Exploring problems encountered in the teaching and learning of statistics in Grade 11

by

Eva Gavhaza Makwakwa

submitted in accordance with the requirements for
the degree of

Master of Science in Mathematics, Science and Technology Education with specialisation in Statistics Education

at the

University of South Africa

Supervisor: Prof. L.D. Mogari

December 2012

Declaration

Student number: 45953163

I declare that

Exploring problems encountered in the teaching and learning of statistics in Grade 11 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

(Ms E. G. Makwakwa)

Abstract

The purpose of the study was to explore problems encountered in the teaching and learning of statistics in grade 11 and to offer ways of addressing them. A convenient sample of 100 grade 11 mathematics teachers and 448 grade 11 mathematics learners participated in the study. A descriptive survey design was used. Data were collected from the teachers using a teacher questionnaire, classroom observation schedule and teacher interview, and from learners through the use of a learner questionnaire, classroom observation schedule, diagnostic test and learner interview. The validity and reliability of all these instruments was established. Descriptive statistics and frequencies were used to analyse learner questionnaires, teacher questionnaires and learners' responses to a diagnostic test. Teacher and learner interviews were transcribed and classified according to themes. Classroom observations were analysed by using themes and checking for similarities and differences.

The results showed that teachers had difficulty with the interpretation and calculation of measures of dispersion; representation and interpretation of data on graphs or plots; determining the five number summary; constructing and interpreting probability diagrams and tables; and interpreting probability terminology.

Also, the results showed that learners experienced difficulties when using graphs to predict the results; interpreting and determining measures of dispersion; computation of quartiles when the total number of data values was even; representing data on graphs or plots; interpreting and determining measures of central tendency; constructing and interpreting probability graphs and tables; and interpreting probability terminology.

The results found possible causes of the teachers' difficulties to be (1) their lack of statistics content knowledge; (2) inadequate textbooks; (3) in-service programmes which did not cover statistical topics, or which did not pay adequate attention to probability; and (4) teachers failure to attend these in-service teacher workshops.

Further, the findings of the study were that the probable causes of learners' difficulties were (1) inadequate teaching of statistics topics in previous grades; (2) teachers' lack of content knowledge in statistics meant they had difficulty explaining concepts to learners; (3)

inadequate learning material and learners’ inability to use the statistics function mode on their calculators; and (4) learners’ lack of conceptual knowledge of certain aspects of statistics.

Lastly, the results of the study found that the possible ways to address the problems in the teaching and learning of statistics in grade 11 to be: (1) teachers should receive financial support from their schools/districts to attend in-service education and training programmes; (2) textbooks should be well written (provide thorough explanations) and contain all the information necessary to teach data handling and probability (i.e. formulae, more examples); (3) in-service teacher programmes should meet the needs of the teachers by offering topics that teachers find difficult to teach; and (4) more and longer inset programmes on probability, preferably five-day workshops, should be arranged.

Key terms

Statistics teachers, data handling, probability, problems, difficulties, in-service education and training, content knowledge, pedagogical knowledge.

Acknowledgements

Firstly, I wish to thank the Almighty God for giving me the wisdom and strength to complete this MSc research project.

The writing of this dissertation was a unique experience to which many people contributed:

My sincere gratitude goes to Professor L. D. Mogari, my supervisor, for his invaluable guidance, constructive criticism, support and encouragement throughout this study.

I am indebted to Professor H.I. Atagana, Professor J. Kriek and Dr E.C. Ochonogor for their support and encouragement.

I am very grateful to Mr J.J. Dhlamini, Dr. J.A. Jimoh and Dr. U.I. Ogbonnaya, who read through the work and gave constructive criticism.

I am also thankful to the school principals, teachers and learners of all the schools that participated in the study.

I would like to thank the National Research Foundation (NRF) and University of South Africa for their assistance in funding all aspects of this research.

Lastly, I would like to offer special thanks to my parents and siblings for their support. I am especially grateful to my brother, Auguston Makwakwa, for his regular motivation and encouragement; I could not have made it without you. Thank you, my brother! Many thanks go to my daughter, Oratilwe: her patience and understanding made it possible for me to complete this study.

MAY GOD BLESS YOU ALL.

Table of contents

List of Abbreviations

FET	-	Further Education and Training
GET	-	General Education and Training
NCS	-	National Curriculum Statement
LO	-	Learning Outcomes
AS	-	Assessment Standards
CAPS	-	Curriculum and Assessment Policy Statement
DoE	-	Department of Education
DoBE	-	Department of Basic Education
INSET	-	In-service Education and Training
B. Tech	-	Bachelor of Technology
STD	-	Secondary Teachers Diploma
ACE	-	Advanced Certificate in Education
B. Ed	-	Bachelor of Education
ICT	-	Information and communication Technology

List of Tables

Page

List of figures

CHAPT ER 1

INTRODUCTION

1.1 Background to the study

The inception of a democratic government in South Africa has been followed by reformations in the education system (Khuzwayo, 2005). This has led to transformation from an unequal and racially segregated system of education to an integrated and equal system. Concomitantly, there has been a spate of curriculum reforms, from Curriculum 2005 (C2005) to the Curriculum and Assessment Policy Statement (CAPS). Prior to 1994, education in South Africa was based on the concept of separate development for different races hence there was an inequitable system of funding, resources and access to education. African learners were not exposed to high quality education; for instance, mathematics was not seen by the apartheid government as part of the education of an African child, as expressed in the words of Dr Verwoerd to parliament on 17 September 1953: "what is the use of teaching a Bantu child mathematics when it cannot use it in practice" (Khuzwayo, 2005). Hence, mathematics was offered by only a few black schools at high school level; and many black learners were prevented from taking mathematics up to high school level. Education was meant to provide black learners with basic knowledge to prepare them for unskilled or semi-skilled jobs (Giliomee, 2009; Macrae, 1994). The situation was even worse for statistics because it was not even taught in schools.

Added to this, there was a scarcity of qualified mathematics teachers which led to poor performance by learners and a general lack of interest in the subject (Khuzwayo, 2005). None of the teachers who obtained their pre-service training from colleges could teach statistics as it was not part of the curriculum at these colleges. The Teachers were trained in poorly equipped and racially segregated colleges of education which were ineffective in providing high quality teacher education. This, coupled with the overcrowded classrooms, often with over 100 learners in a class, and a shortage of teaching and learning resources, led to ineffective methods of teaching mathematics, the ripple effect of which is a generally poor performance in mathematics by school learners today (Ogbonnaya, 2011; Howie & Pietersen, 2001; Adler, 1994).

In an effort to rectify this state of affairs and to meet the country's development needs, the democratic government took a major step in 1994 by starting a process of syllabus revision and subject rationalisation. This led to a single national core syllabus to end a racist and segregated system of education (Department of Education [DoE], 2002). Under the old system, there were no clear educational outcomes; there was a plethora of subjects with little coherence between them; the curricula did not respond to the needs of learners or the country; there was limited mobility across pathways and institutions in the Further Education and Training (FET) band; and the focus was not on a learner acquiring decision-making, problem-solving or critical thinking skills (DoE, 2003). The outgoing C2005 was revised to form the *Revised National Curriculum Statement Grades R-9* and the *National Curriculum Statement Grades 10-12*, later called the National Curriculum Statement for Grades R-12, introduced into schools in 1997 and 2006 respectively. The aims of C2005 were to develop the potential of each learner as a citizen of a democratic South Africa, to seek to create a lifelong learner who is confident and independent, literate, numerate and multi-skilled, compassionate, with a respect for the environment and the ability to participate in society as a critical and active citizen (DoE, 2002).

One notable change in the curriculum was the introduction of Statistics (Data handling) in Mathematical Literacy, Mathematics and Mathematical Science (MLMMS) in the General Education and Training (GET) phase. When the C2005 curriculum was reviewed and later replaced with the National Curriculum Statement (NCS), statistics was also introduced in the Further Education and Training (FET) Mathematics. This was the first time statistics was part of school curriculum. Statistics, in this study, refers to data handling and probability, collectively (see section 1.6). In the NCS Grades 10-12, statistics is incorporated in the fourth learning outcome (LO4: Data Handling and Probability). This revision was made in order to strengthen and streamline the design features of Curriculum 2005 by addressing issues such as simplifying the language it used, aligning curriculum and assessment, and improving teacher orientation and training, learner support materials and provincial support (DoE, 2004). Part of the revision was to change MLMMS to Mathematics.

Data handling and probability are the only aspects of statistics that are taught at FET level. Each of the Learning Outcomes (LO) is linked to a set of Assessment Standards (AS) (i.e. statements that spell out how each LO should be achieved). The Learning Outcomes and Assessment Standards in the mathematics curriculum for the FET band (in grades 11 and 12 only) are divided into Core Assessment Standards and Optional Assessment Standards until the end of year 2012 for grade 11 and until the end of year 2013 for grade 12, as "it is anticipated that the Assessment Standards identified as optional will become compulsory after 2010" (DoE, 2008: 7). This anticipation was proved true because the LO and AS are no longer divided into Core Assessment Standards and Optional Assessment Standards in grade 10. In this grade, the Core Assessment Standards and Optional Assessment Standards have been phased out as a result of the newly introduced curriculum known as Curriculum and Assessment Policy Statements (CAPS), which also advocates the NCS goals. The CAPS Grades 10-12 will replace the NCS Grades 10-12 in the year 2013 in grade 11 and in the year 2014 in grade 12; and that will see the end of the NCS Grades 10-12. The Optional Assessment Standards which have been identified are LO3 and LO4 (DoE, 2008). Even though LO4 is identified as an optional assessment standard in the outgoing NCS Grades 10-12 curriculum, there is a part of Data Handling which is treated as a Core Assessment Standard, while Probability is treated entirely as an optional assessment standard. Also, in the new curriculum (NCS related CAPS) for Grades 10-12, probability and statistics will form part of the 10 main topics (Department of Basic Education [DoBE], 2011: 9). Of concern though was the fact that seemingly the rate of effecting curriculum changes to include statistics did not correspond with the rate of preparing teachers to teach the newly introduced areas of knowledge. It is this apparent discord between making changes in the curriculum and teacher development that necessitates an investigation into the teaching and learning of statistics in schools.

1.2 Study problem

Data handling and probability have been taught in tertiary education only as components of statistics. Unfortunately, in the previous political dispensation they were not taught at colleges of education, where the majority of teachers in classrooms today received their pre-service training. Thus, there was a significant number of teachers who encountered data handling and probability for the first time when it became part of school mathematics in the FET band in 2006 (Makwakwa & Mogari, 2012; Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya & Makwakwa, 2009). These teachers were given until 2010 to gain the necessary knowledge to teach the subject effectively and with confidence (DoE, 2008). The Ministry of Education organised numerous in-service training programmes to develop teachers' content and pedagogical knowledge in statistics (DoE, 2008). Nevertheless, this intervention has had little effect because teachers still have problems that range from having to teach material with which they are unfamiliar to experiencing difficulties in implementing necessary assessment procedures as stipulated by NCS policy guidelines (Carnoy & Chisholm, 2008; Mahlobo, 2009). These problems have been passed on to many learners who performed poorly in statistics related questions in the grade 12 end-of-year mathematics examination in 2011 (Cassim, 2012).

On the other hand, there have been other in-service training initiatives aimed at providing teachers with the necessary content knowledge and skills to teach statistics (data handling implied), organised mainly by non-governmental and governmental organisations such as the South African Statistical Association (SASA), Statistics South Africa (Stats SA) and the Association for Mathematics Education of South Africa (AMESA). According to North and Zewotir (2006), in-service workshops on data handling and probability give teachers the opportunity to upgrade their knowledge so that they can teach statistical literacy to the school leaver in South Africa. AMESA also organised in-service initiatives for subject facilitators or advisors, who were then expected to provide similar training to teachers; it arranged hour-long supportive talks to teachers at its annual conferences and other similar forums as well as demonstration lessons. These stressed the significance of statistics as a problem-solving tool and also demonstrated how to conduct a learner-centred and interactive statistics lesson.

Furthermore, the Institute for Science and Technology Education (ISTE) at the University of South Africa (UNISA) has also offered lessons during school winter vacations to mathematics teachers. These are intended to improve and upgrade teachers' knowledge of data handling and probability content, among others aspects, because teachers consider these topics to be among the most instructionally and conceptually problematic in mathematics (Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya, Dhlamini & Makwakwa, 2010, 2011; Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya & Makwakwa, 2009). Even though ISTE's in-service initiative has been taking place annually since 2009, there are teachers who still have difficulties with teaching data handling and probability (Atagana et al., 2011, 2010).

According to the critical outcomes for mathematics of the National Curriculum Statement (NCS Grades 10-12 and CAPS Grade10-12), at the end of a learning experience, learners must be able to: identify, solve problems and make decisions using critical and creative thinking; collect and use data to establish basic statistical and probability models, solve related problems and critically consider representations provided or conclusions reached, analyse, organise and critically evaluate information, etc. (DoBE, 2011: p.5; DoE, 2003: p.8). It should be noted that the attainment of these learning outcomes by learners depends to a large extent on teachers' effectiveness in teaching data handling and probability. It is therefore essential that problems encountered in the teaching and learning of statistics are investigated and described.

1.3 Purpose of the study

The study seeks to explore problems encountered in the teaching and learning of statistics in grade 11 as well as the sources of these problems. The study will also explore possible ways to address the problems encountered in the teaching and learning of statistics in Grade 11. There is a tendency by authorities to impose ways and means to deal with issues on the ground without seeking the views of those involved, e.g. teachers and/or learners. It is therefore imperative to solicit views and suggestions from those involved with a view to optimising the redress of the identified problems.

1.4 Research questions

i. What problems, if any, are encountered by teachers in the teaching of statistics in grade 11?
ii. What problems, if any, are encountered by learners in the learning of statistics in grade 11?
iii. What are the cause(s) of the problems encountered in the teaching and learning of statistics in grade 11?
iv. What are possible ways of addressing problems encountered in the teaching and learning of statistics in grade 11?

1.5 Significance of the study

It is important for South Africa, as a developing country, to have citizens who are statistically literate, i.e. people who are able to collect data and use it to establish basic statistical and probability models, solve related problems, critically consider representations provided or conclusions reached, analyse, organise and critically evaluate information. In order to achieve this educational goal it is important to determine the nature of the problems associated with the teaching and learning of statistics with a view to finding possible solutions. Even though there have been attempts to develop teachers by providing them with the necessary content knowledge to teach statistics effectively, it seems that these teachers continue to experience difficulties which limit the effectiveness of their classroom practice. It is thus vital that ways are found to improve the teaching of statistics. The situation seems to be aggravated by the fact that some teachers were not taught statistics in their pre-service training. It is important to note that poorly trained teachers are likely to produce weak and ill-prepared learners (Howie & Pietersen, 2001). Hence the need to gain further insight into issues relating to the teaching of statistics. Pursuing such an investigation will hopefully assist in the discovery of ways to improve teachers' content delivery in statistics.

The study is also motivated by Owusu-Mensah's (2008) observation that most existing in-service education and training programmes do not include a formal process of identifying the needs of statistics teachers. In particular, Owusu-Mensah notes that most of these programmes do not conduct surveys or interviews which would reveal topics which learners find difficult or areas of the syllabus which teachers feel uncomfortable about or which they find difficult to teach. This data could be collected prior to or after the commencement of a planned in-service workshop for teachers. It is the goal of this study to explore some of these problems associated with the teaching of statistics.

Teacher morale is a major issue in schools. Low morale decreases proper engagement with colleagues and learners, diminishes productivity, reduces effective learning and breeds cynicism (Gachutha, 2009). When morale is high, teachers are more productive and collaborative, learners excel socially and academically, and the school environment is dynamic and engaging (Hollinger, n.d.). Professional development is one of the major factors affecting morale because it is directly related to learner achievement (Hollinger, n.d.).

Teachers and schools that value professional development or adult learning create the conditions for learners to value learning and when learners make excellent academic progress, teachers feel the rewards of their profession. Providing teachers with meaningful and effective professional development is fundamental for successful schools and high teacher morale (Hollinger, n.d.).

Several authors (Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya, Dhlamini, & Makwakwa, 2011; Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya & Makwakwa, 2009) have reported that learners experience difficulties with statistics (data handling and probability) which adds to their poor performance in mathematics. The current study will also explore problems encountered by learners which are related to content in data handling and probability. By understanding these problems, and by finding ways to address them, poor learner performance in mathematics might be improved.

The National Department of Education has formulated a strategic plan to address challenges and problems associated with classroom implementation of the National Curriculum Statement (NCS) in schools. This project is entitled "Action Plan 2014: Towards the Realisation of Schooling 2025" (Motshekga, 2010). According to Motshekga, the project is aimed at addressing the full gamut of curriculum challenges, some of them being the shortage of resources, inadequate teacher knowledge and a lack of technical skills. This plan makes provision for teacher development, focusing less on pedagogy, as countless "advanced courses" and in-house seminars have in the past, but more on improving teachers' content knowledge (Motshekga, 2010). It is hoped that the results of this study will be presented to the curriculum developers for mathematics or to the review committee which is currently designing the "Action Plan 2014: Towards the Realisation of Schooling 2025". Teachers' observed instructional problems and learners' observed learning problems in statistics could assist curriculum designers and textbook writers in formulating appropriate strategic plans to deal effectively with these issues. It is also hoped that the findings of the study will also contribute a new perspective and knowledge on issues pertaining to statistics teaching and learning.

