will be explored in this study by specifically examining the

rural-urban divide of the South African society. This study will investigate how the education system performs with regard to equity in the provision of education in primary schools. This is particularly pertinent within the current discourse around the national question of the "two economies" co-existing side by side in all the communities across the country (Mbeki, 2004).

The analysis sought to explain the overall performance of learners in the system by exploring those factors that have an effect on performance in numeracy of Grade 3 learners in rural schools as opposed to those in urban schools. Ultimately, a stepwise logistic regression model was used to determine the effect of equity factors on learner performance and how the factors can be modelled to predict the performance of learners in rural and urban primary schools.

1.2 THE SOUTH AFRICAN CONTEXT

South Africa is a land of diversity with its peoples being situated in both rural and urban areas resulting in stark contrasts. The Educational for All (EFA) study conducted in South Africa (2000) found that the spatial population distribution of South Africa shows that 46% of her citizens live in rural areas. The South African EFA report (Department of Education [DoE], 2000a, pp. 5–6) indicates that about 54% of South Africans reside in urban areas, with the numbers increasing annually as people flock to the cities in search of work.

It has been found, through a report on poverty published in May 1998, that just under half of the population live below the poverty line with a large proportion being in the rural areas (Howie, 2002, p. 10). This startling occurrence could have a serious effect on the overall quality of education service delivery of the system. For example, it could pose problems for the system in terms of dealing with the backlog of resources especially in areas previously in the former homelands and in the majority of communities that were marginalised.

In contrast, urbanisation levels, however, vary substantially by province. The percentage of people residing in urban areas varies from only 11% in Limpopo and 35% in North West, to 88 and 97% in the Western Cape and Gauteng, respectively.

The EFA report (DoE, 2000a, pp. 5–6) further reveals that, "the more urbanized provinces tend to have better education facilities and services than the less urbanised provinces". This could have serious implications for service delivery (input) as well as the performance (output) of learners in the less urbanised provinces with regard to the quality of outputs.

However, Statistics South Africa (Statistics South Africa [Stats SA], 2003) warns that the situation in South Africa regarding the classification of the country into urban and rural areas is rather fluid at this stage. There has been a move away from this classification towards an all-inclusive municipal one. A formal urban settlement, unlike a rural settlement, is structured and organised; land parcels (plots or erven) make up a formal and permanent structure; services such as water, electricity and refuse removal are provided; and roads are formally planned and maintained. This category includes suburbs or townlands and townships categorised as urban and semi-urban for the purposes of this study (Stats SA, 2003, p. 187). In contrast, rural settlements are characterised by dirt roads (often making it difficult to access amenities and services), necessitating its residents to travel longer distances to access water, education, health and other basic services.

Nevertheless, in view of important government programmes such as the Integrated Sustainable Rural Development Programme (ISRDP) and the Urban Renewal Programme (URP) strategies, the concept of rural/urban remains an important one for statistical classification (Stats SA, 2003, p. 1). The ISRDP and the URP are part of a targeted intervention strategy of the government to counter poverty and under-development. This should be seen as a measure of redress especially for both the urban and rural areas that have been plagued by the imbalances of the apartheid era.

This study then, in addition to undertaking a comparative analysis of Systemic Evaluation data of Grade 3 learners in urban and rural primary schools, also looks at the extent to which there are equity gaps and any effects on the levels of performance of learners in both urban and rural schools.

1.3 BACKGROUND TO SYSTEMIC EVALUATION (SE)

In this section, a detailed description is given of the Systemic Evaluation (SE) survey process in South Africa. This includes an outline of SE, its purpose and rationale for its introduction.

Legislative framework

The transformation of education in South Africa emphasises the right of quality education for all. The first intention is to redress the discriminatory, unbalanced and inequitable distribution of the education services of the apartheid regime and secondly, to develop a world-class education system suitable to meet the challenges of the 21st century (DoE, 1995). The Systemic Evaluation (SE) Framework (2001) provides the guiding principles for systemic evaluations at the Foundation, Intermediate and Senior Phases of the General Education and Training (GET) Band in South Africa. The Framework begins by establishing the mandate for SE within the National Education Policy Act (NEPA), Act No. 27 of 1996 and the Assessment Policy in the GET Band Grade R to 9 and Adult Basic Education and Training (ABET).

The DoE introduced periodic systemic evaluations with the aim of responding to the transformation agenda of the democratic government in the new dispensation post the apartheid era. Systemic Evaluation was introduced according to the National Education Policy Act (NEPA), Act No. 27 of 1996 and it is meant to assess the extent to which the education system achieves set social, economic and transformational goals of access, redress, equity and quality. Evaluations are conducted at the key transition stages of the schooling system, namely, Grades 3, 6 and 9 (DoE, 2001).

The Purpose of Systemic Evaluation

The main purpose of SE is defined as measuring the "effectiveness of the entire system and the extent to which the vision and goals of the education transformation process are being achieved by it" (DoE, 2003, p. 2). SE uses all relevant information impacting on the system such as resources and facilities, leadership, management and communication, governance and relationships, curriculum provision, teacher characteristics and learner characteristics, where learner performance is a key indicator, amongst others. To this end, learner performance is used as one of the key

indicators of education performance, alongside the contextual factors that impact on teaching and learning (DoE, 2001).

According to the SE Mainstream Report, the first ever South African systemic evaluation in education was conducted through a survey in 2001 at the Grade 3 level that falls under the Foundation Phase (FP) level of the schooling in South Africa. The survey was guided by the transformational goals and a set of twenty-six agreed upon educational indicators to develop the instruments. The survey focused on learner performance and the conditions in which teaching and learning take place. The survey was preceded by a pilot study in 2000 (DoE, 2003).

The SE survey sought to establish factors that might have a bearing on equity and access to quality education by learners. The analysis of the data showed that most of the contextual factors looked at in the study could be correlated with learner performance (DoE, 2003). For example, the findings of the SE survey revealed that there was a correlation between learners' Early Childhood Development (ECD) experience and their performance in the tests: where there was a dearth of resources, learner performance tended to be low and home background factors correlated positively with learner performance (DoE, 2003, p. 23).

1.4 PROBLEM STATEMENT

The demography of the South African society clearly reflects a rural-urban divide across all communities. The legacy of apartheid can be clearly seen through the discrepancies around the so-called ‘digital divide’ in South Africa. Only 56% of South Africans, for example, have access to electricity (Gillwald, 2000, p. 1). Nine percent of households have a telephone, but bear in mind that only one percent of households are rural households. In addition, in excess of 9 million people live in informal dwellings (Gillwald, 2000). The problem is further exacerbated by the fact that just under half of the population (approximately 20 million people), live below the poverty line (i.e. income less than R353 per adult per month).

Poverty is closely correlated with race, gender and the rural-urban divide. Nearly 75% of South Africa’s poor live in rural areas. That translates into two-thirds of South

Africa's poor living in three predominantly rural provinces – Eastern Cape (24%), KwaZulu-Natal (21%) and Limpopo (18%) as recorded in the Globalisation in Education Report (DoE, 2000b). This serves to confirm the notion of the "two economies" existing side by side according to the socio-economic setting of our society (Mbeki, 2004). The parallels of this divide and the two economies are strongly characterised by affluence on the one side and abject poverty on the other.

The rural-urban divide problem is also extended and experienced within the education sector. There is glaring inequity in terms of resourcing of schools from the apartheid era to date. Infrastructure and other facilities still bear testimony to the imbalances and a perpetuation of the past. Communities and schools in rural areas have to contend with the scarcity and lack of adequate resources, for example, lack of running water, electricity, sanitation, communication and dilapidated buildings. Availability of libraries and laboratories is also still a widespread problem for schools in rural areas (Human Sciences Research Council [HSRC], 1997).

In rural areas, both learners and teachers, in most instances, have to travel longer distances to get to school. The likelihood that some learners may come to school without having had something to eat is also high, given the high levels of poverty in rural areas. The Globalisation in Education Report (DoE, 2000b) attests to these problems in that the historical inequities in education have given rise to internal inefficiencies and unequal educational outcomes. More than three years after the first democratic elections, the School Register of Needs Survey (SRN) (HSRC, 1997) reflected these inequities in stark terms. This first national School Register of Needs survey, conducted in 1996, marked a decisive departure from the opaqueness of apartheid education planning which sought to hide the wide inequalities of education in South Africa. By quantifying the provision of physical infrastructure for education, the new democratic government established a rational basis for pursuing a planning strategy based on equity, democracy and justice (DoE, 2000c).

The second School Register of Needs survey was conducted in 2000 in order to update the 1996 data as well as to use the updated data to measure progress and trends between 1996 and 2000 in the provision of physical infrastructure and basic facilities. However, in 2000 about 17% of learners (1.9 million) were still without toilet

facilities. Fifteen percent of toilets were not working at the time of the survey with the majority of those being in the rural areas. Approximately 35% of all schools nationwide had no access to any form of telecommunications and one in every three schools reported that they had no access to water (DoE, 2000c).

The 2000 SRN shows that, in spite of transformation, many schools are still without basic infrastructure such as water, sanitation, power supply and telecommunications. It further points to the fact that there are many small rural schools which are under-resourced (DoE, 2000c). The vast majority of children in South Africa continue to be educated in conditions of extreme neglect. In 1996, one in every four schools had no water within walking distance, and nearly one in ten had to get their water from dams and rivers. Over half (57%) did not have electricity and in addition, over half (52%) had pit latrines for toilets, while 13% had no ablution facilities at all. There was no learning equipment in 73% of schools, and 69% had no learning materials. The level of library provisioning was appalling, with 72% of schools having no library collections. Finally, approximately half of the schools in the most rural provinces had no sports facilities (DoE, 2000b).

The discourse around South Africa as a 'developmental state' is fast gaining momentum both locally and internationally. The demands and challenges of democracy and transformation of education in South Africa are espoused in the Constitution and put to action through the National Education Policy Act (NEPA), Act No. 27 of 1996. These pieces of legislation require that scientific research studies be undertaken to elicit baseline information on the performance of the education system towards meeting the set goals. As a duty of the state to redress the imbalances of the past, it is prudent that data, collected at systems level, is able to generate a debate and provide a feedback loop on issues of equity. The concept of equity is based on the principle that all learners can learn and achieve and that they should be provided with an equal opportunity to do so irrespective of their backgrounds. Achieving equity is vital, as it is indicative of whether the education system is meeting the different needs of all sectors of society. Without addressing issues of equity, the provision of quality education cannot be attained. According to Herman and Klein (1996) equity is defined on the basis of providing opportunities for learning to learners based on access to resources, qualified teachers and instructional tools, access to the types of instructional

content and extra school opportunities, and direct preparations and practice. However, little information is currently available for clear, targeted interventions on the issue of equity for redress in the rural-urban divide of the South African education system and society.

1.5 RATIONALE FOR THE STUDY

There is a growing emphasis on the importance of mathematics, science and technology. At least two studies, namely TIMSS (1999) and SACMEQ II (2004), which compare the performance of South African learners with that of learners from other nations, indicate that South African learners rate poorly in mathematics and science, even in comparison with other developing nations. South Africa's own Grade 6 Systemic Evaluation Report that established a baseline at the Intermediate Phase level and released in December 2005, also tells unpalatable news about educational quality in the country. As expected, Dieltiens (2006, p. 5) states that the lowest performing learners came from township, rural and farm schools.

The rationale behind this study, therefore, is to analyse performance in numeracy of learners from various backgrounds and contexts (rural and urban) within which teaching and learning take place. Khoali (2006, p. 9) argues, "a particular concern is around the huge inequalities still evident in the resource allocation in many black schools, as well as the poor performance of black students in the national examinations and in comparative studies such as TIMSS". Van den Berg (2005) also supports this assertion that learners in formerly black schools still perform poorly as compared to those in formerly white schools. This study will assist in eliciting an underlying understanding of learner performance in numeracy in the South African education system.

This study has importance since valuable information will be generated to help the system focus on the relevant interventions to improve levels of performance in numeracy at the Foundation Phase level of the schooling system. The information generated will be specific and useful in terms of responding to the actual problems faced by learners in both urban and rural settings and contexts.

This study is also unique in the sense that it is the first of its kind to analyse numeracy performance of the Grade 3 learners in a nationally representative sample, which was also the first systemic evaluation conducted in South Africa. It is also unique because it is the first secondary analysis of numeracy learner performance focusing on the rural-urban divides from a nation-wide systemic evaluation conducted in South Africa.

It is important to note that this study will add value in assisting policy- and decision-makers to examine policy achievements and challenges using set transformational goals of access, equity and quality. The study is significant because it will help examine and analyse further the SE 2001 numeracy learner achievement scores and in addition, it will expand on the valuable baseline information yielded by the SE survey.

This study does not only analyse and explain learner performance, but it also further identifies equity factors that could be used to predict the performance of learners from various backgrounds in numeracy. A study conducted by Howie (2002) found that within school-level factors (school location), the location of the school had a significant effect on South African pupils' mathematics achievement. Children in rural schools achieved significantly lower scores than those in urban areas. Whereas Howie's concentrated on the Grade 8 learners at secondary school level, this study will measure the significance of the location of the school on Grade 3 learners' numeracy achievement at primary school level.

Another study, conducted by Young (1997), found that Australian pupils in rural areas also achieved lower scores in mathematics. In the case of South Africa, Howie (2002) asserts that it was not surprising given the remarks made by Hofmeyer and Buckland (1992) that African children in homelands and rural areas were educationally more disadvantaged than those in urban areas. This current study expands on Hofmeyer and Buckland's as well as Howie's studies by explaining the factors specially related to equity affecting rural and urban school learners' achievement in numeracy.

1.6 RESEARCH QUESTIONS

The main research question that is addressed by this study exploring the rural-urban divides of the South African society is:

What are the equity factors that have an effect on the numeracy performance of Grade 3 learners in rural and urban primary schools?

This study forms part of a broad systemic process of monitoring and evaluation within the South African education system. The secondary analysis undertaken in this study was within the context of Grade 3 learners who took part in the SE survey of 2001. This study will undertake an analysis that is both descriptive and exploratory in nature.

In addressing the main research question, this study undertakes a secondary analysis that explains the overall performance of Grade 3 learners in the system. A brief description of the participants in the SE survey will be made. The analysis will also explore how the learners from rural schools fared on the whole in numeracy as compared to their counterparts from urban schools.

Descriptive research questions:

1. *Who are the learners, educators, parents and schools that participated in the SE survey of 2001?*
2. *What is the overall performance of Grade 3 learners in numeracy?*
3. *How does the performance in numeracy of Grade 3 learners in rural schools compare to their counterparts in urban schools?*

This study will explore the contextual factors related to the transformational goal of equity that have an effect on the numeracy performance of Grade 3 learners from both the rural and urban schools. The variables for the equity factors are derived from the contextual questionnaires that were administered during the survey. The purpose was to elicit information on which of these equity factors, from the school and home background levels, had an effect on learner performance in numeracy. On the home background level, an investigation into the effect of equity variables like books, possessions, meals, socio-economic status and education levels of parents and whether they can predict performance in numeracy, will be made. On the school level, pedagogical equity factor variables that relate to the educational methods and principles like availability of learning and teaching support materials, teacher qualifications, facilities, nutrition and funding, will be explored including their predictive ability for performance in numeracy.

Exploratory analysis research questions:

4. *What equity factors in this study have an effect on learner performance in numeracy?*
5. *What home background equity factors predict learner performance in numeracy?*
6. *What pedagogical equity factor predicts learner performance in numeracy?*

As a point of departure, this study is premised on a quantitative research approach due to the nature of Systemic Evaluation survey instruments, analysis methods that were undertaken in this study and the interpretation of the results/findings thereof. This study is anchored on the objectivist epistemology. As a theory of knowledge, objectivism purports that research should be conducted with objectivity and that the findings are comparable with the line of reasoning and knowledge of research conducted before (Guba & Lincoln, 1994).

On the philosophical stance or paradigm regarding methodology, this study is grounded in the post-positivistic theoretical perspective. According to Creswell (2003) underlying post-positivism is a deterministic philosophy with an assumption that outcomes are determined by causes. Therefore, this study seeks to explore the equity factors that have an effect on numeracy achievement. A lucid description and rationale for the research approach, epistemological and paradigmatic assumptions are discussed in detail in Chapter 3.

1.7 STRUCTURE OF THE DISSERTATION

Chapter 2 provides a review of the literature relating to similar studies on learner achievement in numeracy, both nationally and internationally. The review focuses on the performance of rural and urban learners (also including gender, age and language), factors influencing their performance in numeracy/mathematics and the equity indicators that could be used to predict performance in numeracy. In this chapter, a list of key search words is given to guide the literature review for this study. In addition, a review of the literature relating to similar studies on learner achievement in numeracy both nationally and internationally is conducted, guided by the research questions.

In Chapter 3, the research design for SE and this study is discussed. The conceptual framework is also outlined. The research methods chosen to address the research questions for Phases 1 and 2 of this study are discussed in this chapter. The rationale for the choice of these models that have been adopted and adapted for this study is discussed. A link is also drawn between the literature review, the conceptual framework and the research questions. A brief description of the design of the SE study process (sampling, instrument development and data collection) is provided. A description of the design and technical details of the sample for this study are also given in this chapter.

The data analysis methods for this study are discussed in detail in Chapter 4. A post hoc analysis of the data is undertaken as a basis for the secondary analysis in this study with the purpose to scrutinise the data. The preparation of the data for secondary analysis is discussed and the data analysis methods are explained with the rationale for their selection fore grounded. Both the descriptive (e.g. frequencies) and exploratory (reliability, correlation, factor and regression) analyses are presented in this chapter.

The main findings of the study from the descriptive and exploratory analyses are presented in Chapter 5. Here the results of the descriptive statistics and regression analysis are presented, interpreted and discussed.

In Chapter 6 the summary of research questions and approach are given. The summary of the findings and a brief reflection on the study methodology and the conceptual framework are provided. Limitations of the study are discussed in detail here. Finally, the recommendations emanating from the findings and conclusions are presented. This will help to inform future research and interventions to improve on learner performance in numeracy. An overall summary of the study is also made to conclude this chapter.

CHAPTER 2
LITERATURE REVIEW

In this chapter a literature review is conducted to identify information on learner performance in numeracy. The aim of this study is to describe and compare the performance of Grade 3 rural and urban school learners in numeracy, and to explore the influence of contextual equity factors on this performance. Several studies have investigated factors influencing the achievement of learners in mathematics/numeracy, however, not many of these studies were found in developing countries (Howie, 2002), hence this study. An extensive literature review is presented to provide information on how urban and rural learners perform in numeracy and which factors have been found to have an effect on this performance.

2.1 INTRODUCTION

The purpose of this chapter is to present a review of the literature based on the main research question with a view to identify information from research conducted previously that would help inform and guide this study. This literature review would then lead to the conceptual framework for this study (Chapter 3) that seeks to address the main research question.

The main research question for this study is: *What are the main equity factors that influence the performance of Grade 3 rural and urban school learners in numeracy?* This question required that an extensive literature review be done on learner performance in numeracy/mathematics globally. Therefore, it led to information searches largely rooted in research on mathematics achievement. Information on equity and other key factors that were found to influence learner performance in mathematics were also explored. The question of educational indicators also became critical to the classification of the factors regarding the transformational goals of access, redress, equity and quality.

The main sources of literature explored in this section were mainly searches in published hard copy and online research reports and articles in accredited journals using a few key words: numeracy/mathematics achievement, equity factors/indicators, educational indicators and national assessments. The point of departure was the Systemic Evaluation

report that was produced by the DoE on a study conducted at the Foundation Phase (Grade 3) level in 2001. Other sources covered a wide spectrum of contexts and environments ranging from developing and developed countries both regionally and internationally on the subject of mathematics achievement and factors influencing it. A wide range of books on educational research and assessments were also explored. In addition, Internet searches were also conducted using the identified search key words for this study.

Firstly, a brief synopsis of research on mathematics achievement in South Africa is provided (2.2). A discussion of research into international studies on mathematics achievement is presented in 2.3. A focused discussion on research on mathematics achievement of rural and urban school learners follows (2.4). Other research studies that were conducted on factors that were found to have an effect on learner achievement in mathematics locally, regionally and internationally are discussed (2.5) in relation to this study. Equity factors considered to have an effect on mathematics achievement are also explored in this section. A summary of the literature review is presented in 2.6. In conclusion, a critique of the literature is discussed in support of the argument for the rationale behind this study (2.7).

2.2 NATIONAL STUDIES ON MATHEMATICS ACHIEVEMENT IN SOUTH AFRICA

This section focuses on national studies that were previously conducted in South Africa on mathematics achievement. A number of these studies were conducted through NGOs in a number of provinces earmarked for interventions and accelerated improvement in learning. However these studies made use of small samples because they were targeted at improving learning in particular sites. Systemic Evaluation is one of the national studies conducted in the post-apartheid era on numeracy and literacy achievement using a nationally representative sample. The first ever systemic evaluation at the Foundation Phase (Grade 3) level was discussed in Chapter 1. An overview of the baseline SE survey at the Intermediate Phase (Grade 6) level is given in the next subsection followed by other studies conducted on mathematics achievement in the country.

SYSTEMIC EVALUATION (SE)

The DoE conducted a Systemic Evaluation survey at the Intermediate Phase (Grade 6) level in 2004. The purpose of the survey was to establish baseline data on learner achievement and conditions of teaching and learning at the Intermediate Phase (Grade 6) level. The survey involved a sample of approximately 35 000 learners from about 1 000 schools (about 7% of the mainstream public schools with Grade 6) that were selected randomly from all the nine provinces for this purpose. The respondents in the survey were Grade 6 learners, parents/guardians/hostels, teachers of the affected Learning Areas, principals and District officials who all completed specially-designed questionnaires on the contextual factors that may affect teaching and learning in the school system (DoE, 2005).

The main findings show that in general, learners obtained the lowest scores for numeracy (national mean of 30%). The distribution was skewed heavily towards the lower scores, with a high concentration of learners obtaining scores between zero and 40%. These results confirm those of the studies that were conducted in South Africa, viz. MLA (1999) SACMEQ II (2000) and TIMSS (1995, 1999) to show a trend from 1995 immediately after 1994. The general trend indicated that learners' average scores varied according to location of school. This trend showed that learners in rural areas were found to be at a disadvantage regarding their level of performance in mathematics. Average scores ranged from a high of 45% for urban schools, 26% for township schools, 23% for farm schools, and 21% for rural schools and levelled off at 19% for remote rural schools (DoE, 2005).

Academic performance of learners was greatly impacted upon by factors within and outside of the school. The pattern of underachievement was largely aligned to past policies of education provisioning with rural schools worst affected. The majority of the 33 679 Grade 6 learners sampled from 1 000 schools across the country achieved disturbingly low average scores, for example, the language average was 38% while the mathematics average was 27%. A significantly high percentage of learners across all three areas assessed functioned at the "not achieved" level, for example 63% in language and 81% in mathematics (DoE, 2005).

The study found that the performance of learners was similar across all the provinces, and that the average score for girls was slightly higher than that of boys in all three Learning

Areas assessed. Moreover, the scores obtained for the multiple choice questions (MCQs) were significantly higher than those obtained for the Free Response Questions (FRQs). This result indicates that learners performed better on tasks that required them to identify and select a correct response than on tasks that required them to produce their own response. The language in which learners undertook their tests also had a significant effect on their performance. Learners who wrote the tests in their home language obtained significantly higher scores across all learning areas than their counterparts who undertook their assessment tasks in a second or third language (DoE, 2005).

Some of the key findings on the contextual factors impacting on teaching and learning showed that socio-economic status (SES) emerged as an important predictor of learner-achievement with learners in the lower SES categories obtaining significantly lower scores. In scores for mathematics, the difference was approximately 27% and just over 30% for language. Learners in well-resourced schools also obtained significantly higher scores with percentage differences of 49% in language and 27% in mathematics as compared to those in under-resourced schools. Critical additional predictors identified were discipline; learner engagement; homework feedback; parental support; access to information at home and at school; school attendance; teacher resources; school fees and repetition rates (DoE, 2005).

In the 2001 study at the Grade 3 level, equity factors that were found to have influenced learner performance were:

- Discipline, safety and learning atmosphere (largest positive influence),
- Private contributions and utilisation of funds, and
- Functioning of SGBs (explained 9% of the differences in learner performance).

On the other hand, quality indicators influencing learner performance were facilities in school, and teaching and learning materials (DoE, 2003). The availability of these quality indicators are explored in this study as part of the group of equity factors that have an effect on numeracy achievement of learners in rural and urban primary schools.

The wide disparity in the fees levied by schools indicates that schools have inequitable access to private contributions. The norms and standards for the funding of schools had, in the SE study, not been able to demonstrate the extent to which it has been effective as an

instrument of redress. As a result, this study will further explore the effect that this equity indicator might have had on learner performance in numeracy at the Foundation Phase.

QUALITY LEARNING PROGRAMME (QLP)

Funded by the Business Trust, the Quality Learning Programme (QLP) 2000-2004 was implemented in 524 high schools selected by the nine provincial departments of education. The QLP was based on a systemic design, in which training and support programmes were aimed at achieving better management of districts and schools and improved classroom teaching. A longitudinal evaluation (HSRC: 2000 – 2004) found that over the life of the project, QLP schools achieved significantly better results in the matriculation examination than selected control schools, in terms of greater numbers of overall passes, university exemptions and passes in mathematics and English. The evaluation concluded that the project had a significant effect on the performance of these very poor, largely rural, schools. QLP schools showed improvement relative to control schools in a number of areas (Kanjee & Prinsloo, 2005).

At the classroom level, significant improvements were noted in the degree of curriculum coverage completed by QLP classes, teaching to the appropriate level of cognitive demand, and the quantities of reading, writing and homework given. Path analysis modelling revealed that QLP interventions affected the functioning of the system in districts, schools and classrooms, improving indices of functionality relative to those for control schools at all three levels. These improvements, in turn, were associated with improved learner performance. Most notable was the effect of language-across-the-curriculum interventions on the overall matric pass rate: the implication is that good reading and writing skills are a prerequisite to good performance in all subjects and that intervening in this area can affect significant improvements in pupil performance (Kanjee & Prinsloo, 2005).

Although the matric results of most schools in the country improved significantly over this period, the QLP outperformed those of both a set of control schools and the national mean, particularly in mathematics. While improvements in the actual number of passes and the pass rates produced by QLP schools are only slightly better than those of the national mean, the QLP schools show very large differential improvements on three indicators of

quality: number of exemptions, and the number of maths passes at both standard (SG) and higher (HG) grades (Kanjee & Prinsloo, 2005).

QLP focused on high schools whereas this study focuses on learners at the primary school level in both rural and urban areas and in particular, their performance in numeracy. Equity factors that have a significant bearing on the teaching and learning of numeracy, for example, systemic input in terms of resources, educators and funding support will be explored to determine their effect on achievement.

DISTRICT DEVELOPMENT SUPPORT PROGRAMME (DDSP)

The District Development Support Programme (DDSP) on Grade 3 Learner Assessment (2001) was an education improvement initiative of the South African government funded by the United States Agency for International Development (USAID). The programme was implemented in conjunction with the DoE and four provincial departments of education, namely the Eastern Cape, KwaZulu-Natal, Northern Cape and Limpopo.

The District Development and Support Project (2000-2002) was the first initiative based on an explicit systemic design (HSRC, 2003). Working in 453 primary schools in the four poorest provinces, interventions were directed at improving the functionality of districts and schools and improving classroom teaching in language and mathematics. Objective tests of pupil performance in literacy and numeracy at Grade 3 level were conducted during each year of the programme, and again one year later. Significant changes were recorded, and these were holding steady a year after the closure of the DDSP.

The goal of the DDSP was to improve the quality of educational delivery for Grades 1 to 9 in the identified areas. That would entail, amongst others, an improvement in the quality of teaching and learning. In the area of curriculum development, one of the ways the DDSP intended achieving this was through the national Assessment Modelling Initiative (AMI). The fundamental purpose of the AMI was the development and piloting of an assessment model on the district level to contribute towards and inform the development and implementation of a fully functional national assessment system in South Africa (HSRC, 2003).

The Grade 3 numeracy test, adapted from the International Association for the Evaluation of Educational Achievement (IEA) survey conducted in 27 countries, was used to conduct the Grade 3 baseline study during 2000. The tests were translated from English into eight other languages offered by the DDSP schools. The tests were then administered during October 2001. The mean percentage correct for numeracy showed a fairly weak picture of the performance of the learners. It was found that the girls outperformed the boys in the numeracy tasks. The reason for this was not evident and required further investigation (HSRC, 2003).

The findings of the DDSP initiative were based on 610 schools across four of South Africa's poorest provinces. Similarly, an assessment was conducted at Grade 3 in numeracy in the respective languages of learning and teaching at this level. This study compliments the DDSP since the sample used is big enough (5%) to generalise the findings to the Grade 3 population in the rest of the country. A detailed analysis of the performance of Grade 3 learners is undertaken in this study across the rural and urban locales of all the nine provinces in South Africa.

School Effectiveness Studies

The school effectiveness studies were undertaken to identify the mechanisms for effecting the achievements of successful schools and have since become a major research focus in South Africa. Large-scale descriptive studies had failed to make any progress towards this goal, largely because of a lack of data at the school and classroom levels. Small-scale studies, on the other hand, which take these latter aspects as the principal focus for their research, had also not made much headway in revealing this secret. This was in part because they were too small in scale to generalise beyond the very particular circumstances of the respective case (Taylor, Muller & Vinjevold, 2003, p. 66). Such had become the dearth of empirical knowledge at classroom and school level which was needed to improve on teaching and learning that it led to the commissioning of school effectiveness studies.

The first major school effectiveness study, the Pupil Progress Project (PPP), was launched in specific response to the paucity of knowledge in what constitutes an effective school. The PPP was restricted to one province, the Western Cape, where a sample of 90 primary schools was investigated across the full range of socio-economic status, language, learner

performance and geographic location (Van der Berg, Burger & Yu, 2005). Data was collected on SES, family educational practices, school management, classroom teaching and learner performance in language and mathematics. Conventional regression techniques revealed strong associations with learning in the following areas: poverty, language, home educational practices and school management (Taylor, 2006).

The PPP confirmed language as the most powerful influence on learning, after poverty, particularly as children are severely disadvantaged when the home language and the language of instruction (LOLT) do not coincide. In the home, the PPP found that learning is enhanced when parents speak to their children in the LOLT, and where children read and do homework frequently. At the level of school management, the PPP concluded that the deployment of textbooks, instructional leadership and time management all correlated significantly with learning. With regard to instructional leadership, learning is enhanced when principals lead the production of curriculum year plans by teachers and monitor their implementation (Taylor, 2006). The effect of the independent variables namely language, teaching practices and LTSM on numeracy scores of learners in both rural and urban areas, is further explored in this study.

No classroom level factors emerged as significant correlates of learning in the PPP. This is almost certainly due to the research design because the PPP was a snapshot study, with a single point of measurement. It was unable to link the characteristics of any one teacher to student performance, since the learning exhibited by a child at Grade 6 level is the product of six years of schooling, generally involving at least six different teachers. Longitudinal studies are required to isolate the effects of individual teachers and their specific practices on pupil learning (Taylor, 2006). The same can be said about the Systemic Evaluation study whose design could not as well allow for individual educator-learner linkages because of lack of unique educator identification numbers in a school. In this study, an attempt is made to link learners to individual educators at the school level.

A further school effectiveness study conducted in the Western Cape, in 24 poor schools chosen so as to minimise SES differences (Reeves, 2005), examined the effects of pedagogical practices on mathematics learning. The study controlled for prior achievement, administering tests at the start and end of the school year, and was therefore able to link gain scores exhibited by pupils with the classroom practices of individual

teachers. Reeves (2005) found that teaching style (i.e. child-centred vs. teacher-centred) did not matter as much as certain features of pedagogical practices such as teacher responsiveness to learners' levels of ability and progress, and engaging learners at relatively high levels of cognitive demand regarding knowledge. In this study, a once-off 'dipstick-kind' evaluation of the system was done using learner performance and the effect on performance of the contextual equity factors without pre- and post-testing of the learners.

Another important finding in the Western Cape study was that more time spent in mathematics classes is related to achievement gain: simply by attending class, students have a learning opportunity, and inevitably are more advantaged. Reeves' findings also suggest that, although gain scores within any one year depend on both the subject knowledge of the teacher and her ability to understand her pupils and to pace delivery of the curriculum accordingly, the degree of curriculum coverage across grades, from one year to the next, may be a powerful cumulative factor in building pupils' knowledge (Reeves, 2005).

2. 3 INTERNATIONAL STUDIES ON MATHEMATICS ACHIEVEMENT

South Africa has only recently joined a number of international comparative studies in education. To date, South Africa participates in three cross-country comparative studies: TIMSS, PIRLS and SACMEQ. The message coming from all three studies is unambiguous: the country performs poorly compared with many of its more impoverished neighbours, and very poorly in relation to developing countries in other parts of the world. For example, in the latest round of SACMEQ testing conducted in 2000, of the 14 Southern and Eastern African countries participating, South Africa was placed ninth in both reading and mathematics at Grade 6 level (Taylor, 2006).

There have been a number of reports and articles written on the status of mathematics education in South Africa (Taylor & Vinjevold, 1999). Two of these have focused on primary level, Monitoring Learning Achievement (MLA) and Southern Africa Consortium for Monitoring Educational Quality (SACMEQ) and two focused on secondary level (TIMSS and TIMSS-R).

MONITORING LEARNING ACHIEVEMENT (MLA)

The MLA project began in 1992 and was a joint project of UNESCO and UNICEF. The aim was to examine the effectiveness of the basic education provision in terms of actual learning attainment (Chinapah, et al., 1999). The project forms part of the Education for All (EFA) initiative.