1.6 Definition of terms

PROBLEM

In this study, the word "problem" is defined as any obstacle that may hinder teachers' success in generating an environment conducive to the learning and teaching of statistics. These obstacles could include learners' and teachers' difficulties in grasping concepts in statistics, learners' and teachers' inability to reason about the data, learners' and teachers' difficulties in using statistical formulae in calculations, learners' and teachers' inability to choose the correct formulae and learners' and teachers' difficulties in interpreting statistical results.

STATISTICS

Statistics is defined as the art and science of collecting, analysing, presenting and interpreting data. Three terms are used in this description: data description (data handling), statistical inference (probability is the foundation of statistical inference), and statistical modelling. In this study, the term *statistics* is used as a collective term for data handling and probability. Data handling involves collecting data for a particular purpose, sorting data, representing the data in tables, charts or graphs, analysing the results and coming to conclusions. Probability is concerned with determining the likelihood that a given event will occur (see section 2.1.4).

1.7 Structure of Dissertation

The dissertation is structured as follows:

Chapter 1- Introduction

This chapter provides the context of the study by describing its background, the problem, the aim, research questions, and the significance of the study. It also provides the definitions of terms.

Chapter 2-Literature Review and Conceptual Framework

The conceptual framework guiding the study and the review of related literature are presented. The literature is based on the main themes of the problems encountered in the teaching and learning of statistics, possible causes of these problems and recommendations for overcoming difficulties in the teaching and learning of statistics.

Chapter 3-Methodology

This chapter focuses on the methods used in the study, including the research design, sample selection method and data collection instruments. The development of the instruments and the procedures of data collection are discussed, as well as the validity and reliability of instruments, the data analysis and the ethical issues considered in the study.

Chapter 4-Data Analysis

This chapter focuses on data analysis methods and procedures. The results of the data analysis are presented and the findings of the study are used to answer the research questions.

Chapter 5-Summary of the Findings, Discussion, Conclusion and Recommendations

This chapter summarises the study. The findings are discussed together with their implications. Conclusions are reached upon which the recommendations are based.

CHAPTER 2

LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK

2.1 The History of Statistics

The section provides a historical perspective of statistics with a view to shedding light on the genesis and development of statistics as a body of knowledge as well as its utility and importance. The discussion on these aspects helps clarify what statistics all about, justifies the intentions of a study on statistics and draws a line of distinction between statistics and mathematics.

2.1.1 The origin of the word statistics

Pearson's work in the seventies indicates that the term statistics came into use in the 17^{th} century. Its origin is in the Italian word "statista" which means "statist" or "statesman", the "statista" being a man concerned with reasoning about the state, "ragione di stato". In Latin, it was treatises to Disciplina de Ratione status or Disciplina de statu. The work of Pearson reveals that the "Disciplina de statu" was certainly not statistics in its present sense, nor was it concerned with numbers or mathematical theory; rather, it was concerned with constitutional history and descriptions of state constitutions or statecraft.

2.1.2 The birth of present-day statistics

The publication of *Natural and Political Observations Made upon the Bills of Mortality* by John Graunt in 1662 launched the discipline we now call mathematical statistics (Burton, 2007). Even though there is evidence of the use of statistics as far back as before Christ (BC), John Graunt (1620-1674) was the first man to deal with vital statistics; to organise data in the form of descriptive statistics and to draw an extensive set of statistical inferences from mass data (Burton, 2007). The Bills of Mortality were originally weekly and yearly returns of the number of burials in various London parishes. An attempt was made to classify all deaths in London according to several causes (only disease and accident). The data were obtained by searchers, usually "ancient matrons", whose work included viewing the body and inquiring about the disease or casualty that had led to death (Burton, 2007). Among the important statistical regularities Graunt observed were the fact that the male births outnumbered female births, that women lived longer than men, and that the number of people dying from most causes except epidemic diseases remained fairly constant from year to year (Burton, 2007: p. 440).

Burton (2007) explains that John Graunt's methods were applied by Sir William Petty to Irish data. What emerges from Pearson's work is that John Graunt, Gregory King and Sir William Petty never used the term "statistik" nor did they call their data "statistics". Their problems and ideas were taken up by the English mathematicians, Halley and De Moivre, however, and spread through Bernoulli and Euler to continental mathematicians. Pearson explains how these ideas were later expanded in the works of Poisson and Laplace, who introduced the fundamental ideas of calculus of probability into the subject, the foundation of the mathematical science of statistics. But their subject was still not called "statistics" or "statistik". But Pearson's work shows that around 1798, Sir John Sinclair stole the words "statistics" and "statistik" and applied them to the data and methods of "Political arithmetic", which is statistics as we know it today. The scope of the discipline of statistics broadened in the early 19th century to include the collection and analysis of data in general.

Today, statistics is described using three terms: (1) data description (data handling), (2) inference (inferential statistics), and (3) statistical modelling. These are briefly discussed below.

Descriptive statistics (Data handling) involves collecting data for a particular purpose, sorting data, representing it in tables, charts or graphs, analysing the results and coming to conclusions (Wegner, 2007). Data is interpreted by using either pictorial or arithmetic methods. Pictorial methods involve drawing graphs such as bar graphs, histograms, frequency polygons, pie charts, line and broken line graphs. Arithmetic methods involve working out (1) measures of central tendency such as mean, median, or mode; (2) measures of dispersion such as the range, percentiles, quartiles and the interquartile range (Wegner, 2007).

Inferential statistics are used to draw inferences about a population from a sample. That is, inferential statistics are concerned with determining how likely it is that results based on a sample or samples are the same results that would have been obtained for the entire population (Gay, Mill & Airasian, 2011).There are two methods used in inferential statistics: estimation and hypothesis testing. In estimation, the sample is used to estimate a parameter and a confidence interval about the estimate is constructed. Inferential statistics help the student or researcher to make confident decisions in the face of uncertainty (Decaro, 2003; Wegner, 2007). Probability is the foundation of statistical inference (Wegner, 2007). Thus, probability is concerned with determining the likelihood that a given event will occur (Laridon, Barnes, Jawurek, Kitto, Myburgh, Pike, Myburgh, Rhodes-Houghton, Scheiber, Sigabi & Wilson, 2004; Wegner, 2007). Probability helps to measure a decision maker's level of uncertainty about whether some outcome will occur.

Statistical modelling is that area of statistics where mathematical equations are used to build relationships between variables (Wegner, 2007). These equations (called models) are then used to estimate or predict values of one or more of the variables indifferent scenarios (Wegner, 2007).

2.1.3 The origins of probability

Probability theory was inspired by games of chance and gambling during the 16th century. Dicing, card games and lotteries, public and private, were important social and economic activities (Hald, 2003). Gambling originated in the early stages of human history; it was not the invention of a single people but appeared in many places in the world (Burton, 2007). The fundamental principles of probability theory were formulated for the first time in 1654 by two mathematicians, Blaise Pascal and Pierre de Fermat, after a well-known gambler, Chevalier de Méré, consulted Blaise Pascal about some questions concerning a popular game of dice (Burton, 2007). Although there are a number of predecessors (i.e. Girolamo Cardano) of Blaise Pascal and Pierre de Fermat who had solved problems on games of chance during the 15th and 16th centuries, no general theory was developed before Blaise Pascal and Pierre de Fermat (Burton, 2007).

Probability theory became popular, and the subject developed rapidly during the 18th century. The major contributors during this period were Christiaan Huygens, Jakob Bernoulli and Abraham de Moivre (Burton, 2007). In 1812 Pierre Simon Laplace introduced the new ideas and mathematical techniques in his book, *Théorie Analytique des Probabilités*. Before Laplace, probability theory was solely concerned with developing a mathematical analysis of games of chance. In the 18th century, the application of probability moved from games of chance to scientific problems. Laplace applied probabilistic ideas to many scientific and practical problems; mathematical theory of life insurance life tables, the theory of errors, actuarial mathematics and statistical mechanics are all examples of important applications of probability theory developed in the l9th century.

Currently, in the 21st century, probability theory is used in all kinds of risk assessment in the insurance industry, in medical research, in engineering, in finding the genetic makeup of individuals or populations, in quality control, in investment and in virtually every other human endeavour. Probability theory today is particularly valuable as it can be used to determine the expected outcome in any situation, from the chance that a plane will crash to the probability that a person will win the lottery.

2.1.4 What is Statistics?

King and Julstrom (1982) define the term statistics as the science of collection and classification of facts on the basis of relative numbers of occurrences as a ground for inductions; the systematic compilation of instances for the inference of general truth. Mann (1998) defines statistics as a group of methods that are used to collect, analyse, present and interpret data and to make decisions. Moore, McCabe, Duckworth and Alwan (2009) define statistics as the science of collecting, organising and interpreting numerical facts, which we call data. Gravetter and Wallnau (2002) explain statistics as a set of methods and rules for organising, summarising and interpreting information. For Williams, Sweeney and Anderson (2006), statistics is the art and science of collecting, analysing, presenting and interpreting data.

For Steel and Torrie (1980), statistics is the science, pure and applied, of creating, developing and applying techniques such that the uncertanity of inductive inferences may be evaluated. These authors believe that statistics is logic or common sense with a strong mixture of arithmetical procedures. The logic supplies the method by which data are to be collected and determines how extensive they are to be; the arithmetic, together with certain numerical tables, yields the materials on which to base the inference and measure its uncertainity (Steel & Torrie, 1980).

2.1.5 The importance of Statistics

The need for statistics is encountered in everything, ranging from news reports, sports' averages, weather, elections, business reports and stocks to advertisements, economic conditions and so on. Every day, newspapers and other media confront us with statistical information on topics ranging from the economy to education, from films to sports, from food to medicine, and from public opinion to social behaviour (Franklin, Kader, Mewborn, Moreno, Peck, Perry & Scheaffer, 2007). At work, we are presented with quantitative information on budgets, supplies, manufacturing specifications, market demands, sales forecasts, or workloads. At schools, teachers are confronted with educational statistics concerning learner performance.

Statistics play an important role in various fields or in any business where data exists or is generated. For example, Woodward and Francis (1988) observe that the use of statistics in health research is generally extensive since most research requires data collection and subsequent description, summarisation and generalisation from the particular cases observed. Examples include investigations into the relationship between smoking and lung cancer, an evaluation of the efficacy of a treatment for AIDS and a comparison of different methods of screening women for cervical cancer.

Gravetter and Wallnau (2008) believe that statistics is the basis of experimental psychology, used for example to determine whether violence on television has any effect on children's behaviour. In a case like this, data about children's behaviour would be collected; a psychologist would take measurements such as IQ scores, personality scores, reaction time scores, and so on, and would use statistics techniques to make sense of the data collected (Gravetter & Wallnau, 2008).

In weather forecasts, surface weather observations of atmospheric pressure, temperature, wind speed, wind direction, humidity and precipitation are collected in the form of data by a weather observer, or by computer through the use of automated weather stations (Julian & Murphy, 1972). Julian and Murphy explain that the weather forecaster uses weather models called model output statistics (MOS) ─ a technique used to interpret numerical model output and produce site-specific guidance. These computer models are built using statistics that compare prior weather conditions to current weather to predict future weather.

Furthermore, statistics plays an important role in quality testing. That is, companies make thousands of products every day and must make sure that a good quality item is sold. Instead of testing each and every item they ship to consumers, the company uses statistics to test just a few items, called a sample. If the sample passes quality tests, then the company believes that all the items made in the group, called a batch, are good (Hubbard, 2003).

In finance, a stock market analyst uses statistical methods to predict share price movements; financial analysts use statistical findings to guide their investment decisions in bonds, cash, equities, property, etc. (Wegner, 2007).

In 2010, Dr Shashiranjan Yadav, in a speech as a Guest honor at a two day seminar on "Impact of statistics on science & society", emphasised the importance of statistics: "Statistics benefits all of us because we are able to predict the future based on data we have previously gathered" (Yadav, 2010). He stressed that "being able to predict the future not only changes our lifestyle but also helps us to be more efficient and effective in planning and decision making" (Yadav, 2010). Without statistics there would be too much guesswork in our daily lives, which could result in inefficient and/or bad planning, unforeseen shortages of supplies, mismanagement of scarce resources, unnecessary loss of life, waste and many other disadvantages.

2.2 Statistics and Mathematics

According Moore (1988), statistics is as much a distinct discipline as are economics and physics. Although statistics is a mathematical science, it differs from mathematics in fundamental ways; its origins, subject matter, foundation questions and standards are distinct from those of mathematics (Godino, Ortiz, Roa, & Wilhelmi, 2011; Gould & Peck, 2004; Moore, 1988; Nicholson, Ridgway & McCusker, 2006). The origins of statistics lie in official and private data-gathering, in census and tax rolls and mortality tables and its subject matter is data and inference from data (Burton, 2007).

Statistics as a methodological discipline exists not for itself, but rather to offer other fields of study a coherent set of ideas and tools for dealing with data (Cobb & Moore, 1997). The need for such a discipline arises from the omnipresence of variability and "it is this focus on variability in data that sets apart statistics from mathematics" (Franklin, Kader, Mewborn, Moreno, Peck, Perry& Scheaffer, 2007: 6). Franklin et al (2007: 6) also note that the different sources of variability in data such as Measurement Variability, Natural Variability, Induced Variability and Sampling Variability. In addition, Cobb and Moore (1997) articulate that "there is more than just content that distinguishes statistical thinking from mathematics, as statistics requires a different kind of thinking, because data are not just numbers, they are numbers with a context". Cobb and Moore (1997: 803) also believe that

> *although mathematicians often rely on applied context both for motivation and as a source of problems for research, the ultimate focus in mathematical thinking is on abstract patterns: the context is part of the irrelevant detail that must be boiled off over the flame of abstraction in order to reveal the previously hidden crystal of pure structure. In mathematics, context obscures structure. Like mathematicians, data analysts also look for patterns, but ultimately, in data analysis, whether the patterns have meaning, and whether they have any value, depend on how the threads of those patterns interweave with the complementary threads of the story line. In data analysis, context provides meaning.*

To formulate a statistics question requires one to have an understanding of the difference between a question that anticipates a deterministic answer and a question that anticipates an answer based on data that vary (Franklin et al., 2007). Cobb and Moore (1997: 803) emphasise that "to teach statistics well, it is not enough to understand the mathematical theory; it is not even enough to understand also the additional, non-mathematical theory of statistics". They believe that, to teach statistics well one must be like a teacher of literature, have a ready supply of real illustrations, and know how to use them to involve learners in the development of their critical judgement. This is because in mathematics, where applied context is so much less important, improvised examples often work well, and teachers of mathematics become skilful at inventing examples on the spot; conversely, in statistics, improvised examples do not work, because they do not provide authentic interplay between pattern and context (Cobb & Moore, 1997).

Probability is also an important part of any mathematical education and an essential tool in statistics. According to Franklin et al. (2007), the use of probability as a mathematical model and as a tool in statistics takes not only different approaches, but also different kinds of reasoning. Franklin et al. (2007) add that ‘the two important uses of "randomisation" in statistical work occur in sampling and experimental design; when sampling we "select at random" and in experiments we randomly assign individuals to different treatments’. Randomisation thus leads to chance variability in outcomes that can be described with probability models. Therefore, when randomness is present, the statistician wants to know if the observed result is due to chance or something else (Franklin et al., 2007). Even though probability originates within mathematics, research studies have found that probability is more confusing, and have demonstrated that this confusion persists even among those who can recite the axioms of formal probability and do textbook exercises (Cobb & Moore, 1997; Garfield & Ahlgren, 1988). According to Moore (1988), it is unprofessional for mathematicians who lack training and experience in working with data to teach statistics. It is thus important that the problems encountered in the teaching of statistics (data handling and probability) are identified, and that teachers are helped with the teaching of statistics in schools.

2.3 The Teaching of Statistics (data handling and probability) in Schools

Statistics was introduced into the mathematics curriculum at school level in many countries at the beginning of the nineties (Watson, 1998). This was because of statements made in research studies in the past regarding the importance of statistical reasoning or statistical knowledge in society; its instrumental role in other disciplines; and the need for a basic knowledge in many professions (Batanero, Godino & Roa, 2004; Gal, 2002). Recommendations were also made that every high school graduate should be able to use sound statistical reasoning to cope intelligently with the requirements of citizenship, employment and family life and to be prepared for a healthy, happy and productive life (Franklin et al., 2007). Since then, many countries, including but not limited to the United Kingdom (UK), the United States of America (USA), Australia, Italy and New Zealand, have begun to include the teaching of data handling and probability in their school curricula. The introduction of these topics in the mathematics curriculum has posed challenges to teachers as many of them have not completed any courses in statistics in their preparatory degrees or in any pre-service education courses (Batanero, Burrill & Reading, 2011; Watson, 1998).

In this section, the researcher presents the contents of data handling and probability taught at secondary level in both developing and developed countries. The selected countries are South Africa, Australia and Malaysia.

2.3.1 Statistics in the Secondary School Curriculum in South Africa

In South African schools, statistics is not taught as a subject on its own but as an integral part of the mathematics curriculum; data handling and probability are the only two components of statistics taught in South African schools (see DoE, 2003).

2.3.1.1 Data handling

According to DoE (2003), learners in the Further Education and Training (FET) band (Grade 10-12) are required to:

1. Collect, organise and interpret univariate numerical data in order to calculate: measures of central tendency (mean, median, mode) of grouped and ungrouped data; measures of dispersion: range, percentiles, quartiles, interquartiles, variance, standard deviation and semi-interquartile range; errors in measurement; sources of bias; calculations of the regression function which best fits a given set of a bivariate numerical data; and calculation of the correlation coefficient of a set of bivariate numerical data to make relevant deductions.

2. Represent data effectively, choosing appropriately from: bar and compound bar graphs; histograms (grouped data); frequency polygons; pie charts; line and broken line graphs; ogives; box-and-whisker-plots; differentiate between symmetrical and skewed data and make relevant deductions; represent bivariate numerical data as a scatter plot and suggest intuitively whether a linear, quadratic or exponential function would best fit the data (see De Jager, Dewet & Raubenheimer, 2006; Department of Education, 2003; Laridon et al., 2004).

2.3.1.2 Probability

The contents of probability included in the mathematics curriculum in Further Education and Training (FET) band (Grade 10-12) (see DoE, 2003) are as follows:

1. Learners use probability models for comparing the relative frequency of an outcome; learners use Venn diagrams as an aid to solving probability problems, appreciating and correctly identifying: the sample space of a random experiment; an event of the random experiment as a subset of the sample space; the union and intersection of two or more subsets of the sample space; disjoint (mutually exclusive) events; complementary events; identifying dependent and independent events (e.g. from two-way contingency tables or Venn diagrams), and calculating the probability of two independent events occurring by applying the product rule for independent events; using tree diagrams to solve probability problems (where events are not necessarily independent).

2. Generalise the fundamental counting principle (successive choices from *m1* then *m2* then *m3* … options create *m1.m2.m3* … different combined options) and solve problems using the fundamental counting principle (see De Jager, Dewet & Raubenheimer, 2006; Department of Education, 2003; Laridon, et al., 2004).

2.3.2 Statistics in the Secondary School Curriculum in Australia

According to Australian Curriculum, Assessment and Reporting Authority (ACARA) (2 013), the contents of statistics and probability taught in the mathematics curriculum at senior secondary schools in Australia includes the following:

2.3.2.1 Probability

1. Counting and probability

i. Combinations
ii. Language of events and sets
iii. Review of the fundamentals of probability
iv. Conditional probability and independence

2.3.2.2 Statistics

1. Discrete random variables

i. Bernoulli distributions

ii. Binomial distributions

2. Continuous random variables and the normal distribution

i. General continuous random variables
ii. Normal distributions

3. Interval estimates for proportions

i. Random sampling: understand the concept of a random sample; discuss sources of bias in samples, and procedures to ensure randomness; use graphical displays of simulated data to investigate the variability of random samples from various types of distributions, including uniform, normal and Bernoulli.
ii. Sample proportions
iii. Confidence intervals for proportions

2.3.3 Statistics in Secondary School Curriculum in Malaysia

According to the Ministry of Education Malaysia (2004), the statistics taught at secondary level in Malaysia includes the following:

2.3.3.1 Data handling

- Collection of data.
- Frequency, frequency tables and class intervals.
- Pictograph, bar chart, pie chart and line graph.
- Histogram and frequency polygon.
- Cumulative frequency and the ogive.
- Measures of central tendency: mode, mean and median.
- Measures of dispersion: range and inter-quartile range.

2.3.3.2 Probability

- Sample spaces.
- Events.
- The probability of an event.
- The probability of complementary events.

- Combined events.
- The probability of combined events.

It is evident from this information on the teaching of statistics (handling and probability) in developed (i.e. Australia) and developing (i.e. South Africa and Malaysia) countries that there is not much difference in the aspects of statistics taught in senior school classes. For this reason, the current study seeks to explore problems, if any, encountered in the teaching and learning of statistics in schools in South Africa.

2.4 Problems Encountered in the Teaching and Learning of Statistics

2.4.1 Problems encountered in the teaching of statistics

In South Africa, as discussed in section 1.2, some mathematics teachers who are teaching statistics (data handling and probability) have never studied statistics before. Although prospective secondary school teachers may have a major in mathematics, they mostly study only theoretical (mathematical) statistics during their training (Batanero & Diaz, 2010). Nicholson and Darnton (2003) warn that statistics may pose particular problems if the teacher has not studied this before or if the teacher is not a statistics subject specialist; this is because the style of teaching mathematics and statistics differs, mainly with regard to the emphasis on concepts and ways of knowing (Batanero, Burrill & Reading, 2011; Garegae, 2008; Gould & Peck, 2004; Nicholson, Ridgway & McCusker, 2006).

Several studies have been conducted to investigate problems experienced by teachers when teaching statistics (Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya, Dhlamini & Makwakwa, 2010; Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya, & Makwakwa, 2009;Cardoso, 2007; Da Silva & De Queiroz e Silva Countinho, 2008; Garegae, 2008; Groth, 2009; Jacobbe, 2008; Sánchez & García, 2008; Wessels & Niewoudt, 2011). Garegae (2008) studied challenges faced by mathematics teachers when teaching statistics. Her study used an open-ended questionnaire that was completed by 23 senior teachers and 30 ordinary teachers to solicit their experiences in the teaching of statistics in high schools in Botswana. The data collected were analysed using Tesch (1990) and Bogdan and Biklen's (1992) techniques for analysing qualitative data. The results of the study showed that mathematics teachers in high schools faced problems in explaining statistical concepts to learners; solving statistical problems from past examinations papers; designing appropriate activities in statistics for learners; determining syllabus objectives; and relating the teaching of statistics to learners' real-life experiences. Garegae (2008) indicated that these problems were faced particularly by teachers who had not initially been trained to teach statististics.

Atagana et al. (2009) investigated topics perceived as difficult by learners and teachers in the teaching and learning of mathematics. These researchers adopted a descriptive survey design and used an educator questionnaire to collect the data on the sections of the NCS Grades 10-12 and Revised NCS Grades R-9 which teachers found difficult to teach at the Further Education and Training (FET) and General Education and Training (GET) level. One of the objectives of Atagana et al.'s (2009) study was "to identify the topics that teachers have difficulties with in teaching in Mathematics, Science, and Technology at Further Education and Training phase (grades 10 - 12) and General Education and Training phase (grades 8 and 9)" (p.9). The results of the study showed that data handling and probability were two of the topics that teachers experienced difficulties with, but the study involved Grade 12 learners. The current study will focus on grade 11 classes instead.

In Atagana et al.'s (2010) study, a highly improved and adjusted educator questionnaire was administered to a much more geographically diversified sample of teachers from more provinces. The results of the study also revealed data handling and probability as problematic topics for teachers.

Wessels and Nieuwoudt (2011) studied the profile of mathematics teachers' statistical knowledge, beliefs and confidence in order to inform the development of in-service teacher education programmes in statistics for grade 8 and grade 9 teachers. The profiling questionnaire was completed by ninety Grade 8 to 12 teachers from 23 diverse socio-economic schools in a large city in the north of South Africa, all with culturally diverse learner populations (Wessels & Nieuwoudt, 2011). They found that teachers showed high levels of confidence in teaching most statistics topics but low levels of statistical thinking when they had to apply their knowledge of concepts, such as sample and average, in social contexts, including newspaper articles and research reports. Wessels and Nieuwoudt (2011) also reported that more teachers indicated lower levels of confidence in aspects that involved sampling and probability topics (probability language, basic probability data handling calculations).

Da Silva and De Queiroz e Silva Coutinho (2008) explored how in-service secondary school mathematics teachers reasoned about variation of a univariate distribution. Their sample was composed of nine Brazilian teachers taking part in a teacher training course. All but one teacher had a degree in mathematics. All nine teachers taught mathematics in middle or high school. The study found that none of the teachers demonstrated complete reasoning, which would relate the understanding of mean, deviations from the mean, the interval of *k* standard deviations from the mean and estimation of frequency in this interval. The teachers' predominant reasoning about variation was verbal, with the understanding that standard deviation is a measure of sample homogeneity. Da Silva and De Queiroz e Silva Coutinho (2008) suggest that verbal reasoning about variation does not allow mathematics teachers to teach their learners the meaning of measures such as standard deviation, but restricts them instead to the teaching of algorithms. This was emphasised by one of the teachers in the sample who said that he only taught how to calculate the mean, median and the standard deviation, but he had never thought about how these concepts could be related (Da Silva & De Queiroz e Silva Coutinho, 2008). Reading (2004: p85) asserts that the teaching of measures of variariation, such as standard deviation, in schools is notorious among teachers as being particularly complex, resulting in many teachers having difficulty in teaching the concept to learners or simply avoiding it altogether.

Cardoso (2007) applied an activity organised into three stages to 29 high school teachers. A data set, two distributions represented by tables, and two represented graphically. Teachers were asked to analyse the data through association between the mean and standard deviation and between median and quartiles. Teachers were also asked to calculate the mean and standard deviation and to calculate the median and quartiles. The teachers then explained the meaning of the results obtained after calculating the summaries (the mean, standard deviation, median andquartiles). A discussion was established with the teachers on the values of summaries and their meaning. This disscussion revealed that teachers had difficulty in answering and in providing a clear, critical analysis, suggesting that they were unable to make any kind of oral or written analysis to justify their results. Teachers analysed concepts such as mean and median incorrectly. These findings were diagnosed through the observation that some teachers attributed symmetry to all data distributions, leading to a confusion between median and mean.

Da Silva and De Queiroz e Silva Coutinho (2006) investigated how teachers working in Basic Education apply concepts related to the idea of variation. They analysed the concept of variation as it was understood by 10 mathematics teachers, using activities derived from a research methodology conventionally known as Didactic Engineering. Qualitative analysis of the data revealed two different ways of understanding variation: the range and high value of variance or standard deviation as an indicator of a large variation in data. This suggested that participants could not construct the concept of variation around the mean and did not know how to analyse the values of variance or standard deviation, although they knew how to compute these variation measures.

Jacobbe (2008) examined elementary school teachers' understanding of mean and median. Three elementary teachers from the United States participated in the study over the course of 18 months. The participants were interviewed and also completed assessments and questionnaires. The questionnaires consisted of eight questions about the mean and the median. Jacobbe's study revealed that two of the three teachers had procedural knowledge of the measures of centre and one had conceptual understanding of the measure of centre. For example, when the three teachers were asked to describe what the mean represents and the difference between the mean and the median, the two with procedural knowledge could only describe how to calculate the mean and the median; they were unable to explain the conceptual difference between these two measures. On the other hand, the teacher with conceptual knowledge described the difference between the mean and the median but could not calculate the measures of centre. The study also found that teachers did not know when a particular measure was more appropriate or useful than another. The study suggested that teachers did not understand the connection between procedural and conceptual knowledge of measures of centre (Jacobbe, 2008).

Sánchez and García (2008) explored middle school teachers' concepts of statistical dispersion. Data were gathered from six middle school teachers in Mexico, using a questionnaire. The participants were asked (1) to imagine that a die was thrown 60 times and to predict the number of times each of the numbered sides would face up, and (2) to imagine that a die was thrown 1000 times and to predict the number of times each of the numbered sides would face up. Secondly, the teachers carried out a guided activity, answering the same problems from a written questionnaire using computer simulations. Lastly, the teachers were interviewed and videotaped. The findings indicated that these in-service teachers had difficulties in understanding the concept of dispersion. For instance, the study revealed that, before the guided activity, teachers' responses were almost deterministic, while after this activity they realised that each numbered face would appear in a range of values.

Groth (2009) studied characteristics of teachers' conversations about the teaching of mean, median and mode. Nine elementary and middle school teachers were asked to read an article about learners' difficulties in learning mean, median and mode, and then to discuss how the teaching of measures of central tendency could be improved in schools. Each teacher was asked to make at least four posts on an online discussion board and the conversation lasted for one week. Four teachers suggested that the meanings of the measures of centre should not be taught to learners, while two believed that these meanings should be taught. Again, two teachers suggested these measures should be taught using box-and-whisker plots, while a third teacher thought that technology would be needed to teach the meanings of these measures. In other words, the study showed that many believed that learners should not be taught conceptual knowledge about measures of central tendency. Instead, they should be taught procedural knowledge. The findings of Groth's study suggest that if learners have difficulties with understanding measures of central tendency, teachers' beliefs about not teaching the topics conceptually may contribute to their difficulties (Groth, 2009).

Bruno and Espinel (2009) investigated difficulties encountered by future primary school teachers in constructing and interpreting histograms and frequency polygons. Data were collected from 29 pre-service primary education teachers in Spain. The study used a written test with two questions. This test was given at the end of a one-hour mathematics class in a statistics module, during which the students had worked with descriptive statistics for some 10 hours, studying the statistical graphs (frequency histograms, frequency polygons, cumulative frequency polygons, box plots, bar diagrams, frequency polygons and cumulative frequency diagrams, stem-and-leaf plots and box plots, pie charts, bar diagrams and pictograms). The first question in the test asked future teachers to construct a histogram and a frequency polygon. The second question required them to evaluate the five graphs created by students (also future primary school teachers), labelling any mistakes they found and grading each diagram as good, average or poor. The study found that future primary school teachers experienced a great many conceptual and procedural difficulties in the construction of histograms and frequency polygons. These included separating histogram rectangles, inadequate labelling of real numbers on the axes, not considering zero frequency intervals or not completing the frequency polygon. In addition, when the same pre-service teachers were asked to assess the graphs constructed by other pre-service teachers and to identify any mistakes, the errors they had made in constructing their own graphs became evident (Bruno & Espinel, 2009).

It has also been found that teachers still use a traditional theoretical approach to teaching statistics, placing the emphasis on identifying the correct formula and performing calculations, rather than taking a data-driven approach using real data to explain statistical principles, procedures and reasoning (North, Scheiber & Ottaviani, 2010; Wessels & Nieuwoudt, 2011). Groth (2007) observed that, at times, mathematics teachers showed some of the cognitive difficulties experienced by learners in understanding statistical concepts and as a result these teachers "tend to focus on the mechanics of constructing plots rather than on interpretation, and on calculation rather than on data" (Watson, Burrill, Landwehr & Scheaffer, 1992:51). On the other hand, Iversen (1992) warns that statistics has its own subject matter which includes reasoning from uncertain data and making decision in the face of uncertainity and its origins differ from those of mathematics. This study intends to investigate issues embedded in the teaching of data handling.

It is clear from the preceding discussion there are problems associated with the teaching of statistics in schools. It is on this basis that the current study is being conducted with grade 11 learners. If there are any problems found in the teaching of statistics, solutions to these may be found with the view to improving the teaching of statistics.

2.4.2 Problems encountered in the learning of statistics

There is evidence that learners have conceptual difficulties in understanding descriptive statistics and simple statistical ideas such as distribution, average, sample and randomness (see Slauson, 2008; Sharma, 2006; Lee, 2003; Lue, 1998; Pollatsek, Lima & Well,1981). Pollatsek, Lima, and Well (1981) report on students' difficulties in understanding the need to weight data in computing a mean. College students were asked to combine two grade point averages that were based on different numbers of courses into a single average. Pollatsek et al. (1981) believe that "for many students dealing with the mean is a computational rather than a conceptual act" (p. 191) and they also believe that for many students the knowledge "of a computational rule not only does not imply any real understanding of the basic underlying concept, but may actually inhibit the acquisition of more adequate (relational) understanding" (p. 202). Johnson (1985) also found that most senior high school students regarded an average as the usual or typical value. (For example, when asked about the usefulness of the average temperature for a city, many responded that it would tell you what to wear if you went there.)

Behr, Lesh, Post and Silver (1983) mentioned that learners appeared to have difficulties developing intuition about fundamental ideas of probability for the reason that many learners have underlying difficulties with rational number concepts and proportional reasoning, which are used in calculating, reporting and interpreting probabilities (Behr, Lesh, Post & Silver, 1983).

Ben-Zvi and Garfield (2004) listed some of the other reasons for the challenge statistics poses to learners and teachers:

“(1) Many statistical ideas and rules are complex, difficult, and/or counterintuitive. It is therefore difficult to motivate learners to engage in the hard work of learning statistics;
(2) Many learners have difficulty with the underlying mathematics (such as fractions, decimals, proportional reasoning, algebraic formulas and functions) and this interferes with learning the related statistical concepts;
(3) The context in many statistical problems may mislead the learners, causing them to rely on their experiences and often faulty intuitions to produce an answer, rather than selecting an appropriate statistical procedure;
(4) Learners equate statistics with mathematics and expect the focus to be on numbers, computations, formulas and only one right answer. Learners are uncomfortable with the messiness of data, the different possible interpretations based on different assumptions, and the extensive use of writing and communication skills” (Ben-Zvi & Garfield, 2004: 4).

Sharma (2006) explored difficulties with statistical reasoning among form five (14 to 16 year-olds) learners. The study focussed on descriptive statistics, graphical representations and probability. The sample consisted of a class of 29 learners aged between 14 and 16 years, comprising 19 girls and 10 boys. A group of 14 learners participated in the interviews. The purpose of the study was to present and discuss the ways in which learners made sense of information in graphical representations (tables and bar graphs) obtained from the individual interviews. The findings revealed that many of the learners could read and compare data presented in a bar graph, but they were less competent at reading tables. Sharma noted that students' difficulties could be due to instructional neglect of the concepts or linguistic and contextual problems.

Lee (2003: 2330) identified four types of learner difficulties in constructing, interpreting and applying histograms in different real world contexts. These difficulties are:

(1) Learners perceive histograms as displays of raw data, with each bar representing an individual observation, rather than as presenting grouped sets of data,
(2) Learners tend to interpret histograms as two-variable scatter plots or as time sequence plots, (3) Learners tend to look at the vertical axes and compare differences in the heights of the bars when comparing the variation of two histograms,
(4) Learners tend to think deterministically when interpreting a distribution in real world contexts.