Learning attainment of Grade 4 learners was measured by means of achievement scores in literacy, numeracy and life skills and a criterion-referenced approach was used. The results showed that the learners in South Africa performed poorly in numeracy (Howie, 2002, p. 30). No analysis was done of the effect of equity factors on learner performance, as well as disaggregating how the learners from rural areas performed as compared to those from urban areas.

THE SOUTHERN AFRICA CONSORTIUM FOR MONITORING EDUCATIONAL QUALITY (SACMEQ)

SACMEQ was initiated in 1991 when a number of Ministries of Education in the Southern Africa Sub region started working together to address the need for systematic studies of the conditions of schooling and of learner achievement levels. The focus of this work was to establish long-term strategies for building the capacity of educational planners to monitor and evaluate basic education systems. The first two SACMEQ projects (SACMEQ I and SACMEQ II) focused on an assessment of the conditions of schooling and of the quality of education and included achievement data on reading literacy.

South Africa only participated in SACMEQ II, which commenced in mid-1998 (Howie, 2002, pp. 29-30). Data were collected in 2000 and were only reported in late 2004. The learners' responses were scored and a comprehensive analysis was done to establish learners' levels of achievement in reading and mathematics and to examine whether there were differential levels of achievement according to gender, socio-economic level and location of the school (SACMEQ II, 2004). Data on the learners' scores were analysed using computer software that applied the Rasch Model of measurement. The Rasch Model applies a programme that performs a linear transformation of the test scores into a scale on which the pre-determined mean score is 500 and the standard deviation is 100.

The results showed that, for South Africa overall, the Grade 6 learners performed notably lower in mathematics obtaining a mean Rasch score of 486, 2 which was below the pre-determined mean Rasch score of 500. The modal competence obtained by 44,4 % of the learners was at level 2, which was described as emergent numeracy and equated to performance at the level of Grade 3 (SACMEQ II, 2004). South Africa performed significantly behind Swaziland with a Rasch score of 516.5, in comparison with Botswana: 512.9 and Mozambique: 530.0.

The picture to emerge from numerous results similar to these described above, is that South Africa is not getting value for money from its public school system (Taylor, 2006). Taylor further posits that although schooling is accessible to the majority of children, the skills produced are expensive and their quality low. This affects both the trainability of adults in the workplace and the educability of school leavers entering the Further and Higher Education sectors. A prominent feature of the system is the very high degree of inequality between schools, as shown by the differences in scores between high- and low-SES schools on the SACMEQ tests. This is obviously a legacy of the past and one which the present government is finding difficult to reverse (Taylor, 2006). While SACMEQ investigated the degree of inequality between schools using SES, this study explores the comparative effect of SES on the performance of rural and urban school learners and also as a possible predictor of performance in numeracy.

It is important to note that Grade 6 is a key transition stage between the Intermediate and the Senior Phases of the schooling system in South Africa. The significantly low numeracy competencies are cause for serious concern. One of the major challenges facing the education system is the provision of the right number of appropriately qualified mathematics educators in all the schools. It has been argued in the report that this is probably one factor that could have contributed to the particularly low levels of competence in mathematics among the Grade 6 learners.

A prominent feature of the system is the very high degree of inequality between schools, as shown by the differences in scores between high- and low-SES schools on the SACMEQ tests. This study investigates the effect of SES on the performance of learners across the rural-urban divide and further explores predictability of this factor in future performance in numeracy.

THIRD INTERNATIONAL MATHEMATICS AND SCIENCE STUDY (TIMSS)

The International Association for the Evaluation of Educational Achievement (IEA) conducted the Third International Mathematics and Science Study (TIMSS), now known as the Trends in Mathematics and Science Study, between 1990 and 1995, which was repeated as TIMSS-R(epeat) in 1998 and 1999. Internationally, the study covered Grades 3 and 4 but South Africa decided not to participate at this level. This was because of the complexities of the languages in which the tests would have had to be administered. South Africa learners in the Foundation Phase are taught in their mother tongue and this would have called for the translations of the tests from English into ten other official languages. The major findings for TIMSS-South Africa revealed that the results for Grades 7, 8 and 12 were low in comparison with other countries. The performance of the South African learners was uniformly poor across all the content areas in mathematics.

Howie (2002, p. 33) contends that the results of the TIMSS study highlighted the disparities in the schooling system with the pupils in schools from traditionally disadvantaged areas achieving lower scores than pupils from schools in more advantaged areas, although exceptions were found in individual scores. Howie further argues that these differences can be partly explained by unequal provision of resources and facilities in the past, different socio-economic backgrounds, as well as the variation of the home languages of the learners. Location was also found to be a highly significant predictor of achievement with urban learners achieving much higher scores. Similarly, in this study location is also, firstly, investigated in terms of a broad comparison in performance of rural and urban school learners and secondly, as a possible predictor of achievement in numeracy.

2.4 MATHEMATICS ACHIEVEMENT OF RURAL AND URBAN SCHOOL LEARNERS

The following are some studies that were conducted locally and internationally with a view to analyse and report on the achievement of urban and rural school learners in numeracy/mathematics.

SACMEQ II

In the SACMEQ II study, the achievement of Grade 6 learners in South Africa was analysed according to the location of the school that they attended. Using indicators such as availability of amenities and facilities, schools were grouped into designated categories namely city, small town and rural schools. Overall, there were worryingly wide differences in achievement with the worst by learners who attended school in rural settings (Moloi & Strauss, 2005).

It is further reported that a significant majority of rural learners achieved at the pre-numeracy (11,8%) and emergent numeracy (59,6%) levels. In relation to the South African curriculum statement's assessment standards, these levels are at equivalent levels of Grade 2 or lower and Grade 3, respectively. However, learners in urban (city) schools achieved better with just over 17% achieving level 6, concrete problem-solving or the equivalent of Grade 7 (Moloi & Strauss, 2005).

PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT (PISA)

Using PISA 2000 data, Williams (2005) examined cross-national variation in rural mathematics achievement among 15-year-olds in 24 industrialised nations. Rural mathematics scores were significantly lower than scores in urban and medium-sized communities in 14 of 24 countries. However, patterns were complex. Most commonly, a linear relationship was obtained between community size and average maths score. In some countries, however, students in medium-sized communities scored highest, followed by urban then rural locales. Following on William's study (2005), this study focuses on a broad range of learner ages namely, 8 to 13+ year olds at the Foundation Phase of the primary school level in a developing country.

In some countries, such as the United States of America (USA), students in urban communities scored lowest. Williams (2005) contends that one explanation for lower rural achievement was lower SES. Consistent with other studies, the USA showed a marginal raw rural achievement gap, which disappeared when SES was controlled. Once SES was controlled, rural locale predicted mathematics scores in only 4 of 24 countries. Only in Russia was rural locale a statistically significant negative predictor of mathematics achievement, net of socio-economic status. However, the USA showed a substantial gap in

urban achievement but further analysis suggested positive interaction effects between school SES and both urban and rural location (Williams, 2005).

NAEP

The Nation's Report Card Mathematics by Braswell, et al. (2001) indicates that in 2000, National Assessment of Educational Progress (NAEP) conducted a national mathematics assessment of Fourth-, Eighth-, and Twelfth-Grade students in the USA. State-level results were also collected at the fourth and eighth grades within participating states and jurisdictions.

In 2000, fourth-, eighth-, and twelfth-graders in central city schools had lower average scores than their counterparts in urban fringe/large town schools. Fourth- and eighth-graders in central city schools had lower average scores than their counterparts in rural/small town schools. Fourth-graders in urban fringe/large town schools had higher scores than their counterparts in rural/small town schools.

Whereas NAEP again focused on learner performance in mathematics in an industrialised country namely the USA, this study conducted in South Africa, seeks to investigate learner performance in a developing country, but also further explore the equity factors as possible predictors of future numeracy performance of rural and urban school learners.

2.5 FACTORS THAT HAVE AN EFFECT ON LEARNER ACHIEVEMENT IN MATHEMATICS

General factors that have an effect on learner achievement in mathematics

In this section, a review of the literature emerging from other similar and previous studies on mathematics achievement, locally and internationally, is conducted. The purpose is to draw lessons from these on factors that were found to have an effect on learner achievement in mathematics. This assists with the development of the conceptual framework and the research design of this study as well as to enrich it with experiences of what other researchers have done previously.

Socio-economic status (SES)

During the SACMEQ II data analysis for South Africa, Moloi and Strauss (2005) reported that an index for "possessions at home" was constructed and used as a proxy for the SES of the learner. The index comprised of reading and electronic items such as monthly magazines, newspaper, cassette player, video-cassette recorder and a telephone. The index ranged from 1 to 14 and learners whose homes possessed less than half of the items were classified as of a low SES. Those who possessed more than half the items were classified as of a high SES. The SACMEQ II results show that learners from high SES backgrounds achieved higher levels of competence than their counterparts from low SES (Moloi & Strauss, 2005).

Howie (2002) in drawing conclusions regarding factors influencing mathematics achievement of South African pupils, writes that on student-level factors, the home environment has been found to have a positive relationship with mathematics achievement. Socio-economic status (SES) was also found to have a significant relationship with achievement in mathematics (Afrassa, 1998; Howie & Pieterson, 2001; Mohandas, 1999 and Young, 1997) and children from wealthier homes in general achieved higher scores for mathematics. Sojourner and Kushner (1997) found that SES was the most consistent predictor of mathematics achievement of all the factors they examined in their study. On the contrary, in South Africa, a study by Monyana (1996) revealed no significant difference between children from different SES backgrounds. However, several studies have found a positive relationship between the number of books in the home and learner achievement in mathematics (Mullis, et al., 2000; Shen, 2001). Similarly, in this study, SES is explored to investigate its effect on numeracy achievement as a predictor of future performance of Grade 3 learners.

The results of Payne and Biddle's study (1999) on funding, poverty, and mathematics achievement concluded that the level of school funding and child poverty (SES) have substantial and statistically significant net effects on average student achievement among the school districts of America. These effects stand up even when juxtaposed with those of two crucial, district-level control variables. On the other hand, Turner's review (2000) of the same study suggests that students attending schools in districts with more resources should have achievement levels at least as large as students attending schools in districts with fewer resources, if all other characteristics of districts were the same. Turner (2000)

further posits that family income is positively related to student achievement. Funding and provisioning of resources across schools in rural and urban districts of South Africa is researched in this study to determine its effect (as indicators of the extent to which there is equity in the system) on numeracy achievement.

Classroom resources / teaching aids

A study by Mullis (1991) established that on the whole, the use of teaching aids was found to have a positive relationship to achievement. It was also found that where learners who were exposed to working with learning aids such as rulers, blocks, shapes and solids, there was a positive link with geometry achievement in Grade 4 (Howie, 2002). The schools' locality is related to achievement and pupils in rural areas achieved lower scores (Young, 1997). In contrast, school resources were generally found to have no effect on learner achievement in two separate studies conducted in the USA and Hungary (Arnold, 1995; Vari, 1997). The comparative achievement of learners in rural and urban schools is at the heart of this study. Therefore an in-depth investigation will be conducted to determine the net effect of resources and learning and teaching support materials (LTSM) on learner achievement in numeracy.

Language

It is argued that mother tongue/home language is the best medium of education at school because it is the language the child knows well, and in which he or she can express meanings (Pattanayak, 2003). The language of learning and teaching (LOLT) is the official language of instruction in a school decided upon by the school governing body (SGB).

The language used for testing learners in SE was the LOLT at the Foundation Phase as specified by the school. In each case, the LOLT was employed for instruction from Grades 1 to 3. All eleven official languages in South Africa are used as LOLTs in the Foundation Phase. The assessment instruments were administered to learners in the LOLT of the school.

Prior knowledge, skills and achievement

Kurdek and Sinclair (2001) conducted research in the USA on predicting reading and mathematics achievement in Fourth Grade children from kindergarten readiness scores. Their research examined age and gender differences in verbal skills and visual-motor skills at kindergarten, in achievement in reading and mathematics at Grade 4, and the link between skills at kindergarten and later achievement.

Kurdek and Sinclair's study was found to be consistent with findings that younger children in kindergarten perform less well than their older counterparts (Kinard & Reinherz, 1986; Shepard & Smith, 1986; West, Denton & Germino-Hausken, 2000). With controls for age, verbal skills uniquely predicted later mathematics achievement. Readiness in the specific areas of auditory memory, number skills, and visual discrimination predicted later mathematics achievement. However, it is important to note that the sample used in this study was primarily White with regard to ethnicity and middle class, making it unclear whether the findings would generalize to more ethnically diverse and less economically advantaged samples, as to be seen in this study.

Similar to the SE survey findings on ECD, a study conducted in the USA by Denton and West (2002), entitled Children's Reading and Mathematics Achievement in Kindergarten and First Grade in responding to the question, *Do children's knowledge and skills differ by certain child, family, and school characteristics?* concluded that differences in children's achievement (as represented by their overall achievement score) by their family's poverty status, race/ethnicity, and school type, persist from kindergarten through First Grade. However, children's overall reading and mathematics achievement does not vary by their sex.

Parental Education

A study on Students' Mathematics Achievement in Japan was conducted and Sugiyama (2001) writes that a high level of parental education often produces a good educational environment, which in turn has a good influence on children. Almost all households subscribe to newspapers. Children are often exposed to situations in which they must express and deal with something using numbers. Nurtured in a situation where adults are very good at counting and computing, children feel a strong desire to emulate them. The more they are exposed to situations that use numbers in daily life, the better the

environment is for mathematics education. Parents' level of education is one factor that is investigated in this study to determine its effect on learner achievement in numeracy.

In contrast, Prins and Toso (2008), supported by a number of similar studies conducted in the USA on family literacy and parent education, present a view that asserts that the effects of parenting practices and styles on academic outcomes differ considerably by race/ethnicity, SES, generational status, and socio-cultural setting (Kao, 2004; Spera, 2005). This view is also further argued by Valencia and Black (2002, p. 83) that "students, particularly of low SES and of colour, fail in school because they and their families have internal defects, or deficits, that thwart the learning process". This study will thus explore the effect of similar and other equity factors that have an effect on achievement in numeracy for the diverse population of learners in South African public schools.

Equity factors that have an effect on mathematics achievement

Equity factors are derived from a set of education indicators that are statistical values designed to provide information about the condition, stability, functioning or performance of an education system (Strauss, 2005). Implicit in Strauss's definition of indicators is the notion that they deal with the measurable features of an education system. Education indicators show the nature, functioning, efficiency or outcomes of the quantifiable facets of an education system at a national, provincial, regional, school or classroom level, and can thus give an indication of the quality and potential problems of the education system as measured against set standards (Strauss, 2005).

The concept of equity is based on the principle that all learners can learn and achieve and that they should be provided with an equal opportunity to do so, irrespective of their background. Achieving equity is vital, as it is indicative of whether the education system is meeting the different needs of all sectors of society. Without addressing issues of equity, the provision of quality education cannot be attained. According to Herman and Klein (1996) equity is defined on the basis of providing opportunities to learning to learners based on access to resources, qualified teachers and instructional tools, access to the types of instructional content and extra school opportunities, and direct preparations and practice.

In developing countries such as South Africa, stark differences in learner achievement along the urban-rural divide continue to characterise the schooling system and remain a top priority of policy intervention in many countries. In South Africa, the issue of equity is particularly salient given the discriminatory racial policies of the apartheid government, which left a legacy of huge differences in learner achievement between the different racial groups. In the South African context, the achievement of equity is also important because of the gross levels of differential school funding that benefited the White population. Consequently, the present democratic government has made the achievement of equity in education funding a priority (DoE, 2005).

Rural – Urban divide

After 1994, most important initiatives have been introduced in South Africa to develop a fair and equitable society and to amend past inequalities of the apartheid regime. The new education system in South Africa is aimed at developing a fair and equitable system that affords good quality education and training to all learners all over the country. As Pretorius and Lemmer (1998) put it, the main aim is to provide equal educational opportunities for all.

With the introduction of democracy in 1994, the challenge for the education policy makers was to design a system of equitable distribution of opportunities for all learners in the country. The Department of Education developed and implemented a number of education and training policies with a view to transform the education system to make it more inclusive. Nevertheless, Yamauchi (2004) argues that opportunities for education in public schools in South Africa are still unequal between racial groups, even after apartheid.

In the USA, Christie (2001) in her paper titled *Helping state leaders shape education policy,* argues that schools in rural America face an array of problems every bit as daunting and intractable as those confronting schools in urban communities. The disparities between these schools have become worse to the extent that in rural communities, graduating students who see no future locally leave town, and a snowball effect begins. States are beginning to address this problem, however, by expanding educational access for students, grassroots organizing, policy research, and training for rural activities and school boards. These challenges facing rural America are much the same as those experienced here in South Africa, and has resulted in a huge influx towards

the urban areas in search of greener pastures, exacerbating poverty in the rural communities.

Providing equal opportunities for all to learning

On the international front, the International Association for the Evaluation of Educational Achievement (IEA) first introduced the concept of Opportunities to Learn (OTL) as an instrument to validate the differences in student's mathematics achievement across different nations (McDonnell, 1995). The concept was developed to ensure that no schools provide students with unequal opportunities to learn, irrespective of their abilities to learn the subject matter.

There is strong evidence that suggests that OTL is a major equity issue for students who are at risk of not developing academically to their fullest potential (Stevens & Grymes 1993, p. 3). These learners are often classified as poor, disadvantaged or from low socio-economic status. Research has also shown that schools that are highly funded receive a better quality of education and have more advantages than those that receive little funds (McDonnell, 1995; Schwartz, 1995). On the same score, a body of research indicates that unequal provision of education and unfair distribution of resources affects students' learning (Guiton & Oakes; 1995; Oakes, 1990; Oakes, 1992; Oakes, 1985). This study seeks to either confirm or disprove the effect of unequal provision of resources on numeracy achievement in rural and urban schools in South Africa.

Availability of teachers and quality teaching

A study on teaching quality found that students with qualified teachers and who had teaching experience tend to perform better than those with less qualified teachers (Greenberg, Rhodes, Ye & Stancave, 2004). This is also because experienced teachers are more effective with students than inexperienced teachers (Goldhaber, 2002). In order for students to be offered equal opportunities to learning, Oakes and Lipton (1990) maintain that all teachers should have a teaching qualification. They further state that students in poor and minority schools in the USA suffer from poorly qualified teachers because their schools find it difficult to attract qualified and experienced teachers.

Class size

Quality teaching is determined by the size of the class (Pritchard, 1999). This implies that the bigger the class, the less the quality of instruction and the smaller the class, the more the quality of teaching. A number of studies have been conducted on how the class size affects learning. In their studies, Bennet (1996); Pritchard (1999); Rice (1999); Nye, Hedges and Kostantopoulos (2000); Rousseau and Powell (2005) demonstrate that small classes provide better learning environments than large classes. Class size and the number of learners in the classroom impact on the nature of teaching and learning process (Anderson, Ryan & Shapiro, 1989; Bennet, 1996). In large classes, it is difficult for teachers to use small group practices or more investigative approaches as large classes affect the interaction between the teacher and the learner (Anderson, Ryan & Shapiro, 1989).

Furthermore, class size affects particular aspects of classroom practices and learners' classroom characteristics. Smaller classes are perceived as easier to manage and are involving fewer disruptions and behaviour problems (Pritchard, 1999; Rice, 1999:3). Nevertheless Nye, Hedges & Kostantopoulos (2000) and Pritchard (1999) contend that small classes in early grades appear to have lasting benefits. Their follow-up study revealed that learners who experienced more years of small classes in the early grades have higher levels of academic achievement than learners who have fewer years of small classes.

Teaching and learning resources

The availability of resources for teaching and learning are critical in influencing the outcomes. According to Schwartz (1995) students should have access to textbooks and educational facilities so that they are able to learn. Lack of teaching and learning materials affect students and encourages rote-learning techniques by teachers and learners and also interferes with the teachers' capacity to teach well (Elmore, 1997). In addition, Oakes and Lipton (1990) argue that schools should first have enough teachers because fewer teachers mean large classes, which are difficult to control. Inadequacy of teaching and learning resources places a strain on the teaching-learning process.

In another study on OTL issues common to South Africa and the USA, Murray (1997) found that there was a strong relationship between student achievement and race, and

inadequate resources and funding. This implies that students who attend ill-resourced schools are likely to perform poorly as compared to those who attend well-resourced schools. This point will be teased out in this study to see if the different levels of funding and resources available in the school either have a positive or negative effect on learner achievement in numeracy.

2.6 SUMMARY OF KEY FINDINGS FROM THE LITERATURE REVIEW

A number of different issues linked to learners' achievement were researched in this section. Internationally, most research addressing factors related to mathematics achievement were found using data from the USA. No studies were found nationally that attempt to explore the comparative influence of the equity factors on numeracy achievement of rural and urban school learners.

South Africa has participated in a number of international studies namely, TIMSS, TIMSS-R, MLA and most recently, SACMEQ. The results of all these studies on the whole showed that the learners performed comparatively poorly in mathematics. The results of Systemic Evaluation at Grade 6 level confirmed findings of earlier studies. Taylor, Muller and Vinjevold (2003, p. 41) posit that: "Studies conducted in South Africa from 1998 to 2002 suggest that learners' scores are far below what is expected at all levels of the schooling system, both in relation to other countries (including other developing countries) in relation to the expectations of the South African curriculum". With the advent of systemic evaluations, South Africa will now be able to measure how the education system performs.

The findings on equity and other factors that were found to influence learner performance were also presented in this section. These findings warrant in-depth research on how these factors play out on the South African education system given the stark differences and realities of the rural-urban divides.

Some of the equity factors discussed included: the disparities in the schooling system for example unequal provision of resources and facilities; different socio-economic backgrounds; and language were found to account for differences in learner scores. The positive relationship between the number of books in the home and learners' achievement

in mathematics, family's poverty status, race/ethnicity, school type, the level of school funding, educational environment. Most of these factors that are related to equity are explored in this study.

On the classroom and school level factors, the availability and utilisation of teaching and learning resources (including textbooks); school location; teacher qualifications; and the learning environment (including teacher/learner ratios) found to have an effect on learner achievement in mathematics will also be investigated in this study. These findings lay a solid foundation towards the research design and the conceptual framework presented and discussed in the next chapter.

All research studies on the quality of schooling in South Africa concur that poverty remains far and away the most powerful determinant of educational opportunity. The PPP study found that between two-thirds and three-quarters of the variance in pupil scores is explained by socio-economic factors (Van der Berg, Burger & Yu, 2005). Interestingly, the poor appear to be more constrained by their school circumstances in acquiring numeracy skills than in literacy. In other words, the school is more important for learning numeracy than for learning literacy, or literacy is easier to learn at home than numeracy. Effective educational practices occur in the home, the system, the school and the classroom. Such practices can be categorised into five broadly defined factors: language, time management, curriculum coverage, reading and writing, and assessment.

2.7 CONCLUSION

The contrasting findings about the effect of resources on achievement are very critical since the availability of resources is still a contested issue dividing schools in South Africa due to the legacies of the past regime. This study therefore explored their effect on the performance of learners in schools from rural and urban backgrounds.

Family income or SES is one factor or variable that is explored from the SE data to investigate whether it is the case with learners from the rural-urban divide of the South African society that the analysis of resources at home as an index is positively related to achievement as argued by Turner (2000).

On the question of school location, the SACMEQ II findings for South Africa showed that generally, learners in urban schools achieved better than their counterparts in rural schools. A variety of equity factors, for example, availability of resources and SES can be attributed to this achievement.

This study focused mainly on the findings of the collected SE data against the above-mentioned factors explored in Howie's (2002) and other studies, for example home language, SES, gender and school location. It is critical for this study to bring to the fore what the state of affairs is at the Foundation Phase (Grade 1 – 3) of schooling in our system. It is important to note that whereas Howie's (2002) and most of the studies, e.g. SACMEQ II, MLA and TIMSS, that were conducted in South Africa focused on mathematics achievement of learners at the Intermediate and Senior Phase (lower secondary) levels of schooling, this study focuses on the performance of the learners in primary schools at the Foundation Phase (Grade 3) level.

CHAPTER 3
CONCEPTUAL FRAMEWORK AND RESEARCH DESIGN

The conceptual framework and methodology for this study is designed to explore equity factors to determine what effect they have on numeracy achievement of urban and rural primary school learners. The hypothesis for this study is that equity factors have an effect on numeracy performance and can predict it.

In this chapter, the conceptual framework and the research design for Systemic Evaluation (SE) and this study are outlined and discussed. In 3.1 the rationale for the choice of the theoretical framework or the research paradigm on which this study is grounded is discussed. The models adopted and adapted for this study are explained. The SE research design and methodology are presented in 3.2. This is immediately followed by the limitations of the SE survey (3.3). In 3.4 the operational research questions anchored around the main research question for the two phases of this study, namely the descriptive and the exploratory are presented and discussed. The design issues related to the SE sample and the sub-sample for this study are also briefly described and discussed in 3.5. Lastly, the validity and reliability issues relating to the instruments are discussed (3.6).

3.1 CONCEPTUAL FRAMEWORK FOR THIS STUDY

The aim of this study is to describe the performance of the Grade 3 rural and urban school learners in numeracy. The study also seeks to explore the influence of the school and home background factors related to equity on this performance. The objective of the first phase of this study is to provide a description of the participants and the overall performance of Grade 3 learners in the numeracy assessment task from the SE survey. Phase two focuses on the identification and analysis of the main factors that have an effect on learner performance in numeracy. These factors are explored in relation to the background information from the contextual questionnaires that were used during the SE survey data collection.

Theoretical framework

Creswell (2003) posits that there are three critical framework elements to consider when designing a study. These elements are: 1) philosophical assumptions about what constitutes knowledge claims, 2) strategies of inquiry, and 3) methods. To situate this study within the quantitative research approach, the theoretical framework is discussed from the philosophical viewpoint, the strategy/method of inquiry, and the procedure for obtaining collected data and analysis as was undertaken in this study.

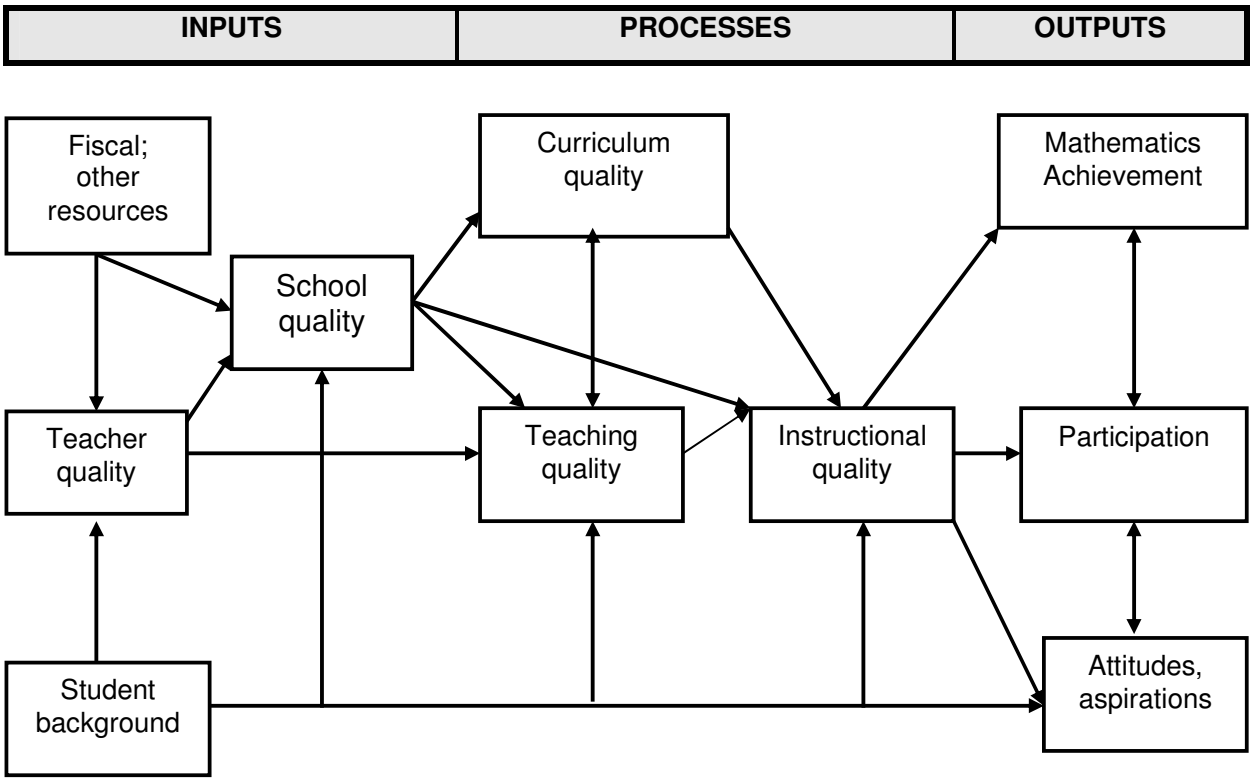
The conceptual framework for the design of this study was developed based on ideas by Crotty (1998) who believes that the paradigm used to situate a study will inform the theoretical or conceptual stance and the methodology. This study concurs with these ideas in that the epistemology (theory of knowledge) that informs this study is objectivism (Creswell, 2003). Objectivism requires that research be conducted with objectivity to ensure that the results fit existing knowledge (Guba & Lincoln, 1994). The theoretical perspective (philosophical stance or paradigm) behind the methodology in the research questions is post-positivistic. A quantitative secondary analysis of data obtained from a survey research (non-experimental design) informed the methodology (strategy that links methods to outcomes) in this study. The methods (techniques and procedures) employed in this study were derived using survey instruments that included analysis of data from questionnaires and learner achievement scores.

Lincoln and Guba (2000) further assert that stating a knowledge claim, which might be called a paradigm, means that a study is undertaken with certain assumptions about what is to be learnt during the research. In the light of that, as stated above, this study follows on the post-positivistic knowledge claim that is sometimes called the “scientific method” research (Creswell, 2003). Creswell goes on further to state that post-positivism reflects a deterministic philosophy in which it is assumed that causes probably determine effects or outcomes. At the heart of this study is the quest to test the hypotheses whether- and which- equity factors 1) have an effect on numeracy achievement and 2) to explore how the identified equity factors can be used to predict performance as reflected in the main research question (Creswell, 2003).

Models

This study is premised on the extent to which the system is responding and meeting the transformational goal of equity in pursuit of redressing the imbalances of the past. The concept of equity is based on the principle that all learners can learn and achieve and should be provided with the opportunity to do so irrespective of their background (DoE, 1995). It is therefore critical that the interplay amongst the various components of the system, namely inputs, processes and outputs is closely monitored. The goal of Systemic Evaluation is to measure and evaluate the critical components of the system in a way that the linkages between these are kept in the balance. This is a breakaway from the previous system of education that was solely focused on the outputs.

The model that was found to be most applicable as a basis for this study is the original framework (Figure 1) derived by Shavelson, McDonnell and Oakes (1987).



Source: Shavelson, McDonnell & Oakes (1987)

Figure 1. Linking elements of the educational system.

The model was adapted accordingly to address the research questions for this study. Howie (2002, p. 71) suggests that to monitor a dynamic system such as education, it is important to depict this in a way that linkages between components of the system can be ascertained and evaluated. The model by Shavelson, McDonnell and Oakes (1987) was developed from an extensive literature review to illustrate the linkages between the social indicators and educational research that are related to educational outcomes.

In order to address the objectives of this study, it is clear that the conceptual framework for this research must inform the research questions. These questions seek to explore the equity indicators related to the factors that influence learner performance in numeracy in the context within which teaching and learning take place. In Figure 2, the model by Shavelson et al. has been adapted. The three main components of the education system namely; inputs, outputs, and processes are integrated within a context where teaching and learning take place in the public schooling system in South Africa.

The context is all-important for the issue of equity and when considering the inputs, processes and outputs as the linking elements of the education system. The disparities (if any) in the system regarding these and differences in performance are explored in this study. These are also used to predict how learners from both the urban and rural contexts are likely to perform in future.

The equity indicators inform the linkages between the components in this model. These equity indicators serve to indicate the possible relationship between the elements of the education system. The linkages are meant to illustrate how the factors presented are inter-linked. This is, therefore, about how the inputs could have an influence on the processes and outputs as well as the influence of the processes on the outputs in a particular context within which teaching and learning takes place. In Figure 2, these have been added to and adapted in Shavelson's original model by depicting the various equity indicators explored in this study that might have an effect on learner performance in numeracy.

Figure 2 presents the linking elements of the education system in terms of inputs, processes and outputs. The arrows in the model point to the direction of the interaction amongst the variables from the inputs, processes and through to the outputs. First and foremost, a learner comes from a home, thus the direction of a similar category of equity

variables flowing from the learner level (at home) and feeding on to the system level (at school). The bold arrows indicate the critical interplay amongst the inputs and processes culminating in the outputs. In simple terms, the school is looked upon to make up for any shortcomings from the learner's home. In the model, the *inputs* are depicted in terms of the equity indicators, which are then analysed in terms of their effect on teaching and learning (processes) and in turn, in terms of achievement (outputs).

Equity is explored in this study in the provision of all the variable factors listed in the model that might have an effect on learner achievement, as shown in the model, and the possible interrelationships. The inputs into the system affect all the processes of education delivery both directly and indirectly.

Inputs

The socio-economic status (SES) is linked to funding provided for both the financial and other resources supplied or demanded at the different levels of the system; for example, school fees, learning and teaching support materials (at school and in the home), provision of adequately qualified teachers, qualifications of parents and nutrition (both at home and at school).