Slauson (2008) investigated students' conceptual understanding of variability by focusing on two numerical measures of variability: standard deviation and standard error. These were two sections of introductory statistics topics taught at a small Midwestern liberal arts college. One section was taught with standard lecture methods and the other by students completing a hands-on active learning laboratory for each of these topics. The students completed assessment questions designed to test conceptual knowledge at the end of each laboratory exercise and both classes completed the Comprehensive Assessment of Outcomes in a first Statistics course (CAOS) as a pre-test and a post-test. A group of students from each section participated in twenty-minute interviews which consisted of statistical reasoning questions. The study showed that even if students passed a standard introductory statistics course, they still lacked the ability to reason statistically. The analysis of the data revealed that students' conceptual understanding of ideas related to standard deviation improved in the active class, but not in the lecture class. The analysis of the qualitative data suggested that understanding the connection between data distributions and measures of variability, and understanding the connection between probability concepts and variability, are very important if students are to successfully understand standard error.

Ghinis, Korres and Bersimis (2009) examined the difficulties Greek senior high school learners identified in learning statistics and how these difficulties were related to the course's level of difficulty. The sample used in the study comprised 163 learners, 64 boys and 99 girls, from three different high schools: Ionnidios School in Piraeus (a sample of 90 students), Evaggeliki School in NeaSmirni (a sample of 38 students) and Hellenic-French School Jeanne D' Arc in Piraeus (a sample of 35 students). Learners completed questionnaires and the study found that learners considered the subject Statistics to be easy or of medium difficulty. Learners identified the difficulties they faced as mainly in the procedure of solving problems, in applying known statistical methodology to unfamiliar, real-life situations and problems and in performing the mathematical operations for extracting results. Learners did not identify difficulties in understanding basic statistical concepts nor in interpreting results. Furthermore, the results revealed that the majority of learners had a positive attitude towards the teachers' methods. However, a percentage of 25.8% were not satisfied. Learners identified the difficulties that teachers face mainly in the presentation and organisation of data in the classroom and in the processing of data. Their most important suggestions for the improvement of teaching included the use of computers, the orientation of the content of Statistics towards learners' preparations for the State Examinations, the use of slides and teachers being better informed about Statistics.

Meletiou-Mavrotheris and Lee (2010) investigated the prior knowledge of graphing that groups of undergraduate Cypriot and United States students brought into the introductory statistics classroom. A sample of 159 students completed a questionnaire designed to assess three aspects of graph comprehension: graph reading and interpretation, graph construction, and graph evaluation. The study findings showed that both Cypriot and United States students had difficulties in tackling more demanding tasks involving group comparison, graph construction, and critical evaluation of information presented graphically. The biggest differences were observed in simple reading and interpretation tasks (Meletiou-Mavrotheris & Lee, 2010).

It will be of interest to see how grade 11 learners at schools in urban townships experience statistics. Clearly, judging by the studies discussed above, learners do have difficulties with statistics. It is against this background that the current study is conducted.

2.5 Possible Causes of Problems Associated with the Learning and Teaching of Statistics

Research studies show that textbooks and curriculum documents prepared for primary and secondary teachers do not offer them enough support in teaching statistics (data handling and probability) (Batanero & Diaz, 2010; Garegae, 2008; Lue, 1998; Ortiz, Cañizares, Batanero & Serran, 2002; Serradó, Cardeñoso & Azcárate, 2005).

Garegae (2008) studied challenges facing mathematics teachers when they teach statistics. Her study used an open-ended questionnaire completed by 23 senior teachers and 30 ordinary teachers to solicit teachers' experiences about the teaching of statistics in high schools in Botswana. The data were analysed using Bogdan Tesch's (1990) and Biklen's (1992) techniques for analysing qualitative data. The results of her study indicated that there was a shortage of appropriate teaching materials. The participants indicated that prescribed textbooks were not only rare but also irrelevant.

Lue (1998) evaluated a descriptive statistics curriculum unit in the high school mathematics curriculum in Taiwan. There were 56 teachers and 301 learners who responded to the questionnaires, 268 learners who took a test and 32 college experts in the Taipei area who responded to general questions about the importance, necessity and appropriateness of a descriptive statistics unit integrated in the mathematics curriculum at high school level. Two of the objectives of the study were to explore the learners' problems and difficulties in learning descriptive statistics, and the major factors which might affect students' learning of descriptive statistics. The results of the study revealed that the major factors affecting the learning of descriptive statistics were lack of transparency of some statistical concepts in some sections in the textbook and the boring nature of the content. The study mentioned three factors that might affect the learning of descriptive statistics (data handling) by high school learners:

(a) Failure to use a calculator makes the calculations complicated and difficult,

(b) Learners find teaching materials boring, and

(c) Some content is presented in inadequate detail (i.e. teacher's manual is supposed to contain enough statistical activities and examples for teachers to use in the classroom).

Lue mention a further four factors that might affect the teaching of descriptive statistics (data handling):

(a) Technical difficulties, such as items needing complicated computations, mean that some statistical questions are left out of tests,

(b) The teaching materials are not adequate, and

(c) Learners' interest is limited.

Serradó, Cardeñoso and Azcárate (2005) carried out a content analysis of lessons dealing with chance and probability in a sample of textbooks aimed at pupils in Spanish Compulsory Secondary Education (12 to 16-year-old learners). Their sample included the full set of textbooks (for all educational levels) in four Spanish series used widely (20 books in total). For each textbook, they analysed the definitions, explanations, examples and activities included under two main topics: a) chance and randomness and b) probability. The analysis and comparison of these textbooks revealed that:

(1) Chance is modelled as basically synonymous with luck and randomness, and is related to the uncertainty of the event only;

(2) Textbooks do not clarify the meaning of terms such as unforeseeable, set, certain, impossible, convergence, etc., with the result that learners might assign inappropriate meanings and thus hinder the construction of probabilistic notions which are described by these terms, such as: random experiment, event and process, sample space, random sequence, stability of relative frequencies; and

(3) These textbooks included neither examples nor activities that would help learners to identify their misconceptions about the occurrence of chance in random series, which appear in the form of heuristics (the gambler's fallacy, representativeness of the sample misconception, "outcome approach", etc.).

It is hoped that the current study will shed some light on the causes of these types of problems.

2.6 Recommendations for Overcoming Difficulties

2.6.1 Overcoming problems in teaching statistics and probability

In order to overcome difficulties encountered in learning statistics and probability, Garfield (1995) recommended that: (1) activity-based courses and the use of small groups could help learners to overcome some misconceptions of probability and enhance the learning of statistics; (2) the use of computer simulations appears to guide learners towards more correct answers to a variety of probability problems; (3) the use of software that allows learners to visualise and interact with data appears to improve learners' understanding of random phenomena.

Garfield and Ahlgren (1988) recommended that teachers should: (1) teach descriptive statistics (also called data handling) on its own without relating it to probability; (2) point out to learners common misuses of statistics (say, in news stories and advertisements); (3) use strategies to improve learners' rational number concepts before approaching proportional reasoning; (4) recognise and confront common errors in learners' probabilistic thinking; (5) create situations requiring probabilistic reasoning that correspond to learners' views of the world. Ghinis, Korres and Bersimis (2009) also suggested that computers be used to improve the teaching of statistics. They also emphasised that teachers should have good content knowledge of statistics.

2.6.2. Training teachers to teach statistics

Khazanon (2005) suggested that the best way to address misconceptions of probability, professional development (In-service Education and Training (INSET) activities) should: (1) provide general information about misconceptions about probability and help teachers to overcome their own misconceptions; (2) help teachers to understand their learners' misconceptions and how these affect the learning of probability and data handling; and (3) show teachers approaches, strategies and techniques which they could use to improve the teaching of probability and to facilitate the resolution of misconceptions among their learners.

Batanero, Godino and Roa (2004), as cited in Papaieronymou (2010: 59) suggested that educators needed to provide better initial training to teachers by offering courses at the college level which are specific to the didactics of probability. Those courses should include: an introduction to the history of probability; information on statistics journals, associations and conferences; the study of fundamental probability concepts; readings from literature on heuristics and biases in probability, as well as learners' difficulties and misconceptions regarding probability; identification of the educational theories and teaching approaches, assessment, teaching resources, and the use of technology; and examples of projects that could be used when teaching probability (Batanero, Godino & Roa, 2004).

Any programme that prepares teachers of statistics should include and be aligned with current guidelines for teaching statistics at K-12 and college level, approved by the American Statistical Association (see The Guidelines for Assessment and Instruction in Statistics Education – GAISE) (Garfield & Everson, 2009). The Guidelines for Assessment and Instruction (GAISE) project has compiled a report on the six recommendations for teaching introductory statistics courses (Aliaga et al., 2010: 1). These six recommendations encourage teachers (educators) of statistics to: (1) emphasise statistical literacy and develop statistical thinking; (2) use real data; (3) stress conceptual understanding rather than mere knowledge of procedures; (4) foster active learning in the classroom; (5) use technology for developing conceptual understanding and analysing data; (6) use assessments to improve and evaluate learner learning (Aliaga et al., 2010). Since learners build their knowledge in an active way by solving problems and interacting with their classmates, Batanero, Godino and Roa (2004) believe that educators or trainers should use a contructivist and social aproach when training teachers, particularly if they want teachers to use the same approach in their teaching.

2.6.3. Knowledge needed to teach statistics

As far as Shulman (1986) is concerned, for teachers to be successful they have to confront both issues (of content and pedagogy) simultaneously, by embodying "the aspects of content most germane to its teachability" (p. 9). Content knowledge and pedagogical knowledge together form the basis of pedagogical content knowledge (PCK) (Erduran, Bravo & Naaman, 2007; Harrison, 2001; Mishra & Koehler, 2006). This knowledge is obtained through pre-service teacher training and in-service education and training.

Papaieronymou (2010) identified the suggested probability knowledge for secondary mathematics teachers through an examination of recommendations from four professional organisations, namely the American Mathematical Society (AMS), the American Statistical Association (ASA), the Mathematical Association of America (MAA) and the National Council of Teachers of Mathematics.

With respect to pedagogical content knowledge specific to probability, here commended that (1) teachers need to acquire an awareness and the ability to confront common probabilistic misconceptions and student difficulties relative to probability concepts (as suggested by the ASA, the MAA and the NCTM); (2) teachers need to be able to use technology to carry out simulations in order to illustrate probabilistic concepts (as recommended by all four of the professional organisations) and should also be able to use concrete objects such as dice, cards, and spinners to demonstrate probability concepts to students in the mathematics classroom (as suggested by the ASA and the NCTM); (3) secondary mathematics teachers should be able to represent probabilities using various models such as the area model (as suggested by the NCTM) (Papaieronymou, 2010: 363).

Again, with respect to curricular knowledge specific to probability, Papaieronymou (2010: 363) recommends that (1) secondary school mathematics teachers should be aware that they can use various computer programs (i.e. Fathom and Data Scope) in their mathematics classrooms when working with probability concepts (as suggested by the AMS). They should also understand the power of simulation as a technique which can be used to solve probability problems (as recommended by the MAA and the NCTM); (2) these teachers must be able to plan and conduct experiments and simulations, distinguish between experimental and theoretical probability, determine experimental probabilities, use experimental and theoretical probabilities to formulate and solve probability, and use simulations to estimate the solution to problems of chance; (3) secondary school mathematics teachers need to understand the law of large numbers and to be able to illustrate this using simulations, to help learners develop an understanding of probability as a long-run relative frequency; (4) with regard to theoretical probability, teachers should know about and be able to use both discrete and continuous probability distributions, and to understand probability distributions and the normal distribution; (5) they should also be able to use simulations to study probability distributions; (6) teachers should be introduced to fair games and be able to understand expected value; (7) secondary mathematics teachers should understand the concept of statistical significance, including significance levels and p-values, as well as the concept of confidence intervals, including margin of error; (8) secondary school mathematics teachers should be able to use counting techniques such as permutations and combinations to determine such (theoretical) probabilities; (9) such teachers should also understand and be able to calculate the probabilities of independent and dependent events, and of compound events made up of independent and dependent events; (10) teachers should use various representations such as area models and tree diagrams to aid learners in forming a better understanding of compound events; (11) teachers should be aware of the uses of probability in many fields and its misuse in newspapers and magazines. The current study is being pursued to expose problems encountered in the teaching and learning of statistics. It is hoped that recommendations will be made once the study has determined the problems that teachers and learners have with statistics.

2.7 Conceptual Framework

The current study follows the Systems Approach suggested by Joyce and Weil (1980). This approach postulates that the teaching and learning process has input and output goals. In order to achieve good results, the input must include suitable materials. Croninger, Rice, Rathbun and Nishio (2007) argue that the single most significant factor affecting the learner is the teacher; that the influences of teachers on learner achievement are both additive and cumulative. Croninger et al. (2007) also indicate that the impact of a well-prepared teacher on learner outcomes can outweigh factors such as learners' socioeconomic background, language background and minority status. The current study investigates the problems encountered in the teaching of a relatively newly introduced topic in the curriculum. In order to facilitate learning, teachers require both content knowledge; pedagogical knowledge (Flores et al., 2000; Iszak & Sherin, 2003; Wilen et al., 2004) and pedagogical content knowledge. When a new topic is included in the curriculum, teachers must receive the necessary training if they are to teach effectively.

In the current study, it is argued that teacher knowledge influences the depth of teaching and, in turn, the quality of learning. This study assumes a position that the knowledge of teachers who are under qualified can be upgraded through in-service education and training programmes, the use of high quality textbooks and by elevating teacher background (qualifications, subject major, teaching experience). The framework of this study is highlighted in figure 1. This figure illustrates the relationship between variables which determine the cause of problems encountered in the teaching and learning of statistics in grade 11. It shows that problems experienced by both teachers and learners during instruction are the result of teacher knowledge (i.e. the teacher's pedagogical knowledge, content knowledge and pedagogical content knowledge) which is, in turn, influenced by various factors. These factors include: in-service education and training programmes, quality of teaching and learning materials (textbooks) and teacher background (qualification, subject major, teaching experience). These are extraneous variables which must be controlled if the problems encountered in the teaching and learning of statistics are to be addressed.

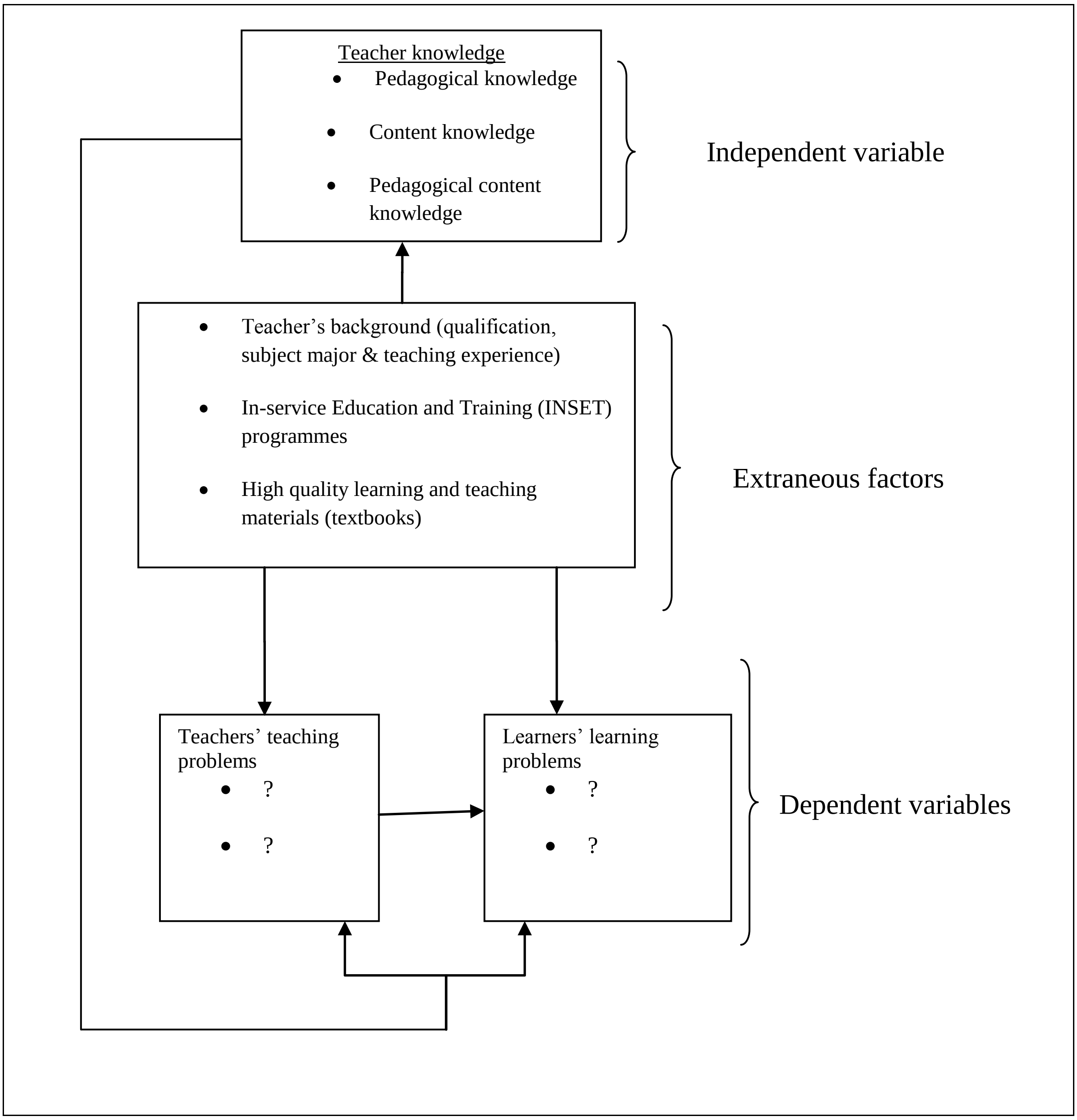


Figure 1: The conceptual framework highlighting problems in the teaching and learning of statistics

2.7.1 Statistics Teaching

The study has identified teachers' in-service education and training, teacher background and the quality of textbooks used as paramount in effective teaching. The discussion in this section focuses on each of these factors.

2.7.1.1 In-service Education and Training (INSET) programmes

RangaRao and BhaskaraRao, (2004) define in-service education and training (INSET) as continuing form of education offered to teachers after their initial professional training and leads to the improvement of their professional competence. Inset programmes may include courses and activities that incumbent teachers may undertake to improve their skills and knowledge. Inset helps teachers not only to learn new skills but also develop new insights into pedagogy and their own practice; provide knowledge and skills relating to emerging curricular changes - content, process and evaluation and to overcome knowledge gaps and deficiencies of pre-service education; improve quality and efficiency of the new educational systems and promote the scientific and technological advances and innovations; and prepare teachers for new roles (Ogbonnaya, 2007; National Council for Teacher Education, n.d.). It is argued that the introduction of statistics in the curriclum necessitates inset for teachers so that they could be effective when teaching it. Literature shows that teachers who attend inset tend to teach much more effectively and this, in turn, impacts positively on learner achievement (see, for example, Jamil, Atta, Ali, Balock & Ayaz, 2011; Naoreen, Aslam, Arshad & Nausheen, 2011; Khurshid, 2008; Farooq & Shahzadi, 2006; Angrist & Lavy, 2001; Parsad, Lewis, Farris & Greene, 2001). It is against this background that the current study is pursued.

Jafri and Shahzadi (2002) insist that poorly trained teachers can cause undue disturbance in the educational system for not being abreast with modern knowledge and latest trends regarding the educational methods. However, there are inset programmes that do not adequately address the developmental needs of teachers. These inset programmes tend to focus on lesson preparation (Moeini, 2008; Owusu-mensah, 2008) and learner discipline (Liston & Zeichner, 1990) without improving and updating the teachers' content knowledge. Shulman (1986) has suggested a framework for teacher development programme that includes three major domains, namely content knowledge (involves the mastery of specific content being taught); pedagogical knowledge (involves understanding the theories and principles of teaching and learning, understanding the learner, and the knowledge and principle of classroom behaviour and management); and pedagogical content knowledge (involves the ability to blend technique and content, including understanding how the given topics are related to one another and how they are most effectively organized and presented in the classroom). There have also been instances where inset programmes are criticized for not first identifying the needs of teachers before organising them (Owusu-mensah, 2008). Effective in-service education and training begins with an understanding of teachers' needs (Gaible & Burns, 2005; Moenie, 2008; Yigit, 2008) and then work towards addressing them.

In summary, an in-service education and training programme that is linked to subject matter, pedagogical knowledge and pedagogical content knowledge of teachers; has been designed according to the needs of teachers and runs over a reasonable period tends to be effective. In turn, effective inset improves teachers' professional competency. Hence, inset is an issue of interest in the current study.

2.7.1.2 Teachers' background

Teachers' background is another essential factor in the effective teaching of statistics. In this study teachers' background is used to encompass the teachers' qualifications (highest level of qualification in Mathematics and Statistics) and years of teaching experience. Teachers need to be appropriately qualified to teach statistics and also teachers' teaching improves as they get experienced.

2.7.1.2.1 Teachers' qualifications

Teachers' qualifications in this study were evaluated in terms of the education level of the teachers. This refers to the highest level of qualification obtained by the teacher in statistics. There are studies that have found that subject-specific teacher qualifications correspond positively with learners' performance. For instance, Goldhaber and Brewer (1998) found that subject specific degrees acquired have a positive impact on learner test scores. Armstrong (2009) found that subject specific degrees earned contribute most substantially to the overall performance of learners. Rice (2003) found that learners taught by a teacher with an advanced degree in a specific subject tend to achieve better in that subject. Furthermore, underlining the importance of subject-specific credentials, Goldhaber and Brewer (1997) found that mathematics scores of learners taught by teachers with master's degrees or certification in subjects other than mathematics were not different those of learners taught by teachers with lesser qualifications. It is on this basis that teachers' qualification are thought to contribute to their teaching. Hence, the current study has interest in teachers' qualification.

2.7.1.2.2 Teachers' teaching experience

Teachers were requested to indicate their years of teaching experience. The number of years of teaching is categorized into: Less than a year; 1 to 2 years; 3 to 4 years; and more than 5 years. Newly appointed teachers are less effective than those with some experience (Clotfelter, Ladd, & Vigdor, 2007; Betts, Zau, & Rice, 2003). Boyd, Lankford, Loeb, Rockoff and Wyckoff (2008) found that the largest gain in achievement is attributable to teacher experience. Ladd (2008, quoted in Rice, 2010) found that teachers with more than twenty years of experience are more effective than teachers with no experience.

In his study, Ost (2009) found that teachers with more grade-specific experience perform better in terms of their learners' test score gains. Furthermore, Ost (2009) found that a teacher who teaches the same grade for each of her first five years helps learners perform 0.137 standard deviations better than learners with a novice teacher, and that a teacher who teaches different grades every year for her first five years that teacher helps learners perform 0.0985 standard deviations better than a novice teacher. This implies that a teacher who always repeats grade assignments improves 39% more than a teacher who never repeats grade assignments. Rockoff (2004) found that teacher experience is a significant predictor of test scores for both reading subjects and mathematics computation, but not mathematics concepts. Thus, teachers' years of experience seem to relate to learners' achievement.

2.7.1.2.3 Textbooks

A textbook is one of the most important orientations for the teacher and learner and has a way of influencing teacher's work in its entirety (Haggarty & Pepin, 2002). Textbooks make it possible to establish a connection between the curriculum intentions and classroom activities (Schmidt, McKnight, & Raizen, 2002). Textbooks have an important role as a material in classroom context, because they are the major resource for content and pedagogical approaches, and it is also commonly assumed that textbooks (with accompanying teacher guides) are one of the main sources for the content covered and the pedagogical styles used in classrooms (Haggarty & Pepin, 2002). The teachers use the textbooks as a source of context and the way of teaching, and the learners use for classroom exercise and homework assignment (Fan, Chen, Zhu, Qiu, & Hu, 2004).

Remillard and Bryans (2004) indicated that most inexperienced and un(der)-qualified teachers, who lack subject matter knowledge, tend to rely on upon explicit guidance from textbooks about what and how to teach. Swanepoel (2010) stated three reasons why a teacher opts to follow a textbook rigidly: (i) an unqualified or inexperienced teacher can use a textbook to survive, because it can provide guidance, support and security, (ii) a well-qualified, experienced teacher can choose to make extensive use of a textbook because it is consistent with his or her own and researched-based views on education and (iii) a busy teacher may use it to save time. An effiecient textbook can play a valuable role in inexperienced and (un)der-qualified teachers' professional development as it can help teachers to keep ahead of most of thier learners and learn as they go (Hubisz, 2003; Newton & Newton, 2006). Lemmer, Edwards and Rapule (2008) showed that one of the most cost-effective ways to improve academic performance of learners and teachers is to improve the quality of textbooks. Most of the inexperienced and un(der)-qualified teachers consider the textbook as the correct and sometimes even as the only source of knowledge and follow it rigidly (Ogan-Bekiroglu, 2007; Arriassecq and Greca, 2007; Tarr, Ch´avez, Reys and Reys, 2006; Pepin and Haggerty, 2003).

Textbooks are expected to provide a framework for what is taught, how it may be taught, the sequence in which it can be taught, how it can be thoroughly explained (i.e. clarify the meaning of terms, provide examples and activities that could enable the learners to become aware of their misconceptions) (Lemmer, Edwards, & Rapule, 2008). In sum, a textbook is an important resource to promote teaching and learning. The unavailability of quality textbooks tends to impede effective teaching and meaningful learning. Hence the current study seeks to determine the extent to which textbooks promote or impede the teaching and learning of statistics.

2.7.2 Statistics Learning

It is generally accepted that good teaching leads to good learning (Ogbonnaya, 2011). Hence in most cases the focus is on ensuring that effective teaching takes place in our classrooms. That is teachers are adequately trained; appropriately qualified and provided with necessary materials to teach effectively. Even though teacher readiness to teach effectively may be the focus as in the current study, it also suffices to determine how learning takes place. In particular, the current study's interest is on how learners cope with statistics content and the extent to which they have learned statistics. It is argued that if teachers encounter problems with the teaching of statistics, learners in turn will have problems learning statistics. This argument is consistent with the views of Groth (2009), Da Silva and De Queiroz e Silva Coutinho (2008), Sharma (2006) and Reading (2004). It is on the basis of this argument that the current study determines how learners cope with the learning of statistics and extent to which the knowledge of statistics has developed.

2.7.3 Reflections on the Conceptual Framework

In sum, the study has been conceptualized around the notion that problems found in the teaching and learning of statistics emanate largely from the deficient teachers' content knowledge, pedagogical knowledge and pedagogical content knowledge. The deficiency seems to come about as a result of inadequate or lack of necessary in-service education and training for teachers when there are changes in the curriculum and the supply of requisite teaching and learning resources such as textbooks. When a changes are effected in a curriculum it is of paramount importance that teachers are familiarized with the amended aspects of the curriculum so that they can successfully implement newly look curriculum. In particular, teachers have to be properly qualified and necessary textbooks have to be supplied to optimize the successful implementation of the newly introduced aspects in the curriculum. Measures of successful implementation entail effective teaching and meaningful learning of newly introduced aspects. Otherwise, there is likelihood that there might be problems encountered in the teaching and learning of those new aspect(s) of the curriculum. In the current study the relatively new aspect introduced in the curriculum is statistics. Furthermore, teachers tend to teach better as they accumulate more experience. It should also be noted that learners tend to perform better if they are well taught. In particular, they encounter minimal problems in the learning of statistics content and their level of statistics improves vastly. It is against this background that the current study is pursued.

CHAPTER 3

METHODOLOGY

3.1 Research Approach and Design

This study adopted a quantitative research approach and followed a descriptive survey research design. In a survey study information is assessed on “attitudes, opinions and behaviour or characteristics of a population” (Creswell, 2011: 376). In the current study, teachers’ and learners’ views and opinions on problems encountered in the teaching and learning of statistics are explored. This design allowed the researcher to observe, interview and ask participants to complete a questionnaire and write a diagnostic test to uncover problems encountered in the teaching and learning of statistics in schools (Creswell, 2011; Cohen, Manion & Morrison, 2007). Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya & Makwakwa (2010) used this design to document what educators and learners perceived as difficult topics in mathematics, physical science and life science at the Further Education and Training (FET) Band. According to Creswell (2011: 376) descriptive survey designs “differ from experimental research in that they do not involve a treatment to participants by researcher. Because, survey researchers do not experimentally manipulate the conditions, and they cannot explain cause and effect”. Creswell (2011: 376) also indicates that survey research designs differ from correlational designs because their “focus is directed more toward learning about the population and less on relating variables or predicting outcomes, as is the focus in correlational research”.

3.2 Sampling

3.2.1 Population

The target population was grade 11 mathematics learners and their respective teachers from public[1] schools in the Gauteng Province of South Africa. Several studies (e.g. Khuzwayo, 2005) have reported that many learners who perform poorly in mathematics at grade 12 level come from public schools. The rationale for the population selection was: (1) no new content on data handling was covered in grade 12 mathematics; (2) it would be easier to conduct a

[1] Public schools are government-aided to some extent, where government provides the minimum, and parents contribute to basics and extras in the form of school fees (Education, 2010). Government also assists teachers from public schools to attend educational development programmes to improve their content knowledge.

study with grade 11 learners since grade 12 learners would be preparing for the end-of-the-year national examinations; and, other than that, probability topics are optional to all grade 12 mathematics learners and this might have limited the size of the sample in the study because some of these learners might have dropped probability as a topic in grade 12. It was on this basis that grade 11 was chosen for the study.

3.2.2 Sample and sampling technique

A convenient sample was used in this study, and consisting of 448 learners and 100 teachers who were based in 10 schools in the Tshwane South District of the Gauteng Province of South Africa. A convenient sampling technique was chosen for this study because it is normally a problem to get teachers, learners or schools to participate in a study of this nature. In convenient sampling "the researcher selects participants because they are willing and available to be studied" (Creswell, 2011: 145). Therefore a sample selected for this study is not a representative of the population (Cohen, Manion & Morrison, 2007). Nevertheless, the sample used in the current study is sufficient for the purpose of the study. A questionnaire (see section 4.1.2) was completed by all 448 learners, whilst 269 learners took part in classroom observations (see section 4.2), 248 learners wrote a diagnostic test and 10 learners were interviewed (see section 4.4). In the case of the teachers, 100 completed a questionnaire (see section 4.1.1), seven were observed teaching (see section 4.2) and six were interviewed (see section 4.4). Classroom observations were conducted in eight classrooms at four of the 10 schools involved in the study. Semi-structured interviews (both learners and teachers) were conducted from the four schools which were selected for classroom observations.

3.3 Instruments

This study used questionnaires (a teacher questionnaire and a learner questionnaire), a diagnostic test, a classroom observation schedule and semi-structured interviews (teacher interview and learner interview) to collect the data (see appendices 1, 2, 3, 4, 5 and 6).

3.3.1 Purpose of the instruments

The section outlines the purpose of each data collection instrument used in the study.

3.3.1.1 Questionnaires

Two forms of questionnaires were used in the study, namely learner questionnaires and teacher questionnaires. The subsection discusses the purpose of each of the questionnaires.

3.3.1.1.1 Learner questionnaire

A learner questionnaire was used to collect data on the problems learners encountered in the learning of statistics. It also provided learners with an opportunity to reflect on the difficulties their teachers face in the teaching of statistics. Ghinis, Korres and Bersimis (2009) used the same procedure to collect data from learners on problems which teachers faced in the teaching of statistics.

3.3.1.1.2 Teacher questionnaire

The purpose of the teacher questionnaire was to collect data on (1) the problems encountered in the teaching of statistics; (2) the causes of these problems; (3) teachers' suggestions on how to alleviate the problems encountered in the teaching of statistics.

3.3.1.2 Classroom observation

According to Cohen, Manion & Morrison (2007:398), observation enables researchers to gather "live data from naturally occurring social situations". Robson (2002:310) explains that observation "provides a reality check" because "what people do may differ from what they say they do". In this study, observation of learners and teachers in the classroom was conducted for the following reasons: (1) to identify the kind of difficulties encountered in the learning and teaching of statistics (data handling implied); (2) to check whether all the topics on data handling were being taught at schools; and (3) to discover things that learners and teachers do not feel free to mention in questionnaires (Cohen, Manion& Morrison, 2007). Probability lessons were not observed as probability was treated as an optional assessment topic and these topics are taught in the fourth semester (see Appendix 8), a time of the year when researchers are not allowed to collect data at schools (see the conditions in Appendix 9).

3.3.1.3 Diagnostic test

Diagnostic tests are designed to identify areas where a learner is experiencing difficulty with an academic skill (Johnson & Christensen, 2004). In this study, a diagnostic test was used to identify difficulties learners experienced in (1) applying the appropriate statistical methodology and formulae (these questions elicited data on whether learners were able to apply the appropriate methods and formulae for solving the statistical problems); (2) interpreting statistical results (these questions investigated the ability of learners to interpret plots and graphs); and (3) interpreting or reasoning about graphs, measures of central tendency and measures of dispersion etc.

3.3.1.4 Interviews

Teacher and learners were interviewed according the relevant schedule, respectively. The subsection presents descriptions of each schedule.

3.3.1.4.1 Learner interview schedule

Learner interviews were used to obtain further information and clarification on learners' responses to the questionnaires and diagnostic test. The interview helped to ascertain learners' feelings and concerns about the problems they experienced while learning statistics. Baloyi (2011:189) believes that, "unlike researchers, who in effect spend very little time with teachers through interviews or observations, learners usually spend a minimum period of a year in the company of their teacher". This means that learners can provide valuable information on the events within classroom environments. In this regard, the interview was confined to a group of ten learners.

3.3.1.4.2 Teacher interview schedule

Teacher interviews were also used as follow-up on certain responses to the teacher's questionnaire to obtain further information and clarification these answers. The interviews were confined to a group of seven teachers.

3.3.2 Development of Instruments

The section discusses the development of each data collection instrument used in the study.

3.3.2.1 Questionnaires

The subsection discusses the development of each questionnaire used in the study.

3.3.2.1.1 Learner questionnaire

The researcher developed a learner questionnaire consisting of sections A, B, C and D. Section A focuses on extracting learners' demographic information – gender and location (area). Section B deals with content related to data handling. Section C deals with the content related to probability which learners find difficult to master. Section D concentrates on the causes or reasons for the problems data handling and probability pose for learners (see Appendix 2).

3.3.2.1.2 Teacher questionnaire

The researcher developed a teacher questionnaire comprising five sections, A, B, C, D and E. Section A focuses on extracting teachers' demographic information – gender, grades taught, location (area), qualifications, experience teaching data handling and probability. Section B deals with content related to data handling and probability which teachers find difficult to teach. Section C covers content related to data handling and probability which learners find difficult and the causes of these problems. Section D focuses on extracting information about teachers' professional development. Section E concentrates on suggestions on how to solve the problems encountered in the teaching of statistics (see Appendix 1).

3.3.2.2 Diagnostic test

The researcher adapted and modified previous mathematics examination papers to develop a diagnostic test. The test consisted of long questions on data handling only (see Appendix 4). Probability questions did not form part of the diagnostic test as this topic is treated as optional and many schools do not teach it.