Both the inputs from home and the school (Figure 2), are reflective of the SES of the parents and that of the surrounding school community. SES refers to the level of affluence, poverty or in simple terms, the standard of living of the people determined by income and possessions in the home. In this regard, the levels of funding from the Department as well as the parents' contributions through school fees, are linked to SES. The ability of the parents to pay school fees with or without exemption is usually indicative of their SES. The funding also provided to schools by government, that is the national norms and standards for school funding, is one critical input factor disbursed on a pro-poor equitable share formula not only to redress the imbalances of the past, but also to offset the levels of poverty of the formerly disadvantaged schools and communities. In this regard, this study explored the effect of SES on numeracy achievement of rural learners as compared to their urban counterparts.

Learning and Teaching Support Materials (LTSM) refers to those possessions that are available in the home that provide educational support to learners on what is learnt at

school. Materials in the home include amongst others television (TV), video recorders (VCR), radio, computer, books, textbooks and newspapers or magazines. These are also highly dependent on the SES of the parents, as the higher the SES of parents, the more likely the higher their levels of income, literacy and education and therefore, the higher the likelihood to find more LTSM in the home.

At the school level, the capacity of the school to source quality learning and teaching support materials (LTSM) is also dependent on funding available for such. Some examples of the LTSM variables identified in this study for that purpose are learner exercise books, collection of books, numeracy teaching aids, calculators, computers, worksheets or textbooks, radio, television, overhead projectors and the Internet.

Meals and the Primary School Nutrition Project (PSNP) are essentially meant to provide nutrition and are available for learners at both the home and school respectively. The introduction of the latter, which is the school nutrition programme in primary schools, was meant to alleviate the plight of those learners whose socio-economic circumstances predispose them to hunger and starvation (RSA, 1994). The programme is a good initiative with the potential to reduce suffering and enhance active participation by all learners, rich and poor, in the education system. The PSNP is implemented in all public primary schools with the aim of providing feeding for needy children.

Qualifications refer to the highest education level attained by the parent or guardian of the Grade 3 learner. Parent qualification does not flow into the teacher qualification at the system level because it has no influence on it. The line linking the two is only meant to show similarity of the variable. Sheldon and Epstein (2005) argue that national and international studies have made student performance in mathematics a high priority in schools. They indicate that after control for prior levels of mathematics achievement, analyses indicated that effective implementation of practices that encouraged families to support their children's mathematics learning at home was associated with higher percentages of students who achieved proficiency or more on standardised mathematics achievement tests. Historically, parents have played important roles in some aspects of education reform (Bloch & Tabachnick, 1994, p. 196). On the part of teacher qualification variables, reference is made to their highest school and academic qualifications, level of professional training and highest level in mathematics.

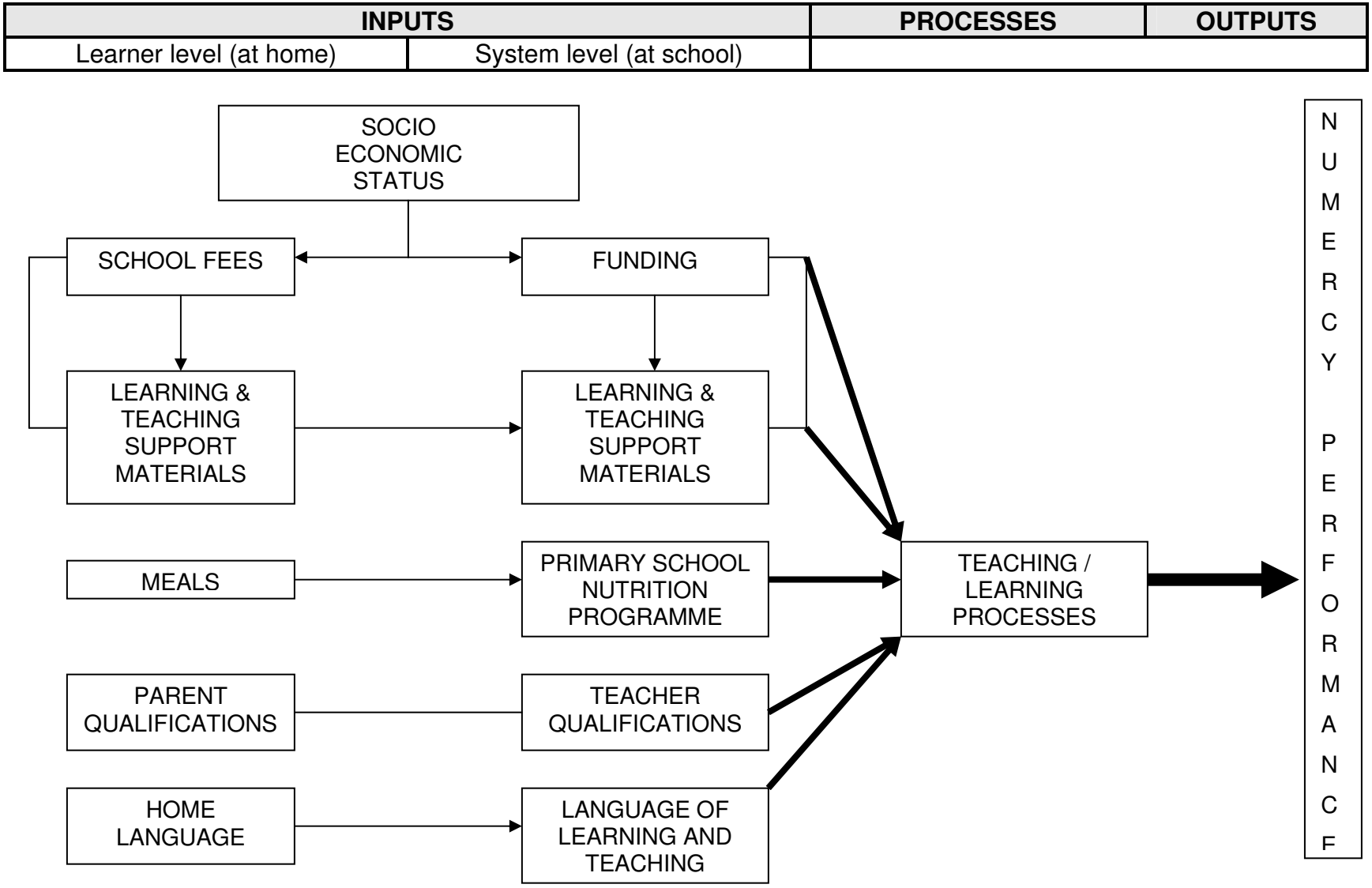


Figure 2. Equity factors that have an effect on numeracy achievement, adapted from Shavelson, et al. (1987)

Language is also an important input factor as a medium through which knowledge can be imparted, accessed and communicated. In this conceptual framework, the home language of the learner (also referred to as the mother tongue), the language of the parent and the language frequently spoken at home are explored. The language of learning and teaching (LOLT) adopted by the school, namely the medium through which the curriculum is delivered, is also explored. The LOLT is supposed to be a medium through which all the learners will be provided an equal opportunity to learn. The LOLT, also the language of the test, in the school is also explored against the home language of the learners.

Howie (2002, p. 35) asserts that with eleven official languages in South Africa recognised by the Constitution, the language of instruction is a complex and sensitive issue. The President's Education Initiative (PEI) research (Taylor & Vinjevold, 1999) suggested that few schools have developed formal language policies. It was found that learning through the primary (home) language in the early years of schooling was on the decrease. In general Non-English speaking parents see English as a means to social and economic advancement and as a language of power. As a result, they put pressure on schools to offer English as the language of learning from the onset of schooling and send their children to pre-primary schools where English is spoken. This might have had an influence on learner achievement for those learners who wrote the systemic evaluation test in English, which may be their second or third language. Therefore, the influence of language on numeracy achievement is also explored in this study.

The survey by the Joint Education Trust in 1998 on teaching in multicultural classrooms found that language plays an important role in the formulation and expression of concepts in mathematics (Howie, 2002, p. 36). This study takes forward the JET survey in that it explores the extent to which language, acquired by the learners at Grade 3, can be used to predict performance in numeracy.

Processes

Different *processes* relating to what is taught, how it is taught and its quality take place within the schools, and inside the classroom (Howie, 2002). These processes are essentially about

the teaching and learning processes in terms of the support given through the inputs and the medium through which the curriculum is delivered. They culminate in the outcome of education processes giving an indication of what the learners have been able to achieve from the system. In this regard, the processes culminate in the numeracy scores attained by the learners in the SE test.

Outputs

The *outputs* generally refer to the outcome of all the education processes. In this study, the only measure of outputs chosen for the adapted model is achievement or performance. This achievement is reflected by the outcome in terms of the aggregate numeracy scores from the results of the SE survey. The final aggregate score for numeracy that was explored is captured in the data as 'RNUM', which is the sum of the recoded values. The maximum possible score for 'RNUM' was 40 and the percentages corresponding to 'RNUM' are labelled 'PNUM'.

3.2 SYSTEMIC EVALUATION RESEARCH DESIGN AND METHODOLOGY

In this section, a number of design issues of the SE survey will be highlighted that are important for this study. The sample, instrument development, data collection and processing will be described and discussed in this section. The assessment instruments designed for literacy, numeracy and life skills were administered to approximately 54 000 Grade 3 learners. In addition, questionnaires were given to learners, their teachers, principals and parents or guardians to complete.

Sampling

The target population for the SE survey was all Grade 3 learners in public schools in South Africa. Learners were, therefore, the intended unit of analysis. According to the SE Mainstream Report (DoE, 2003, p. 9) a sample of five percent of all the Grade 3 learners from all districts/regions in the country was included in the survey, as reflected in Table 1 below. The school was the sampling unit. The list of all South African public schools with Grade 3 was used as the sampling frame. At the first stage of sampling, schools were categorised

according to the administration unit, which is the region in some provinces or districts in other provinces.

A cluster sampling technique was used to select the schools that participated in the survey. The number of schools to cover the sampled learners was randomly chosen from an alphabetic list of schools having at least 30 learners each. Five percent (5%) of the total number of Grade 3 learners was selected in each of the districts/regions. A fixed number of learners (40 pupils) within each school were randomly selected across the grade using simple random sampling during the second stage of sampling.

Table 1

SE school sampling, 2001

Province	Total Number of Learners	Intended sample of learners (5% of Total)	Achieved sample (Number of Learners Who Actually Participated)	Percentage representing achieved sample
Eastern Cape	213 915	10 696	9 456	4,4
Free State	57 699	2 885	2 889	5,0
Gauteng	126 321	6 316	6 220	4,9
KwaZulu-Natal	245 038	12 252	11 115	4,5
Mpumalanga	84 725	4 236	4 048	4,8
Northern Cape	16 397	820	913	5,6
Limpopo	164 577	8 229	8 062	4,9
North West	88 979	4 449	4 524	5,1
Western Cape	8 1601	4 080	4 080	5,0
Total	1 079 252	53 963	51 307	4,8

Source: SE Mainstream Report (DoE, 2003 p. 9)

Instruments

Officials of the DoE from both the national office and provinces largely developed the instruments used to assess learner performance in three learning areas in collaboration with the consortium of service providers namely, the Human Sciences Research Council (HSRC),

the Centre for Research in Education Policy and Development (CEPD) of the University of Witwatersrand and the Research Institute for Education Planning (RIEP) of the University of Free State (DoE, 2003). Learner assessment tasks comprised Numeracy, Literacy and Life skills.

The DoE officials and the consortium developed the numeracy instrument collaboratively. The assessment instruments were based on frameworks drafted by members of the consortium together with the DoE Foundation Phase specialists for the three Learning Programmes assessed during the survey. The frameworks took into consideration the intended (school system) curriculum in terms of what was considered to be the essential competencies that a Grade 3 learner should be able to master.

The numeracy assessment instrument comprised 20 questions that were scored out of a total of 40 points. The items were both closed and open-ended types of questions. Six of the closed-ended items were of multiple-choice type. The multiple-choice items consisted of a question and four choices for an answer, namely three distracters and the correct answer. Fourteen of the items were open-ended (13 short answers), where learners were required to complete a word or number sentence by circling, colouring in or writing only one word or numerical response and one extended response, where learners were required to show their working to get to an answer.

Schools were asked to indicate their languages of learning and teaching (LOLTs) in order to administer the tests in the language of learning at the particular school. The tasks were set in English and later translated into all the other ten official languages that were indicated by schools as their LOLTs. National and provincial officials of the DoE undertook the translation processes. These officials came from regions in which the languages that they had to translate the tasks into, were predominantly spoken, e.g. Tshivenda in Limpopo and Setswana in North West, and were considered to be competent in those languages. No evidence of how the translations were verified could be found in the report. All the instruments were piloted and suitable items selected for the final instrument (DoE, 2003, p. 8).

The contextual questionnaires were developed based on the constructs derived from the twenty-six (26) educational quality indicators that were adopted by the DoE. These indicators – developed from the context, input, process and output models of schooling – were placed against the education transformation goals of access, redress, equity and quality in a matrix format as captured in the Framework for Systemic Evaluation (DoE, 2001). The matrix offers a clear perspective on the place of the different indicators within a dual framework (DoE, 2003, p. 8). The questionnaires were administered at the schools to provide the context within which teaching and learning take place.

Six sets of questionnaires were administered during the survey. The *Principal* questionnaire dealt with the principal's background information and school background factors including enrolment, management and administration and curriculum. Grade 3 *teachers* completed the second set of questionnaires. These questionnaires included questions about the teacher's background, qualifications, teaching and curriculum support. The *learners* completed the third set of questionnaires. In this questionnaire, questions were asked about the learners' background and conditions about teaching and learning. The heads of *districts/circuits* completed the fourth set of questionnaires where questions were asked about management, development and support given to schools. The *parents* or *guardians* of the Grade 3 learners completed the fifth set. The questionnaire dealt with home background factors, satisfaction rates with children's learning and support given to the Grade 3 learner. The *hostel* master or mistress completed the last set of questionnaires. This questionnaire dealt with conditions in the hostel where learners resided during a normal school week. The response rate in this questionnaire was very poor such that no meaningful analysis could be undertaken for the SE study.

Data collection

The survey was conducted across all the nine provinces of the Republic of South Africa. The DoE officials from each province were trained to conduct the survey. Training was organised nationally and took place at six venues to ensure uniformity of administration. Furthermore, a data collection manual was given to individual data collectors as a resource to refer to during data collection. The SE data were collected over a period of two days in each school during

the months of September and October in 2001. To enhance national monitoring, each province had a coordinator and monitors who quality controlled the entire process and reported accordingly. The learner assessment tasks were coded and scored by officials of the DoE representing all official language groups. The HSRC provided the technical assistance during the coding and scoring process (DoE, 2003, p. 10).

Data coding and scoring, entry and cleaning

According to the SE Mainstream Report (DoE, 2003) trained officials of the DoE from provinces undertook the coding and scoring of the learner assessment tasks in November 2001. The HSRC provided the training of scorers and moderators. Both the HSRC and DoE Quality Assurance section officials moderated approximately five percent of the instruments scored.

In order to reduce entry errors, the data entry of the contextual questionnaires was done on a prepared template, which limited the entries to those responses appropriate for each question. For the assessment tasks, data was captured and cleaned according to a pre-specified format with the assistance of consortium members. The cleaning and verification process was conducted to ensure that the error rate was less (DoE, 2003, p. 10).

Limitations of SE survey

The information was gathered using a paper-and-pencil mode of assessment. The report is, therefore, limited to competencies (Assessment Criteria) that can be assessed only in written form.

The time it would take to complete the assessment tasks and the contextual questionnaires dictated that a limited number of questions could be included in the instrument. Competencies that, by their nature, 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 exhaustive nor conclusive in its coverage of all the competencies that learners in this phase are expected to achieve. The report only provides a ‘snapshot’ of the competencies that learners could demonstrate in a written form.

There was a reliance on self-report data on the questionnaires used in the SE study and therefore the findings on background factors should be viewed with this limitation in mind. Given that the provincial and district officials collected the data, teachers and principals may have under-reported certain types of behaviour in order to avoid sanction. They may also have over-reported in cases where it was assumed that this could attract positive attention (e.g. over-reporting resource deprivation in order to attract more resources to the school) due to social desirability.

3.3 RESEARCH QUESTIONS

The research design for this study is anchored around the main and operational research questions.

Main research question

The main research question for this study is: *What are the equity factors that have an effect on the numeracy performance of Grade 3 learners in rural and urban primary schools?* Arising from the literature review and the conceptual framework adapted for this study, a number of specific operational research questions for this study are fore grounded.

This study is divided into two phases. In phase one, the participants in the SE survey are described and include the performance of the Grade 3 learners in numeracy. Phase two explores and focuses on the identification, selection and analysis of the main equity factors that have an effect on learner performance in numeracy and how these could be used to predict performance.

Specific research questions

Phase 1: Descriptive research questions

The operational questions in this phase are:

1. *Who are the learners, teachers, parents and schools that participated in the SE survey of 2001?*

2. *What is the overall performance of Grade 3 learners in numeracy? How does the performance in numeracy of Grade 3 learners in rural schools compare to their counterparts in urban schools?*

The specific research questions flowing from the conceptual framework direct this study in Phase 1 and they are addressed in Chapter 4 and the results reported in Chapter 5:

1. *Who are the learners, teachers, parents and schools that participated in the SE survey of 2001?*

To answer this question, information concerning the background of the learners, of the teachers, of the parents and background information about the schools is provided. This question relates directly to the equity variables in the conceptual framework relating to learners' home background, teachers' qualifications, parents' qualifications as well as their SES and system level (school) factors. The description of the profile of the population that participated in the SE survey provides for a lucid context of the exploration in Phase 2.

2. *What is the overall performance of Grade 3 learners in numeracy?*

To address this question, it is important to provide a description of the overall performance in numeracy by gender, age, LOLT, home language of the learners, province and location of school.

3. *How did the Grade 3 learners in rural schools perform, compared to their counterparts in urban schools?*

This question is related to the outputs as illustrated in Figure 2. The overall results of the learners in rural schools are compared to those of the learners in urban schools. These overall results for numeracy are described in general by gender, by age, by language and by location of school.

Phase 2: Exploratory research questions

As indicated earlier, in this study Phase 2 is exploratory and is driven by the main research question, namely: *What are the main equity factors that have an effect on the numeracy*

performance of Grade 3 learners in rural and urban primary schools? The methodology to explore the specific research questions in this phase is addressed in Chapter 4 and the results of the analysis are reported in Chapter 5. A number of specific questions guide the research in this phase:

4. *What equity indicators in this study have an effect on learner performance in numeracy?*
In this section, the various factors that were found to have an effect on learner achievement in mathematics during the literature review are lifted from the SE study data with a view to explore them further for selection in this study as predictors in preparation for the regression analysis.

5. What i) home background and ii) pedagogical equity indicators predict learner performance in numeracy?
The selected indicator variables related to equity, for example parents and teacher qualifications, LTSM, facilities (both at home and in school) and funding are explored to measure their effect on achievement. This will, in turn, determine how they could be used to predict performance of learners in rural and urban schools.

3.4 SAMPLE

SE sample
The SE sampling process was discussed in 3.2 above. It is important to remember that the thrust of this study lies in the urban-rural divide. From the SE study, the categories for urban are subdivided into urban and semi-urban while the categories for rural are semi-rural and remote-rural respectively. The SE categories were recoded as urban and rural for this study.

On the whole, the sample for SE involved 51 307 Grade 3 learners randomly sampled in about 1 000 schools from across nine (9) provinces. Amongst other participants involved in the SE study were 1 309 principals, 2 508 educators and 47 353 parents/guardians of the Grade 3 learners.

Secondary analysis sample (Sub-sample)

The original SE sample could be used in its entirety since a problem was encountered during the data files matching of the various respondents. This was due to a large number of educators in the more than one-educator schools who were not assigned unique numbers, and could therefore not be linked to the individual learners. As a consequence, the target population for the secondary analysis was the learners in the one-educator rural and urban primary schools. As presented earlier in 3.2, learners were the intended unit of analysis. The procedures and steps followed to arrive at this study sample, are discussed in detail later in Section 4.5 that deals with the selection of the variables for regression modelling. This was important to consider since it would be scientifically sound to work with a matched dataset containing learners with complete cases for the variables to be included in the model.

The sample used in the model excludes all the schools with more than one educator. It therefore did not make sense to apply the weights that were calculated for the original sample including all the schools. For purposes of this study's sub-sample, there were 478 one-educator schools with 15 805 learners (about 40% of learners in the original sample). Of the 15 805 learners, only 8 833 made it into the final model due to large missing data or incomplete cases for variables to be used to undertake a meaningful regression analysis.

The final number of participants, which could be used for modelling during the equity variables selection and the subsequent regression analysis, informed the sub-sample for this study.

There were 2 508 educators in the original SE sample, 19% of which are from one-educator schools and could be linked to individual learners that they taught. About 80% of the educators could not be meaningfully included in the secondary analysis sample because they could not be linked to individual learners. A glaring problem with the data was that these educators were not allocated unique ID numbers during the administration of the study and the subsequent capturing of the data. In this study, the sample is taken from the SE survey for the descriptive phase and a sub-sample is drawn for the exploratory phase. The selection of the participants in the sub-sample for this study is discussed in detail in Chapter 4.

3.5 VALIDITY AND RELIABILITY ISSUES

In this section validity and reliability issues pertaining to the instruments used during the SE survey are discussed. Secondly, a post hoc analysis is done in this study to test the reliability of the scales used in the contextual questionnaires. Therefore, by exploring the post-positivistic principles of validity and reliability, generalisability and prediction are critical outcomes that this study sought to achieve (Dawes, Horan & Hackett, 2000).

Validity

Gay and Airasian (2003) state that validity refers to the degree to which the instrument measures what it purports to measure. A measure should reflect that which it intends to measure in a valid way. According to Gay and Airasian (2003), the most important form of validity for a scholastic achievement assessment is content validity. It is always critical to consider whether the test instrument assesses all aspects of the curriculum appropriately and to what extent the content for the phase is covered.

Content validity refers to the adequacy with which a measure has sampled from the intended universe or domain of content (Gay & Airasian, 2003). The SE learner assessment instruments were designed to cover most aspects of the new curriculum of Outcomes-Based Education (OBE), (DoE, 2003). Most of the content was based on Curriculum 2005 (C2005). Key knowledge and skills were assessed at the appropriate levels for a Grade 3 learner.

The numeracy learner assessment task comprised a pencil and paper test taken by the learners aimed at assessing their mathematical abilities. It was developed with C2005 in mind. No record of the assessment frameworks that were developed to ensure that the items were fairly spread across the Specific Outcomes (SOs) could be found. In Table 2, a post hoc analysis of the classification of the items in the numeracy test is given according to the SOs as reflected in C2005.

Table 2

Classification of the items in the numeracy instrument

Specific Outcomes (SOs)	Description of SO	No. of Items
1 & 2	Development of number concept	7
4	Mathematical relationships	2
5	Measurement	1
6	Data management	3
7 & 8	Shapes and space	4
9	Communication	3
Total		**20**

Table 2 shows that seven items in the test covered SOs 1 and 2. SO 4 was covered by two items, SO 5 one, SO 6 three, SOs 7 & 8 four and SO 9 was covered by three items. Only one SO (3), that deals with the historical development of mathematics, was not covered. No explanation could be provided as to the least number of items for SOs 4 (2 items) and 5 (1 item) respectively since no record of the assessment task frameworks was found. Even though it was not stated or reported, not all the assessment criteria that can be assessed through pencil and paper could be covered. This could have been due to the limited number of items that learners would be able to handle in the allocated time of the test. Although there could be doubts about the quality of the test, it is important, however, to note that the test still reflects the appropriate competencies and content knowledge expected of a Grade 3 learner.

Reliability

According to Gay and Airasian (2003) reliability refers to the attribute of consistency of measurement. It indicates the degree to which the total score obtained in a test is free of measurement error. The degree of internal homogeneity of the questions in a test can be used as an indication of the reliability of a test. Gay and Airasian (2003) further posit that the Alpha coefficient provides an indication of the homogeneity of a set of questions. The stronger the interrelationship between questions, the higher Alpha will be to a maximum of 1.

The reliability of a scale indicates how free it is from random error. Internal consistency can be used to assess the reliability of a scale. It refers to the degree to which the items that make

up the scale are all measuring the same underlying attribute (that is the extent to which the items hang together. Cronbach's Alpha Coefficient measures internal consistency. It provides an indication of the average correlation among all the items that make up the scale (Gay & Airasian, 2003).

The reliability alpha coefficients were calculated to determine the reliability of the SE learner assessment instruments. The alpha coefficient was used as it is widely accepted as a fundamental index of reliability. The alpha coefficient for the numeracy instrument was 0.84 of which the results could be regarded as highly reliable (DoE, 2003, p. 57).

The consortium that included the HSRC and the CEPD, validated the SE learner assessment instruments. Contextual factors instruments were validated by the RIEP. The items were piloted for suitability and refined before the main SE survey was conducted. Various stakeholders, such as curriculum specialists from the DoE as well as international specialists from, amongst others, the International Centre for Assessment Research at the University of London, the Foundation for Education Research in the United Kingdom, and individual experts from the United States of America and Canada, also exposed the instruments to rigorous scrutiny (DoE, 2003).

Conclusion

The rationale behind the choice of the conceptual framework for this study was discussed in this chapter. The theoretical framework was discussed to situate the study paradigmatically and within the scientific approach. The research design, based on the research questions, was also presented. In Chapter 4, the descriptive and exploratory phases identified for this study are investigated and further analysed, with the results reported in Chapter 5.

CHAPTER 4
DATA ANALYSIS METHODS

In this study, a number of data analysis methods were used in order to adequately address the research questions. The data analysis methods used in this study and their rationale are discussed first (4.1). This is followed by a discussion on the preparation of the data for secondary analysis (4.2). Next, the descriptive analysis methods are presented and discussed (4.3). A brief discussion on the exploratory analysis methods (reliability, factor and correlation analyses) follows not only to explore the contextual factors but also to select the variables for the analysis of equity factors that have an effect on numeracy achievement (4.4). Lastly, the logistic regression analysis is done to determine the effect of the equity factors on learner achievement and also to predict performance in numeracy (4.5)

4.1 INTRODUCTION

The purpose of the secondary data analysis for this study is to describe, explore and report on: 1) the overall performance of learners in numeracy, 2) the equity factors that have an effect on that performance, and 3) equity factors that predict the performance in numeracy of rural and urban primary school learners.

This study seeks to describe and analyse the performance of the South African Grade 3 learners in numeracy using a secondary analysis. The purpose of the secondary data analysis for this study is to describe, explore and report on: 1) the overall performance of learners in numeracy, 2) the equity factors that have an effect on that performance, and 3) equity factors that predict the performance in numeracy of rural and urban primary school learners.

The performance of the learners in rural schools was juxtaposed against that of learners in urban schools. A further analysis was undertaken to explain factors that affect learner performance in terms of the equity indicator variables and how they can be used to predict the performance of these learners in numeracy. Regression analysis modelling was applied in order to predict one variable from the knowledge of one or more variables explored in this study (Howell, 1999).

Data analysis took place in three phases.

Phase 1

In the first phase, a post hoc analysis was done to closely look at the content of the questionnaires to identify the items with variables that contributed to the equity factors. The purpose was to lift the equity factors, from the SE survey contextual questionnaires data, which have an effect on learner performance for further investigation, so as to be able to answer the research questions for this study.

In the end, the variables were selected in line with the SE Framework (DoE, 2001) matrix of indicators that had them classified into the transformational goals of access, equity, redress and quality. There was no explanation provided for the categorisation and classification of the indicators in the matrix. A post hoc re-classification was done for some of the variables to fit in the model for this study.

Phase 2

An analysis of the equity factors used to describe the performance of the Grade 3 learners in the numeracy test was conducted next using descriptive statistics. Figure 3 points to the level of analysis at the heart of the subsystems that was the basis for the analysis in this study. The unit of analysis at this level was the learner. An analysis was done of how learners performed nationally using the selected variables for this study, namely gender, age, location and language. A further analysis was also undertaken using the selected variables.

Analysis of the data was undertaken by running the descriptive statistics; for example, mean scores for the numeracy task according to the selected variables for this study. Frequencies were generated using SAS and SPSS software programmes.

Graphs were used to check the normality of the data in terms of the overall distribution of the learners' scores regarding their characteristic; that is, whether they are unimodal, symmetrical or bell-shaped (van Lill & Grieve, 1990). Checking the normality of the data also provided a useful guide regarding the analysis of the scores in numeracy. Howell (2001); Easton and McColl (1997); van Lill & Grieve (1990) assert that a skewed distribution is characterised by an accumulation of scores at the left or right hand side and

thus it is asymmetrical. It therefore follows that a positively skewed distribution would mean that a considerable number of respondents obtained high scores possibly due to the easiness of the test. A negatively skewed distribution means that a considerable number of respondents obtained low scores possibly due to a number of factors, for example, difficulty of the test or language.

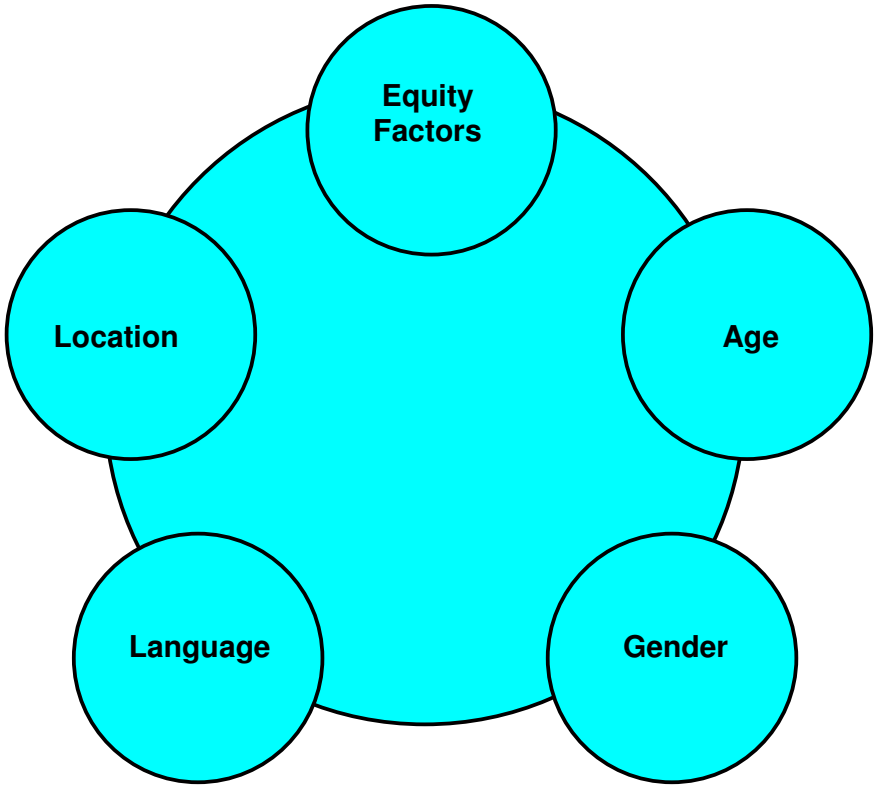


Figure 3. Unit of analysis (Learner)

On the other hand, scatterplots were used to establish the linearity of the relationships in the data through visual inspection only. Cross tabulations were done using contingency tables. These were done to examine and analyse the relationship between the independent variables (gender, age, location, province and language) and the dependent (numeracy scores) variable at the same time.

Phase 3

In the third phase, a statistical scale analysis of the contextual factors was done to analyse the equity factors that could be used in the model to predict numeracy performance. Cross tabulations were used to capture all the selected variables to indicate how learners in

different locations performed at various levels of the system, both at home and in the school.

Regression modelling was also pursued in this study to predict the performance of learners in numeracy. Howell (1999, p. 173) indicates that with regression, one wants to write an equation that explains how differences in one variable relate to differences in another or more and that allows us to predict a person's score on one variable from knowledge of that person's score on another variable. Regression analysis is a technique used to predict the value of the dependent variable using one or more independent variables (Howell, 1999). In this study, a prediction of performance in numeracy (the dependent variable) of learners, against gender, age, location (rural/urban) and home language (the independent variables) and the equity factors, was explored.

Using stepwise logistic models to explore their suitability for a regression analysis to predict the performance of the learners in numeracy, the Cramer's V statistic was applied to select the variables to be included in the model. The Cramer's V statistic gives a measure of the correlation between the selected independent variables and the dependent variable namely numeracy scores (above or below the median for the sample) as shown by RNUM (Cramer, 1999).

4.2 PREPARING THE DATA FOR ANALYSIS

It was important to first make sure that the data that was made available could be usable during the secondary analysis. It is important to ascertain that a number of steps need to be followed to be able to later use the data in a user-friendly way and format which is accessible.

Step I: Inspecting the dataset

A post hoc preliminary scanning and checking of the raw datasets for incomplete, inaccurate, inconsistent and irrelevant data (Vithal & Jansen, 1997, p. 27) was conducted. This meant going through the collected data in the form of responses and records captured. Both the assessment and contextual equity datasets were analysed. Only the principal questionnaire contained school classification data about location. For this, the data was

recoded to distinguish clearly between rural and urban schools to address the research questions.

Step II:Reporting on the quality of the dataset

To start with, there was a problem with not being able to link learners with individual educators where schools in the original SE sample had more than one Grade 3 educator. No unique identification numbers were allocated to educators in these schools. It was difficult to analyse the actual relationship between teacher inputs and learner outputs, without connecting the learners to their educators. In the end, and for the scope of this study, meaningful data analysis could only be done for those schools that had only one educator since they could be clearly linked to learners in this study.

In reflecting on the quality of the SE data, Gustafsson (2003) and Oosthuizen (2005) wrote in detail on the merging process of the datasets. Briefly, there were data files that corresponded to data captured from each of the questionnaires. That is, there were separate files for learner numeracy test scores, and then for each of the questionnaires with contextual data for learners, educators, principals and parents. To work with these as a coherent set, they had to be matched or merged according to a common identifier.

There were two learner files, one containing the achievement data and one containing learner contextual data, from the questionnaires completed by each learner that participated in the study. In total, there were learner achievement data captured for 52 482 learners and contextual data for 51 316 learners. Matching the data resulted in 50 154 observations for which there were learner data in both files. There was a significant match of 93.5% and that indicates a loss of 6.5% of observations in the learner base dataset.

More data could not merge successfully when other files were merged into the matched learner base dataset. The original parent file had responses from 47 353 parents of which 45 684 merged into the matched “learner base” dataset file. This meant that nine percent of learners in the learner base file did not have parent data to link with. Only for 45 097 learners out of 48 514, could the principal data be successfully merged with the matched learner dataset file. For all the matches, the number of learners for which there were learner performance and contextual data for the learner, parent, educator and principal data amounted to 40 478 observations. This is 80.7% of all learners in the learner dataset file.

The process of checking the representivity of data in the case of large quantities of missing data is important to reflect statistics accurately, argues Poswell (2006). It became evident that with large amounts of missing data occurring, it was going to be difficult to retain a substantial number of observations later when analysing a number of variables simultaneously through regression analysis. Regression analysis requires all observations to have data in each specified variable otherwise all observations without complete data would be dropped.

There were large amounts of missing data (see Table 3 below) to adequately impute missing values and although the use of dummy variables would have been the preferred option, the requirement would have meant a dummy variable for nearly every variable of interest, which was not a viable option given the scope of this study.

Table 3

Percentages of missing values for all variables in each dataset

	Minimum	**Maximum**	**> 10% Missing values % of all Variables**	**> 20% Missing values % of all Variables**
Learner Performance	0.0	1.4	0.0	0.0
Learner Context	5.0	28.0	80.6	17.2
Principal Context	10.4	99.3	100.0	29.7
Parent Context	13.5	67.1	100.0	81.6

Source: Systemic Evaluation 2001 Datasets (Poswell, 2006)

Note: Missing values exclude skip questions.

In the end, only the merged dataset containing complete cases for participants in one-educator schools only could be used for the secondary analysis in this study. The dataset and the sub-sample are described in detail in Section 4.5.

4.3 DESCRIPTIVE ANALYSIS METHODS

Two descriptive statistical methods that were used to provide a brief description of the participants in the SE survey of 2001 and in this study sub-sample, are discussed in this

section. Frequencies were done to describe how the participants were spread in the sample in terms of demographic information, background, personal characteristics and overall learner performance in numeracy. Cross tabulations were also done to describe how the population was spread according to location and also the availability of resources in different types of location.

Frequencies

A descriptive statistical analysis was run after cleaning the data to provide a description of the participants from the original SE study sample. The purpose was to profile the population so as to provide a thorough understanding of the context for this study leading to the exploratory and regression analyses. The characteristics of the sub-sample population were also analysed through frequencies. Personal characteristics (age, gender and language), qualifications or education levels, SES and location of schools were analysed using data from the contextual questionnaires to respond to the research questions to profile the participant; for example, learners, educators, parents and school principals.

Cross tabulations

A descriptive analysis was done for the sub-sample using cross tabulations in order to provide information on the overall performance of learners in numeracy. The analysis was also done to elicit more information as to how learners performed comparatively in numeracy by location (rural versus urban) against the availability of resources as identified by the selected equity indicator variables; for example, books in the home, computers for teaching and numeracy reference materials.

4.4 EXPLORATORY ANALYSIS METHODS

In this section, the exploratory analysis methods using SPSS programme, namely reliability, factor, correlation and multiple regression analyses that were applied in this study, are discussed in more detail. Firstly, a description will be given to provide information on the reliability of the scales used in the contextual questionnaires administered in the SE survey. This will then be followed by the description and the discussion of the factor analysis. Correlation analysis will also be discussed in this section and finally, a description is given of multiple regression modelling applied in this study. The results of the exploratory analyses are presented in the findings chapter (Chapter 6).

Reliability analysis of scales

In checking the reliability of a scale, Pallant (2001, p. 45) indicates that the internal consistency of a scale refers to the degree to which the items that make up a scale "hang together". The advantage of this method is that it allows one to check whether the items all measure the same underlying construct. Cronbach's alpha coefficient is a statistic, used as an indicator of internal consistency of a scale, and should be above 0.7.

In the case where a scale has less than 10 items, it may be more appropriate to report the mean inter-item correlation for the items. Briggs and Cheek (1986) recommend an optional range for the inter-item correlation of 0.2 to 0.4. The procedure for checking the reliability of the scales of the data in the SE contextual questionnaires was applied using SPSS. The reliability analysis of the scales was only done on selected equity indicator variables identified for this study in the conceptual framework to explore the internal consistency of some of the scales in the principal as well as the parent questionnaires.

Correlation analysis

Correlation analysis is used to describe the strength and direction of the linear relationship between two variables. Pearson product-moment correlation coefficient is designed for interval level (continuous) variables. It can also be used when there is one continuous variable (e.g. numeracy scores of Grade 3 learners) and one dichotomous variable (e.g. gender: male/female) (Pallant, 2001, p. 115).

Pearson correlation coefficients (r) can only take on values from –1 to +1. Pallant (2001) states that the sign on the front indicates whether there is a positive correlation (as one variable increases, the other decreases). The size of the absolute value (ignoring the sign) provides an indication of the strength of the relationship. A perfect correlation of 1 or –1 indicates that the value of one variable can be determined exactly by knowing the value on the other variable. A scatterplot of this relationship would show a straight line. On the other hand, a correlation of 0 indicates no relationship between the two variables. Knowing the value on one of the variables provides no assistance in predicting the value on the second variable. A scatterplot would show a circle of points, with no pattern evident (Pallant, 2001, p. 115).

Factor analysis

Factor analysis is a collection of methods used to examine how underlying constructs influence the responses on a number of measured variables (DeCoster, 1998). Pallant (2001, p. 151), on factor analysis posits that it takes a large set of variables and looks for a way that the data may be summarised using a smaller set of factors or components. It can also be used to reduce a large number of related variables to a more manageable number, prior to using them in other analyses such as simple multiple regression.

De Coster (1998) further states that factor analysis is performed to examine the pattern of correlations (or covariances) between the observed measures. Measures that are highly correlated (either positively or negatively) are likely influenced by the same factors, while those that are relatively uncorrelated are likely influenced by different factors. Again De Coster (1998) goes on to argue that the common factors, sometimes called latent variables, are hypothetical variables that explain why a number of variables are correlated with each other. As is the case with the large number of contextual equity indicator variables in the SE study, it became important to apply a factor analysis to summarise the data before applying the logistic regression to predict performance in numeracy in this study.

4.5 LOGISTIC REGRESSION ANALYSIS

A regression analysis procedure was followed in this study using the SAS programme. It is posited that regression analysis is used when one wants to predict a continuous dependent variable from a number of independent variables (Tabachnick & Fidell, 1989). Logistic regression is specifically a variation of ordinary regression which is used when the dependent (response) variable is a dichotomous variable (i.e. it takes only two values, which usually represent the occurrence or non-occurrence of some outcome event, usually coded as 0 or 1) and the independent (input) variables are continuous, categorical, or both (Tabachnick & Fidell, 1989).

Unlike ordinary linear regression, logistic regression does not assume that the relationship between the independent variables and the dependent variable is a linear one. It also does not assume that the dependent variable or the error terms are distributed normally. Logistic regression estimates the probability of a certain event occurring (Tabachnick & Fidell, 1989).

Cramer's V Statistic

Following a careful study of literature on regression analyses, a conclusion was made to pursue applying the Cramer's V statistic for initial selection of equity variables for logistic regression modelling to be done later in this study. Cramer's V is a statistic measuring the strength of association or dependency between two (nominal) categorical variables in a contingency table (Agresti, 2002). It is a correlation coefficient that indicates the relationship between two categorical variables. It is similar to Pearson's correlation coefficient for two quantitative variables. Like Pearson's coefficient, Cramer's V ranges from -1 to 1, with 0 indicating no relationship and -1 or 1 indicating a perfect relationship. Also like Pearson's coefficient, the square of Cramer's V indicates the proportion of the total possible association (i.e., the maximum possible value of the chi-square statistic) that is present in the data (Cramer, 1999).

Selecting the variables for initial logistic regression modelling

With regard to this secondary analysis of the SE dataset, a process of initial variable selection and manipulation is described. The use of a new reduced set of variables in the regression modelling, of how equity indicator inputs have an effect on the output (numeracy scores), at the Grade 3 level of the South African schooling system, is also described.

The research questions being addressed by this analysis are *What 1) home background and 2) pedagogical equity indicators predict learner performance in numeracy?* In selecting the variables to answer these research questions for the regression analysis to explore equity factors that could predict performance in numeracy, the following steps were undertaken.

The initial process of extracting the selected equity variables from the original SE dataset files involved merging the file records in SAS. To merge the files successfully the data were cleaned first to check for and deal with invalid codes where applicable. To start the process, two-way tables were used to see which of the selected variables were associated with numeracy or could predict numeracy achievement. After identifying the relevant variables associated with numeracy, a decision was taken to focus on those variables that had a strong association. These were included in the model to be explored for the

regression analysis. Thereafter, an analysis was run to create the model. The results of the analysis and the model are discussed in Chapter 6.

The original Excel file contained a summary of the number of schools where there were data from more than one educator. Using the School ID and the Province and District fields to merge different files, a number of learners with data for all five instruments, that is learner achievement data (LAD), learner questionnaire (LEQ), parent questionnaire (PARQ), principal questionnaire (PRQ), and the educator questionnaire (EDQ), emerged which could be utilized. The schools were then sorted by the number of educators and those with more than one educator (i.e. those that had 2-5 educators) were highlighted.

There were 630 schools with more than one educator. The total number of educators in these schools was 1 686. These schools combined had 22 512 (57%) learners that participated in the SE study. In contrast, there were 478 schools with only one educator, thus 478 educators. Approximately 40% of the learners (15 805) could be found in these schools. However, there were 21 schools with no educator data containing 640 (1.6%) learners.

Initially there were 40 variables from the EDQ questionnaire that were considered for the model with a variety of response types. Most were of the Yes/No type of responses. It was difficult to aggregate some of these in a meaningful and consistent way, for example, on the variables of home language, age category, qualification and training. It was also difficult to aggregate variables meaningfully where the 4-point Likert scale was used; for example, for limits to teaching and learning (EDQ.38). It was at this stage that a decision was taken to exclude schools with multiple educators and to go ahead with the analysis with schools that had only one educator since it would be possible to link the educators to the learners.

Sifting the data for regression modelling

Data files were created in SAS combining all five out of six data files from the original Excel dataset. The sixth file contained data from district officials and was not considered within the scope of this study. Schools with more than one educator were excluded. Frequencies and percentages were run to check for invalid codes and low frequencies (Appendix 1).

A total number of 15 680 learners had data for all five instruments, RNUM scores and one educator per school. There were 478 schools, some with only six learners and 50 districts, some with only 14 learners, which would make them difficult to model. This group of schools with very few learners would automatically be discarded during regression modelling.

Data cleaning and recoding

The district and school variables were excluded. The ages of the learners were recoded into the three age groups that is 1) 7-9, 2) 10, and 3) >10. Invalid codes for example, learner language code 0, were deleted. Educator's highest qualifications were combined (Honours and Master's groups in EDQ3).

Selection of consolidated variables for final regression modelling

The Cramer's V statistic was applied to select the individual variables with the highest correlations to RNUM or PNUM to include in the model. The Cramer's V statistic gives a measure of the "correlation" between the selected independent variables and numeracy (RNUM above or below the median for the sample i.e. 10).

The logistic models were then run to provide the consolidated variables (see Cramer.xls file, Appendix 2). Three worksheets were generated. The first worksheet (Appendix 2a) is about the Cramer's V statistic and it lists the statistics for all 103 variables. In the second worksheet (Appendix 2b), the variables are sorted from highest to lowest. Correlations were generated between the variables and the numeracy scores. Variables with the Cramer's V statistic highlighted in green indicate V>= 0.1. The MD and MD% are the counts and percentages for missing data. The highlighted rows are "combination" variables that were created during the exploration process but were not used in the models. The third worksheet (Appendix 2c) is about the initial model results showing variables with Cramer's V>= 0.1, which became possible candidates to include in the model. The variables with strong correlations with numeracy are in row 4, bold and black. On the other hand, the co-linear independent variables with V> 0.5 are highlighted in bold and red. These variables are correlated with each other and only one of a related pair could be included in the model.

Stepwise logistic models

The models were run to explore which of the independent variables best predict numeracy. Baker (2000, p. 82) asserts that many analysts do not support the stepwise selection approach, largely due to the fact that the method has been used irresponsibly, or whilst ignoring the social and economic dynamics of the system at hand, in our case the schooling system. The use of the SE framework and matrix of a set of 26 agreed upon indicators in the selection of equity variables helped to prevent this problem, especially as the South African education system is still largely riddled with inequities and social ills of the past.

Initial model (Appendix 3a)

All the 23 selected independent variables with Cramer's V >= 0.1 are given. There were only 288 learners with complete data (i.e. no missing data). A quasi-complete separation of data points was done. It was found that the maximum likelihood estimate may not exist and therefore the validity of model fit was questionable.

Second model (Appendix 3b)

Only 13 variables remained in the model with V >=0.1. There was low co-linearity with each other that is V<0.5. Only 1 108 learners with complete data could go into the model. Again in this model there was a quasi-complete separation of data points. The maximum likelihood estimate may not exist as well in this model and the validity of model fit was questionable.

Third model (Appendix 3c)

In this model, only nine variables with V>=0.1 survived. In this case, the problem of low co-linearity also surfaced. There was a significantly low missing data percentage. There was a marked increase of learners with complete data totalling 8 833. The model was probably very limited by the fact that it included few of the original 103 independent variables due to many co-linear or related variables and much missing data.

Fourth model (Appendix 3d)

In this model, the nine variables with V>=0.1 were included in the model with the 10^{th} additional variable to disaggregate the analysis by the location of the school. Once again a problem of low co-linearity was experienced. There was a significantly low missing data percentage. The number of learners with complete data was slightly lower than for model 3

at 8 820. The model was probably also limited by the fact that it included few of the original 103 independent variables. However, the model was accepted for application to predict the performance in numeracy of learners from various backgrounds explored in this study. A brief discussion of the logistic procedure applied is presented and discussed below.

Model information

The data set used in this procedure was for the numeracy scores (RNUM) response variable obtained from 8 820 observations of learners from schools with only one educator. The observations lost due to the deletion of incomplete cases amounted to 6 985. There were two (2) levels of response for the response variable (RNUM). The type of regression model that was fit for this data was the binary logit. The "Optimisation Technique" refers to the iterative method of estimating the regression parameters in terms of the regression coefficient and standard errors (UCLA: Academic Technology Services, 2007). The technique applied in default in this procedure was the Fisher's scoring method. The "Ordered Value" indicates how the levels of the dependent variable (numeracy scores) were ordered/modelled (UCLA, 2007). Total Frequency distribution of the response variable was 5 000 observations with an 'above median' score and 3 820 with a 'below median' score. The probability modelled was 'above median'.

Model Fit Statistics

The Model Convergence Statistics describe whether the maximum-likelihood algorithm has converged or not, and what kind of convergence criterion was used to asses convergence. The default criterion is the relative gradient convergence criterion (GCONV), UCLA (2007). The various "Criterion" measurements used to assess the model fit, i.e. The Akaike Information Criterion (AIC) and Schwarz Criterion (SC) are deviants of negative two times the Log-Likelihood (-2 Log L), UCLA (2007). The "Intercept Only" column refers to the respective criterion statistics with no predictors in the model, i.e., just the response variable. Conversely, the "Intercept and Covariates" column corresponds to the respective criterion statistics for the fitted model. A comparison of the values in this column can be made with the criteria corresponding "Intercept Only" value to assess model fit/significance (UCLA, 2007).

The “Test” shows the three asymptotically equivalent Chi-Square tests, namely Likelihood Ratio, Score and Wald. They test against the null hypothesis that at least one of the predictors' regression coefficient is not equal to zero in the model. Chi-Square, DF and Pr > ChiSq are the Chi-Square test statistic, Degrees of Freedom (DF) and associated p-value (PR>ChiSq) corresponding to the specific test that all of the predictors are simultaneously equal to zero (UCLA, 2007).

Analysis of Maximum Likelihood Estimates

In the “Parameter” column are the predictor variables in the model and the intercept. The column “DF” gives the degrees of freedom corresponding to the Parameter. The “Estimate” values are the binary logit regression estimates for the Parameters in the model. The logistic regression model models the log odds of a positive response (probability modelled is ‘above median’ =1) as a linear combination the predictor variables. The “Standard Errors” of the individual regression coefficients are used in both the 95% Wald Confidence Limits and the Chi-Square test statistic (UCLA, 2007).

The “Effect” refers to the predictor variables that are interpreted in terms of odds ratios. Therefore, the “Point Estimate” values interpret the odds ratio such that for a one unit change in the predictor variable, the odds ratio for a positive outcome is expected to change by the respective coefficient, given the other variables in the model are held constant. The “95% Wald Confidence Limits” show that for a given predictor variable with a level of 95% confidence, it can be said that there is 95% confidence that upon repeated trials, 95% of the confidence intervals would include the "true" population logistic regression coefficient (UCLA, 2007).

Association of Predicted Probabilities and Observed Responses

A pair of observations with different observed responses is said to be concordant if the observation with the lower ordered response value (‘above median’ = 0) has a lower predicted mean score than the observation with the higher ordered response value (‘above median’ = 1). On the contrary, if the observation with the lower ordered response value has a higher predicted mean score than the observation with the higher ordered response value, then the pair is discordant. If a pair of observations with different responses is neither concordant nor discordant, it is a tie (UCLA, 2007).

The total number of distinct pairs with one case having a positive response ('above median' = 1) and the other having a negative response ('above median' = 0) is what is referred to as "Pairs" in the output. Somer's D is used to determine the strength and direction of relation between pairs of variables (UCLA, 2007). In this example, it equals the difference between the percent concordant and the percent discordant divided by 100: (69.2 – 30.4)/100 = 0.388.

The values for the Goodman-Kruskal Gamma method range from -1.0 (no association) to 1.0 (perfect association). Somer's D is used to determine the strength and direction of relation between two pairs of variables. Kendall's Tau-a is a modification of Somer's D that takes into the account the difference between the number of possible paired observations and the number of paired observations with a different response. Finally, the "c" statistic values range from 0.5 to 1, where 0.5 corresponds to the model randomly predicting the response, and a 1 corresponds to the model perfectly discriminating the response (UCLA, 2007).

Final model applications for predicting numeracy achievement of urban and rural school learners

The fourth and final model was applied exploring the different variable combinations within the context of the learners in both rural and urban primary schools in the sample. Various model applications were tested to lift any meaningful comparisons that could emerge to predict the performance of these learners in numeracy. This analysis sought to answer the research question, *What i) home background and ii) pedagogical equity indicators predict learner performance in numeracy?*

The core of this analysis was to compare how learners from different backgrounds (rural and urban) would fare using the selected equity indicator variables in the final regression model. The various model applications' results are presented in the next chapter.

4.6 CONCLUSION

Three phases of data analysis to conduct a secondary analysis of the SE study were presented in this chapter. The phases involved, amongst others, a post hoc analysis that

resulted in the re-classification of the variables to match the selected indicators for the transformational goal of equity. Further analysis was conducted using descriptive statistics to establish the authenticity of the data. To this end, the normality of the data was checked to establish the overall distribution of the numeracy scores. The descriptive analysis phase sought to describe the participants in SE study sample as well the sub-sample according to their background and personal characteristics.

The last phase of data analysis involved the exploratory statistical scale analysis to select equity factors which had strong correlation (Cramer's V) values with numeracy. Reliability, correlation and factor analyses were undertaken to validate and to strengthen the rationale for the selection of variables that were considered in the final logistic regression modelling. Various stepwise logistic models were explored taking into account statistical considerations regarding the population in the model sub-sample being representative. Statistical measures like Model information, Fit statistics and the analysis of Maximum Likelihood Estimate were used to validate and interpret the final model and its applications. Findings on the final model are presented and discussed in detail in the next chapter.

CHAPTER 5
EQUITY FACTORS AND NUMERACY ACHIEVEMENT OF LEARNERS IN URBAN AND RURAL PRIMARY SCHOOLS

This study is a secondary analysis of the 2001 Systemic Evaluation survey numeracy achievement of Grade 3 learners in rural and urban primary schools. Findings on the analysis of performance of learners in numeracy and the effect of equity factors on this performance are reported on. A brief description of the participants and their overall performance in the SE survey numeracy assessment is provided. The results of the descriptive analysis lay the foundation for the exploratory as well as the regression analyses undertaken in this study. The effect of equity factors on the performance of participants in the sub-sample is reported on in terms of how it plays itself out on future numeracy performance of learners in rural and urban schools.

In this chapter, the results of the descriptive, exploratory and regression analyses are presented and discussed. First, the results of the descriptive statistics are presented (5.1). Then the overall performance in numeracy of Grade 3 learners in rural schools is compared to their counterparts in urban schools (5.2). This is followed by results of the exploratory analysis (5.3) of the equity factors that were found to have an effect on learner performance in numeracy based on the conceptual framework adapted for this study. The results of the regression analysis of how these factors predict the performance of rural and urban school learners in numeracy are also presented and discussed (5.4).

5.1 PROFILES OF PARTICIPANTS IN THE STUDY

The background data obtained from contextual questionnaires on the profiles of learners, educators, parents and principals of schools for both the SE survey and this study are discussed in this section. A comparative analysis of the results of the numeracy achievement of Grade 3 learners in rural and urban primary schools is also presented. The results of the descriptive analysis are presented using the selected independent variables, namely gender, age, location, language and numeracy achievement scores as the dependent variable. The results are also presented by location for the selected equity indicator variables that went into the final regression modelling.

The data about who the participants were in the SE survey of 2001 were already available, thus a secondary analysis thereof was conducted. Overall 52 482 learners, 2 508 educators, 1 309 school principals and 53 755 parents participated in the SE survey.

Profiles of learners that participated in the SE survey of 2001

This section answers the specific research question:

Who are the learners, educators, parents and schools that participated in the SE survey of 2001?

Demographics

The results of the descriptive analysis of the sub-sample for this study are presented. Of the original 52 482 learners that took part in the SE survey, 15 805 from schools with only one Grade 3 educator participated in this secondary analysis of which only 8 820 were considered for cases that went into the final model. According to the response profile, the numeracy scores of 5 000 learners were above the median while those of 3 820 were below the median (see Appendix 4). It must be noted that 6 985 observations were deleted due to missing values for the response or exploratory variables.

Personal characteristics

In the SE sample, the learners were aged from seven (7) to ten (10) years and above. The mean age was 9.4 years. Approximately 62.8% of the learners were of the appropriate age for the grade in terms of the Admission Policy of the Department of Education (DoE, 1998) namely 7 years to 9 years (DoE, 1998). More than a third (37.2%) of the learners were found to be over the grade age norm mainly due to retention (DoE, 2003).

Similarly, in this study sub-sample (n = 8 820), about 61% of the learners were in the appropriate grade age category of 7 to 9 years (see Appendix 4) and the mean age (9.5 years) was very similar to that of the SE sample. Approximately 21% of the learners were found to be 10 years old. It is highly probable for the learners in the latter age category to have repeated a grade, at least once in the phase. This is the maximum number of years that can be repeated by a learner in one phase according to the policy regarding progression (DoE, 1998). Almost every one out of five learners (about 18%) was not of the appropriate grade norm age, which means that they were above 10 years of age.

There were 882 urban learners (18%) and 2 870 semi-rural (58%) in the 7 to 9 years age category (n = 4 966) overall (see Appendix 4). In this age category the urban learners comprise 78% (n = 1 132) of learners urban schools overall. It is important to note that the vast majority of learners in rural schools combined, 78% (n = 6 549), are in the older age category of learners, that is 10 years and above.

On gender, the relative percentages for girls and boys who participated in the SE study (n = 52 482) were similar at 49.9% and 50.1% respectively. No statistically significant differences were observed in percentages between participating boys and girls (n = 8 820) within SE and this study overall. Each of the samples had the same pattern with negligible differences between boys and girls (see Appendix 4).

Home background characteristics

The language spoken by most learners in the SE sample (n = 52 482) was IsiZulu (24%) followed by IsiXhosa (20%), whilst the least spoken was IsiNdebele (1.7%), (DoE, 2003). In the sub-sample (n = 8 820) the language spoken by most learners was IsiZulu (30.4%) followed by Sepedi (15.4%) whilst the least spoken was IsiNdebele (1.03%), see Appendix 4. The difference in the sub-sample (n = 8 820) indicates a significant loss of IsiXhosa speaking learners in the analysis.

Profiles of educators that participated in the SE survey of 2001

Grade 3 educators were asked to complete a questionnaire on the conditions within which teaching and learning were taking place in their schools and classrooms. The numbers of educators who responded vary across the schools in the original SE sample. Some of the schools had only one educator, while others had more than one educator but allocated the same ID number in the same school.

Personal characteristics

The majority (98%) of educators in SE survey (n = 2 458) were females. This could be expected in this early phase of schooling as in late schooling there tends to be more male teachers. The SACMEQ II study found that 58% of Grade 6 teachers were female, (SACMEQ II, 2004). The ages of the educators in the SE sample ranged between less than 20 years and 51 years and over. Quite a large percentage of educators (43.2%) were between 31 and 40 years while only about 10% were aged between 21 and 30 years and

9.5% were 51 years and above (Table 4). The low percentage in the younger teachers' age category of 21 – 30 years could be a sign of young people in the country not opting for teaching as a profession. The mean age of SACMEQ II Grade 6 teachers was 39 years (SACMEQ II, 2004).

Table 4

Distribution of educators' ages (2001 SE sample dataset)

	N	% of educators
Under 20 yrs	3	.1
21 - 30 yrs	257	10.4
31 - 40 yrs	1 067	43.2
41 - 50 yrs	910	36.8
51 yrs & over	234	9.5
Total	2 471	100.0
Total	**2 508**	

Missing data N = 37

Qualifications and experience

In terms of academic education, 10% of teachers in the SE survey (n = 250) indicated that they had obtained or completed a degree (Table 5a).

Table 5a

Distribution of educators' highest academic qualifications (2001 SE sample dataset)

	N	% of educators
None	166	7.0
Diploma	1 730	73.0
Degree (BA, BSc, etc.)	237	10.0
Honours	40	1.7
Masters (MA, MSc, etc.)	6	.3
Other	192	8.1
Total	2 371	100.0
Total	**2 508**	

Missing data N = 137

Approximately 97% indicated that they had obtained at least a Matric or Grade 12 certificate meaning that three percent did not. In the sub-sample, 73% of the teachers

indicated that they had a total equivalent of three years of teacher training, which is either a junior or senior primary teaching diploma (Tables 5a & 5b).

Table 5b

Distribution of educators' professional training qualifications (2001 SE sample dataset)

	N	% of educators
None	35	1.5
L/J PTC	362	15.1
J PTD	1 016	42.4
H PTC	275	11.5
S PTD	497	20.8
Other	209	8.7
Total	2 394	100.0
Total	**2 508**	

Missing data N = 114

Educators were also asked about the years of teaching experience they had altogether in their careers but also in Grade 3 specifically. In teaching Grade 3 specifically, educators reported years of experience ranging from one to 42 years.

Profiles of parents that participated in the SE survey of 2001

The parent of a learner in Grade 3 randomly sampled to participate in the SE survey completed a parent questionnaire. In the absence of the parent(s), a guardian or any member of the family or relatives that the learner stayed with during a normal school week, completed the questionnaire.

Parents education levels

Approximately 26% of the learners' parents in SE (n = 53 755) reported that they had training beyond Grade 12 (Table 6). There was a significant number (14 681) of parents with missing data on this item. This could have been the case because of the low levels of literacy of the majority of parents in South Africa or simply because the respondents were not willing to disclose their education status.

Table 6

Distribution of parents' education levels (2001 SE sample dataset)

	N	% of parents
Training after Grade 12	10 068	25.8
Grade 12 but no further training	7 200	18.4
Grade 7 but no Grade 12	8 469	21.7
Lower than Grade 7	13 337	34.1
Total	39 074	100.0
Total	**53 755**	

Missing data N = 14 681

Socio-economic status

In the SE study a list of items was included in the learner and parent questionnaires that were a measure of socio-economic status. The list included items such as TV, radio, video recorder, computer, telephone, electricity, running tap water, motor vehicle, land and livestock. Similarly, for SACMEQ II the socio-economic backgrounds of pupils were assessed through the application of a "Home Possessions Index" (HPI) in both the SACMEQ I and II projects. This index was constructed by asking learners if they had some specific items at home like daily newspaper, weekly or monthly magazine, radio, television set, video cassette recorder, telephone, electricity, piped water, and a table to write on (SACMEQ II, 2004).

In the SE study, as many as 29% of the learners (n = 33 812), according to their parents, lived in homes without electricity and 40% where there was no running water (n = 32 546). A further disaggregation by location showed that about 12% of learners in urban areas (n = 15 006) lived in homes that did not have running water while 57% of learners in rural areas (n = 17 540) were without it (Table 7). Eleven percent in urban areas (n = 15 054) indicated that they did not have electricity available while 41% of their counterparts in rural areas (n = 18 758) were without electricity (Table 8). These statistics confirm the low socio-economic status of the majority of people living in the rural areas of our country and the hardships they have to endure almost on a daily basis.

Table 7

Number of urban and rural learners in homes with running water (2001 SE sample dataset)

		Location of school		Total
		urban	rural	urban
Home facilities availability –	Yes	13 198	7 455	20 653
Running tap water	No	1 808	10 085	11 893
Total		15 006	17 540	32 546

Missing data N = 21209

Table 8

Number of urban and rural learners in homes with electricity (2001 SE sample dataset)

		Location of school		Total
		urban	rural	urban
Home facilities availability -	Yes	13 468	11 156	24 624
Electricity	No	1 586	7 602	9 188
Total		15 054	18 758	33 812

Missing data N = 19 943

Books in the home may also be used in contexts like South Africa as an indicator of socio-economic status as well as parental education levels (Howie, 2002, p. 128). The SE results show that according to parents, a majority 62% of the learners (n = 53 755) had few or no books (fewer than 10) at home. More than 66% of the learners in this study's sub-sample (n = 8 820) had fewer than 10 books at home (see Appendix 4).

Other SES variables (items) that were listed in the SE study that were used as a measure of socio-economic status, for example TV, radio, video recorder, computer, telephone, motor vehicle, land and livestock at home did not make it into the final model that included the equity home background factors, thus they are not reported here.

Profiles of schools that participated in the SE survey of 2001

It should be noted that the classifications of schools as rural, remote rural, urban and semi urban should be treated with care. Schools principals were asked to provide a subjective response based on what they considered to be the extent of 'rural ness' or 'urbanity' of the setting in which their school was located. An example is that the size and characteristics of

a school in a semi urban area might have a lot in common with the size and characteristics of a school in an urban area.

Geographic location of schools

This study focuses on the comparative analysis of the performance in numeracy of rural and urban school learners. As can be seen in Table 9, in the majority of the SE schools (61%) principals reported that they were located in rural areas while 39% are located in urban areas (n = 1309). The split of the schools in the sample reflects the broader population of the country, as approximately half of the country's people live in the rural areas. It is reported in the most recent data on children's urban/rural status, that more than half of South Africa's children (54%) live in rural areas (Stats SA, 2005)

Table 9

Geographic location of schools (2001 SE sample dataset)

	N	% of schools
Urban	315	24.4
Semi-urban	188	14.6
Rural	650	50.3
Remote-rural	138	10.7
Total	1 291	100.0
Missing System	18	
Total	**1 309**	

Missing data N = 18

In the sub-sample (n = 8820), the majority of learners (78%) were found in schools that were located in rural areas (see Appendix 4). A further disaggregation according to location shows that the majority, 62%, of learners were in semi rural schools, 16% remote rural, 13% urban and 9% semi urban.

Management and leadership of schools

Principals were also asked to complete a questionnaire on the conditions within which teaching and learning were taking place in their schools. Howie (2002, p. 150) argues that in order for schools to function well it is important that the leadership of the school is effective in managing the school.

Personal characteristics of principals

No data was available in the dataset to report on the gender of school principals that participated in the SE survey. About 81% of principals in the SE sample were relatively old, that is, above 41 years (Table 10).

Table 10

Distribution of principals' ages (2001 SE sample dataset)

	N	% of principals
Under 20 yrs	1	.1
21 - 30 yrs	5	.4
31 - 40 yrs	239	18.4
41 - 50 yrs	691	53.1
51 yrs & over	365	28.1
Total	1 301	100.0
Total	**1 309**	

Missing data N = 8

Qualifications and teaching experience

Principals were also asked about the years of teaching experience as principals at their current schools. The longest service recorded as a principal at the present school was 35 years. Approximately 42% of principals reported to having a teacher college training qualification while 25% had obtained a university degree (Table 11).

Table 11

Distribution of principals' highest academic qualifications (2001 SE sample dataset)

	N	% of principals
Primary	116	9.3
Secondary	48	3.8
Teacher college training	527	42.2
University degree	308	24.7
Post graduate	227	18.2
Other	23	1.8
Total	1 249	100.0
Total	**1 309**	

Missing data N = 60

5.2 SOUTH AFRICAN GRADE 3 LEARNERS' PERFORMANCE IN NUMERACY

The achievement as well as the background data obtained from contextual questionnaires, administered on Grade 3 learners and their educators, school principals and parents or guardians, are presented and discussed in this section.

Overall learners' performance by Learning Programme

The national average scores for SE show that in general, learners obtained the lowest scores for numeracy of all the three Learning Programmes that were assessed (see Figure 4). The national mean for numeracy was 30%. The distribution of numeracy scores was negatively skewed as most of the learners scored below 40% (DoE, 2003).