3.3.2.3 Interviews

The interview schedule for both teachers and learners was semi-structured (see appendices 5 and 6). Semi-structured questions do not offer choices from which the respondent can select an answer; rather, the question is phrased to allow for individual responses. The approach of such an interview provides a high degree of objectivity and uniformity and also allows probing and clarification (McMillan & Schumacher, 2001). Learner and teacher interview schedules were also developed from scratch by the researcher. The teacher and learner interview schedules elicited information from teachers and learners on the problems they encountered in the teaching and learning of data handling and probability; the causes of problems they encountered in the teaching and learning of statistics; and suggestions on how to address these problems.

3.3.2.4 Classroom observation

This study used a structured classroom observation schedule (see Appendix 3). The schedule was adapted and modified from a study by Community College Survey of Student Engagement (CCSSE) (CCSSE, 2006). The data collected from classroom visits corroborated the data gathered through other means.

3.4 Validity and Reliability of Instruments

The section discusses the validity and reliability of the data collection instruments used in the study.

3.4.1 Validity

3.4.1.1 Questionnaires

3.4.1.1.1 Teacher questionnaire

Content validity was verified by five experts in statistics and statistics education. Each expert was given a validity form (for example, see Appendix 7) to judge whether the questionnaire measured the intended content area of the study (Gay & Airasian, 2003). The general view of the experts was that the researcher should change the four point scale (No Problems, Few Problems, Average Problems, and Serious Problems) in questionss 8.1 – 11.13, as “It is unusual that teachers admit their weaknesses”. The four point scale in these questions was thus changed to Very competent, Competent, Slightly competent and Not at all competent. Except for this concern, the overall impressions were that the instrument would be able to serve its intended purpose. The changes were effected in the final teacher questionnaire (see Appendix 1).

3.4.1.1.2 Learner questionnaire

Content validity was assured by five experts in the statistics education and statistics field. Again, each expert was given a validity form to judge whether the questionnaire measured the intended content area of the study (cf. Gay & Airasian, 2003). Their suggestions were as follows: (1) “I suggest that you avoid double-barrelled questions in 3.1 – 3.8 by separating the concepts so that they are randomly distributed in the questionnaire”, and (2) “I suggest that you put Neutral in the middle to conform to Likert scale format”. Even though these suggestions were made, the experts’ overall impression was that the instrument was sound and covered all aspects necessary for the teaching and learning of statistics at Grade 11 level. The changes were effected in the final learner questionnaire (see Appendix 2).

3.4.1.2 Diagnostic test

The content validity of the diagnostic test was ascertained by the same five experts in the statistics education and statistics field. These experts determined the degree to which the test measured the intended content area (Gay & Airasian, 2003). The experts suggested that “Instructions, mark allocation for each question be shown”. The experts also suggested that “in order to establish more of the learners’ problems, it would be better to ask learners to give reasons for their choices”. Despite these suggestions, the experts’ overall impressions were that the diagnostic test covered all aspects necessary for the teaching and learning of statistics at Grade 11 level, that the time allocation was fair and the content covered adequately. The suggestions were implemented in the final diagnostic test instrument (see Appendix 4).

3.4.1.3 Interviews

The teacher interview instrument was validated by comparing it to a teacher questionnaire instrument which had already been shown to be valid, and the learner interview instrument was validated in the same way. This kind of comparison is known as convergent validity. Since the two instruments (teacher interview and teacher questionnaire; learner interview and learner questionnaire) were equivalent, it was assumed that the interviews were valid as they were comparable to the proven validity of the corresponding instrument (Cohen, Manion & Morrison, 2007).

3.4.1.4 Classroom observation

The classroom observation instrument was validated by comparing it to the teacher questionnaire instrument and learner questionnaire, that is, convergent validity was established. Since the teacher questionnaire and learner questionnaire had already been shown to be valid, it was assumed that the classroom observation instrument was valid (Cohen, Manion & Morrison, 2007).

3.4.2 Reliability

The learner questionnaire was piloted with a purposive sample of three hundred grade11 mathematics learners from another district in Gauteng Province. The aim of conducting a pilot study was to determine the feasibility of the study and to check the suitability of the data collection technique and instruments.

3.4.2.1 Questionnaires

The reliability of the teacher and learner questionnaire was determined by computing Cronbach's (1951) alpha as a measure of the internal consistency of scores from the questionnaires. The Cronbach's alpha (α) of the teacher questionnaire as a whole was found to be 0.723 (see Appendix 10), and Cronbach's alpha (α) of the learner questionnaire as a whole was found to be 0.938 (see Appendix 11). The values of the reliability coefficients ranged from 0 to 1.0. A coefficient of 0 means no reliability while 1.0 indicates perfect reliability. Generally, if the reliability of a test is above 0.80, it is said to have very good reliability; if it is below .50, it would not be considered a very reliable test (George & Mallery, 2003). The values of 0.723 and 0.938 were regarded as excellent, based on the rule of thumb.

3.4.2.2 Diagnostic test

Stability, also called test-retest reliability, was used to determine the reliability of the diagnostic test. Stability reliability is the degree to which scores on the same test are consistent over time (Gay & Airasian, 2003). The diagnostic test was piloted with 40 grade 11 mathematics learners from another district of Gauteng Province. Stability reliability was determined by administering the test to 40 learners, and again to the same group after 15 days. The two sets of scores were correlated. The correlation coefficient is significant at 95%

or a higher interval (Cohen, Manion, & Morrison, 2007). The test showed significant correlation of 0.782 obtained at 99% confidence interval (see Appendix 12). The Spearman Brown formula, $R=2r/(1+r)$, was used and gave a reliability of 0.878. This showed that the diagnostic test was reliable.

3.4.2.3 Interviews

3.4.2.3.1 Learner interview schedule

Test-retest reliability was used to determine the reliability of the learner interview instrument. This was established by interviewing eight learners from another district in Gauteng Province. The consistency of their responses established that the learner interview instrument was reliable.

3.4.2.3.2 Teacher interview schedule

Test-retest reliability was also used to determine the reliability of the teacher interview instrument. This was established by interviewing four teachers from another district in Gauteng Province. The teachers' responses were consistent, indicating that the teacher interview instrument was reliable.

3.4.2.4 Classroom observation

Reliability of the classroom observation was determined through a process of repeated usage of the classroom observation schedule.

3.5 Data Collection

Data were collected in two phases:

Phase One: Questionnaires

Firstly, data were collected from teachers' and learners' responses to the questionnaires. The researcher took the questionnaires to the 10 schools where they were completed. This was done (1) to ensure that all the questionnaires reached schools at the same time and (2) to ensure that all the questionnaires were given to the right people at the schools. All the questionnaires (for teachers and learners) had a cover page stating the purpose of the study and requesting the participants not to write their names on the questionnaires (see Appendices 1 and 2).

Phase Two: **Classroom observation, Diagnostic test and Interviews**

Phase two consisted of three successive stages of data collection as described below. The researcher visited the schools at the time that data handling was being taught. According to the 2011 Gauteng Province Teacher Work Schedule for grade 11 mathematics, data handling was supposed to be taught over a two-week period, a week after the schools had reopened for term 3 (see Appendix 8). The researcher timed her visits to coincide with the statistics lessons. The stages were as follows:

Stage1: **Classroom observation**

The classroom observation schedule was used to check the extent to which teachers' content knowledge was evident, to observe teachers' method of instruction, to observe how they presented their lessons and their method of teaching, etc. (see Appendix 3). It was also used to observe learners' behaviour, to observe whether learners asked questions or participated in the lesson; and also how learners went about lesson activities (see Appendix 3). Classroom observations were conducted in eight classrooms at four of the 10 schools involved in the study. The schools were those that were easily and conveniently accessible to the researcher. Field notes were recorded during these classroom observations. Only seven aspects of data handling (ogive, measure of central tendency, five number summary, stem-and-leaf plot, box-and-whisker plot, measure of dispersion, and scatter plot and line of best fit) were observed. Each aspect of data handling was observed by the researcher until it was agreed between the learners and teacher that it had been exhausted. Therefore, the total time spent during the classroom observations was 2208 minutes (see table 16).

Stage 2: Diagnostic test

Once a classroom observation had been completed, the researcher administered a diagnostic test to the learners to determine the extent of their content knowledge. Two hundred and forty-eight (248) learners wrote the test once the teaching of data handling had been completed by all teachers in the sample. The researcher administered it herself to ensure that it was written according to the requirements. A diagnostic test was administered in the same seven classrooms at four of the 10 schools in which the classroom observations had been conducted. All four schools wrote the test on the same day.

Stage 3: **Interviews**

After the diagnostic test stage, a group of six teachers and 10 learners from the four schools where lessons had been observed were interviewed to elicit their views on the problems they encountered in statistics. These groups of six teachers and 10 learners were chosen because they were willing and available to be interviewed. The Researcher conducted the interviews in separate room provided. For teachers interviews were done in English while for learners a combination of the learners' home language and English was used as this enabled effective communication. A Dictaphone was used to record these interviews, which took place during break or at lunchtime.

3.6 The Ethics of Research

Ethics are generally concerned with beliefs about what is right or wrong, proper or improper, good or bad. In research, the ethical issues are dealt with in such a way as to protect the rights and welfare of subjects (McMillan & Schumacher, 2001). The foremost rule of ethics in research is that subjects (participants) should not be harmed in anyway, either physically or mentally (Gay & Airasian, 2003).

In order to ensure that I complied with the code of ethics, all participants were informed about the nature and purpose of the study. They were assured that their participation in the study was voluntary, and that they had the right to refuse to participate and the right to end their involvement at any time without any penalty. Participants were informed of the purpose of the study before they agreed to take part.

Furthermore, to avoid possible harm to the participants, data about them remained confidential. The names of participants and the names of their schools were not revealed in the study; rather, codes were used instead of names. Confidentiality and anonymity was explained to them in the letters of informed consent.

In addition, protocol was observed both in obtaining permission to conduct the research and when requesting participants' agreement to be part of the study. Because the study involved teachers and learners at various secondary schools, formal procedures were followed by requesting permission from the Department of Education (see Appendix 9) to conduct research in schools. The researcher wrote letters to the District Office and to school principals to seek their permission to conduct the study (see Appendices 13 and 14). All the letters, together with the study proposal, were submitted to the university's Ethics Committee for final scrutiny and approval. Lastly, data were not falsified and all findings reported in the study were as revealed by the results.

3.7 Method of data analysis

To determine the descriptive statistics (frequencies, percentages, Mean and standard deviation) of the learner questionnaires, teacher questionnaires and to compute the scores of the diagnostic tests, the researcher used the computer program Statistical Package for Social Sciences (SPSS) version 20.0. The data generated from the use of the questionnaires were analysed using frequencies, percentages, mean and standard deviation to answer research questions. For research question one, the mean of the four point scale was used to determine level of competency as indicated below.

Response Categories	**Value**	**Limit scale**
Very well competent	4	3.50 - 4.00
Well competent	3	2.50 - 3.49
Slightly competent	2	1.50 - 2.49
Not at all competent	1	1.00 - 1.49

NOTE: For instance any mean that falls within 3.50 – 4.00 was classified as very well competent.

For research question two, the four point scale was used to determine level of difficulties as indicated below.

Response Categories	**Value**	**Limit scale**
Highly difficult	4	3.50 - 4.00
Difficult	3	2.50 - 3.49
Less difficult	2	1.50 - 2.49
Not difficult	1	1.00 - 1.49

NOTE: For instance any mean that falls within 1.50 - 2.49 was classified as less difficult.

Teacher and learner interviews were transcribed and classified according to themes. Classroom observations were also analysed using themes, checking similarities and differences with the themes emerging from the interviews. A detailed explanation of the analysis of the data from the classroom observations and interviews is provided in section 4.2 and section 4.4 respectively.

CHAPTER 4

DATA ANALYSIS

This chapter presents the analysis of the data collected from the learner and teacher questionnaires, the classroom observations, learner and teacher interviews and the diagnostic test. Data from these instruments were analysed independently.

4.1. Questionnaires

4.1.1. Teacher questionnaires

4.1.1.1 Teachers' demographic information

A total of 100 grade 11 teachers completed the questionnaire. Table 1 reflects that the majority (58%) of the teachers were males. It was also observed that 59% of the teachers had a diploma. Almost 65% had a minimum of college mathematics, implying that about 35% of the teachers could be regarded as not having adequate knowledge of mathematics. The table also shows that 36% of the teachers taught data handling and probability at grades 10, 11 and12 levels. Furthermore, it was evident that 53% of teachers had not studied statistics during their pre-service training, while 30% had studied elementary statistics during their pre-service training. It can be seen from the table that 41% of teachers had been teaching data handling for more than five years, compared to probability, where only 27% of the teachers had taught it for five years. The results indicated that 19% of the teachers had taught probability for less than a year, compared to data handling, where only 6% of the teachers had taught for less than a year. It can be concluded from these findings that more than half the teachers did not have the knowledge necessary to teach statistics.

Table 1: Data reflecting teachers' demographic information (n=100)

	Percentage	
Gender		
Male	58.0	
Female	42.0	
Highest Qualification		
Certificate	-	
Diploma	59.0	
Bachelor's Degree	39.0	
Master's Degree	2.0	
Doctorate	-	
Highest Mathematics Qualification		
Grade 12 or lower	8.0	
Mathematics diploma	46.0	
Mathematics I	7.0	
Mathematics II	19.0	
Mathematics III	18.0	
Postgraduate	2.0	
Currently teaching Data handling and Probability to		
Grade 10, grade 11 & 12	36.0	
Grade 10 & 12 only	-	
Grade 10 & 11 only	28.0	
Grade 11 & 12 only	17.0	
Grade 10 only	-	
Grade 11 only	19.0	
Grade 12 only	-	
Highest qualification in statistics		
Never studied statistics	53.0	
Statistics I	30.0	
Statistics II	8.0	
Statistics III	8.0	
Postgraduate	1.0	
Years teaching statistics topics		
	Probability	**Data handling**
Less than a Year	19.0	6.0
1 to 2 years	24.0	22.0
3 to 4 years	30.0	31.0
More than 5 years	27.0	41.0

4.1.1.2Problems experienced by teachers when teaching data handling

Table 2 explores the problems teachers encountered when teaching data handling. A four-point scale was used where teachers indicated their degree of competency in teaching aspects of data handling such as (1) interpretation and determination of measures of central tendency (items 8.1, 8.2, 8.3 and 8.4); (2) interpretation and determination of measures of dispersion (items 8.5, 8.6, 8.7 and 8.8); (3) representation of data using graphs or plots (items 8.9, 8.10, 8.11, 8.12, 8.13, 8.14, 8.15, 8.16 and 8.17). Then overall means of teacher responses were determined for each degree of competency and the limit scale as indicated in section 3.7 determined whether each of the teachers responses were not at all competent, slightly competent, well competent and very well competent.

Table 2: Teachers' problems in teaching data handling

Items	Category1: Interpretation and determination of measures of central tendency	Frequencies of teachers				Mean	Standard deviation (SD)	Remark
		Not at all Competent	Slightly competent	Well Competent	Very well competent			
8.1	Explaining the meaning of measures of central tendency (mean, mode, median) to learners.	1	4	32	63	3.58	.589	Very well competent
8.2	Teaching learners to do the calculations of measures of central tendency (mean, mode, median) with ungrouped data.	1	7	34	58	3.48	.703	Well competent
8.3	Teaching learners to do the calculations of measures of central tendency (mean, mode, median) with grouped data.	1	8	46	45	3.35	.672	Well competent
8.4	Interpreting the measures of central tendency (mean, mode, median) for learners.	3	11	52	34	3.18	.716	Well competent
Category 2: Interpretation and determination of measures of dispersion								
8.5	Explaining the meaning of measures of dispersion (standard deviation, range, and variance) to learners.	1	30	42	27	2.94	.802	Well competent
8.6	Teaching learners to do the calculations for measures of dispersion (standard deviation, variance, range).	2	23	47	28	3.01	.772	Well competent
8.7	Interpreting the measures of dispersion (standard deviation, variance, range) for learners.	2	31	46	21	2.87	.747	Well competent
8.8	Teaching learners the calculations of contents of five number summary (lower quartile, middle quartile, upper quartile).	1	11	31	57	3.44	.729	Well competent
Category 3: Representation of data using graphs or plots								
8.9	Teaching learners to make drawings of box-and-whisker diagrams on a number line.	1	13	32	54	3.39	.751	Well competent
8.10	Teaching learners to construct stem-and-leaf plot.	2	11	33	54	3.39	.764	Well competent
8.11	Teaching learners to use stem-and-leaf plot to determine the quartiles (lower quartile, median and upper quartile).	3	18	41	38	3.14	.817	Well competent

8.12	Constructing cumulative frequency tables for learners.	2	12	38	48	3.32	.764	Well competent
8.13	Making a drawing of an ogive (cumulative frequency curve) for learners.	2	18	40	40	3.18	.796	Well competent
8.14	Interpreting an ogive (cumulative frequency curve) for learners.	3	26	41	30	2.98	.829	Well competent
8.15	Teaching learners to represent bivariate numerical data as a scatter plot.	8	29	34	29	2.84	.940	Well competent
8.16	Determining the line of best fit on a scatter plot for learners.	8	31	30	31	2.84	.961	Well competent
8.17	Teaching learners to select a function that best fits the data.	9	32	34	25	2.75	.936	Well competent

The data presented in category 1 of table 2 indicates that teachers were well competent with the interpretation and determination of measures of central tendency that is: Explaining the meaning of measures of central tendency (mean, mode, median) to learners; teaching learners to do the calculations of measures of central tendency (mean, mode, median) with ungrouped data; teaching learners to do the calculations of measures of central tendency (mean, mode, median) with grouped data; and interpreting the measures of central tendency (mean, mode, median) for learners.