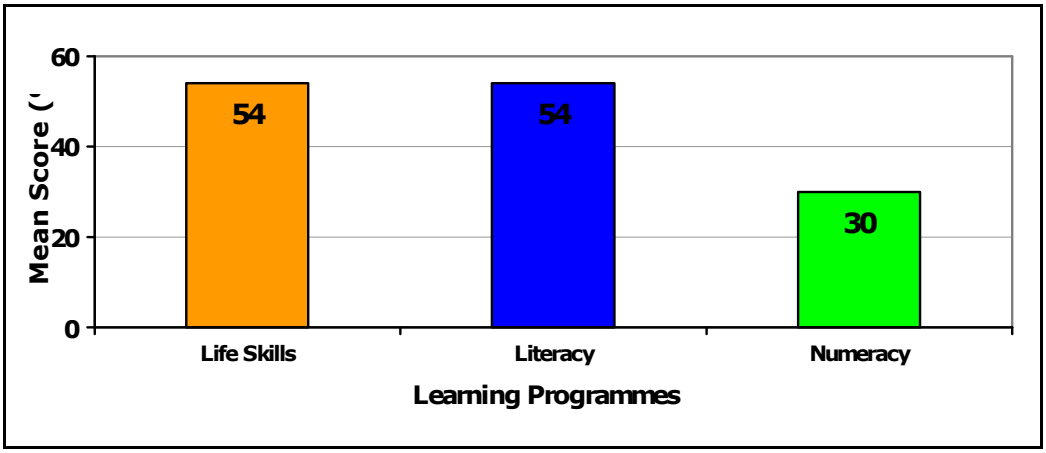


Figure 4. Overall learners' performance by Learning Programme

Source: SE Mainstream Report (DoE, 2003)

Overall learners' performance by gender

The average scores for girls were slightly higher than those of boys in all three learning programmes assessed during the SE survey (see Figure 5). The pattern of performance, however, is similar for both boys and girls in that the lowest scores were obtained for numeracy, that is 29% and 31% respectively (DoE, 2003).

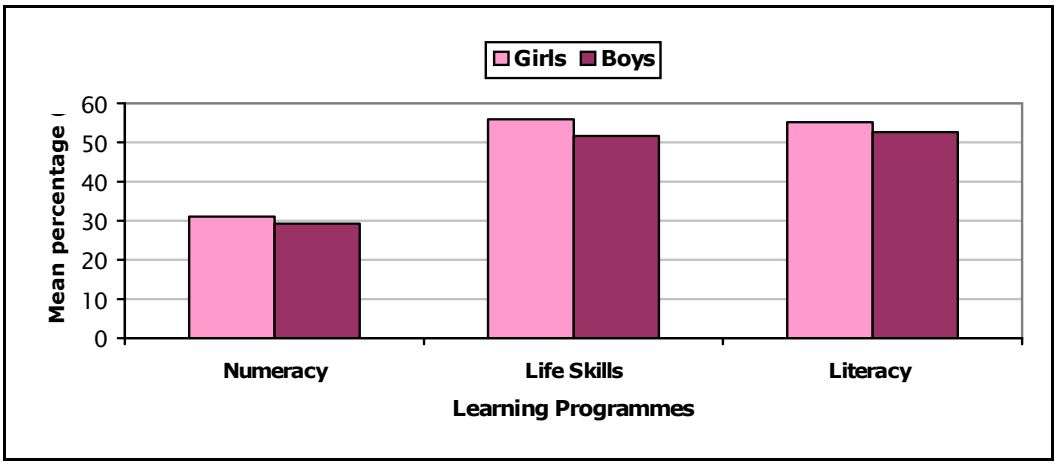


Figure 5. Overall learners' performance by gender on Learning Programmes assessed
Source: SE Mainstream Report (DoE, 2003)

Overall learners' performance by language of learning and teaching (LOLT) and home language (HL)

The SE survey data indicates that the majority of learners (75%) responded to the assessment tasks in their home language (DoE, 2003). However, there were large variations between provinces. For example, in Gauteng, learners who took the instrument in their home language were in the minority (approximately one-third). In other words, about 64% of the learners responded to the assessment tasks in a second or third language. This could be attributed to the movement of learners to English-medium schools in urban areas (DoE, 2003).

Analysis of the effect of home language on learner scores revealed that learners who took the test in their home language obtained significantly higher scores than their counterparts who had to respond to the assessment tasks in their second or third language (DoE, 2003).

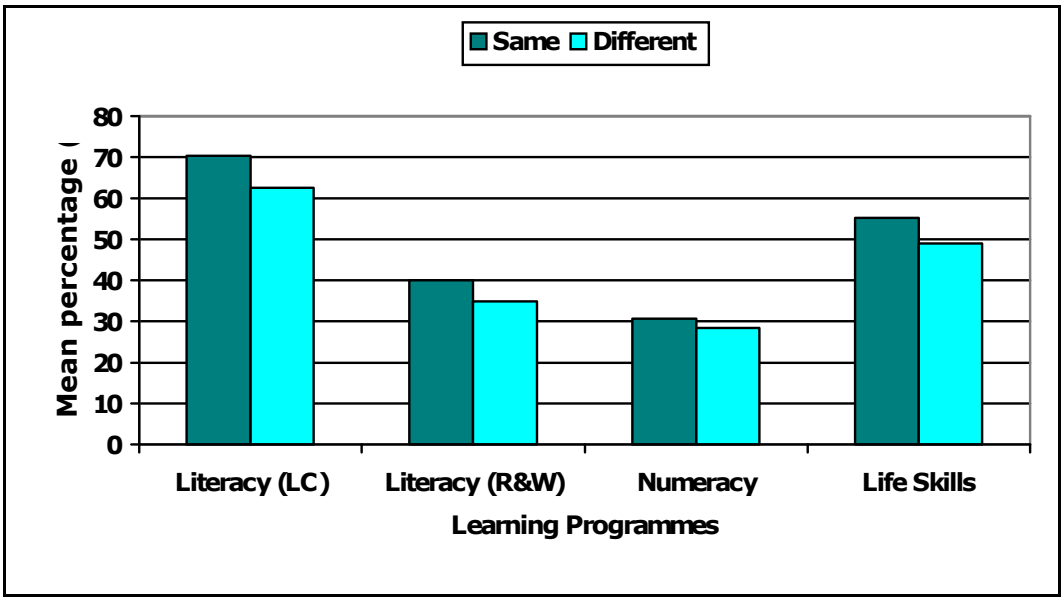


Figure 6. Overall learners' performance by language of learning and teaching (LOLT) and home language (HL) on Learning Programmes assessed

Source: SE Mainstream Report (DoE, 2003)

5.3 THE COMPARISON IN NUMERACY PERFORMANCE OF GRADE 3 LEARNERS IN RURAL AND URBAN SCHOOLS

The descriptive analysis results of the sub-sample data are presented in this section to address the research question on how learners in rural schools performed in numeracy as compared to the urban learners. It should be noted that the four categories for location, that is, urban, semi urban, rural and remote rural, were retained for this analysis. The results are interpreted within the context of the rural-urban divide of the two strands of classification for juxtaposition in this study. It is also important to note that the majority of learners in the sub-sample making it into the final model overall were from rural schools, that is 6 683 rural and 1 937 urban (see Appendix 4).

Cross tabulations were run to determine the number of learners who fitted both characteristics for location of school, namely urban and rural (see Appendix 4). To compare performance by location, there were 779 urban and 744 remote rural learners with the numeracy scores above median. The Row Pct - percentage of all the learners in the row (Appendix 4) shows that 779 urban learners make up 15.58% of the total 5 000 learners in

the above median category overall whereas the remote rural make up only 14.88%. In the Col Pct - percentage of all the learners in the columns for above median, the 779 urban learners make up 66.41% of the 1132 urban while the remote rural are 51.20% of the total 1 483. These results show that learners in remote rural schools are not performing as well as their counterparts in urban schools.

The availability of extra resources might be expected to be associated with school location and performance (Guiton & Oakes, 1995). The underlying assumption here is that such resources are more likely to be provided in urban areas – where it is known from other studies that parents, governments and business tend to be more supportive of the schools. On the same score, the results summarised in the sub-sample for the availability of the computer and OHP respectively for teaching numeracy did support this line of argument (see Appendix 4). The difference on these resources being 'always' available between the semi rural/remote rural areas (about 13%) and urban/semi urban areas (about 87%) was very huge. This further confirms the digital divide gap in the country in with regard to the equity of provisioning is concerned especially considering the backlogs of information technology infrastructure for schools and communities in rural areas.

5.4 MAIN EQUITY FACTORS THAT AFFECT NUMERACY PERFORMANCE OF LEARNERS IN RURAL AND URBAN AREAS

The exploratory data analysis methods for the results in this section were discussed in Chapter 4 (4.4). The focus of this section is to report the results of the exploratory analysis of the equity factors variables that were selected for this study. Originally 103 equity indicator variables were selected for this study. Not all the variables could make it into the regression modelling due to the minimal number of observable cases with complete data from the five datasets in SE. Therefore, the variable selection model is discussed in the sections that follow from the results of the analysis using the Cramer's V statistic.

Results of the reliability analysis

A post hoc analysis of the SE contextual questionnaires selected equity indicator variables was done to check the reliability of scales. The purpose was to check the internal consistency of the scales used.

In the first example (see Appendix 5), two items from the principal's questionnaire were analysed for internal consistency of the scale to measure the availability of school facilities or resources. The resources were linked to numeracy teaching, namely computer for teaching and OHP as part of the selected equity indicator variables. The Cronbach's alpha coefficient for this scale was .647 slightly below .7. The Corrected Item Total Correlation, which gives an indication of the degree to which each item correlates with the total score (Pallant, 2001), was .502 for each of the items and higher than the .3 threshold for low values sufficing to consider keeping the items intact. However, the Alpha If Item Deleted value was negative violating reliability model assumptions.

Overall, most of the scales from the selected equity indicator variables had less than 10 items and it therefore became necessary to report the mean inter-item correlation for the items as reflected in the results that follow. The other problem worth noting is that most of the scales had very low Alpha values (less than .7) thus making it very difficult to ascertain the reliability of the scales in the questionnaires.

Results of the correlation analysis

As reported in Chapter 4 during the data analysis that the Cramer's V statistic was used to select variables to be included in the model (see Appendix 2). Three worksheets were generated providing for all the variables selected through their strong association and correlation with numeracy until finally arriving at those that made it into the final model (see Appendix 2a, 2b and 2c). The variables that were selected from the original 103 to be included in the model were categorised according to list A, taking into account the strength of their association and correlation with numeracy as follows:

List A – Re-worked categories of independent and dependent variables

- Numeracy category
- Learner's home language
- School's section 21 status
- Annual school fees
- Computers available for teaching
- Impact of inadequate physical facilities
- Number of books in the home
- Availability of numeracy reference materials

- Calculators in the classroom
- OHP available for teaching

Results of the factor analysis (Combining several variables into one factor)

Where it was clear that several closely related original variables were linkable to one equity indicator, factor analysis was used to extract a single variable, or 'factor', that synthesised the values of the several variables. This statistical method is commonly used in education input-output analyses (Willms & Somers, 2001, p. 415; Hungi, 2005, p. 2).

The new learner SES, parent SES and school funding variables were derived using factor analysis. For example, the variable learner/parent_SES was derived from six original variables relating to the physical condition of the learner's home and the presence of the three household items, namely availability of a computer, video and number of books in the home that emerged as significant from the Cramer's V statistical correlation analysis (see Appendix 2c). Some further examples include the collapsing of the 13 funding variables in the Principal's data file into one factor. Such variables may not relate to numeracy individually but en masse they may.

Eight codes were assigned the age of learners in the original SE data. To consolidate these, age groups of the learners were determined so that they could be categorized. Three categories were determined based on variations on compliance or non-compliance to the admission policy of the DoE (DoE, 1998). The recoded age categories are 1 = 7-9 years (grade-appropriate), 2 = 10 years (still grade-appropriate having repeated only once in the phase as per policy) and 3 = 10^{+} years (not grade-appropriate)

The location of schools was recoded into two categories namely Urban and Rural from the original four in the Principal's data file (PR5), that is, value label codes1 (urban) and 2 (semi-urban) into Urban while 3 (rural) and 4 remote rural into Rural. There were a few invalid codes in the data, which were omitted. The results of this recoding of location did not come out as significant in the Cramer's V statistical analysis (v = -1.0), see Appendix 2a.

On the educator's highest academic qualification (EDQ3), there were few educators with Honours and Master's degrees and these were combined into one group. The results show

that these combined qualifications variables did not survive due to low correlation (v = 0.08), see Appendix 2a.

5.5 EQUITY INDICATORS THAT PREDICT LEARNERS' PERFORMANCE IN NUMERACY

The selection of variables for regression modelling was explained in Chapter 4. The findings and the process of developing a model to predict performance in numeracy using stepwise and logistic regression models will be discussed in this section. The systematic selection of variables for the final model is presented. Finally, the results of the logistic regression model are presented and discussed.

Stepwise logistic regression analysis was used to determine the significance of each explanatory variable and to eliminate any variable that did not have a significant effect on the model. To explore which of the independent equity variables best predict numeracy in this study, the stepwise regression analysis was applied and yielded the results presented in the next section. The analysis ended up with the 10 variables in the final model, namely 1) Learner's home language, 2) School's section 21 status, 3) Annual school fees, 4) Computers available for teaching, 5) Impact of inadequate physical facilities, 6) Number of books in home, 7) Numeracy reference material available, 8) Calculators in the classroom, 9) Location of school, and 10) OHP available for teaching.

Results of the stepwise logistic regression modelling

The four models which are a culmination of the stepwise logistic regression analysis are presented below. A systematic stepwise elimination of the equity variables was done to select those with complete cases to make it into the final model (Appendices 3a – 3d).

Stepwise logistic model 1 (see Appendix 3a)

From the original 103 variables (Appendix 2a) all the 23 independent variables with Cramer's V >= 0.1 were put into model 1 (Table 12). School location classification variables were recoded from the original four into two (1 = urban and 2 = rural) This resulted in only one percent of learners with complete data (n = 288), in one-educator schools remaining in the model. There was a quasi-complete separation of data points. It could therefore be established with certainty that the Maximum Likelihood Estimate

(MLE) might not exist. The MLE is used for fitting a mathematical model to the data and it offers a way of tuning the free parameters of the model to provide a good fit (Aldrich, 1997). Also critical for this result was that the validity of model fit was questionable. This model was not accepted due to a very low number of learners from the sub-sample but also due to the validity of the model fit.

Stepwise logistic model 2 (see Appendix 3b)

In the second model (Table 12) only 13 variables of the original 103 remained with V >=0.1. The variables had a low "co-linearity" with each other with the value of V<0.5. It means that there was no linear relationship between the independent variables. Therefore, they could not be used together as predictors in the regression model. There were only 7% of learners (n = 1 108) with complete data from the original one-educator schools dataset. There was a quasi-complete separation of data points, which therefore meant that the maximum likelihood estimate might not exist. Validity of model fit was also questionable, thus the model could not be accepted for this study.

Stepwise logistic model 3 (see Appendix 3c)

This model (Table 12) yielded only nine variables with V>=0.1. The variables also had low co-linearity. The positive thing about this result was the low missing data percentage. Again the strength here was that 56% of learners (n = 8833) had complete data. However, it should be noted that the model would probably be limited by the fact that it included so few of the original 103 independent variables due to many co-linear/related variables and missing data.

Stepwise logistic model 4 (see Appendix 3d)

The model was an extension of model 3 that had 9 variables with V>=0.1. The tenth variable (Availability of OHP for teaching) made it into the final model. Table 12 below shows the equity variables that made it from the initial model to this final model (model 4) during the stepwise selection. Another critical variable of location of the school was included with the original four classification area codes (1 = urban, 2 = semi-urban, 3 = semi-rural and 4 = remote-rural). The purpose was to answer the research question on comparisons of performance between learners in rural and urban schools from the original setting as was reported by the principals to explore any distinct differences that might occur. The results are discussed in detail in the next section.

Table 12

Equity variables in the stepwise logistic regression models

	Variables	Source	Model			
		Questionnaire	1	2	3	4
1.	Learner's language	Learner	X	X	X	X
2.	Parent's language	Parent	X			
3.	Educator's language	Educator	X			
4.	Language parent speaks to learner - Setswana	Parent (12H)	X			
5.	Language parent speaks to learner - Sepedi	Parent (12F)	X			
6.	Educator's highest maths level	Educator (5A)	X	X		
7.	School's section 21 status	Principal (57)	X	X	X	X
8.	Annual school fees	Principal (52)	X	X	X	X
9.	Computer – principal	Principal (25I)	X	X	X	X
10.	OHP available	Principal (25K)	X			X
11.	Language parent speaks to learner - Xitsonga	Parent (12K)	X			
12.	Home language books appropriate	Educator (23Bco)	X	X		
13.	Condition of numeracy materials	Educator (23Ago)	X	X		
14.	Physical facilities	Educator (38D)	X	X	X	X
15.	Number of books at home	Parent (17)	X	X	X	X
16.	Numeracy references available	Educator (23Emo)	X	X	X	X
17.	Numeracy materials available	Educator (23Amo)	X			
18.	Home language books condition	Educator (23Bgo)	X			
19.	Video available at home	Parent (7C)	X	X		
20.	Calculators available in class	Educator (19Q)	X	X	X	X
21.	Language parent speaks to learner - Sesotho	Parent (12G)	X			
22.	Location – urban/rural	Principal (5)	X	X	X	X
23.	Educator's qualification <gr12.	Educator (2A)	X			

The fourth model was adopted for this study within the context of the limitations cited earlier in Chapter 4. Table 12 above is a summary of the variables that went into the final model. The groups of equity variables combined for modelling are presented with their probabilities to either increase (shaded) or decrease (non-shaded) the odds of obtaining a score above the median indicated.

NB: *The results of the regression analysis below are interpreted by school location to compare and predict the performance of rural and urban primary school learners in numeracy (Table 13). It is interesting to note that for location, the urban and remote rural were not far apart in that both had indices increasing the odds probability to obtain a score above the median.*

Table 13

Final logistic regression model summary

	Intercept = 1.1268			**Average odds** = 3.085766232
Nr	**Variable**	**Level**	**Par**	**Index**
1	Learner's home language	English	1.0822	**2.951164977**
1	Learner's home language	IsiXhosa	0.6694	**1.95306513**
1	Learner's home language	Siswati	0.4143	**1.513311052**
1	Learner's home language	Tshivenda	0.3677	**1.444408651**
1	Learner's home language	IsiZulu	0.2872	**1.332690725**
1	Learner's home language	Sesotho	**0.085**	**1.088717067**
1	Learner's home language	Sepedi	-0.0725	**0.930065747**
1	Learner's home language	Xitsonga	-0.0973	**0.90728378**
1	Learner's home language	Setswana	-0.4551	**0.634384526**
1	Learner's home language	Afrikaans	-0.512	**0.599295788**
1	Learner's home language	Other	-0.5561	**0.573441129**
1	Learner's home language	Isindebele	-1.2128	**0.297363495**
2	School's section 21 status (PR57)	3 Full	0.4115	**1.509079708**
2	School's section 21 status (PR57)	1 No	**0.0993**	**1.104397569**
2	School's section 21 status (PR57)	2 Partly	-0.5108	**0.600015374**
3	Annual school fees (PR52)	7 501-1000	0.7963	**2.217321642**
3	Annual school fees (PR52)	5 301-400	0.4677	**1.596318435**
3	Annual school fees (PR52)	9 >R2000	0.19648	**1.217110978**
3	Annual school fees (PR52)	8 1001-2000	0.1362	**1.145911053**
3	Annual school fees (PR52)	3 101-200	**-0.00918**	**0.990862008**
3	Annual school fees (PR52)	1 <R50	-0.0832	**0.920167095**
3	Annual school fees (PR52)	2 50-100	-0.1133	**0.892882755**
3	Annual school fees (PR52)	6 401-500	-0.4822	**0.617423565**
3	Annual school fees (PR52)	4 201-300	-0.9088	**0.403007543**
4	Computers available for teaching (PR25I)	3 Always	0.404	**1.497803947**
4	Computers available for teaching (PR25I)	2 Sometimes	0.3635	**1.438354857**
4	Computers available for teaching (PR25I)	1 Not available	**-0.7675**	**0.464172049**
5	Impact of inadequate facilities (ED38D)	2 A little	**0.3149**	**1.370122292**
5	Impact of inadequate facilities (ED38D)	4 Great deal	-0.0219	**0.978338064**
5	Impact of inadequate facilities (ED38D)	1 Not at all	-0.1174	**0.88922943**
5	Impact of inadequate facilities (ED38D)	3 Quite a lot	-0.1756	**0.838953498**
6	Number of books in home (PAR17)	3 >20	0.1715	**1.187084143**
6	Number of books in home (PAR17)	2 10-20	-0.0273	**0.973069277**
6	Number of books in home (PAR17)	1 <10	**-0.1442**	**0.865714588**
7	Number of books in home (PAR17)	Yes	0.1352	**1.144765715**
7	Number of books in home (PAR17)	No	**-0.1352**	**0.873541186**
8	Calculators in classroom (ED19Q)	Yes	0.1376	**1.147516452**
8	Calculators in classroom (ED19Q)	No	**-0.1376**	**0.871447201**
9	Location of school (PR5)	3 Rural	**0.106**	**1.111821877**
9	Location of school (PR5)	2 Semi-urban	0.0508	**1.05211245**
9	Location of school (PR5)	4 Remote-rural	-0.0079	**0.992131123**
9	Location of school (PR5)	1 Urban	-0.1489	**0.861655276**
10	OHP available for teaching (PR25K)	3 Always	0.1646	**1.178921456**
10	OHP available for teaching (PR25K)	2 Sometimes	0.1618	**1.175625093**
10	OHP available for teaching (PR25K)	1 Not available	**-0.3264**	**0.721516523**

Estimated indices:

Above 1: Increase odds (probability of score above median

Below 1: Decrease odds (probability of score above median

HOME BACKGROUND EQUITY FACTORS THAT PREDICT LEARNERS' PERFORMANCE IN NUMERACY

After the stepwise selection of the variables to explore which had an effect on numeracy achievement, it became prudent to further to investigate those equity factors that could be used for logistic regression modelling. The purpose was to determine whether these factors could predict learner performance in numeracy. There were only two variables out of the 10 equity variables in the final logistic regression model that could be classified as home background equity factors, namely home language (HL) of the learner and the number of books in the learner's home. Five of the 11 official languages, that is Sepedi, Xitsonga, Setswana, Afrikaans and IsiNdebele had average odds indices of p value below 1 indicating a decrease in the probability of obtaining a score above the median (Table 14).

Model 4 Application I – English

The model application in the table below (Table 14 - English) indicates that being a learner whose home language is English and attending an urban school, doubles the chances of obtaining a score above the median by about 95%. For the same learner in this school from a home with more than 20 books, the odds of obtaining a score above the median increase by 19%. The chance or likelihood of obtaining a score above or below the median is measured by the impact on odds. The 'odds' refers to the likelihood of something happening or not to impact on an outcome or result. A formula to calculate the odds change factor is $p = \sigma / (\sigma + 1)$ where (p) is the percentage odds change factor, σ is impact on odds and 1 is a constant.

Example:

$$p = \sigma / (\sigma + 1)$$

$$= 18.97 / (18.97 + 1)$$

$$= 0.9499$$

$$= 94.99\%$$

The odds change factor calculated through the formula above indicates that it can be predicted that, overall, under the same conditions of the variables as listed in the model, the odds of obtaining a score greater than the median to obtaining a score not greater than median translates into a probability of 95%. This is indicated by the odds change factor (probability – greater than median). It can therefore be said that the probability of a child obtaining a score above the median in these conditions is 95%.

Table 14

Example 1: Model 4 Application 1 – English (Home background factors)

Variable	Level of significance (p-value)	Source	Selection appropriate category:	Impact on odds	Increase / decrease in odds	
Learner's home language	<0.0001	LQLANG	**ENGLISH**	**2.951**	195.12%	**Increase**
School's section 21 status	<0.0001	PR57	3 FULL	**1.509**	50.91%	**Increase**
Annual school fees	<0.0001	PR52	2 50-100	**0.893**	10.71%	**Decrease**
Computers available for teaching	<0.0001	PR25I	2 SOMETIMES	**1.438**	43.84%	**Increase**
Impact of inadequate physical facilities	<0.0001	ED38D	1 NOT AT ALL	**0.889**	11.08%	**Decrease**
Number of books in home	<0.0001	PAR17	**3 >20**	**1.187**	18.71%	**Increase**
Numeracy reference material available	<0.0001	ED23EMa	NO	**0.874**	12.65%	**Decrease**
Calculators in classroom	<0.0001	ED19Q	YES	**1.148**	14.75%	**Increase**
Location of school	0.0219	PR5	**1 URBAN**	**0.862**	13.83%	**Decrease**
OHP available for teaching	<0.0001	PR25K	3 ALWAYS	**1.179**	17.89%	**Increase**
Odds - greater than median			**18.97**			
Probability - greater than median			**94.99%**			

Model 4 Application II – IsiNdebele

Looking at a different language and with variations around location and the number of books in the home, a different picture is painted. The variables (Table 15) show that for a learner from an IsiNdebele speaking home in a remote-rural school decreases the odds of obtaining a score above the median by 70%. The results further show that for the same learner in this school (remote-rural) from a home with less than 10 books, it decreases the odds of obtaining a score above the median by 13%. The odds change factor for this model predicts overall that under the same conditions of the variables as listed in the model, the probability of obtaining a score above the median in future is 14%. This is in contrast to a learner speaking English in an urban school presented in the first model application (Table 14) whose probability of obtaining a score above the median (95%) is far above in this case (14%).

CHAPTER 6
CONCLUSIONS AND RECOMMENDATIONS

The aim of this study was to undertake a comparative secondary data analysis of numeracy achievement of Grade 3 learners in urban and rural primary schools. The extent to which there are gaps as well as similarities with regard to equity factors that have an effect on performance of learners in these schools was also explored with a view to ultimately predict performance in numeracy.

In this chapter, a summary of the research questions is provided (6.1) followed by the summary of the main findings[1] (6.2). A reflection is made on the methodology (6.3) and the conceptual framework (6.4) for this study. Recommendations on the key findings (6.5) are made with some thoughts for further research fore grounded in conclusion (6.6).

6.1 SUMMARY OF RESEARCH QUESTIONS

The main research question for this study was: *What are the equity factors that have an effect on the numeracy performance of Grade 3 learners in rural and urban primary schools?* This study sought to explore the rural-urban divides in the South African education system. This was done with a view to investigate the extent to which equity is being implemented as a way of redressing the imbalances of the past. A comprehensive literature review was done to find out which generic as well as equity factors, overall were found in previously conducted studies on learner achievement in numeracy/mathematics.

This study was divided into two phases to address the main research question. Firstly, it became prudent to profile the participants in the Systemic Evaluation (SE) survey and also the performance of Grade 3 learners in numeracy. Therefore three specific questions

[1] The results and their interpretation for the population in the sub-sample should be treated with caution. The results for the sub-sample apply only to respondents in the one-educator schools that made it into the final model (n = 8820). Large quantities of data were lost or not used during regression modelling due to recorded high values of missing data, low co-linearity and correlation of selected variables for this study.

arose and these were: 1) *Who are the learners, teachers, parents and schools that participated in the SE survey of 2001?* 2) *What is the overall performance of Grade 3 learners in numeracy? and 3) How does the performance in numeracy of Grade 3 learners in rural schools compare to their counterparts in urban schools?*

The second phase of this study sought to address the specific research questions, *What equity factors in this study have an effect on learner performance in numeracy?* and *What i) home background and ii) pedagogical equity factors predict learner performance in numeracy?* An exploration of the main equity factors that have an effect on learner performance in numeracy was undertaken for identification, selection and analysis in preparation for regression modelling. In the end, a regression analysis of the equity factors is done to predict performance.

6.2 SUMMARY OF THE MAIN FINDINGS

In response to the research questions in 6.1 above, the main findings for the first phase of this study provide a description of the participants in the SE survey of 2001 that was conducted by the Department of Education. In addition, the overall performance of Grade 3 learners in numeracy is described according to selected variables that included school location. The summary of the findings to the second phase research questions are presented on the contextual and background data preparation, selection and analysis of the identified equity factors. To conclude, the results of the selected equity factors that were modelled through a logistic regression to predict performance in numeracy are discussed.

Below are the research questions and the summary of results from secondary data analysis of the original SE sample as well the sub-sample for this study.

1. ***Who are the learners, educators, parents and schools that participated in the SE survey of 2001?***

In the original 2001 SE sample datasets obtained there were 52 482 learners, 2 508 educators, 1 309 school principals and 53 755 parents. Information to this effect was deciphered from contextual questionnaires. The results of the descriptive analysis of the sub-sample for this study showed that only 15 805 learners from schools with only one

Grade 3 educator were eligible for the secondary analysis. It was decided to proceed with the analysis with schools that had only one educator since it would be possible to link the educators to the learners.

In the end, only the merged dataset containing complete cases for participants in one-educator schools only could be used for the secondary analysis in this study. It is recommended that the design of future SE studies be improved so that more data obtained from the survey are useful. After careful elimination and consideration of cases with complete data, only 8 820 learners were considered for cases that went into the final logistic regression model.

The majority of learners in the both the SE sample and this study sub-sample were of the appropriate grade norm age of 7 years to 9 years (62.8% and 61% respectively). A further disaggregation of the sub-sample by location showed that 28% of the learners in the 7 to 9 years age category were from urban and semi-urban schools combined while the majority (78%) were in rural and remote-rural schools combined. Eighty nine percent (89%) of the learners in the older age category, that is 10 years and above, were from rural (rural and remote-rural combined) schools (see Appendix 4).

Girls and boys who participated in the SE study were similar at almost 50 percent for each (see Appendix 1). Both the SE sample and this study sub-sample had the same pattern of participation between boys and girls with negligible differences between both (see Appendix 4). Paviot, Heinsohn and Korkman (2007) argue that in many African communities, girls are responsible for household duties with the result that enrolment rates of girls generally tend to be lower than those for boys. Many other studies confirm the same, for example those conducted by Fuller, Singer and Keiley (1995) and Buchmann (2002). In this study, sub-sample, the number of boys (4 215) who participated was slightly higher than for girls (4 192).

The language spoken by most learners in both the SE sample and this study was IsiZulu with IsiNdebele the least spoken (see Appendix 4).

The majority (98%) of educators in the SE survey were females. Approximately 43.2% of the educators were aged between 31 and 40 years while only 9.5% were 51 years and

above (see Table 4). The low percentage of young teachers in the 21 – 30 years (10%) age category should be investigated and more graduates in this age category should be recruited. Ten percent of the educators in the SE survey had obtained or completed a degree and close to 97% had obtained at least a Matric or Grade 12 certificate. Just over 63% of the educators had a total equivalent of three years of teacher training (see Table 5b).

In the SE sample, just over a quarter (26%) of the parents had training beyond Grade 12 (see Table 6.). Many studies have shown that parents' education levels influence the learner's performance at school (Davies, 2004; Duru-Bellat, 2004; Glasman, 2004). Research in this area also shows that parents with higher levels of education are more likely to support and encourage their child's schooling because as Davies, 2004 and Ireson and Rushforth, 2005 argue, these parents value and understand the economic and social benefits that better education will provide. There is often a link between parental education levels and learner achievement. As Howie (2002, p. 126) argues, parental education may influence the parental aspirations for the child with the result that parents with higher levels of education may provide additional means and resources to assist their children at school as well as providing a more stimulating learning environment at home.

On resources in the home, most of the learners in urban areas were better off than their counterparts in rural areas. About 88% of learners in urban areas lived in homes with running water while 57% of learners in rural areas were without it (see Table 7). A further 89% in urban areas indicated that they had electricity available while 41% of their counterparts in rural areas were without electricity (see Table 8). The SE results show that, according to parents, the majority (62%) of the learners had few or no books (less than 10) at home. On the same score, over two thirds of the 8 820 learners in the sub-sample had less than 10 books at home (see Appendix 4).

The majority of the SE schools (61%) were located in rural areas. In the sub-sample, the majority of learners were also found in schools that were located in rural areas. A total of 6 883 (78%) learners from the 8 820 in the final model were in rural schools (see Appendix 4). It is important to note that no data was available to report on the gender of school principals in the original SE sample. Just over half (53%) of principals in the SE sample were aged between 41 and 50 years and over (see Table 10). Only 42.2% of

principals had a teacher college training qualification while just under a quarter (24.7%) had obtained a university degree (see Table 11).

2. *South African Grade 3 learners' performance in numeracy*

The SE survey results on learner performance show that learners obtained the lowest score for numeracy overall. The national mean for numeracy was 30%. Girls performed slightly higher than boys. The majority of learners (75%) wrote the numeracy test in their home language (DoE, 2003).

Overall, the analysis of the effect of home language on learner scores revealed that learners who took the test in their home language obtained significantly higher scores than their counterparts who had to respond to the assessment tasks in their second or third language (DoE, 2003). Ouane (2003) and others have pointed out that learners who have little opportunity to speak their home language at school tend to encounter difficulties in learning. Therefore, it might be expected that learners whose home language is not the same as the Language of Learning and Teaching (LOLT) will not perform well. On the other hand, learners whose home language is the same as the LOLT might be expected to perform well (Paviot, Heinsohn & Korkman, 2005).

3. *The comparison in numeracy performance of Grade 3 learners in rural and urban schools*

It was reported earlier in Chapter 5 (5.3) that overall, the majority (75.7%) of the learners in the sub-sample making it into the final model were from rural schools. On the whole, there were 779 urban and 744 remote rural learners with the numeracy score above median (see Appendix 4). In comparison to their counterparts in urban schools, learners in remote rural schools were not performing as well as their counterparts in urban schools.

4. *Main equity factors that affect numeracy performance of learners in rural and urban areas*

In analysing the reliability of scales in the contextual questionnaires through checking their internal consistency, it was found that most of the scales had very low Alpha values (less than 0.7) thus making it very difficult to ascertain the reliability of the scales in the questionnaires (see Appendix 5).

To measure the correlation between the independent and dependent variables, the Cramer's V statistic was used to select variables to be included in the model (see Appendix 2a). In the initial stage, all 103 variables were listed (see Appendix 2b). The list contained variables sorted from highest to lowest with variables with Cramer's V>= 0.1, which became possible candidates to be included in the model. The variables selected from the original 103 to be included in the model were categorised for convenience into ten categories that were user-friendly easy to work with (see list A of categories in 6.2.2). Results of the factor analysis further simplified the process of combining several equity variables into one factor. The new learner SES, parent SES and school funding variables were derived using factor analysis (see Appendix 2c).

5. *Home background and pedagogical equity factors that predict learners' performance in numeracy*

To answer this question, logistic regression analysis had to be done first by modelling the factors.

Home background predictors

To predict learners' performance in numeracy, only two variables from a list of 10 equity variables had a significant impact in the final logistic regression model to respond to the question of the home background factors. These factors were home language and the number of books in the learner's home.

The results (Table 13) showed that for home language, five official languages namely Sepedi, Xitsonga, Setswana, Afrikaans and IsiNdebele had average odds indices of p value below 1 indicating a decrease in the probability of obtaining a score above the median. Tables 14 and 15 point to the effect of these equity factors in predicting the performance in numeracy of learners in rural and urban primary schools. Not all the variables could make it into the regression modelling due to the minimal number of observable cases with complete data from the five datasets in SE. It is important to note that these results only refer to the population in the sub-sample and not the original SE sample.

Pedagogical factor predictors

Data pertaining to the pedagogical factor were explored during the regression analysis (see Table 16 and 17). A combination of pedagogical equity variables had a significant impact in the model. The variables were availability of computers for teaching, inadequate physical facilities, availability of numeracy reference materials, calculators, etc. in classroom and OHP available for teaching. There is no doubt that with the availability of more resources there is likely to be an increase in the odds (the likelihood of something happening or not to impact on an outcome) of learners in both rural and urban schools to achieve better in future.

6.3 REFLECTION ON METHODOLOGY

The limited number of the variables due to high missing data values had a significant impact in reducing the sample from the original SE in respect of the educator and learner populations. Only schools with one educator could be considered for the final model thus the findings of this study cannot be generalised to the entire Grade 3 learner population which was the object of the study.

It was decided to proceed with the analysis involving only one-educator schools in spite of the high missing data values because individual learners could be linked to their educators. It would not have been possible to link individual learners with their educators in schools with more than one educator since no unique ID numbers were assigned to them. This is one aspect that should be improved in the design of future SE studies so that much more data can be used meaningfully. There was a significant number of learners (15 805 approximately 40% of the original SE sample) and schools (478, over 40%) to work with, thus the decision to go ahead with the analysis was made.

The first phase of this study drew largely on the application of descriptive analysis methods to establish a context for the exploratory phase. A post hoc analysis was done on the contextual questionnaires data to establish and determine the content validity and reliability of the equity factors leading to the exploratory phase. Correlation, factor and reliability analyses of scales were done to select variables for regression modelling. This led to some of the factors in the conceptual framework, for example nutrition, falling off during the stepwise logistic regression analysis. The variables were selected from the SE

framework matrix of indicators that was based on a set of 26 agreed upon indicators classified under the transformation goal of equity.

The choice to undertake a descriptive analysis was important, firstly to verify the validity and authenticity of the statistics as reported in the SE study. A rigorous process of data, capturing, entry, cleaning and verification is strongly recommended for future SE studies to avoid too much loss of data and for cost effectiveness. Secondly, this stage of analysis was necessary to ensure that the data were clean in preparation for further exploratory analyses using the data with complete cases for the sub-sample. Initially there was also a need to recode the school location variables to address the main research question with only a dichotomous (rural and urban) classification whereas in the original questionnaire, there were four categories, namely urban, semi-urban, rural and remote-rural.

The effect of equity factors on numeracy achievement was explored in the second phase of analysis, using quantitative research methods that framed the methodological structure. Research questions were posed to explore the effect of equity factors on performance of rural and urban primary school learners and how these could predict performance. A stepwise logistic regression analysis was chosen for a number of reasons. Firstly, it was because with Regression modelling one would be able to predict the performance of learners in numeracy. Regression analysis made it possible to predict the value of the dependent variable, that is numeracy scores of learners, against gender, age, location (rural/urban) and home language (the independent variables). The quantitative methodology explored the central research hypothesis contending that equity factors have an effect on numeracy performance and further that these factors could predict performance. The hypothesis was confirmed in this study though with a limited number of variables from the conceptual framework in the final model.

6.4 REFLECTION ON THE CONCEPTUAL FRAMEWORK

The aim of this study was to conduct a secondary analysis of numeracy achievement of learners in rural and urban primary schools from a SE survey. Therefore, the findings and results of this study should be viewed and understood within this context.

This study was grounded on a post-positivistic theoretical perspective. According to Guba and Lincoln (1994) post-positivism requires objectivity to ensure that the results fit existing knowledge. Therefore, this approach was pursued to ensure that the findings from this study are well grounded scientifically and in existing knowledge theories. The findings reveal that equity factors have an effect on numeracy achievement and they can be used to predict performance thereof.

Further, the original model by Shavelson, McDonnell and Oakes (1987) was chosen and adapted for this study to monitor the education system in a way that the linkages between inputs, processes and outputs could be monitored. Although all attempts were made in this study to use as much data as were available, not all of it could be meaningfully used. It was reported earlier that this was due to high missing data values as well as low-co linearity and correlation of selected variables in the adapted model for this study's conceptual framework.

Initially about 103 variables (Appendix 2a) were identified to fit into the factors (elements) in the components (inputs-processes-outputs) of the conceptual framework. Almost all of the equity factors identified could be located in the literature review. However, only a small number of the original variables (about 9) made it into the final logistic regression model. Most of the variables were lost due to very low reliability alpha coefficients or correlation p values and high missing data values. Parent education levels, teacher qualifications and meals (both at home and at system level) are some examples of the elements in the conceptual framework (Figure 2) that lost more variables during the reliability and correlation analyses in preparation for regression modelling. Very high values of missing data were also some of the glaring problems of the SE data that made it difficult to work with the data. In the final regression analysis model, 10 variables could be mapped on the remaining factors within the components of the adapted conceptual framework.

The *input* variables (Figure 2), both at home and system levels came out as very important. SES, funding, language and LTSM stood out as critical equity factors that had an effect on numeracy achievement of learners according to specific area in which the school was located. As reported earlier no factors were found to be significant for both parent and teacher qualifications. On *processes,* only the LOLT was found to be a

significant equity factor and no other direct variables were found in the contextual questionnaires that could be significantly linked to achievement. As a result of the factor analysis, language was dealt with in the analysis as an input factor and the results reported in section 5.2. Finally, an analysis of the *outputs* in the form of the dependent variable was explored and results reported with regard to the effect of school location, input and process equity factors on predicting performance in numeracy. Strangely, the impact on odds to obtain a score above the median in future for semi-urban and semi-rural schools increased, whereas it was the contrary with the urban schools.

6.5 CONCLUSIONS AND RECOMMENDATIONS

In this section, the recommendations emanating from the findings in this chapter are presented. The purpose of this chapter is to help inform future research and interventions to improve the design of future Systemic Evaluation studies and learner performance in numeracy. The recommendations are made with a view to propose concrete intervention strategies for the improvement of the lot of our children in all types of settings, rural or urban.

CONCLUSIONS AND RECOMMENDATIONS REGARDING OVERALL LEARNERS' PERFORMANCE IN NUMERACY

Key Finding 1. Learners achieve lower scores for numeracy than literacy.

Learners do not perform well in numeracy as compared to other Learning Programmes. The results of the SE survey showed that learners' scores were lower than could have been intuitively expected. The national average scores for SE presented in Chapter 5 showed that in general, learners obtained the lowest scores for numeracy of all the three Learning Programmes that were assessed (see 5.2, Figure 4). This was particularly the case with numeracy. The national mean for numeracy was 30% while literacy was 54%.

The most probable reason for the low performance in numeracy could be that the majority of the learners in the sample came from previously disadvantaged schools. Their teachers, in many cases, did not receive proper training in the teaching of mathematics during the their teacher training in the apartheid era. As recorded in several national and international studies, for example TIMSS (1995, 1999), SE (2004) and SACMEQ II

(2004), South African learners are still under-performing in numeracy and mathematics overall.

Recommendation: *Teaching practices and resources should be stepped up with a well thought through interventionist support mechanism and plan in place to monitor implementation and evaluate progress if any. The Foundations for Learning Campaign recently launched by the Minister of Education should be implemented vigorously.* The Foundations for Learning Campaign was designed as a measure of support to schools to improve on numeracy and literacy levels in the primary system (DoE, 2008).

Key Finding 2. Learners who wrote the test in their home language obtain significantly higher scores.

Language was found to have an effect on learner scores. The language of the test viś a viś the home language of the learner came out as important on the results reported earlier in the findings chapter. The SE results revealed that learners who took the test in their home language obtained significantly higher scores than their counterparts who wrote the test in their second or third language (see 5.2, Figure 6).

Taylor (2006) argues that children are severely disadvantaged when the home language and the language of instruction are not the same. Further, on secondary analysis of TIMSS-R data, Howie (2002), found the LOLT to be a significant predictor of learner achievement in mathematics.

Recommendation: *The implementation of the language policy should be enforced especially in the Foundation Phase of schooling. It is also critical that the school reinforces the shortfall of learning resources available for learners from especially poor families, e.g. by augmenting reading books in the LOLT of the school. Parents should be encouraged to speak more often to their children in the LOLT when reading and doing homework (Taylor, 2006).*

Key Finding 3. Learners in remote rural schools are not performing as well as their counterparts in urban schools.

The majority of learners in this study's sub-sample were from rural schools. This study found that learners in remote-rural schools do not perform as well as their counterparts in

urban schools. The results in 5.3 show that only about 15% of learners from remote-rural schools in the sub-sample obtained a numeracy score above median (see Appendix 4).

Moloi and Strauss (2005) found that during the SACMEQ II study, learners in rural schools fared the worst. The low performance by learners in rural schools was also confirmed by a study conducted by Williams (2005) using PISA 2000 data. The study found that rural mathematics scores were significantly lower than scores in urban and medium-size communities in 14 out of 24 countries It is clear that with unequal conditions, for example, availability of learning and teaching support materials (LTSM) and facilities, still prevalent between rural and urban schools, the status quo of perpetual underperformance for the poor and marginalized rural communities, will continue.

Recommendation: *A pro-poor and rural biased programme of intervention and support like the Quality Improvement, Development, Support and Upliftment Programme (QIDSUP) should be rolled out as a matter of urgency to all identified schools with a view to improve on numeracy and literacy achievement.* The QIDSUP programme is an initiative of the Department of Education to resource schools earmarked as the poorest of the poor in quintiles 1 – 3 with a view to improve the quality of teaching and learning. A basic minimum resource package to support the teaching of numeracy and literacy is made available to the identified schools.

CONCLUSIONS AND RECOMMENDATIONS REGARDING EQUITY FACTORS AND LEARNERS' PERFORMANCE IN NUMERACY

Key Finding 4. Equity factors have an effect on learner performance in numeracy.
Equity factors impact on learners' performance in numeracy. This study found that 10 equity factor variables had an effect on learners' performance in numeracy by location of school (see 5.4, list A of final model variables). The variables included language, SES (e.g. school fees) and LTSM (e.g. books in home, computers, OHP.).

This finding confirms earlier results from a SACMEQ II study that showed that SES has an effect on numeracy performance. The study found that learners from high SES backgrounds achieved higher levels of competence than their counterparts from low SES (Moloi & Strauss, 2005). On the other hand, school resources were generally found to

have no effect on learner achievement in two separate studies conducted in the USA and Hungary (Arnold, 1995; Vari, 1997).

Recommendation: Schools can make a large difference when they are better resourced in poor and isolated communities. *Schools should make up for some of the learning support materials and resources that are not always available at home and in the community in general, for example books, library and computers. Qualified teachers should also be available to teach in rural areas with some incentives.*

Key Finding 5. Home background equity factors predict learner performance in numeracy.

The home background of learners and the availability of equity variables were found to have a role to play on learner performance in numeracy. Home language and the number of books in the learner's home had a significant and positive effect in predicting the performance of learners in numeracy (see 5.5). There was a marked decrease in the probability of obtaining a score above the median as compared to others for five out of the 11 official languages (Afrikaans, IsiNdebele, Sepedi, Setswana and Xitsonga).

A finding from another baseline systemic evaluation study at the Grade 6 level also confirms that learners, who wrote the tests in their home language, obtained significantly higher scores across all learning areas than their counterparts who wrote the tests in a second or third language (DoE, 2005). Pattanayak (2003) posits that home language is the best medium of education at school because it is the language the child knows well, in which he or she can express meanings.

Recommendation: *Teachers who are adequately qualified and trained to teach in the various LOLTs in the Foundation Phase should be produced by institutions of higher learning. Learning and teaching resource materials should be made available in all the 11 official languages.*

Several studies have found a positive relationship between the number of books in the home and learners' achievement in mathematics (Mullis, et al., 2000; Shen, 2001). In this study, there was a significant increase in the probability to obtain a score above the median when there were more books (>20) in the home.

Recommendation: *A system should be devised to make reading books and other library material available for learners for during and after school use through book corners, classroom libraries, public libraries and mobile libraries. Parents should also be encouraged to collect reading materials, model reading and encourage a love and culture for reading even if they start with some scripture reading*

Key Finding 6. Pedagogical equity factor predicts learner performance in numeracy.
The equity variables that are available or provided for by the school are significant predictors of the performance of learners in numeracy. In this study, availability of computers for teaching, lack of inadequate physical facilities, availability of numeracy reference materials, calculators in the classroom and overhead projector were found to have a positive effect on numeracy achievement (see 5.4).

Howie (2002) supports this finding by arguing that where learners in Grade 4 had experienced learning using learning aids such as rulers, blocks, shapes and solids, there was a positive link with geometry achievement.

Recommendation: *A basic minimum resource package targeted for QIDSUP schools should be made available to all schools prioritising those in remote-rural areas as a benchmark.*

6.6 LIMITATIONS OF THE STUDY

This study was designed to give insight into numeracy achievement and the state of affairs regarding equity in South African schools in both rural and urban areas. A few limitations are discussed in this section.

- The study was *not designed for making causal statements* pertaining to how specific factors, for example equity factors, impact on learner achievement.

- The analysis of contextual data is based on self-reported data obtained from principals, parents, educators, and learners during the main SE survey in 2001. There is, therefore, an inherent danger for participants to either over- or under-report subjectively.

- During the preparation of the data for secondary analysis, some variables, especially in the contextual questionnaires, were eliminated since the rate of response was too low, for example missing data of over 15%. This could adversely affect the analysis on the effect of the equity factors on learner achievement.

- Scale analysis during the secondary exploratory analysis The other problem worth noting is that most of the scales had very low Alpha values (less than 0.7) thus making it very difficult to ascertain the reliability of the scales in the questionnaires.

6.7 CONCLUSION

This study found that fourteen years into democracy, substantial gaps in achievement and provisioning are even now clearly evident. Learners in rural schools are at grave risk of under-performing, dropping out, falling pregnant and experiencing an unsatisfactory learning experience at school because of the backlogs in addressing issues of equity in the education system.

The gaps between rural and urban schools still continue unabated despite some policy interventions that are targeted towards mostly rural communities. Initiatives of the Department of Education (DoE) still have to bear fruit in terms of improving educational quality and in particular literacy and numeracy. The QIDSUP initiative as well as the Foundations for Learning, and the Quality Teaching campaigns are some of the programmes that are intended to bridge the divide between rural and urban schools with at least what is considered to be minimum requirements and resources for a school to do its core task, namely teaching and learning.

On the whole, it can be concluded that if more resources and qualified teachers (who are motivated and actually go to school on time and teach) can be made available, in the form of mediated pedagogical equity variables, to rural primary schools in particular, the odds for success will be increased. Trends will have to be established during the next cycle of SE to ascertain whether there are gains being made on performance as well as on equity regarding the provisioning of resources. In conclusion, the different impact on odds, i.e. the likelihood of something happening/being available to impact on an outcome, below or above the median scores for the various languages requires further investigation.

Table 15

Example 2: Model 4 Application II – IsiNdebele (Home background factors)

Variable	Level of significance (p-value)	Source	Selection appropriate category:	Impact on odds	Increase / decrease in odds	
Learner's home language	<0.0001	LQLANG	**ISINDEBELE**	**0.297**	70.26%	**Decrease**
School's section 21 status	<0.0001	PR57	1 NO	**1.104**	10.44%	**Increase**
Annual school fees	<0.0001	PR52	2 50-100	**0.893**	10.71%	**Decrease**
Computers available for teaching	<0.0001	PR25I	1 NOT AVAIL	**0.464**	53.58%	**Decrease**
Impact of inadequate physical facilities	<0.0001	ED38D	3 QUITE ALOT	**0.839**	16.10%	**Decrease**
Number of books in home	<0.0001	PAR17	**1 <10**	**0.866**	13.43%	**Decrease**
Numeracy reference material available	<0.0001	ED23EMa	NO	**0.874**	12.65%	**Decrease**
Calculators in classroom	<0.0001	ED19Q	NO	**0.871**	12.86%	**Decrease**
Location of school	0.0219	PR5	**4 REMOTE RUR**	**0.992**	0.79%	**Decrease**
OHP available for teaching	<0.0001	PR25K	1 NOT AVAIL	**0.722**	27.85%	**Decrease**
Odds - greater than median			**0.17**			
Probability - greater than median			**14.25%**			

PEDAGOGICAL EQUITY FACTOR THAT PREDICTS LEARNER PERFORMANCE IN NUMERACY

A number of variables on the pedagogical factor of availability of resources for teaching were analysed by regression analysis. Five of the final variables in the model account for this pedagogical factor namely availability of computers for teaching, inadequate physical facilities, availability of numeracy reference materials, calculators in the classroom and OHP available for teaching. The following are the results on how the pedagogical factor predicts numeracy performance by location as well.

Model 4 Application III – IsiXhosa

As shown in the model application (Table 16 - IsiXhosa), a remote rural school is highly likely to experience a decrease on the impact of the odds to obtain a score above the median given the five statuses of the selected pedagogical factor variables. The range of decrease in impact on odds is 13% for non-availability of numeracy reference materials to 54% for non-availability of computer for teaching. Overall under all the variable conditions listed in the model, the probability of obtaining a score above the median in future is 47% as indicated by the odds change factor.

Table 16

Example 3: Model 4 Application III – IsiXhosa (Pedagogical factor)

Variable	Level of significance (p-value)	Source	Selection appropriate category:	Impact on odds	Increase / Decrease in odds	
Learner's home language	<0.0001	LQLANG	ISIXHOSA	**1.953**	95.31%	**Increase**
School's section 21 status	<0.0001	PR57	2 PARTLY	**0.600**	40.00%	**Decrease**
Annual school fees	<0.0001	PR52	3 101-200	**0.991**	0.91%	**Decrease**
Computers available for teaching	<0.0001	PR25I	1 NOT AVAIL	**0.464**	53.58%	**Decrease**
Impact of inadequate physical facilities	<0.0001	ED38D	3 QUITE ALOT	**0.839**	16.10%	**Decrease**
Number of books in home	<0.0001	PAR17	3 >20	**1.187**	18.71%	**Increase**
Numeracy reference material available	<0.0001	ED23EMa	NO	**0.874**	12.65%	**Decrease**
Calculators in classroom	<0.0001	ED19Q	NO	**0.871**	12.86%	**Decrease**
Location of school	0.0219	PR5	4 REMOTE RUR	**0.992**	0.79%	**Decrease**
OHP available for teaching	<0.0001	PR25K	1 NOT AVAIL	**0.722**	27.85%	**Decrease**
Odds - greater than median			**0.90**			
Probability - greater than median			**47.44%**			

Model 4 Application IV – IsiXhosa

The example in the table below (Table 17 - IsiXhosa) indicates that for a learner in a remote rural school with five of the pedagogical factor variables available as identified in the selection column, the odds of obtaining a numeracy score above the median is likely to increase. As can be seen from the table, the odds increased for all the pedagogical factor variables except for the impact of inadequate physical facilities (impact on odds decrease by 11%). The model predicts overall that the probability of a learner in a school of these conditions obtaining a score above the median in future is 96%.

Table 17

Example 4: Model 4 Application IV – IsiXhosa (Pedagogical factor)

Variable	Level of significance (p-value)	Source	Selection appropriate category:	Impact on odds	Increase / Decrease in odds	
Learner's home language	<0.0001	LQLANG	ISXHOSA	**1.953**	95.31%	**Increase**
School's section 21 status	<0.0001	PR57	3 FULL	**1.509**	50.91%	**Increase**
Annual school fees	<0.0001	PR52	3 101-200	**0.991**	0.91%	**Decrease**
Computers available for teaching	<0.0001	PR25I	3 ALWAYS	**1.498**	49.78%	**Increase**
Impact of inadequate physical facilities	<0.0001	ED38D	1 NOT AT ALL	**0.889**	11.08%	**Decrease**
Number of books in home	<0.0001	PAR17	3 >20	**1.187**	18.71%	**Increase**
Numeracy reference material available	<0.0001	ED23EMa	YES	**1.145**	14.48%	**Increase**
Calculators in classroom	<0.0001	ED19Q	YES	**1.148**	14.75%	**Increase**
Location of school	0.0219	PR5	2 SEMI-URBAN	**1.052**	5.21%	**Increase**
OHP available for teaching	<0.0001	PR25K	3 ALWAYS	**1.179**	17.89%	**Increase**
Odds - greater than median			**23.22**			
Probability - greater than median			**95.87%**			

CONCLUSION

The findings on the results of the descriptive, exploratory and regression analyses were presented and discussed. The overall performance in numeracy of Grade 3 learners in rural schools compared to their counterparts in urban schools was not satisfactory. Emerging from the exploratory analysis and the conceptual framework adapted for this study, language, funding and LTSM were found to have an effect on learner performance in numeracy according to the location of the school.

In conclusion, the results of the regression analysis using the impact on odds of the equity variables to predict the performance of rural and urban school learners in numeracy were also presented and discussed. In the final chapter, a reflection on both the methodology and conceptual framework, a summary of the key findings and recommendations are presented.

REFERENCES

Afrassa, T.M. (1998). *Mathematics achievement in the lower secondary school stages in Australia and Ethiopia: a comparative study of standards of student achievement and student-level factors influencing achievement.* Unpublished PhD Thesis. Adelaide: Flinders University of South Australia, School of Education.

Agresti, A. (2002). *Categorical Data Analysis, 2nd Edition.* New York, NY: John Wiley & Sons

Aldrich, J (1997). R.A. Fisher and the making of maximum likelihood 1912 – 1922. A paper on the history of Maximum Likelihood: *Statistical Science 12* (3): 162–176. doi:10.1214/ss/1030037906

Anderson, L.W., Ryan, D., & Shapiro, B. (1989). The IEA classroom environment study. NY: Pergamon

Arnold, C.L. (1995). *Using HLM and NAEP data to explore school correlates of 1990 mathematics and geometry achievement in Grades 4, 8 and 12. Methodology and Results.* Washington: National Centre for Education Statistics.

Baker, B.D. (2000). Can flexible non-linear modelling tell us anything new about educational productivity? *Economics of Education Review, 20*(1), 81 - 92

Balli, S.J. (1998). When mom and dad help: Student reflections on parent involvement with homework. *Journal of Research and Development in Education, 31*(3), 142 – 146.

Bennett, N. (1996). Class size in primary schools: Perceptions of head teachers, chairs of governors, teachers and parents. *British Educational Research Journal, 22*(1), 33-55.

Bloch, M. N., & Tabachnick, B. R. (1994). Improving parent involvement as school reform. In K. M. Borman & N. P. Greenman (Eds.). *Changing American education: Recapturing the past or inventing the future?* (pp. 261-293). Albany: State University of New York Press.

Braswell, J.S., Lutkus, A.D., Grigg, W.S., Santapu,S.L., Tay-Lim, B.S.-H., & Johnson, M.S. (2001). *The nation's report card: Mathematics 2000*, Washington, DC: National Center for Education Statistics.

Briggs, S.R., & Cheek, J.M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality, 54,* 106 – 148.

Buchmann, C. (2002). *Getting ahead in Kenya: social capital, shadow education, and achievement.* In Fuller, B., Hannum, E. (Eds.), Schooling and Social Capital in Diverse Cultures. JAI Press, Amsterdam, pp. 133 – 159.

Chinapah, V., H'dggui, E.M., Kanjee, A., Falajayo, W., Fomba, C.O., Hamissou, O., et al. (1999). *Towards Quality Education for All.* Pretoria: HSRC.

Christie, K. (2001). "Stateline" *Phi Delta Kappan*. Retrieved from www.ecs.org Reprinted with permission in 2005 Education Commission of the States "Helping state leaders shape education policy" Denver, Colorado

Cramer, H. (1999). Mathematical Methods of Statistics, Princeton University Press.

Creswell, J.W. (2003). *Research Design. Qualitative, quantitative, and mixed methods approaches (2nd ed.)* Thousand Oaks, CA.: Sage Publications.

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process.* Thousand Oaks, CA: Sage.

Davies, S. (2004). School choice by default? Understanding the demand for private tutoring in Canada. *American Journal of Education 110*(3), 233 – 255.

Dawes, M.E., Horan, J.J., & Hackett, G. (2000). Experimental evaluation of self-efficacy treatment on technical/scientific career outcomes. *British Journal of Guidance & Counselling, 28*(1), 87 – 99.

DeCoster (1998). *Overview of Factor Analysis.* Retrieved July 7 2006 from http://www.stat-help.com/notes.html

Denton, K., & West, J., (2002). *Children's reading and mathematics achievement in kindergarten and first grade* (NCES 2002). Washington, DC: US Department of Education, National Center for Education Statistics.

Department of Education (1995). *White Paper on Education and Training. Government Gazette, Vol. 357, No. 16312.* Pretoria: Government Printers.

Department of Education (1998). *Admission Policy.* Pretoria.

Department of Education (2000a). *Education for All: The South African Assessment Report.* Pretoria.

Department of Education (2000b). *Education in a global era: Challenges to equity, opportunities for diversity.* Country paper: South Africa delivered at the 14th Conference of Commonwealth education Ministers, Halifax, Nova Scotia, Canada, 27 – 30 November 2000

Department of Education (2000c). *Schools Register of Needs.* Pretoria.

Department of Education (2001). *Framework for Systemic Evaluation.* Pretoria.

Department of Education (2003). *National Report on Systemic Evaluation*: *Mainstream Education, Foundation Phase.* Pretoria.

Department of Education (2005). *National Report on Systemic Evaluation*: *Grade 6, Intermediate Phase.* Pretoria.

Department of Education (2008). *Foundations for Learning Gazette, Vol. 306, No. 30880.* Pretoria: Government Printers.

Dieltiens, V. (2006). Systemic Evaluation and the National Protocol on Assessment. *Quarterly Review of Education & Training in South Africa, 13*(1), 5 – 6.

Duru-Bellat, M. (2004). *Social inequality at school and educational policies.* Fundamentals of Educational Planning. Paris: UNESCO/IIEP.

Easton, V.J., & McColl, J.H. (1997). *Statistics glossary. Presenting data [online glossary]* Retrieved October 16, 2003 from http://www.stats.gla.ac.uk/steps/glossary/presenting_data.html#freqtab

Elmore, R. (1997) *Investing in teacher learning: staff development and instructional improvement in Community School District 2, New York City*. New York: CPRE

Fuller B., Singer J., & Keiley M. (1995). Why do daughters leave school in Southern Africa? Family economy and mothers' commitments. *Social Forces, 74*(2), 657-681.

Gay, L. R., & Airasian, P. (2003). *Educational Research: Competencies for analysis and application*. Upper Saddle River: Merrill Prentice Hall.

Glasman, D. (2004). Le Travail des Eléves pour lĔcole en Dehors de lĔcole [The School Work of Pupils Outside School]. Savoie, France: Inversité de Savoie, Faculté de Lettres, Langues et Sciences Humaines.

Gillwald, A. (2000). *South Africa and the Information Age.* A response paper: Castells Tour, CHET: Pretoria.

Goldhaber, D. (2002). "The Mystery of Good Teaching: Surveying the Evidence on Student Achievement and Teachers' Characteristics." *Education Next 2(1): 50 – 55.*

Greenberg, E., Rhodes, D., Ye, X., & Stancave, F. (2004). Prepared to Teach: Teacher preparation and student achievement in eighth-grade mathematics. American educational research association, 2004 annual meeting, San Diego, Calif.

Guba, E.G., & Lincoln, Y.S. (1994). Competing paradigms in qualitative research. In N.K. Denzin & Y.S. Lincoln (Eds). *Handbook of Qualitative Research.* Thousand Oaks, CA: Sage.

Guiton, G., & Oakes, J. (1995). "Opportunity to learn and conceptions of educational equality." *Educational Evaluation and Policy Analysis, 17*(3), 323-336.

Gustafsson, M. (2003). *Technical Characteristics of the 2001 Systemic Evaluation Data (2003a) To what extent can the 2001 Systemic Evaluation dataset be used to make inferences about the schooling system (2003b).* Unpublished Reports, Department of Education, Pretoria.

Herman, J.L., & Klein, D.C.D. (1996). Evaluating Equity in Alternative Assessment: An Illustration of Opportunity-to-Learn Issues. *Journal of Educational Research, 89*(4), 246-56.

Hofmeyer, J., & Buckland P. (1992). *Education System Change in South Africa.* In R. McGregor & McGregor, A. (Eds.). McGregor's Alternatives. Kenwyn: Juta's & Co. Ltd.

Howell, D.C. (1999). *Fundamental statistics for the behavioural sciences.* 4th edition. University of Dermont. Duxberry Press, International Thompson Publishing Inc.

Howell, D.C. (2001). *Fundamental statistics for the behavioural sciences; 4th edition & Statistical methods for psychology; 4th edition. [Online glossary]* Retrieved October 16, 2003, from http://www.uvm.edu/~dhowell/StatPages/Glossaries/GlossaryA-D.html

Howie, S.J., & Pieterson, J.J. (2001). Mathematics Literacy of Final Year Students: South African Realities. *Studies in Educational Evaluation,* 27(1), 7 – 26.

Howie, S. (2002). *English language proficiency and contextual factors influencing mathematics achievement of secondary school pupils in South Africa.* Published Thesis University of Twente. PrintPartners Ipskamp, Enschede.

Human Sciences Research Council (1997). The Education Foundation and Research Institute for Education Planning (1997). Schools Register of Needs Survey 1996: CEM/HEDCOM Presentation. Pretoria: Department of Education.

Human Sciences Research Council (2003). The District Development Support Programme. Pretoria.

Hungi, N. (2005). *Factors influencing Standard 6 pupils achievement in Kenya: A multilevel analysis.* Harare, SACMEQ. Available online from: <http://www.sacmeq.org> [Accessed May 2006].

Ireson, J., & Rushforth, K. (2005). *Mapping and evaluating shadow education.* ESRC Research Project RES-000-23-0117. End of Award Report. Institute of Education, University of London, London.

Kao, G. (2004). Parental influences on the educational outcomes of immigrant youth. *International Migration Review, 38,* 427 – 449.

Kanjee, A. & Prinsloo, C.H. (2005) Improving learning in South African schools: the Quality Learning Project (*QLP*). *Summative evaluation* (2000 to 2004).

Khoali, T. (2006). Quality education: The problem of definition. *Quarterly Review of Education & Training in South Africa, 13*(1), 5 – 6.

Kinard, E. M., & Reinherz, H. (1986). Birthdate effects on school performance and adjustment: A longitudinal study. *The Journal of Educational Research, 79*, 366-372.

Kurdek, L.A., & Sinclair, R.J. (2001). Predicting reading and mathematics achievement in fourth grade children from kindergarten readiness scores. *Journal of Educational Psychology 93*(3), 451 – 455.

Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 163-188). Thousand Oaks, CA: Sage.

Mbeki, T. (2004). "Two Worlds: two economies". *South African Labour Bulletin, 28*(2), 10–11.

McDonnell, L. M. (1995). "Opportunity to learn as a research concept and policy instrument." *Educational Evaluation and Policy Analysis, 17*(3), 305-322.

Mohandas, R. (1999). *Mathematics and Science achievement of junior secondary students in Indonesia,* Unpublished PhD. Thesis. Adelaide: Flinders University of South Australia, School of Education.

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

Monyana, H.J. (1996). *Factors related to mathematics achievement of secondary school pupils.* M.Ed. Thesis. University of South Africa. Unpublished.

Mullis, I.V.S. (1991). *The state of mathematics achievement: NAEP's assessment of the nation and the trial assessment of the states.* Washington, DC.: GPO.

Mullis, I., Martin, M., Gonzalez, E.J., Gregory, K.D., Garden, R.A., O'Connor, K.M., et al. (2000). *TIMSS 1999 International Mathematics Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade.* Chestnut Hill: Boston College.

Murray, G. (1997). Opportunity-to-Learn Issues Common to South Africa and the United States. *Journal of Negro Education, 66*(4), Education in a New South Africa: The Crises of Conflict, the Challenges of Change, pp. 376-382 doi:10.2307/2668165

Nye, B., Hedges, L.V., & Konstantopoulos, S. (2000) The Effects of Small Classes on Academic Achievement: The Results of the Tennessee Class Size Experiment. *American Educational Research Journal*, *37*(1), 123-151 doi:10.2307/1163474

Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.

Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science*. Santa Monica, CA: The Rand Corporation.

Oakes, J. (1992). On tracking and individual differences. *Educational leadership 50*(2), 18 – 22.

Oakes, J., & Lipton, M. (1990). *Making the best of schools: A handbook for parents, teachers, and policymakers.* New Haven: Yale University Press

Oosthuizen (2005). 'Technical Report on Data Sources', Draft Report 1: *Development Policy Research Unit for the Education Foundation Trust,* Unpublished.

Ouane, A. (2003). The impossible debate about the use of mother tongues in education. In A. Ouane (Ed.), *Towards a multilingual culture of education* (pp.51-86). Hamburg: UNESCO Institute for Education.

Pallant, J. (2001). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows (versions 10 and 11).* Buckingham, UK: Open University Press.

Parsons, J.E., Adler, T.F., & Kaczala, C.M. (1982). Socialisation of achievement attitudes and beliefs: Parental influences. *Child Development*, *Vol.53,* 310 – 321.

Pattanayak, D.P. (2003). Mother tongues: The problem of definitions and the educational challenge. In A. Ouane (Ed.), *Towards a multilingual culture of education* (pp.23-28). Hamburg: UNESCO Institute for Education.

Paviot, L., Heinsohn, N., & Korkman, J. (2007). Extra Tuition in Southern and Eastern Africa: Coverage, growth and linkages with pupil achievement. *International Journal of Educational Development (2007),* doi:10.1016/j.ijedudev.2007.02.003.

Payne, K.J., & Biddle, B.J. (1999). Funding, poverty, and mathematics achievement (Turner, 2000) in *Educational Researcher, 29*(7), 27 – 29

Poswell, L. (2006). Educational Outcomes in South Africa: A production function approach. *Technical Report on Systemic Evaluation for Grade 3, 2001.* DoE and SISERA, UCT.

Pretorius, F. & Lemmer, E.M. (1998) *South African Education and Training: Transition in a Democratic Era.* (Eds). Johannesburg: Hodder & Stoughton, 1998.

Prins, E., & Toso, B.W. (2008) Defining and measuring parenting for educational success: A critical discourse analysis of the parent education profile. *American Educational Research Journal, 45*(3) 555 – 596.

Pritchard, I. (1999). Reducing class size: What do we know? [Online]. Available: http://www.ed.gov/pubs/ReducingClass/Class_size.html Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Reeves, C. (2005). The effects of pedagogical practices on mathematics learning Review Committee (2000) A South African Curriculum for the Twenty First Century: Report of the Review Committee on Curriculum 2005. Pretoria: Ministry of Education.

Rice, J.K. (1999). The impact of class size on instructional strategies and the use of time in high school mathematics and science courses. *Educational Evaluation and Policy Analysis, 21*(2), 215-130.

Rousseau, C.K., & Powell, A. (2005). "Understanding the Significance of Context: A Framework to Examine Equity and Reform in Secondary Mathematics". *The High School Journal – 88*(4), 19-31.

Republic of South Africa (1994). *RDP White Paper. Government Gazette, Vol. 2, No. 16085.* Pretoria: Government Printers

SACMEQ II (2004). *South Africa Report*. http://www.sacmeq@iiep.unesco.org

Schwartz, W. (1995). Opportunity To Learn Standards: Their Impact on Urban Students. ERIC/CUE Digest No. 110. File retrieved 13102006 from http://eric.ed.gov/ERICWebPortal/Home.portal?_nfpb=true&_pageLabel=RecordDetails&ERICExtSearch_SearchValue_0=ED389816&ERICExtSearch_SearchType_0=eric_accno&objectId=0900000b801b5132

Shavelson, R.J., McDonnell, L.M., & Oakes, J. (1987). *Indicators for monitoring mathematics and science education: a sourcebook.* Santa Monica (CA, USA): The RAND Corporation.

Sheldon, S.B., & Epstein, J.L. (2005). Involvement Counts: Family and Community Partnerships and Mathematics Achievement. *The Journal of Educational Research, (98)*4, 196 – 206.

Shen, C. (2001). Social values associated with cross-national differences in mathematics and science achievement: A cross-national analysis*, Assessment in Education, 8*(2), 198-223.

Shepard, L.A., & Smith, M.L. (1986). "Synthesis of Research on School Readiness and Kindergarten Retention.", *Educational Leadership, 44*(3), 78-86.

Sojourner, J., & Kushner, S.N. (1997). *Variables that impact the education of African-American students: parental involvement, religious socialization, socio-economic status, self-concept and gender.* Paper presented at the Annual Meeting of the American Educational Research Association in Chicago, Illinois, March 24 - 28, 1997.

Spera, C. (2005). A review of the relationship among parenting practices, parenting styles, and adolescent school achievement. *Educational Psychology Review, 17,* 125 – 146.

Statistics South Africa (2003). *Census 2001: Investigation into appropriate definitions of urban and rural areas for South Africa: Discussion document/ Statistics South Africa.* Pretoria: Statistics South Africa, 2003195p. [Report No. 03-02-20 (2001)]

Statistics South Africa (2005) *General Household Survey 2004*. Pretoria, Cape Town: Statistics South Africa

Stevens, F. I., & Grymes, J. (1993). Opportunity to learn: Issues of equity for poor and minority students. Washington, DC: U.S. Department of Education, National Center for Education Statistics. (ED 356 306)

Strauss, J.P. (2005). *A model for evaluating South Africa's education system based on SACMEQ II research data.* R.I.E.P, University of the Free State, Bloemfontein, South Africa.

Sugiyama, Y. (2001). On Students' Mathematics Achievement in Japan *Online article* http://www.nctm.org/dialogues/2001-11/20011103.htm

Tabachnik, B. G., & Fidell, L. S. (1989). Using multivariate statistics. (2nd edition). New York: HarperCollins

Taylor, N. (2006). *Equity, Efficiency and the Development of South African Schools.* File retrieved 29092006 from http://64.233.167.104/search?q=cache:0Bnlx7Lz2aQJ:jet.org.za/documents/Tay

Taylor, N., & Vinjevold, P. (Eds.) (1999). *Getting Learning Right*. Johannesburg: Joint Education Trust

Taylor, N.C., Muller, J.P., & Vinjevold, P. (2003). *Getting schools working: research and systemic school reform in South Africa.* Cape Town: Pearson Education Publishers.

TIMSS (1999) *International Mathematics Report*

http://www.timss.bc.edu/timss1999i/math_achievement_report.html

Turner, S.E. (2000). A comment on "Poor School Funding, Child Poverty, and Mathematics Achievement". *Educational Researcher,* 29(5), 15 –18.

UCLA: Academic Technology Services, Statistical Consulting Group. (2007). *Introduction to SAS.*

from http://www.ats.ucla.edu/stat/sas/notes2/ (accessed November 24, 2007).

Valencia, R. R., & Black, M. S. (2002). "Mexican Americans don't value education!" On the basis of myth, mythmaking, and debunking. *Journal of Latinos and Education, 1*(2), 81 – 103.

Van den Berg, S. (2005). Apartheid's Enduring Legacy: Inequalities in Education. Paper presented to the Oxford University/University of Stellenbosch Conference on South African Economic Policy Under Democracy: A 10-Year Review, Stellenbosch, 27-28 October 2005.

Van der Berg, S., Burger, R., & Yu, D. (2005) *Determinants of Education Quality: A report on the Western Cape Primary School Pupil survey 2003*. Mimeo.

Van Lill, J.B., & Grieve, K.W. (1990). *Descriptive statistics for students in the human sciences.* Pretoria. Human Resources Information (CC).

Vari, P. (1997). *Monitor'95. National assessment of student achievement.* Budapest: National Institute of Public Education.

Vithal, R., & Jansen, J. (1997). Designing your first research proposal. Cape Town: Juta

West, J., Denton, K., & Germino-Hausken, E. (2000). *America's kindergarteners: Findings from the Early Childhood Longitudinal Study, kindergarten class of 1998 – 1999, Fall 1998.* Washington, DC: U.S. Department of Education.

Williams, J. H. (2005). Cross-national variations in rural mathematics achievement: A descriptive overview. *Journal of Research in Rural Education, 20*(5). Retrieved [6 September 2006] from http://www.umaine.edu/jrre/20-5.htm

Willms, J.D., & Somers, M. (2001). Family, classroom and school effects on children's educational outcomes in Latin America. *School effectiveness and school improvement, 12*(4), 401-445.

Yamauchi, F. (2004). Race, equity and public schools in post-apartheid South Africa: is opportunity equal for all kids? FCND Discussion Paper No. 182. Washington, D.C.: International Food Policy Research Institute.

Young, D.J. (1997). *A multilevel analysis of science and mathematics achievement.* Paper presented at the Annual Meeting of the American Educational Research Association in Chicago, Illinois, March 24 - 28, 1997.

Appendix 1 DEMOGR1

1 COMBINE

14:27 Tuesday, March 20, 2007 1

PHILLIP TSHABALALA T06075 DPG9079 OD427014

DEMOGRAPHICS - ORIGINAL LAD LEARNERS

The FREQ Procedure

Cumulative Frequency	Cumulative Percent	AGECAT	Frequency	Percent
		1 7-9	30423	62.89
30423	62.89			
		2 10	9878	20.42
40301	83.31			
		3 10+	8073	16.69
48374	100.00			

Frequency Missing = 4108

GENDER

Cumulative Frequency	Cumulative Percent	GENDER	Frequency	Percent
		1	25021	50.80
25021	50.80			
		2	24232	49.20
49253	100.00			

Frequency Missing = 3229

LANGUAGE

Cumulative Frequency	Cumulative Percent	LANGUAGE	Frequency	Percent
		1	6418	12.53
6418	12.53			
		2	3165	6.18
9583	18.72			
		3	869	1.70
10452	20.41			
		4	10039	19.61
20491	40.02			
		5	12084	23.60
32575	63.62			
		6	4890	9.55
37465	73.17			
		7	3804	7.43
41269	80.60			
		8	4503	8.79
45772	89.39			
		9	1452	2.84
47224	92.23			
		10	1745	3.41

48969	95.64	11	2144	4.19
51113	99.82	12	91	0.18
51204	100.00			

Frequency Missing = 1278

SUMMARY STATISTICS AGE - ORIGINAL LAD LEARNERS

The MEANS Procedure

Analysis Variable : AGE AGE

N	Mean	Median	Std Dev	Minimum	Maximum
48374	9.4192128	9.0000000	1.2238355	7.0000000	13.0000000

LANGUAGE FREQUENCIES - ORIGINAL LEQ LEARNERS

The FREQ Procedure

LQLANG

LQLANG	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	5716	12.33	5716	12.33
2	2663	5.74	8379	18.07
3	679	1.46	9058	19.54
4	9354	20.18	18412	39.71
5	10821	23.34	29233	63.05
6	4265	9.20	33498	72.25
7	3579	7.72	37077	79.97
8	4398	9.49	41475	89.46
9	1130	2.44	42605	91.90
10	1620	3.49	44225	95.39
11	2046	4.41	46271	99.80
12	91	0.20		

46362 100.00

Frequency Missing = 4954

1 COMBINE

14:27 Tuesday, March 20, 2007 4

PHILLIP TSHABALALA T06075 DPG9079 OD427014

SCHOOL AREA FREQUENCIES - LAD MERGED WITH PRQ

The FREQ Procedure

PR5

PR5	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	11482	26.16	11482	26.16
2	7078	16.13	18560	42.29
3	21150	48.19	39710	90.48
4	4180	9.52	43890	100.00

Frequency Missing = 503

Appendix 1 DEMOGR2

1 PROG1
14:12 Tuesday, March 20, 2007 1
PHILLIP TSHABALALA T06075 DPG9079
OD427014
DEMOGRAPHICS FOR FINAL COMBINED FILE -
1 EDUCATOR

The FREQ Procedure

PR5

PR5	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1846	11.73	1846	11.73
2	1133	7.20	2979	18.92
3	10156	64.51	13135	83.43
4	2608	16.57	15743	100.00

Frequency Missing = 62

AGECAT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1 7-9	8468	58.78	8468	58.78
2 10	3196	22.18	11664	80.96
3 10+	2743	19.04	14407	100.00

Frequency Missing = 1398

GENDER

GENDER	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	7494	50.54	7494	50.54
2	7334	49.46	14828	100.00

Frequency Missing = 977

1 PROG1
14:12 Tuesday, March 20, 2007 2
PHILLIP TSHABALALA T06075 DPG9079
OD427014
DEMOGRAPHICS FOR FINAL COMBINED FILE -
1 EDUCATOR

Appendix 1 DEMOGR2

The FREQ Procedure

LQLANG

LQLANG	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1073	7.61	1073	7.61
2	347	2.46	1420	10.06
3	125	0.89	1545	10.95
4	2639	18.70	4184	29.65
5	4086	28.96	8270	58.62
6	2092	14.83	10362	73.44
7	668	4.73	11030	78.18
8	1411	10.00	12441	88.18
9	328	2.32	12769	90.50
10	639	4.53	13408	95.03
11	687	4.87	14095	99.90
12	14	0.10	14109	100.00

Frequency Missing = 1696

LANGUAGE

LANGUAGE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1227	7.97	1227	7.97
2	481	3.13	1708	11.10
3	181	1.18	1889	12.27
4	2650	17.22	4539	29.49
5	4325	28.10	8864	57.59
6	2411	15.66	11275	73.26
7	662	4.30	11937	77.56
8	1399	9.09	13336	86.65
9	625	4.06	13961	90.71
10	676	4.39	14637	95.10
11	740	4.81	15377	99.91

12 14 0.09
15391 100.00

Frequency Missing = 414

1 PROG1
14:12 Tuesday, March 20, 2007 3
PHILLIP TSHABALALA T06075 DPG9079
OD427014

SUMMARY STATISTICS FOR AGE - FINAL COMBINED FILE

The MEANS Procedure

Analysis Variable : AGE AGE

N	Mean	Median	Std Dev	Minimum	Maximum
14407	9.5100298	9.0000000	1.2658629	7.0000000	13.0000000

1 PROG1
14:12 Tuesday, March 20, 2007 4
PHILLIP TSHABALALA T06075 DPG9079
OD427014

DEMOGRAPHICS FOR CASES IN MODEL

The FREQ Procedure

PR5

PR5	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1173	13.30	1173	13.30
2	764	8.66	1937	21.96
3	5430	61.56	7367	83.53
4	1453	16.47	8820	100.00

AGECAT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1 7-9	4966	60.76	4966	60.76
2 10	1748	21.39	6714	82.15
3 10+	1459	17.85	8173	100.00

Frequency Missing = 647

Appendix 1 DEMOGR2

GENDER

GENDER	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	4215	50.14	4215	50.14
2	4192	49.86	8407	100.00

Frequency Missing = 413

LQLANG

LQLANG	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	709	8.04	709	8.04
2	319	3.62	1028	11.66
3	91	1.03	1119	12.69
4	1289	14.61	2408	27.30
5	2682	30.41	5090	57.71
6	1358	15.40	6448	73.11
7	370	4.20	6818	77.30
8	969	10.99	7787	88.29
9	209	2.37	7996	90.66
10	400	4.54	8396	95.19
11	417	4.73	8813	99.92
12	7	0.08	8820	100.00

1 PROG1
14:12 Tuesday, March 20, 2007 5
PHILLIP TSHABALALA T06075 DPG9079
OD427014
DEMOGRAPHICS FOR CASES IN MODEL

The FREQ Procedure

LANGUAGE

LANGUAGE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	768	8.90	768	8.90
2	367	4.25		

Appendix 1 DEMOGR2

1135	13.15				
1237	14.33	3	102	1.18	
2514	29.12	4	1277	14.79	
5059	58.60	5	2545	29.48	
6420	74.37	6	1361	15.77	
6739	78.06	7	319	3.70	
7641	88.51	8	902	10.45	
7830	90.70	9	189	2.19	
8222	95.24	10	392	4.54	
8622	99.87	11	400	4.63	
8633	100.00	12	11	0.13	

Frequency Missing = 187

1 PROG1

14:12 Tuesday, March 20, 2007 6

PHILLIP TSHABALALA T06075 DPG9079 OD427014

SUMMARY STATISTICS FOR AGE - CASES IN MODEL

The MEANS Procedure

Analysis Variable : AGE

AGE

N	Mean	Median	Std Dev	Minimum	Maximum
8173	9.4642114	9.0000000	1.2543015	7.0000000	13.0000000

Description	Variable	Cramer's V
L language	LANGUAGE	0.23
L gender	GENDER	-0.04
LQ language	LQLANG	0.24
use calculators	LQ12I	0.06
use computers	LQ12J	0.09
worksheets work	LQ12L	0.08
place to write	LQ13A	0.08
L library	LQ13B	0.06
library frequency	LQ13C	0.03
library homework	LQ13D	0.04
read paper/mag	LQ16	0.05
numeracy workbook	LQ21	0.01
frequency workbook	LQ22	0.07
meal before	LQ25	0.07
bring food to school	LQ26	0.04
feeding scheme	LQ27	0.06
meal after	LQ28	0.07
P language	PARLANG	0.24
P radio	PAR7A	0.03
P tv	PAR7B	0.03
video	PAR7C	0.10
P computer	PAR7D	0.09
P school fees	PAR10A	0.01
books	PAR10C	0.01
stay away fees	PAR11	0.00
speak Afrikaans	PAR12A	0.02
speak English	PAR12B	0.06
speak Ndebele	PAR12C	0.09
speak Xhosa	PAR12D	0.09
speak Zulu	PAR12E	0.02
speak Sepedi	PAR12F	0.20
speak Sesotho	PAR12G	0.10
speak Setswana	PAR12H	0.20
speak Swati	PAR12I	0.07
speak Venda	PAR12J	0.05
speak Tsonga	PAR12K	0.12
speak Other	PAR12L	0.03
P education	PAR14	0.05
buy paper/mag	PAR16	0.07
number books	PAR17	0.10
writing material	PAR24A	0.02
text/workbooks	PAR24B	-0.01
feeding scheme	PAR24H	0.05
Pr qualification	PR2	0.07
building conditions	PR21	0.07
Pr radio	PR25E	0.02
Pr computer	PR25I	0.12
Pr tv	PR25J	0.08
ohp	PR25K	0.12
internet	PR25M	0.07
Pr school fees	PR51A	0.07
community levies	PR51B	0.02

govt grants	PR51C	0.00
donor funds	PR51D	0.02
NGOs	PR51E	-0.03
fundraising functions	PR51F	-0.01
annual school fees	PR52	0.14
% pay fees	PR53	0.03
partial exemption	PR54E	0.04
full exemption	PR54F	0.01
norms & stds	PR55	0.03
poverty ranking	PR56	0.07
section 21	PR57	0.15
E language	EDLANG	0.22
E age	ED1	0.08
E qualification<G12	ED2A	-0.17
E qualification>G12	ED2B	0.00
E highest qualification	ED3	0.08
highest prof training	ED4	0.08
highest maths level	ED5A	0.17
multigrade teach	ED16	0.05
share classroom	ED18	0.00
exercise books	ED19K	-0.02
classroom books	ED19M	0.00
Tsquare/metre stick	ED19N	-0.05
wall charts	ED19O	0.06
calculators	ED19Q	0.10
duplicator	ED20C	0.08
photocopier	ED20D	0.08
E computer	ED20E	0.07
typewriter	ED20F	-0.07
electricity	ED21	0.04
numeracy mat avail	ED23AMA	0.10
numeracy mat cond	ED23AGO	0.11
numeracy mat approp	ED23ACO	0.05
hlang books avail	ED23BMA	0.07
hlang books cond	ED23BGO	0.10
hlang books approp	ED23BCO	0.12
numeracy refs avail	ED23EMA	0.10
numeracy refs cond	ED23EGO	0.08
numeracy refs approp	ED23ECO	0.05
resource centre	ED25A	0.07
E library	ED25B	0.03
shortage	ED38C	0.06
physical facilities	ED38D	0.11
large classes	ED38E	0.04
leaners SES	ED38I	0.04
LOLT	ED38N	0.04
multicultural classes	ED38O	0.07
multigrade classes	ED38P	0.08
L age category	AGECAT	0.01
Urban/Rural	AREA	-0.10

Description	Variable	Cramer's V	MD	%MD
LQ language	LQLANG	**0.24**	1678	10.6
P language	PARLANG	**0.24**	1085	6.9
L language	LANGUAGE	**0.23**	414	2.6
E language	EDLANG	**0.22**	509	3.2
speak Setswana	PAR12H	**0.20**	8513	**53.9**
speak Sepedi	PAR12F	**0.20**	8396	**53.1**
highest maths level	ED5A	**0.17**	10991	**69.5**
section 21	PR57	**0.15**	1474	9.3
annual school fees	PR52	**0.14**	223	1.4
Pr computer	PR25I	**0.12**	148	0.9
ohp	PR25K	**0.12**	217	1.4
speak Tsonga	PAR12K	**0.12**	9105	**57.6**
hlang books approp	ED23BCO	**0.12**	7340	**46.4**
numeracy mat cond	ED23AGO	**0.11**	5199	**32.9**
physical facilities	ED38D	**0.11**	1235	7.8
num ref mat score JEMS		**0.1048**	699	4.4
num materials score JEMS		**0.1041**	654	4.1
number books	PAR17	**0.10**	2622	16.6
numeracy refs avail	ED23EMA	**0.10**	815	5.2
numeracy mat avail	ED23AMA	**0.10**	892	5.6
hlang books cond	ED23BGO	**0.10**	6007	**38.0**
hlang book score JEMS		**0.1001**	373	2.4
video	PAR7C	**0.10**	4798	**30.4**
calculators	ED19Q	**0.10**	794	5.0
speak Sesotho	PAR12G	**0.10**	8644	**54.7**
speak Xhosa	PAR12D	**0.09**	7970	**50.4**
P computer	PAR7D	**0.09**	5126	**32.4**
speak Ndebele	PAR12C	0.09	9213	**58.3**
use computers	LQ12J	0.09	3300	20.9
worksheets work	LQ12L	0.08	3031	19.2
place to write	LQ13A	0.08	2464	15.6
Pr tv	PR25J	0.08	187	1.2
sources of funding JEMS		0.0822	3064	19.4
numeracy refs cond	ED23EGO	0.08	6921	**43.8**
E highest qualification	ED3	0.08	863	5.5
E qualification JEMS		0.0804	240	1.5
multigrade classes	ED38P	0.08	921	5.8
E age	ED1	0.08	315	2.0
duplicator	ED20C	0.08	1205	7.6
photocopier	ED20D	0.08	601	3.8
highest prof training	ED4	0.08	668	4.2
building conditions	PR21	0.07	297	1.9
multicultural classes	ED38O	0.07	745	4.7
Pr qualification	PR2	0.07	801	5.1
internet	PR25M	0.07	187	1.2
hlang books avail	ED23BMA	0.07	526	3.3
resource centre	ED25A	0.07	1718	10.9
meal after	LQ28	0.07	2827	17.9
poverty ranking	PR56	0.07	6997	**44.3**
frequency workbook	LQ22	0.07	2537	16.1
food count JEMS		0.0684	4426	28.0
Pr school fees	PR51A	0.07	110	0.7

meal before	LQ25	0.07	2671	16.9
E computer	ED20E	0.07	1009	6.4
buy paper/mag	PAR16	0.07	1205	7.6
speak Swati	PAR12I	0.07	9151	**57.9**
L library	LQ13B	0.06	2487	15.7
wall charts	ED19O	0.06	704	4.5
use calculators	LQ12I	0.06	3252	20.6
shortage	ED38C	0.06	353	2.2
feeding scheme	LQ27	0.06	2244	14.2
speak English	PAR12B	0.06	6341	**40.1**
feeding scheme	PAR24H	0.05	2498	15.8
multigrade teach	ED16	0.05	187	1.2
numeracy mat approp	ED23ACO	0.05	6589	**41.7**
read paper/mag	LQ16	0.05	2063	13.1
numeracy refs approp	ED23ECO	0.05	7842	**49.6**
speak Venda	PAR12J	0.05	9286	**58.8**
P education	PAR14	0.05	2972	18.8
LOLT	ED38N	0.04	902	5.7
large classes	ED38E	0.04	626	4.0
leaners SES	ED38I	0.04	1104	7.0
library homework	LQ13D	0.04	2628	16.6
partial exemption	PR54E	0.04	886	5.6
electricity	ED21	0.04	126	0.8
bring food to school	LQ26	0.04	2314	14.6
speak Other	PAR12L	0.03	9726	**61.5**
% pay fees	PR53	0.03	142	0.9
P radio	PAR7A	0.03	1734	11.0
E library	ED25B	0.03	1147	7.3
library frequency	LQ13C	0.03	2462	15.6
norms & stds	PR55	0.03	503	3.2
P tv	PAR7B	0.03	3375	21.4
community levies	PR51B	0.02	2308	14.6
speak Afrikaans	PAR12A	0.02	7533	**47.7**
writing material	PAR24A	0.02	2188	13.8
speak Zulu	PAR12E	0.02	6291	**39.8**
donor funds	PR51D	0.02	2209	14.0
Pr radio	PR25E	0.02	384	2.4
L age category	AGECAT	0.01	1398	8.8
books	PAR10C	0.01	4689	29.7
P school fees	PAR10A	0.01	1165	7.4
full exemption	PR54F	0.01	1080	6.8
numeracy workbook	LQ21	0.01	2181	13.8
classroom books	ED19M	0.00	708	4.5
govt grants	PR51C	0.00	1773	11.2
stay away fees	PAR11	0.00	2367	15.0
share classroom	ED18	0.00	148	0.9
E qualification>G12	ED2B	0.00	1308	8.3
text/workbooks	PAR24B	-0.01	2346	14.8
fundraising functions	PR51F	-0.01	883	5.6
exercise books	ED19K	-0.02	391	2.5
NGOs	PR51E	-0.03	2550	16.1
L gender	GENDER	-0.04	977	6.2
Tsquare/metre stick	ED19N	-0.05	486	3.1

typewriter	ED20F	-0.07	964	6.1
Urban/Rural	AREA	**-0.10**	62	0.4
E qualification<G12	ED2A	**-0.17**	12645	80.0

Appendix 3c MODEL3

1 PROG1

09:00 Monday, March 5, 2007 1

PHILLIP TSHABALALA T06075 DPG9079

OD427014

THIRD STEPWISE MODEL - OMIT VARIABLES WITH HIGH MD

The LOGISTIC Procedure

Model Information

Data Set	DTA.D10
Response Variable	NUMERACY
NUMERACY SCORE ABOVE OR BELOW MEDIAN	
Number of Response Levels	2
Number of Observations	8833
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	NUMERACY	Total Frequency
1	ABOVE MEDIAN	5011
2	BELOW MEDIAN	3822

Probability modeled is NUMERACY='ABOVE MEDIAN'.

NOTE: 6972 observations were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Class Level Information

Class	Value	Design Variables					
		1	2	3	4	5	6
LQLANG	0	1	0	0	0	0	0
	AFRIKAANS	0	1	0	0	0	0
	ENGLISH	0	0	1	0	0	0
	ISINDEBELE	0	0	0	1	0	0
	ISIZULU	0	0	0	0	1	0
	ISXHOSA	0	0	0	0	0	1
	OTHER	0	0	0	0	0	0
	SEPEDI	0	0	0	0	0	0
	SESOTHO	0	0	0	0	0	0

Appendix 3c MODEL3

	0	0	1	0	0	0						
				SETSWANA			0	0	0	0	0	0
	0	0	0	1	0	0						
				SISWATI			0	0	0	0	0	0
	0	0	0	0	1	0						
				TSHIVENDA			0	0	0	0	0	0
	0	0	0	0	0	1						
				XITSONGA			-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1						
			PR57	1	NO		1	0				
				2	PARTLY		0	1				
				3	FULL		-1	-1				
			PR52	1	<R50		1	0	0	0	0	0
	0	0										
				2	50-100		0	1	0	0	0	0
	0	0										
				3	101-200		0	0	1	0	0	0
	0	0										
				4	201-300		0	0	0	1	0	0
	0	0										

The LOGISTIC Procedure

Class Level Information

Class	Value	Design Variables					
		1	2	3	4	5	6
	5 301-400	0	0	0	0	1	0
	6 401-500	0	0	0	0	0	1
	7 501-1000	0	0	0	0	0	0
	8 1001-2000	0	0	0	0	0	0
	9 >R2000	-1	-1	-1	-1	-1	-1
PR25I	1 NOT AVAIL	1	0				
	2 SOMETIMES	0	1				
	3 ALWAYS	-1	-1				
ED38D	1 NOT AT ALL	1	0	0			
	2 A LITTLE	0	1	0			
	3 QUITE ALOT	0	0	1			
	4 GREAT DEAL	-1	-1	-1			
PAR17	1 <10	1	0				
	2 10-20	0	1				
	3 >20	-1	-1				
ED23EMa	NO	1					
	YES	-1					
ED19Q	NO	1					
	YES	-1					

Appendix 3c MODEL3

AREA	RURAL	1
	URBAN	-1

Step 0. Intercept entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
984.7758	32	<.0001

Step 1. Effect LQLANG entered:

The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12086.601	11592.073
SC	12093.688	11684.194
-2 Log L	12084.601	11566.073

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	518.5283	12	<.0001
Score	492.9044	12	<.0001

	Wald	443.4420	12

<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
540.2997	20	<.0001

Step 2. Effect PR57 entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12086.601	11367.048
SC	12093.688	11473.341
-2 Log L	12084.601	11337.048

THIRD STEPWISE MODEL - OMIT VARIABLES WITH HIGH MD

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	747.5536	14	<.0001
Score	701.5075	14	<.0001
Wald	626.5803	14	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
342.2627	18	

<.0001

Step 3. Effect PR25I entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12086.601	11263.365
SC	12093.688	11383.832
-2 Log L	12084.601	11229.365

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	855.2360	16	<.0001
Score	778.1702	16	<.0001
Wald	664.5867	16	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
248.8451	16	<.0001

Step 4. Effect ED38D entered:

The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Appendix 3c MODEL3

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12086.601	11186.031
SC	12093.688	11327.756
-2 Log L	12084.601	11146.031

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	938.5702	19	<.0001
Score	856.0418	19	<.0001
Wald	733.6553	19	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
167.4244	13	<.0001

Step 5. Effect PR52 entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12086.601	11127.243
SC	12093.688	11325.658
-2 Log L	12084.601	11071.243

Appendix 3c MODEL3
PHILLIP TSHABALALA T06075 DPG9079

OD427014

THIRD STEPWISE MODEL - OMIT VARIABLES

WITH HIGH MD

The LOGISTIC Procedure

Testing Global Null Hypothesis:

BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1013.3587	27	<.0001
Score	905.0401	27	<.0001
Wald	756.4824	27	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
93.6961	5	<.0001

Step 6. Effect ED19Q entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12086.601	11081.906
SC	12093.688	11287.407
-2 Log L	12084.601	11023.906

Testing Global Null Hypothesis:

BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1060.6954	28	<.0001
Score	945.3755	28	<.0001
Wald	786.3736	28	<.0001

Appendix 3c MODEL3

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
46.7901	4	<.0001

Step 7. Effect ED23EMa1 entered:

THIRD STEPWISE MODEL - OMIT VARIABLES WITH HIGH MD

The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12086.601	11056.474
SC	12093.688	11269.062
-2 Log L	12084.601	10996.474

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1088.1269	29	<.0001
Score	968.1967	29	<.0001
Wald	803.5299	29	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
19.3136	3	0.0002

Step 8. Effect PAR17 entered:

Appendix 3c MODEL3

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12086.601	11041.141
SC	12093.688	11267.901
-2 Log L	12084.601	10977.141

1 09:00 Monday, March 5, 2007 8 PROG1
PHILLIP TSHABALALA T06075 DPG9079 OD427014
THIRD STEPWISE MODEL - OMIT VARIABLES WITH HIGH MD

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1107.4603	31	<.0001
Score	984.7733	31	<.0001
Wald	815.3291	31	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
0.0700	1	0.7913

NOTE: No (additional) effects met the 0.05 significance level for entry into the model.