The data in category 2 of table 2 also shows that teachers were well competent with the Interpretation and determination of measures of dispersion that is: Explaining the meaning of measures of dispersion (standard deviation, range, and variance) to learners; teaching learners to do the calculations for measures of dispersion (standard deviation, variance, range); interpreting the measures of dispersion (standard deviation, variance, range) for learners; and teaching learners the calculations of contents of five number summary (lower quartile, middle quartile, upper quartile).

Further, the data in category 3 of table 2 shows that teachers were well competent with the representation of data using graphs or plots that is: Teaching learners to make drawings of box-and-whisker diagrams on a number line; teaching learners to construct stem-and-leaf plot; teaching learners to use stem-and-leaf plot to determine the quartiles (lower quartile, median and upper quartile); constructing cumulative frequency tables for learners; making a drawing of an ogive (cumulative frequency curve) for learners; interpreting an ogive (cumulative frequency curve) for learners; teaching learners to represent bivariate numerical data as a scatter plot; determining the line of best fit on a scatter plot for learners; and teaching learners to select a function that best fits the data.

Therefore, the results in table 2 suggest that teachers had no problems with the teaching all the topics listed in table 2. That is teachers did not experience any problems with interpretation and determination of measures of central tendency; interpretation and determination of measures of dispersion; and (3) representation of data using graphs or plots.

4.1.1.3 Problems experienced by teachers when teaching probability

Table 3 explores the problems teachers encountered when teaching probability. A four-point scale was used where teachers indicated their degree of competency in teaching aspects of probability such as: (1) construction and interpretation of probability diagrams and tables (items 9.1, 9.2, 9.3, 9.4, 9.5 and 9.6) and (2) the understanding or interpretation of probability terminology (items 9.7, 9.8, 9.9, 9.10, 9.11, 9.12 and 9.13). Then overall means of teacher responses were determined for each degree of competency and the limit scale as indicated in section 3.7 determined whether each of the teachers responses were not at all competent, slightly competent, well competent and very well competent

Table 3: Teachers' problems in teaching probability

Items	Category 1: Construction and interpretation of probability diagrams and tables	Frequencies of Teachers				Mean	Standard deviation (SD)	Remark
		Not at all Competent	Slightly competent	Well competent	Very well Competent			
9.1	Teaching learners to construct Venn diagram from a given word problem.	17	37	29	17	2.46	.968	Slightly competent
9.2	Teaching learners to construct two-way contingency tables from a given word problem.	20	48	21	11	2.23	.897	Slightly competent
9.3	Teaching learners to construct tree diagrams from a given word problem.	22	41	21	16	2.31	.992	Slightly competent
9.4	Teaching learners to use two-way contingency tables for problem solving.	24	48	18	10	2.14	.899	Slightly competent
9.5	Teaching learners to use Venn diagrams for problem solving.	23.0	41.0	22.0	14.0	2.27	.973	Slightly competent
9.6	Teaching learners to use tree diagrams for problem solving.	21.0	41.0	25.0	13.0	2.30	.948	Slightly competent
Category 2: The understanding or interpretation of probability terminology								
9.7	Teaching learners to identify dependent and independent events from Venn diagrams.	22	35	34	9	2.30	.916	Slightly competent
9.8	Teaching learners to identify dependent and independent events from two-way contingency tables.	20	42	29	9	2.27	.886	Slightly competent
9.9	Teaching learners to use Venn diagrams to solve probability problems where events are not necessarily independent.	24	44	25	7	2.15	.869	Slightly competent
9.10	Teaching learners to use tree diagrams to solve probability problems where events are not necessarily independent.	25	40	27	8	2.18	.903	Slightly competent
9.11	Teaching learners to calculate the probability of two independent events by applying product rule for independent events: P(A and B) = P(A). P(B)	24	42	19	15	2.25	.989	Slightly competent

9.12	Teaching learners to identify mutually exclusive events from Venn diagrams.	27	40	21	12	2.18	.968	Slightly competent
9.13	Teaching learners to differentiate between independent and dependent events.	25	44	19	12	2.18	.947	Slightly competent

The data in category 1 of table 3 above revealed that teachers were slightly competent with the construction and interpretation of probability diagrams and tables that is: Teaching learners to construct Venn diagram from a given word problem; teaching learners to construct two-way contingency tables from a given word problem; teaching learners to construct tree diagrams from a given word problem; teaching learners to use two-way contingency tables for problem solving; teaching learners to use Venn diagrams for problem solving; and teaching learners to use tree diagrams for problem solving.

Also, the data in category 2 of table 3 indicates that teachers were slightly competent with the following: Teaching learners to identify dependent and independent events from Venn diagrams; teaching learners to identify dependent and independent events from two-way contingency tables; teaching learners to use Venn diagrams to solve probability problems where events are not necessarily independent; teaching learners to use tree diagrams to solve probability problems where events are not necessarily independent; teaching learners to calculate the probability of two independent events by applying product rule for independent events: P(A and B) = P(A). P(B); teaching learners to identify mutually exclusive events from Venn diagrams; and teaching learners to differentiate between independent and dependent events.

Therefore, the results in table 3 suggest that teachers had problems with all the topics listed in table 3. That is teachers had problems with the construction and interpretation of probability diagrams and tables and also with interpretation of probability terminologies.

4.1.1.4 The cause(s) of problems teachers encounter in teaching statistics

Table 4 explores the cause of the problems related to the teaching of statistics (data handling and probability). A five point likert scale was used where teachers indicated of agreement or disagreement with the issues linked to (1) teachers' content knowledge of statistics (items

10.1, 10.4, 10.5, 10.13), (2) teachers' attitudes toward the teaching of statistics (items 10.2, 10.3, 10.6, 10.12), (3) in-service education and training programmes (items 10.7, 10.8) and (4) teaching material (items 10.9, 10.10,10.11). Then the overall percentages of teacher responses were determined for each level of agreement or disagreement.

Table 4: Cause of problems in teaching statistics

Causes of problems in the teaching of data handling and probability		Percentage of Teachers				
		Strongly disagree	**Disagree**	**Neither disagree nor agree**	**Agree**	**Strongly Agree**
Items	**Category 1:Teachers' content knowledge of statistics**					
10.1	I do not have statistics content knowledge/ I did not study statistics.	23.0	35.0	13.0	22.0	7.0
10.4	I do not have enough experience in the teaching of statistics.	20.0	32.0	13.0	28.0	7.0
10.5	There are different types of problems in statistics and I lack the problem-solving skills to deal with them.	14.0	32.0	14.0	30.0	10.0
10.13	I am still using a teacher-dominated method to teach as opposed to the recommended learner-centred approach.	20.0	46.0	16.0	18.0	0.0
	TOTAL ROW % OF CATEGORY 1	**19.25**	**36.25**	**14**	**24.5**	**6**
	Category 2:Teachers' attitudes toward the teaching of statistics					
10.2	I do not like teaching statistics.	44.0	44.0	8.0	3.0	1.0
10.3	I do not see the importance of statistics (data handling and probability) in the syllabus.	51.0	42.0	4.0	2.0	1.0
10.6	I cannot give myself enough time to prepare properly for my teaching of statistics because I am committed in other learning areas.	32.0	43.0	12.0	11.0	2.0
10.12	My learners do not pay enough attention when I am teaching statistics.	22.0	51.0	22.0	4.0	1.0
	TOTAL ROW % OF CATEGORY 2	**18.33**	**39.67**	**17**	**20.67**	**4.33**
	Category 3: in-service education and training programmes					
10.7	Most of the in-service training programmes that I have attended did not cover statistics topics	15.0	39.0	14.0	24.0	8.0
10.8	My school does not support my attendance of teacher development programmes to improve my statistics knowledge.	24.0	44.0	14.0	16.0	2.0
	TOTAL ROW % OF CATEGORY 3	**19.5**	**41.5**	**14.0**	**20.0**	**5.0**
	Category 4: Teaching material					
10.9	The textbooks do not explain thoroughly (i.e. formulae are missing) and do not provide enough examples.	15.0	34.0	18.0	24.0	9.0

10.10	I do not seek assistance from other teachers when I experience problems.	25.0	50.0	12.0	12.0	1.0
10.11	I do not have sufficient teaching resources to teach statistics.	15.0	35.0	21.0	26.0	3.0
	TOTAL ROW % OF CATEGORY 4	**18.33**	**39.67**	**17**	**20.67**	**4.33**

The results reflected in category 4 of table 4 reveal that 25 % (i.e. 20.67% of *agree* + 4.33% of *strongly agree*) of teachers agreed that they encountered problems because of inadequate teaching material. In this category, 33% (i.e. 24% of *agree* + 9% of *strongly agree*) of them claimed that the textbooks they were using did not explain thoroughly and did not provide enough examples; 29% (i.e. 26% of *agree* + 3% of *strongly agree*) of the teachers agreed that they did not have adequate teaching resources to teach statistics; and 13% (i.e. 12% of *agree* + 1% of *strongly agree*) of teachers agreed that they did not talk to other teachers when they found themselves experiencing difficulties.

It is also revealed in category 1 of table 4 that 30.5% (i.e. 24.5% of *agree* + 6% of *strongly agree*) of the teachers agreed that they encountered problems because they lacked content knowledge in statistics. In this category, 29% (i.e. 22% of *agree* + 7% of *strongly agree*) of teachers said this was because they had not studied statistics; 35% (i.e. 28% of *agree* + 7% of *strongly agree*) did not have enough experience of teaching statistics; 40% (i.e. 30% of *agree* + 10% of *strongly agree*) lacked problem solving skills; and 18% (i.e. 18% of *agree*)were yet to adopt the recommended learner-centred approach.

In addition, the data in category 2 of table 4 show that 6.5% (i.e. 5% of *agree* + 1.5% of *strongly agree*) of teachers encountered problems resulting from their own attitude towards teaching of statistics. In this category, 4% (i.e. 3% of *agree* + 1% of *strongly agree*) of teachers did not enjoy teaching statistics; 3% (i.e. 24.5% of *agree* + 6% of *strongly agree*) did not see the importance of teaching it; 13% (i.e. 11% of *agree* + 2% of *strongly agree*)admitted that they did not allow enough time to prepare properly for their teaching of statistics because of their commitments in other learning areas; and 5% (i.e. 4% of *agree* + 1% of *strongly agree*)of them encountered problems because their learners did not pay attention in class.

Lastly, the data in category 3 of table 4 show that 25% (i.e. 20% of *agree* + 5% of *strongly agree*) of the teachers encountered problems regarding the in-service education training programmes. In this category, 32% (i.e. 24% of *agree* + 8% of *strongly agree*) of the teachers agreed that they encountered problems because the programmes they had attended did not cover statistics topics, and 18% (i.e. 16% of *agree* + 2% of *strongly agree*) said they were not encouraged by their schools to attend these programmes to improve their statistics knowledge.

Therefore, the results in table 4 point to the major causes of problems encountered in the teaching of statistics as being: (1) teachers' lack of statistical content and pedagogical knowledge, (2) inadequate in-service education training programmes, as most of these programmes did not cover statistics topics, (3) teaching material. Teachers claimed that the textbooks were not explicit enough and that they lacked examples.

4.1.1.5 Teachers' opinions about learners' problems with data handling

In order to gather information on the problems encountered in the learning of data handling, the researcher asked teachers to indicate in the questionnaire, as indicated in table 5 whether learners had problems with: (1) the interpretation and determination of measures of central tendency (items 11.1, 11.2, 11.3 and 11.4), (2) the interpretation and determination of measures of dispersion (items 11.5, 11.6, 11.7 and 11.8), and (3) the use of graphs for the prediction of results (items 11.9, 11.10, 11.11, 11.12 and 11.13). A four point scale was used where teachers indicated their learners' level of difficulty in learning aspects of data handling.

Table 5: Teachers' opinions about learners' problems in the learning of data handling

Items	Category 1: The interpretation and determination of measures of central tendency	Frequencies of Teachers				Mean	Standard deviation (SD)	Remark
		Not difficult	Less difficult	Difficult	Highly difficult			
11.1	Understanding the meaning of measures of central tendency (mean, mode, median).	51	33	16	0	1.65	.744	Less difficult
11.2	Calculating measures of central tendency (mean, mode, median) with ungrouped data.	44	42	13	1	1.71	.729	Less difficult
11.3	Calculating measures of central tendency (mean, mode, median) with grouped data.	28	48	20	4	2.00	.804	Less difficult
11.4	Interpreting measures of central tendency (mean, mode, median).	22	38	31	9	2.27	.908	Less difficult
Category 2: The interpretation and determination of measures of dispersion								
11.5	Understanding the meaning of measures of dispersion (standard deviation, range and variance).	12	32	40	16	2.60	.899	Difficult
11.6	Calculating the measures of dispersion (standard deviation, variance, range).	21	29	32	18	2.47	1.020	Less difficult
11.7	Interpreting measures of dispersion (standard deviation, variance, range).	5	28	40	27	2.89	.863	Difficult
11.8	Understanding the content and doing calculations for five number summary (lower quartile, middle quartile, upper quartile).	40	32	25	3	1.91	.877	Less difficult
Category 3:The use of graphs for the prediction of results								
11.9	Drawing box-and-whisker diagrams on a number line.	45	31	19	5	1.84	.907	Less difficult
11.10	Constructing cumulative frequency table.	32	37	26	5	2.04	.887	Less difficult
11.11	Making a diagram of an ogive (cumulative frequency curve).	31	35	24	10	2.13	.971	Less difficult
11.12	Representing bivariate numerical data as a scatter plot.	19	25	36	20	2.57	1.018	Difficult
11.13	Selecting a function that best fits the data for linear, quadratic and exponential.	16	28	30	26	2.64	1.030	Difficult

The data reflected in category 1 of table 5 shows that the teachers believed that their learners had less difficulty with the following topics: Understanding the meaning of measures of central tendency (mean, mode, median); calculating measures of central tendency (mean, mode, median) with ungrouped data; calculating measures of central tendency (mean, mode, median) with grouped data; and interpreting measures of central tendency (mean, mode, median).

The data in category 2 of table 5 shows that the teachers indicated that their learners experienced less difficulty with the following topics: Calculating the measures of dispersion (standard deviation, variance, range); understanding the content and doing calculations for five number summary (lower quartile, middle quartile, upper quartile). Meanwhile the teachers indicated that their learners experienced difficulty with the following topics: Understanding the meaning of measures of dispersion (standard deviation, range and variance); and interpreting measures of dispersion (standard deviation, variance, range).

Furthermore, table 5 in category 3 shows that the teachers indicated that their learners had less difficulty with the following topics: Drawing box-and-whisker diagrams on a number line; making a diagram of an ogive (cumulative frequency curve); and constructing cumulative frequency table. Meanwhile the teachers indicated that their learners experienced difficulty with representing bivariate numerical data as a scatter plot, and selecting a function that best fits the data for linear, quadratic and exponential.

Therefore the data in table 5 suggest that learners had problems with all the topics listed in table 5 with more difficulty in understanding the meaning of measures of dispersion (standard deviation, range and variance); interpreting measures of dispersion (standard deviation, variance, range); and representing bivariate numerical data as a scatter plot, and selecting a function that best fits the data for linear, quadratic and exponential.

4.1.1.6 Teachers' opinions about problems encountered in the learning of probability

In order to collect information on the problems encountered in the learning of probability, the researcher asked teachers to indicate in the questionnaire whether learners had problems with:

(1) construction and interpretation of probability diagrams and tables (items 12.1, 12.2, 12.3, 12.6 and 12.7), and (2) understanding or interpretation of probability terminology (items 12.4, 12.5, 12.8 and 12.9). A four point scale was used where teachers indicated their learners' level of difficulty in learning aspects of probability. The results of this section of the questionnaire are reflected in table 6.

Table 6: Teachers' opinions about problems in the learning of probability

Category 1: Construction and interpretation of probability diagrams and tables		Frequencies of teachers				Mean	Standard deviation (SD)	Remark
		Not difficult	Less difficult	Difficult	Highly difficult			
12.1	Using and constructing Venn diagrams from a given word problem.	8	27	31	34	2.91	.965	Difficult
12.2	Using and constructing two-way contingency tables from a given word problem.	10	20	38	32	2.92	.961	Difficult
12.3	Using and constructing tree diagrams from a given word problem.	11	28	28	33	2.83	1.016	Difficult
12.6	Using Venn diagrams to solve probability problems where events are not necessarily independent.	8	24	29	39	2.99	.980	Difficult
12.7	Using tree diagrams to solve probability problems where events are not necessarily independent.	7	30	29	34	2.90	.959	Difficult
Category 2: Understanding or interpretation of probability terminology								
12.4	Identifying dependent and independent events from Venn diagrams.	7	28	28	37	2.95	.968	Difficult
12.5	Identifying dependent and independent events from two-way contingency tables.	9	25	30	36	2.93	.987	Difficult
12.8	Calculating probability of two independent events by applying a product rule for independent events: P (A and B) = P(A). P(B).	8	27	24	41	2.98	1.005	Difficult
12.9	Understanding the difference between independent and dependent events.	10	23	28	39	2.96	1.014	Difficult

The data in category 1 of table 6 shows that the teachers believed that their learners had difficulty with the following topics: Using and constructing Venn diagrams from a given word problem;

using and constructing two-way contingency tables from a given word problem; using and constructing tree diagrams from a given word problem; using tree diagrams to solve probability problems where events are not necessarily independent; using Venn diagrams to solve probability problems where events are not necessarily independent.

It is also clear from table 6 in category 2 that the teachers felt that their learners had difficulty with the following topics: Identifying dependent and independent events from Venn diagrams; identifying dependent and independent events from two-way contingency tables; understanding the difference between independent and dependent events; and calculating probability of two independent events by applying a product rule for independent events: P (A and B) = P(A). P(B).