Summary of Stepwise Selection

Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq	Variable Label

Appendix 3c MODEL3

	1	LQLANG	12	1	492.9044	.
<.0001	LEARNERS HOME LANGUAGE					
	2	PR57	2	2	223.5905	.
<.0001	SECTION 21 STATUS OF SCHOOL					
	3	PR25I	2	3	101.6524	.
<.0001	AVAILABILITY OF COMPUTER FOR TEACHING					
	4	ED38D	3	4	82.8945	.
<.0001	INADEQUATE PHYSICAL FACILITIES					
	5	PR52	8	5	73.1190	.
<.0001	ANNUAL SCHOOL FEES					
	6	ED19Q	1	6	47.1986	.
<.0001	CALCULATORS ETC. IN CLASSROOM					
	7	ED23EMa	1	7	27.4955	.
<.0001	NUMERACY REFERENCE MATERIALS AVAILABLE					
	8	PAR17	2	8	19.2427	.
<.0001	NUMBER OF BOOKS AT HOME					

Type III Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
LQLANG	12	271.4118	<.0001
PR57	2	114.1179	<.0001
PR52	8	65.0749	<.0001
PR25I	2	58.8518	<.0001
ED38D	3	78.9079	<.0001
PAR17	2	19.1755	<.0001
ED23EMa	1	27.0035	<.0001
ED19Q	1	32.8770	<.0001

THIRD STEPWISE MODEL - OMIT VARIABLES WITH HIGH MD

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	1.1828	0.1386	72.7877	<.0001
LQLANG	0	1	1.5995	0.7164	4.9843	0.0256
LQLANG	AFRIKAANS	1	-0.4897	0.1304	14.1074	0.0002
LQLANG	ENGLISH	1	1.0957	0.2152	25.9206	<.0001

Appendix 3c MODEL3

LQLANG	ISINDEBELE	1	-1.3873	0.2644	27.5292 <.0001
LQLANG	ISIZULU	1	0.1069	0.1016	1.1064 0.2929
LQLANG	ISXHOSA	1	0.5018	0.1107	20.5468 <.0001
LQLANG	OTHER	1	-0.5936	0.7703	0.5938 0.4409
LQLANG	SEPEDI	1	-0.2939	0.1081	7.3973 0.0065
LQLANG	SESOTHO	1	0.000640	0.1384	0.0000 0.9963
LQLANG	SETSWANA	1	-0.6129	0.1144	28.7192 <.0001
LQLANG	SISWATI	1	0.2629	0.1627	2.6090 0.1063
LQLANG	TSHIVENDA	1	0.1442	0.1357	1.1293 0.2879
PR57	1 NO	1	0.0753	0.0390	3.7257 0.0536
PR57	2 PARTLY	1	-0.5189	0.0487	113.3903 <.0001
PR52	1 <R50	1	-0.1364	0.0638	4.5695 0.0325
PR52	2 50-100	1	-0.1176	0.0704	2.7921 0.0947
PR52	3 101-200	1	0.0701	0.1495	0.2199 0.6391
PR52	4 201-300	1	-0.8482	0.1551	29.9083 <.0001
PR52	5 301-400	1	0.3586	0.1889	3.6039 0.0576
PR52	6 401-500	1	-0.4297	0.1994	4.6426 0.0312
PR52	7 501-1000	1	0.8608	0.1968	19.1324 <.0001
PR52	8 1001-2000	1	0.1144	0.1314	0.7569 0.3843
PR25I	1 NOT AVAIL	1	-0.7734	0.1143	45.8006 <.0001
PR25I	2 SOMETIMES	1	0.3794	0.1769	4.5998 0.0320
ED38D	1 NOT AT ALL	1	-0.1064	0.0516	4.2462 0.0393
ED38D	2 A LITTLE	1	0.3313	0.0412	64.8312 <.0001
ED38D	3 QUITE ALOT	1	-0.2044	0.0380	28.9815 <.0001
PAR17	1 <10	1	-0.1438	0.0346	17.2506 <.0001
PAR17	2 10-20	1	-0.0251	0.0428	0.3456 0.5566
ED23EMa	NO	1	-0.1280	0.0246	27.0035 <.0001
ED19Q	NO	1	-0.1535	0.0268	32.8770 <.0001

Odds Ratio Estimates

Effect			Point Estimate	95% Wald Confidence Limits
LQLANG	0	vs XITSONGA	6.916	1.504 31.797
LQLANG	AFRIKAANS	vs XITSONGA	0.856	

		LQLANG	ENGLISH vs XITSONGA	4.179
0.646	1.135			
		LQLANG	ISINDEBELE vs XITSONGA	0.349
2.630	6.641			
		LQLANG	ISIZULU vs XITSONGA	1.555
0.199	0.612			
		LQLANG	ISXHOSA vs XITSONGA	2.307
1.251	1.932			
		LQLANG	OTHER vs XITSONGA	0.772
1.824	2.919			
		LQLANG	SEPEDI vs XITSONGA	1.041
0.149	3.987			
		LQLANG	SESOTHO vs XITSONGA	1.398
0.829	1.308			
		LQLANG	SETSWANA vs XITSONGA	0.757
1.042	1.875			
0.592	0.968			

The LOGISTIC Procedure

Odds Ratio Estimates

Effect					Point Estimate	95% Wald Confidence Limits	
LQLANG	SISWATI	vs	XITSONGA		1.817	1.288	2.563
LQLANG	TSHIVENDA	vs	XITSONGA		1.614	1.211	2.150
PR57	1 NO	vs	3 FULL		0.692	0.583	0.822
PR57	2 PARTLY	vs	3 FULL		0.382	0.313	0.466
PR52	1 <R50	vs	9 >R2000		0.768	0.663	0.889
PR52	2 50-100	vs	9 >R2000		0.782	0.665	0.920
PR52	3 101-200	vs	9 >R2000		0.944	0.670	1.330
PR52	4 201-300	vs	9 >R2000		0.377	0.262	0.541
PR52	5 301-400	vs	9 >R2000		1.259	0.820	1.934
PR52	6 401-500	vs	9 >R2000		0.573	0.364	0.899
PR52	7 501-1000	vs	9 >R2000		2.081	1.336	3.241
PR52	8 1001-2000	vs	9 >R2000		0.986	0.730	1.333
PR25I	1 NOT AVAIL	vs	3 ALWAYS		0.311	0.225	0.429
PR25I	2 SOMETIMES	vs	3 ALWAYS		0.985	0.563	1.726
ED38D	1 NOT AT ALL	vs	4 GREAT DEAL		0.918	0.787	1.070
ED38D	2 A LITTLE	vs	4 GREAT DEAL		1.422	1.252	1.614
ED38D	3 QUITE ALOT	vs	4 GREAT DEAL		0.832	0.741	0.935
PAR17	1 <10	vs	3 >20		0.731		

Appendix 3c MODEL3

PAR17	2 10-20 vs 3 >20		0.824
ED23EMa	NO vs YES		0.774
ED19Q	NO vs YES		0.736

0.632	0.846		
0.696	0.975		
0.703	0.853		
0.662	0.817		

Association of Predicted Probabilities and Observed Responses

Percent Concordant	68.5		
Percent Discordant	31.0	Gamma	0.375
Percent Tied	0.5	Tau-a	0.376
Pairs	19152042	c	0.184
			0.687

Appendix 3d final MODEL4

1 PROG1

07:41 Tuesday, March 20, 2007 1

PHILLIP TSHABALALA T06075 DPG9079 OD427014

FOURTH STEPWISE MODEL - OMIT HIGH MD, LANG0, PR5 IN

The LOGISTIC Procedure

Model Information

Data Set	DTA.D10
Response Variable	NUMERACY NUMERACY SCORE ABOVE OR BELOW MEDIAN
Number of Response Levels	2
Number of Observations	8820
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	NUMERACY	Total Frequency
1	ABOVE MEDIAN	5000
2	BELOW MEDIAN	3820

Probability modeled is NUMERACY='ABOVE MEDIAN'.

NOTE: 6985 observations were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Class Level Information

Class	Value	Design Variables									
		1	2	3	4	5	6	7	8	9	10	11
LQLANG	AFRIKAANS	1	0	0	0	0	0	0	0	0	0	0
	ENGLISH	0	1	0	0	0	0	0	0	0	0	0
	ISINDEBELE	0	0	1	0	0	0	0	0	0	0	0
	ISIZULU	0	0	0	1	0	0	0	0	0	0	0
	ISXHOSA	0	0	0	0	1	0	0	0	0	0	0
	OTHER	0	0	0	0	0	1	0	0	0	0	0
	SEPEDI	0	0	0	0	0	0	1	0	0	0	0
	SESOTHO	0	0	0	0	0	0	0	1	0	0	0
	SETSWANA	0	0	0	0	0	0	0	0	1	0	0

Appendix 3d final MODEL4

0	0	0	1	0	0	SISWATI	0	0	0	0	0
0	0	0	0	1	0	TSHIVENDA	0	0	0	0	0
0	0	0	0	0	1	XITSONGA	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1						
				PR57	1	NO	1	0			
					2	PARTLY	0	1			
					3	FULL	-1	-1			
				PR52	1	<R50	1	0	0	0	0
0	0	0									
					2	50-100	0	1	0	0	0
0	0	0									
					3	101-200	0	0	1	0	0
0	0	0									
					4	201-300	0	0	0	1	0
0	0	0									
					5	301-400	0	0	0	0	1
0	0	0									

1 PROG1

07:41 Tuesday, March 20, 2007 2

PHILLIP TSHABALALA T06075 DPG9079 OD427014

FOURTH STEPWISE MODEL - OMIT HIGH MD, LANG0, PR5 IN

The LOGISTIC Procedure

Class Level Information

Class	Value	Design Variables									
		1	2	3	4	5	6	7	8	9	10	11
	6 401-500	0	0	0	0	0	1	0	0			
	7 501-1000	0	0	0	0	0	0	1	0			
	8 1001-2000	0	0	0	0	0	0	0	1			
	9 >R2000	-1	-1	-1	-1	-1	-1	-1	-1			
PR25I	1 NOT AVAIL	1	0									
	2 SOMETIMES	0	1									
	3 ALWAYS	-1	-1									
ED38D	1 NOT AT ALL	1	0	0								
	2 A LITTLE	0	1	0								
	3 QUITE ALOT	0	0	1								
	4 GREAT DEAL	-1	-1	-1								
PAR17	1 <10	1	0									
	2 10-20	0	1									
	3 >20	-1	-1									
ED23EMa	NO	1										
	YES	-1										
ED19Q	NO	1										
	YES	-1										
PR25K	1	1	0									

Appendix 3d final MODEL4

2		0	1	
3		-1	-1	
PR5	1 URBAN	1	0	0
	2 SEMI-URBAN	0	1	0
	3 SEMI-RURAL	0	0	1
	4 REMOTE RUR	-1	-1	-1

Step 0. Intercept entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
1022.1330	35	<.0001

FOURTH STEPWISE MODEL - OMIT HIGH MD, LANG0, PR5 IN

The LOGISTIC Procedure

Step 1. Effect LQLANG entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	11578.911
SC	12077.858	11663.928
-2 Log L	12068.773	11554.911

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	513.8629	11	<.0001

	Score	488.6706	11
<.0001			
	Wald	439.8522	11
<.0001			

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
583.3130	24	<.0001

Step 2. Effect PR57 entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	11353.486
SC	12077.858	11452.673
-2 Log L	12068.773	11325.486

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	743.2876	13	<.0001
Score	697.6605	13	<.0001
Wald	623.3346	13	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
385.5841	22	

Appendix 3d final MODEL4

<.0001

Step 3. Effect PR25I entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	11249.834
SC	12077.858	11363.190
-2 Log L	12068.773	11217.834

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	850.9397	15	<.0001
Score	774.2784	15	<.0001
Wald	661.2323	15	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
293.5863	20	<.0001

Step 4. Effect ED38D entered:

The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Appendix 3d final MODEL4

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	11172.701
SC	12077.858	11307.312
-2 Log L	12068.773	11134.701

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	934.0722	18	<.0001
Score	851.9785	18	<.0001
Wald	730.1684	18	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
213.0682	17	<.0001

Step 5. Effect PR52 entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	11113.665
SC	12077.858	11304.954
-2 Log L	12068.773	11059.665

Appendix 3d final MODEL4

PHILLIP TSHABALALA T06075 DPG9079

OD427014

FOURTH STEPWISE MODEL - OMIT HIGH MD, LANG0, PR5 IN

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1009.1084	26	<.0001
Score	901.1841	26	<.0001
Wald	753.1383	26	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
138.7387	9	<.0001

Step 6. Effect ED19Q entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	11068.470
SC	12077.858	11266.844
-2 Log L	12068.773	11012.470

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1056.3030	27	<.0001
Score	941.3652	27	<.0001
Wald	782.8691	27	<.0001

Appendix 3d final MODEL4

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
92.1837	8	<.0001

Step 7. Effect PR25K entered:

The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	11034.452
SC	12077.858	11246.995
-2 Log L	12068.773	10974.452

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1094.3213	29	<.0001
Score	972.8173	29	<.0001
Wald	806.2103	29	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
54.7943	6	<.0001

Step 8. Effect ED23EMa entered:

Appendix 3d final MODEL4

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	11009.299
SC	12077.858	11228.927
-2 Log L	12068.773	10947.299

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1121.4743	30	<.0001
Score	995.5597	30	<.0001
Wald	823.8146	30	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
27.6743	5	<.0001

Step 9. Effect PAR17 entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Appendix 3d final MODEL4

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	10995.168
SC	12077.858	11228.965
-2 Log L	12068.773	10929.168

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1139.6058	32	<.0001
Score	1010.9817	32	<.0001
Wald	834.3707	32	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
9.6574	3	0.0217

Step 10. Effect PR5 entered:

The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	12070.773	10991.499
SC	12077.858	11246.551
-2 Log L	12068.773	

Appendix 3d final MODEL4

10919.499

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1149.2740	35	<.0001
Score	1022.1330	35	<.0001
Wald	846.1209	35	<.0001

NOTE: All effects have been entered into the model.

Summary of Stepwise Selection

Step	Effect Entered	Removed	Number DF	In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq	Variable Label
1	LQLANG		11	1	488.6706	.	<.0001	LEARNERS HOME LANGUAGE
2	PR57		2	2	223.9903	.	<.0001	SECTION 21 STATUS OF SCHOOL
3	PR25I		2	3	101.6240	.	<.0001	AVAILABILITY OF COMPUTER FOR TEACHING
4	ED38D		3	4	82.6980	.	<.0001	INADEQUATE PHYSICAL FACILITIES
5	PR52		8	5	73.3627	.	<.0001	ANNUAL SCHOOL FEES
6	ED19Q		1	6	47.0586	.	<.0001	CALCULATORS ETC. IN CLASSROOM
7	PR25K		2	7	37.6896	.	<.0001	AVAILABILITY OF OHP FOR TEACHING
8	ED23EMa		1	8	27.2090	.	<.0001	NUMERACY REFERENCE MATERIALS AVAILABLE
9	PAR17		2	9	18.0486	.	0.0001	NUMBER OF BOOKS AT HOME
10	PR5		3	10	9.6574	.	0.0217	SCHOOL LOCATION

Type III Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
LQLANG	11	269.0973	<.0001
PR57	2	109.8299	<.0001
PR52	8	68.7290	<.0001
PR25I	2	54.5578	<.0001
ED38D	3	65.2425	<.0001

Appendix 3d final MODEL4

OD427014

FOURTH STEPWISE MODEL - OMIT HIGH MD, LANG0, PR5 IN

The LOGISTIC Procedure

Type III Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
PAR17	2	19.2386	<.0001
ED23EMa	1	29.3647	<.0001
ED19Q	1	25.1771	<.0001
PR5	3	9.6427	0.0219
PR25K	2	38.3381	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	1.1268	0.1332	71.5818	<.0001
LQLANG	AFRIKAANS	1	-0.5120	0.1234	17.2123	<.0001
LQLANG	ENGLISH	1	1.0822	0.2106	26.4093	<.0001
LQLANG	ISINDEBELE	1	-1.2128	0.2590	21.9312	<.0001
LQLANG	ISIZULU	1	0.2872	0.0893	10.3353	0.0013
LQLANG	ISXHOSA	1	0.6694	0.0991	45.6666	<.0001
LQLANG	OTHER	1	-0.5561	0.7676	0.5249	0.4688
LQLANG	SEPEDI	1	-0.0725	0.0973	0.5557	0.4560
LQLANG	SESOTHO	1	0.0850	0.1313	0.4191	0.5174
LQLANG	SETSWANA	1	-0.4551	0.1029	19.5743	<.0001
LQLANG	SISWATI	1	0.4143	0.1564	7.0165	0.0081
LQLANG	TSHIVENDA	1	0.3677	0.1277	8.2904	0.0040
PR57	1 NO	1	0.0993	0.0397	6.2651	0.0123
PR57	2 PARTLY	1	-0.5108	0.0492	107.6378	<.0001
PR52	1 <R50	1	-0.0832	0.0659	1.5954	0.2066
PR52	2 50-100	1	-0.1133	0.0736	2.3701	0.1237
PR52	3 101-200	1	-0.00918	0.1565	0.0034	0.9532
PR52	4 201-300	1	-0.9088	0.1594	32.4972	<.0001

PR52	5	301-400	1	0.4677	0.1942	5.7996	0.0160
PR52	6	401-500	1	-0.4822	0.2022	5.6861	0.0171
PR52	7	501-1000	1	0.7963	0.1963	16.4455	<.0001
PR52	8	1001-2000	1	0.1362	0.1320	1.0645	0.3022
PR25I	1	NOT AVAIL	1	-0.7675	0.1173	42.8278	<.0001
PR25I	2	SOMETIMES	1	0.3635	0.1777	4.1846	0.0408
ED38D	1	NOT AT ALL	1	-0.1174	0.0525	4.9973	0.0254
ED38D	2	A LITTLE	1	0.3149	0.0419	56.5056	<.0001
ED38D	3	QUITE ALOT	1	-0.1756	0.0392	20.1060	<.0001
PAR17	1	<10	1	-0.1442	0.0348	17.1813	<.0001
PAR17	2	10-20	1	-0.0273	0.0429	0.4053	0.5244
ED23EMa		NO	1	-0.1352	0.0249	29.3647	<.0001
ED19Q		NO	1	-0.1376	0.0274	25.1771	<.0001
PR5	1	URBAN	1	-0.1489	0.0694	4.5998	0.0320
PR5	2	SEMI-URBAN	1	0.0508	0.0752	0.4567	0.4991
PR5	3	SEMI-RURAL	1	0.1060	0.0446	5.6353	0.0176
PR25K	1		1	-0.3264	0.0577	31.9860	<.0001
PR25K	2		1	0.1618	0.0845	3.6612	0.0557

The LOGISTIC Procedure

Odds Ratio Estimates

Effect				Point Estimate	95% Wald Confidence Limits	
LQLANG	AFRIKAANS	vs	XITSONGA	0.661	0.491	0.889
LQLANG	ENGLISH	vs	XITSONGA	3.253	2.023	5.232
LQLANG	ISINDEBELE	vs	XITSONGA	0.328	0.186	0.578
LQLANG	ISIZULU	vs	XITSONGA	1.469	1.179	1.831
LQLANG	ISXHOSA	vs	XITSONGA	2.153	1.697	2.731
LQLANG	OTHER	vs	XITSONGA	0.632	0.121	3.304
LQLANG	SEPEDI	vs	XITSONGA	1.025	0.815	1.289
LQLANG	SESOTHO	vs	XITSONGA	1.200	0.883	1.631

			0.544 0.898
			1.177 2.364
			1.190 2.130
			0.615 0.871
			0.325 0.486
			0.652 0.876
			0.621 0.867
			0.565 1.172
			0.227 0.482
			0.846 2.033
			0.321 0.802
			1.168 2.841
			0.695 1.275
			0.222 0.433
			0.547 1.686
			0.779 1.061
			1.232 1.592
			0.762 0.965
			0.630 0.844
			0.692 0.971
			0.692 0.841
			0.682 0.846
			0.702 1.075
			0.846 1.329
			0.986 1.274
			0.517 0.724
			0.762 1.305

LQLANG	SETSWANA	vs	XITSONGA	0.699
LQLANG	SISWATI	vs	XITSONGA	1.668
LQLANG	TSHIVENDA	vs	XITSONGA	1.592
PR57	1 NO	vs	3 FULL	0.732
PR57	2 PARTLY	vs	3 FULL	0.398
PR52	1 <R50	vs	9 >R2000	0.756
PR52	2 50-100	vs	9 >R2000	0.734
PR52	3 101-200	vs	9 >R2000	0.814
PR52	4 201-300	vs	9 >R2000	0.331
PR52	5 301-400	vs	9 >R2000	1.311
PR52	6 401-500	vs	9 >R2000	0.507
PR52	7 501-1000	vs	9 >R2000	1.822
PR52	8 1001-2000	vs	9 >R2000	0.941
PR25I	1 NOT AVAIL	vs	3 ALWAYS	0.310
PR25I	2 SOMETIMES	vs	3 ALWAYS	0.960
ED38D	1 NOT AT ALL	vs	4 GREAT DEAL	0.909
ED38D	2 A LITTLE	vs	4 GREAT DEAL	1.400
ED38D	3 QUITE ALOT	vs	4 GREAT DEAL	0.857
PAR17	1 <10	vs	3 >20	0.729
PAR17	2 10-20	vs	3 >20	0.820
ED23EMa	NO	vs	YES	0.763
ED19Q	NO	vs	YES	0.759
PR5	1 URBAN	vs	4 REMOTE RUR	0.869
PR5	2 SEMI-URBAN	vs	4 REMOTE RUR	1.060
PR5	3 SEMI-RURAL	vs	4 REMOTE RUR	1.121
PR25K	1	vs	3	0.612
PR25K	2	vs	3	0.997

Association of Predicted Probabilities and Observed Responses

Percent Concordant	69.2	
Percent Discordant	30.4	Gamma
Percent Tied	0.5	Tau-a
Pairs	19100000	c
Somers' D	0.388	
	0.390	
	0.191	
	0.694	
1		

Appendix 3d final MODEL4

Appendix 4 CROSSTAB

1 PROG1C
13:17 Monday, April 23, 2007 1
PHILLIP TSHABALALA T06075 DPG9079
OD427014

CROSSTABS FOR LEARNERS IN MODEL - BY SCHOOL LOCATION

The FREQ Procedure

Table of AGECAT by PR5

AGECAT	PR5(SCHOOL LOCATION)			
Frequency| Row Pct | Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE RUR
1 7-9	882	532	2870	682
	17.76	10.71	57.79	13.73
	77.92	73.48	57.68	50.86
2 10	160	115	1101	372
	9.15	6.58	62.99	21.28
	14.13	15.88	22.13	27.74
3 10+	90	77	1005	287
	6.17	5.28	68.88	19.67
	7.95	10.64	20.20	21.40
Total	1132	724	4976	1341

	Total
1 7-9	4966
	13.73
	50.86
2 10	1748
	21.28
	27.74
3 10+	1459
	19.67
	21.40
Total	8173

Frequency Missing = 647

Table of GENDER by PR5

GENDER(GENDER)	PR5(SCHOOL LOCATION)			
Frequency| Row Pct | Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE RUR
MALE	595	329	2561	730
	14.12	7.81	60.76	17.32
	52.28	45.69	49.64	

	Total
MALE	4215

52.52 |

---------+--------+--------+--------+--------+

FEMALE | 543 | 391 | 2598 |
660 | 4192

| 12.95 | 9.33 | 61.98 |
15.74 |

| 47.72 | 54.31 | 50.36 |
47.48 |

---------+--------+--------+--------+--------+

Total 1138 720 5159
1390 8407

Frequency Missing = 413

CROSSTABS FOR LEARNERS IN MODEL - BY SCHOOL LOCATION

The FREQ Procedure

Table of LQLANG by PR5

LQLANG(LEARNERS HOME LANGUAGE)
PR5(SCHOOL LOCATION)

| Frequency |
| Row Pct |
| Col Pct |1 URBAN |2 SEMI-U|3 SEMI-R|4 REMOTE| Total
| |RBAN |URAL | RUR |

---------+--------+--------+--------+--------+

AFRIKAANS | 291 | 135 | 275 |
8 | 709

| 41.04 | 19.04 | 38.79 |
1.13 |

| 24.81 | 17.67 | 5.06 |
0.55 |

---------+--------+--------+--------+--------+

ENGLISH | 194 | 62 | 47 |
16 | 319

| 60.82 | 19.44 | 14.73 |
5.02 |

| 16.54 | 8.12 | 0.87 |
1.10 |

---------+--------+--------+--------+--------+

ISINDEBELE | 3 | 20 | 64 |
4 | 91

| 3.30 | 21.98 | 70.33 |
4.40 |

| 0.26 | 2.62 | 1.18 |
0.28 |

---------+--------+--------+--------+--------+

ISXHOSA | 133 | 77 | 827 |
252 | 1289

| 10.32 | 5.97 | 64.16 |
19.55 |

| 11.34 | 10.08 | 15.23 |

17.34 |

ISIZULU	150	184	1940	408	2682
	5.59	6.86	72.33	15.21	
	12.79	24.08	35.73	28.08	
SEPEDI	63	67	933	295	1358
	4.64	4.93	68.70	21.72	
	5.37	8.77	17.18	20.30	
Total	1173	764	5430	1453	8820

(Continued)

CROSSTABS FOR LEARNERS IN MODEL - BY SCHOOL LOCATION

The FREQ Procedure

Table of LQLANG by PR5

LQLANG(LEARNERS HOME LANGUAGE) PR5(SCHOOL LOCATION)

Frequency / Row Pct / Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE	Total
SESOTHO	159	17	180	14	370
	42.97	4.59	48.65	3.78	
	13.55	2.23	3.31	0.96	
SETSWANA	140	71	612	146	969
	14.45	7.33	63.16	15.07	
	11.94	9.29	11.27	10.05	
SISWATI	8	59	81	61	209
	3.83	28.23	38.76	29.19	
	0.68	7.72	1.49	4.20	

Appendix 4 CROSSTAB

TSHIVENDA	2	63	268	67	400
	0.50	15.75	67.00	16.75	
	0.17	8.25	4.94	4.61	
XITSONGA	27	8	201	181	417
	6.47	1.92	48.20	43.41	
	2.30	1.05	3.70	12.46	
OTHER	3	1	2	1	7
	42.86	14.29	28.57	14.29	
	0.26	0.13	0.04	0.07	
Total	1173	764	5430	1453	8820

CROSSTABS FOR LEARNERS IN MODEL - BY SCHOOL LOCATION

The FREQ Procedure

Table of NUMERACY by PR5

NUMERACY(NUMERACY SCORE ABOVE OR BELOW MEDIAN)

PR5(SCHOOL LOCATION)

Frequency / Row Pct / Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE	Total
ABOVE MEDIAN	779	493	2984	744	5000
	15.58	9.86	59.68	14.88	
	66.41	64.53	54.95	51.20	
BELOW MEDIAN	394	271	2446	709	3820
	10.31	7.09	64.03	18.56	
	33.59	35.47	45.05	48.80	
Total	1173	764	5430	1453	8820

Appendix 4 CROSSTAB

Table of PR57 by PR5

PR57(SECTION 21 STATUS OF SCHOOL)

PR5(SCHOOL LOCATION)

Frequency Row Pct Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE	Total
1 NO	601 9.61 51.24	479 7.66 62.70	3946 63.13 72.67	1225 19.60 84.31	6251
2 PARTLY	206 14.54 17.56	127 8.96 16.62	887 62.60 16.34	197 13.90 13.56	1417
3 FULL	366 31.77 31.20	158 13.72 20.68	597 51.82 10.99	31 2.69 2.13	1152
Total	1173	764	5430	1453	8820

PHILLIP TSHABALALA T06075 DPG9079 OD427014

CROSSTABS FOR LEARNERS IN MODEL - BY SCHOOL LOCATION

The FREQ Procedure

Table of PR52 by PR5

PR52(ANNUAL SCHOOL FEES) PR5(SCHOOL LOCATION)

Frequency Row Pct Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE	Total
1 <R50	224 5.12	269 6.15	3055 69.86	825	4373

18.87			19.10	35.21	56.26
56.78					
	2 50-100	405	30	1358
205	1998				
		20.27	1.50	67.97
10.26					
		34.53	3.93	25.01
14.11					
	3 101-200	77	118	23
0	218				
		35.32	54.13	10.55
0.00					
		6.56	15.45	0.42
0.00					
	4 201-300	100	92	27
37	256				
		39.06	35.94	10.55
14.45					
		8.53	12.04	0.50
2.55					
	5 301-400	58	0	26
89	173				
		33.53	0.00	15.03
51.45					
		4.94	0.00	0.48
6.13					
	6 401-500	39	24	71
0	134				
		29.10	17.91	52.99
0.00					
		3.32	3.14	1.31
0.00					
	7 501-1000	55	72	35
30	192				
		28.65	37.50	18.23
15.63					
		4.69	9.42	0.64
2.06					
	8 1001-2000	156	66	129
43	394				
		39.59	16.75	32.74
10.91					
		13.30	8.64	2.38
2.96					
	9 >R2000	59	93	706
224	1082				
		5.45	8.60	65.25
20.70					
		5.03	12.17	13.00
15.42					

Appendix 4 CROSSTAB

Total	1173	764	5430	1453	8820

1 PROG1C
13:17 Monday, April 23, 2007 6
PHILLIP TSHABALALA T06075 DPG9079 OD427014
CROSSTABS FOR LEARNERS IN MODEL - BY SCHOOL LOCATION

The FREQ Procedure

Table of PR25I by PR5

PR25I(AVAILABILITY OF COMPUTER FOR TEACHING)

PR5(SCHOOL LOCATION)

Frequency / Row Pct / Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE RUR	Total
1 NOT AVAIL	714	554	5386	1422	8076
	8.84	6.86	66.69	17.61	
	60.87	72.51	99.19	97.87	
2 SOMETIMES	138	37	0	0	175
	78.86	21.14	0.00	0.00	
	11.76	4.84	0.00	0.00	
3 ALWAYS	321	173	44	31	569
	56.41	30.40	7.73	5.45	
	27.37	22.64	0.81	2.13	
Total	1173	764	5430	1453	8820

Table of PR25K by PR5

PR25K(AVAILABILITY OF OHP FOR TEACHING)
PR5(SCHOOL LOCATION)

Frequency / Row Pct / Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE RUR	Total
1	263	275	4794		

1361	6693		3.93	4.11	71.63
20.33			22.42	35.99	88.29
93.67					
		2	175	62	171
41	449		38.98	13.81	38.08
9.13			14.92	8.12	3.15
2.82					
		3	735	427	465
51	1678		43.80	25.45	27.71
3.04			62.66	55.89	8.56
3.51					
		Total	1173	764	5430
1453	8820				

CROSSTABS FOR LEARNERS IN MODEL - BY SCHOOL LOCATION

The FREQ Procedure

Table of ED38D by PR5

ED38D(INADEQUATE PHYSICAL FACILITIES) PR5(SCHOOL LOCATION)

Frequency Row Pct Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE RUR	Total
1 NOT AT ALL	287	292	716	154	1449
	19.81	20.15	49.41	10.63	
	24.47	38.22	13.19	10.60	
2 A LITTLE	452	137	1550	247	2386
	18.94	5.74	64.96	10.35	
	38.53	17.93	28.55	17.00	
3 QUITE ALOT	285	114	1649	657	2705
	10.54	4.21	60.96		

Appendix 4 CROSSTAB

24.29 |
| 24.30 | 14.92 | 30.37 |
45.22 |

---

4 GREAT DEAL | 149 | 221 | 1515 |
395 | 2280
| 6.54 | 9.69 | 66.45 |
17.32 |
| 12.70 | 28.93 | 27.90 |
27.19 |

Total 1173 764 5430
1453 8820

Table of PAR17 by PR5

PR5(SCHOOL LOCATION) PAR17(NUMBER OF BOOKS AT HOME)

Frequency|
Row Pct |
Col Pct |1 URBAN |2 SEMI-U|3 SEMI-R|4 REMOTE| Total
|RBAN |URAL | RUR |

---

1 <10 | 568 | 407 | 3866 |
1019 | 5860
| 9.69 | 6.95 | 65.97 |
17.39 |
| 48.42 | 53.27 | 71.20 |
70.13 |

---

2 10-20 | 262 | 159 | 993 |
255 | 1669
| 15.70 | 9.53 | 59.50 |
15.28 |
| 22.34 | 20.81 | 18.29 |
17.55 |

---

3 >20 | 343 | 198 | 571 |
179 | 1291
| 26.57 | 15.34 | 44.23 |
13.87 |
| 29.24 | 25.92 | 10.52 |
12.32 |

Total 1173 764 5430
1453 8820

CROSSTABS FOR LEARNERS IN MODEL - BY SCHOOL LOCATION

The FREQ Procedure

Table of ED23EMa by PR5

ED23EMa(NUMERACY REFERENCE MATERIALS

Appendix 4 CROSSTAB

AVAILABLE)

PR5(SCHOOL LOCATION)

Frequency Row Pct Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE RUR	Total
YES	888 17.55 75.70	459 9.07 60.08	2884 57.00 53.11	829 16.38 57.05	5060
NO	285 7.58 24.30	305 8.11 39.92	2546 67.71 46.89	624 16.60 42.95	3760
Total	1173	764	5430	1453	8820

Table of ED19Q by PR5

ED19Q(CALCULATORS ETC. IN CLASSROOM)

PR5(SCHOOL LOCATION)

Frequency Row Pct Col Pct	1 URBAN	2 SEMI-URBAN	3 SEMI-RURAL	4 REMOTE RUR	Total
YES	783 23.56 66.75	324 9.75 42.41	1882 56.64 34.66	334 10.05 22.99	3323
NO	390 7.09 33.25	440 8.00 57.59	3548 64.54 65.34	1119 20.36 77.01	5497
Total	1173	764	5430	1453	8820

APPENDIX 5: RELIABILITY ANALYSIS

Example 1: Reliability analysis of scales (PRQ – Availability of resources/facilities)

Case Processing Summary

		N	%
Cases	Valid	1135	97.8
	Excluded(a)	26	2.2
	Total	1161	100.0

a Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.647	2

Item Statistics

	Mean	Std. Deviation	N
School Facilities/Resources – Computer for teaching	1.28	.672	1135
School Facilities/Resources – OHPs	1.71	.921	1135

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
School Facilities/Resources – Computer for teaching	1.71	.848	.502	.(a)
School Facilities/Resources – OHPs	1.28	.452	.502	.(a)

a The value is negative due to a negative average covariance among items. This violates reliability model assumptions. You may want to check item codings.

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
2.99	1.921	1.386	2

Example 2: *Reliability analysis of scales (LAD-LQ-PARQ - Learner language)*

Case Processing Summary

		N	%
Cases	Valid	44614	83.0
	Excluded(a)	9141	17.0
	Total	53755	100.0

a Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.822	4

Item Statistics

	Mean	Std. Deviation	N
Learner Home Language	5.1826	2.64470	44614
language of learning and teaching	4.5506	2.70923	44614
home language	5.1050	2.64636	44614
language of instruction versus home language	1.7855	.41047	44614

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Learner Home Language	11.4411	25.641	.873	.651
language of learning and teaching	12.0731	26.525	.794	.699
home language	11.5187	25.532	.879	.648
language of instruction versus home language	14.8382	55.868	-.001	.927

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
16.6237	56.030	7.48530	4