Therefore, the data in table 6 suggests that the learners experienced difficulties with the interpretation of probability terminology and also with construction and interpretation of probability diagrams and tables.

4.1.1.7 Teachers' opinions about the causes of learners' problems in statistics

In order to gather information on what might be the cause of learners' problems, teachers were asked to indicate their opinions about learners' problems in the questionnaire, as reflected in table 7. The study explored issues linked to: (1) teachers' content knowledge of statistics (items 13.1, 13.2, 13.3, 13.4 and 13.5), (2) learning material (items 13.6, 13.7 and 13.8), and (3) the teaching of statistics topics in previous grades (items 13.9, 13.10, 13.11, and 13.12).

Table 7: Teachers' opinions about the causes of learners' problems in the learning of statistics

	The causes of problems in the learning of statistics	Percentage of Teachers				
		Strongly disagree	**Disagree**	**Neither disagree nor agree**	**Agree**	**Strongly agree**
	Category 1: Teachers' content knowledge of statistics					
13.1	Learners encounter problems because I experience problems in explaining concepts to learners.	21.0	36.0	19.0	22.0	2.0
13.2	Learners encounter problems because I do not understand some of the topics in the syllabus.	22.0	47.0	11.0	18.0	2.0
13.3	Learners encounter problems because I do not know which method of instruction I should use to teach statistics (data handling and probability).	21.0	39.0	24.0	15.0	1.0
13.4	Learners encounter problems because I did not study statistics at tertiary level.	21.0	38.0	14.0	19.0	8.0
13.5	Learners encounter problems because I do not have enough teaching experience.	33.0	41.0	12.0	13.0	1.0
	TOTAL ROW % OF CATEGORY 1	**23.6**	**40.2**	**16**	**17.4**	**2.8**
	Category 2: Learning material					
13.6	Learners encounter problems because I do not have enough class time to assist them individually.	18.0	28.0	20.0	26.0	8.0
13.7	Learners encounter problems because they do not have a mathematics textbook.	24.0	45.0	12.0	17.0	2.0
13.8	Learners encounter problems because they do not know how to use the statistics function mode on a calculator.	10.0	30.0	21.0	29.0	10.0
	TOTAL ROW % OF CATEGORY 2	**17.33**	**34.33**	**17.67**	**24**	**6.67**
	Category 3: The teaching of statistics topics in previous grades					
13.9	Learners encounter problems because they do not give themselves enough time to practise data handling and probability problems.	7.0	21.0	16.0	42.0	14.0
13.10	Learners encounter problems because statistics (data handling and probability) is too difficult for them.	10.0	47.0	29.0	10.0	4.0
13.11	Learners encounter problems because certain sections of data handling and probability were not properly taught in previous years.	10.0	13.0	30.0	35.0	12.0

13.12	Learners encounter problems because certain sections of data handling and probability were never taught in the lower classes.	10.0	12.0	30.0	38.0	10.0
	TOTAL ROW % OF CATEGORY 3	**9.25**	**23.25**	**26.25**	**31.25**	**10**

The data in category 1 of table 7 shows that 20.2% (i.e. 17.4% of *agree* + 2.8% of *strongly agree*) of the teachers indicated that learners encountered problems because their teachers lack content knowledge in statistics. In this category, 24% (i.e. 22% of *agree* + 2% of *strongly agree*) of teachers felt that learners encountered problems because teachers themselves have problems in explaining concepts to learners; 20% (i.e. 18% of *agree* + 2% of *strongly agree*)indicated problems arise because teachers do not understand some of the topics in the syllabus; 16% (i.e. 15% of *agree* + 1% of *strongly agree*) of teachers indicated that learners encountered problems because teachers were uncertain about the methods they should use to teach statistics; 27% (i.e. 19% of *agree* + 8% of *strongly agree*)indicated that learners encountered problems because teachers did not have further qualifications in statistics; and 14% (i.e. 13% of *agree* + 1% of *strongly agree*) indicated that learners' problems stemmed from teachers' lack of experience.

It is also clear from table7 in category 2 that 30.67% (i.e. 24% of *agree* + 6.67% of *strongly agree*) of teachers believed learners encountered problems because of inadequate learning material. In this category, 19% (i.e. 17% of *agree* + 2% of *strongly agree*) indicated that learners had no mathematics textbook; 39% (i.e. 29% of *agree* + 10% of *strongly agree*) that learners did not know how to use the statistics function mode on the calculator; 34% (i.e. 26% of *agree* + 8% of *strongly agree*) of teachers that learners encountered problems because teachers did not have enough class time to assist them individually.

The data in category 3 of table 7 shows that 41.25% (i.e. 31.25% of *agree* + 10% of *strongly agree*) of teachers agreed that learners' problems were the result of poor teaching of statistics topics in previous grades. In this category, 47% (i.e. 35% of *agree* + 12% of *strongly agree*) of teachers believed problems occurred because certain sections of data handling and probability had not been adequately taught in previous years; 48% (i.e. 38% of *agree* + 10% of *strongly agree*) believed problems occurred because certain sections of data handling and probability were not taught at all in previous years; 56% (i.e. 42% of *agree* + 14% of *strongly agree*) of teachers indicated that learners did not give themselves enough time to practise data handling and probability problems; 14% (i.e. 10% of *agree* + 4% of *strongly agree*) indicated that learners found statistics too difficult.

The results in table 7 suggest that the major causes of learners' problems in statistics are: (1) the teaching of statistics topics in the previous grades. Certain sections of data handling and probability were not taught properly, or not taught at all in the lower grades; (2) learning material. Learners do not know how to use the statistics function mode on the calculator; (3) teachers' inadequate content knowledge in statistics. Teachers experience problems when explaining concepts to learners.

4.1.1.8 In-service education and training (inset) programmes

In-service workshops (inset) are intended to broaden teachers' content knowledge and to help them tackle their classroom problems (Ogbonnaya, 2007). Teachers who spend more time in in-service training programmes are more likely to improve their instruction (Parsad, Lewis, Farris & Greene, 2001). It is against this background that the following issues are explored to ascertain whether these programmes improved teachers' content and pedagogical knowledge of statistics: (1) the number of days teachers spent in inset programmes over the past 24 months, (2) the extent to which the topics of statistics (data handling and probability) were covered during inset workshops, (3) the extent to which teachers' subject and pedagogical knowledge improved after attending the inset programmes, and (4) factors that prevented teachers from attending more inset programmes.

4.1.1.8.1 Number of days spent in inset programmes

Table 8 reflects the number of days spent by teachers in inset programmes over the last 24 months.

Table 8: Period spent in inset programmes

Days	Percentage of teachers
None	32.0
Fewer than three days	7.0
Three to seven days	15.0
Eight to 14 days	19.0
Fifteen to 22days	16.0
Twenty-two to 31 days	5.0
More than 31 days	6.0

Table 8 indicates that more than half (54%) of the teachers spent at most seven days in inset programmes during the 24 months preceding the study; 32% of teachers did not attend inset programmes during this period; 7% spent fewer than three days in inset programmes in the past 24 months; 15% spent between three to seven days in inset programmes during this period. This implies that more than half (54%) of the teachers had spent less time in inset programmes to upgrade their statistics knowledge during the 24 months prior to the study.

4.1.1.8.2 The extent to which topics on probability and data handling were covered in workshops

Table 9 explores the coverage of probability and data handling in the workshops. A five-point scale was used; teachers indicated the extent to which probability and data handling were covered during inset programmes they had attended over the past 24 months. The overall percentages of teacher responses were determined for each degree of facilitation.

Table 9 : Emphasis on probability and data handling in inset programmes

	Percentage of Teachers					
	Did not attend	Not at all	Slightly	Moderately	Largely	Total
Probability	30.0	23.0	7.0	23.0	17.0	100.0
Data handling	30.0	16.0	11.0	21.0	22.0	100.0

Table 9 indicates that 30% of the teachers questioned had not attended an inset programme on the topic of probability in the 24 months preceding the study. Of the 70% of teachers who had attended programmes during this time, 30% (i.e. 23% of *not at all* + 7% of *slightly*) claimed that the probability topic had not been adequately dealt with.

Of the 70% of teachers who had attended inset programmes, 27% (i.e. 16% of *not at all* + 11% of *slightly*) said that the topic of data handling had not been well handled.

4.1.1.8.3 The extent to which teachers' subject and pedagogical knowledge improved after attending inset programmes

Table 10 explores the extent to which teachers' subject and pedagogical knowledge improved after attending inset programmes. A five-point scale was used, with teachers indicating the extent to which they believed their subject and pedagogical knowledge had improved after attending these programmes. The overall percentages of teacher responses were determined for each degree of improvement.

Table 10: Improvement of teachers' subject matter and pedagogical knowledge

		Percentage of teachers				
		Did not attend	Not at all	Not very well	Pretty Well	Completely
Subject matter (content Knowledge)	Probability	30.0	22.0	15.0	26.0	7.0
	Data handling	30.0	9.0	9.0	41.0	11.0
Method of teaching (pedagogical knowledge)	Probability	30.0	21.0	17.0	22.0	10.0
	Data handling	30.0	9.0	14.0	35.0	12.0

Table 10 shows that 30% of the teachers had not attended inset programmes on the topic of probability during the last 24 months. Of the 70% of teachers who had attended such programmes, 37% (i.e. 22% of *not at all* + 15% of *slightly*) believed that their subject matter knowledge had not improved. Only a few 18% (i.e. 9% of *not at all* + 9% of *slightly*) of those who had attended these inset programmes on data handling believed that their subject matter knowledge had not improved.

Again, table 10 reflects that 38% (i.e. 21% of *not at all* + 17% of *slightly*) of the teachers who had attended inset programmes on probability during this period said that their methods of teaching (pedagogical knowledge) had not improved, and 23% (i.e. 9% of *not at all* + 14% of *slightly*) of those who had attended these inset programmes on data handling believed that their method of teaching (pedagogical knowledge) had not improved. Data in table 10 suggests that teachers' content knowledge and pedagogical knowledge on the topic of probability did not improve as much as their content knowledge and pedagogical knowledge on the topic of data handling.

4.1.1.8.4 Reasons that prevented teachers from attending inset programmes regularly

The reasons that may have prevented teachers from attending more inset workshops are presented in table 11.

Table 11: Reasons that prevented teachers from attending more inset workshops

Items	Reasons	Percentage of teachers				
		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
17.1	The inset programmes were too expensive; I could not afford them as I was expected to pay for myself.	31.0	31.0	22.0	11.0	5.0
17.2	There was a lack of financial support from our school.	25.0	33.0	22.0	12.0	8.0
17.3	The inset programmes were arranged during teaching time.	20.0	30.0	22.0	22.0	6.0
17.4	The inset programmes were arranged over weekends.	17.0	25.0	26.0	27.0	5.0
17.5	I did not have time because of family responsibilities.	34.0	38.0	21.0	4.0	3.0
17.6	There were no suitable inset programmes for me to attend.	20.0	32.0	22.0	15.0	11.0
17.7	The inset programmes were presented by incompetent facilitators and were not productive.	29.0	32.0	30.0	8.0	1.0
17.8	The inset programmes were offered very far from where I teach.	19.0	35.0	28.0	11.0	7.0

Table 11 shows that 32% (i.e. 27% of *agree* + 5% of *strongly agree*) of the teachers were unable to attend many inset programmes because these were held over weekends; 28% (i.e. 22% of *agree* + 8% of *strongly agree*) could not attend because the programmes took place during teaching time; 26% (i.e. 15% of *agree* + 11% of *strongly agree*) said there were no suitable inset programmes for them to attend; 20% (i.e. 12% of *agree* + 8% of *strongly agree*) were prevented from attending because there was a lack of financial support from their schools.

4.1.1.8.5 How can problems in the teaching of statistics be addressed?

Suggestions on how to address the problems encountered in the teaching of statistics (data handling and probability) obtained from teachers are presented in table 12.

Table 12: Suggestions on how to address problems relating to the teaching of statistics

Items	Suggestions	Percentage of teachers				
		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
18.1	Teachers should attend inset programmes regularly prior to the teaching of the topics in which they encounter problems.	0.0	3.0	6.0	41.0	50.0
18.2	Teachers should get financial support from their schools/districts to attend inset programmes.	1.0	2.0	4.0	29.0	64.0
18.3	Schools should plan ahead to allow their teachers to attend inset programmes.	1.0	2.0	2.0	34.0	61.0
18.4	Inset programmes should be arranged and organised by reputable and competent service providers.	1.0	1.0	2.0	38.0	58.0
18.5	Inset programmes should meet the needs of the teachers by offering the topics in which they encounter difficulties.	0.0	1.0	3.0	36.0	60.0
18.6	Inset programmes should be organised in the same circuit/district as the teachers' place of work.	0.0	4.0	10.0	34.0	52.0
18.7	It is recommended that teachers take a formal tertiary course in statistics to improve their content knowledge and teaching.	0.0	2.0	12.0	43.0	43.0
18.8	Textbooks should be well written (explained thoroughly) and contain all necessary information to teach data handling and probability (i.e. formulae, more examples, etc.)	0.0	0.0	6.0	33.0	61

Table 12 indicates that 64% of the teachers believed strongly that they should receive financial support from their schools/districts to attend inset programmes; 61% felt strongly that textbooks should be well written and contain all information necessary to teach data handling and probability; 61% strongly agreed that schools should plan ahead for their teachers to attend inset programmes; 60% strongly agreed that inset programmes should offer the topics that teachers found difficult to teach; and 58% felt strongly that inset programmes should be arranged and organised by reputable and competent service providers.

4.1.1.9 Summary of analysis of teacher questionnaires

This section provides a brief summary of the findings from the analysis of teacher questionnaires in the light of the research questions.

4.1.1.9.1 Problems encountered by teachers when teaching certain topics in statistics

The analysis of the data obtained from the teacher questionnaires found that teachers had not encountered problems with the interpretation and determination or calculation of measures of dispersion, the representation and interpretation of data on graphs or plots, and the interpretation and determination of measures of central tendency. The findings also revealed that teachers had problems with construction and interpretation of probability diagrams and tables (i.e. Venn and tree diagrams, two-way contingency tables), and with understanding and interpreting probability terminology (i.e. mutually exclusive events, independent and dependent events etc.).

4.1.1.9.2 Causes of teachers' problems when teaching statistics

This analysis also revealed that the problems encountered by teachers are mostly caused by their lack of content knowledge of statistics. Most of these high school teachers of statistics had either never studied statistics during their pre-service training, or only in the first year of their degree, where most of the time was spent on an introduction to statistics. The findings also suggest that many problems are caused by inadequate or vague textbooks. Certain formulae are missing from these books and not enough examples are provided. Furthermore,the study found that teachers encountered problems because they lacked other resources essential to teaching statistics.

In addition, teachers mentioned that they encountered problems because they lacked experience in teaching statistics. For instance, it emerged from the analysis that 22% of the teachers had been teaching data handling for less than three years; 24% had taught probability for a similar period. Teachers also encountered difficulties because most of the in-service teacher training programmes they attended did not cover statistics topics, and those that did failed to pay adequate attention to probability, and as such made no significant impact on improving the teachers' content knowledge. Lastly, the results also showed that more than half of the teachers (54%) had spent at most seven days in in-service teacher programmes in a two-year period.

4.1.1.9.3 Possible ways to address the problems encountered by teachers when teaching statistics

The questionnaire included a request to teachers for possible solutions to their instructional problems. They provided the following suggestions: (1) teachers should receive financial support from their schools/districts to attend more in-service teacher programmes, (2) in-service teacher programmes should meet the needs of the teachers by offering those topics which teachers find particularly problematic, and (3) textbooks should be more explicitly written, provide more detailed explanations and also include other essential information on how to teach data handling and probability.

4.1.2 Learner questionnaires

A total of 448 grade 11 mathematics learners completed the learner questionnaire. Most learners who responded to the questionnaire were female (235), accounting for 52.5% of all learners who participated in the study. The learner questionnaire sought to investigate problems learners had with data handling and probability.

4.1.2.1Problems in learning data handling

Table 13 explores the problems encountered in the learning of data handling. A four point scale was used where learners indicated their level of difficulties in learning aspects of data handling such as (1) interpreting and determining measures of central tendency (items 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7 and 3.8), (2) interpreting and determining measures of dispersion (items 3.9, 3.10, 3.11, 3.12, 3.13 and 3.14), (3) representing data on graphs/plots (items 3.15, 3.16, 3.17, 3.18, 3.26 and 3.27), and (4) using graphs to predict results (items 3.19, 3.20, 3.21, 3.22, 3.23, 3.24, 3.25 and 3.28). Then overall means of learner responses were determined for each level of difficulty and the limit scale as indicated in section 7 determined whether the learner responses were not difficult, less difficult, difficult and highly difficult.

Table 13: Learners' problems in learning data handling

Items	Category 1: Interpreting and determining measures of central tendency	Frequencies of learners				Mean	Standard deviation (SD)	Remark
		Not difficult	Less difficult	Difficult	Highly difficult			
3.1.	Understanding the meaning of the mean as a measure of central tendency.	274	115	40	19	1.56	.825	Less difficult
3.2.	Understanding the meaning of the median as a measure of central tendency.	262	113	58	15	1.61	.836	Less difficult
3.3.	Calculating the median of ungrouped data.	227	140	66	15	1.71	.839	Less difficult
3.4.	Calculating the mean of ungrouped data.	279	112	40	17	1.54	.810	Less difficult
3.5.	Calculating the median of grouped data.	279	112	45	12	1.53	.782	Less difficult
3.6.	Calculating the mean of grouped data.	273	81	80	14	1.63	.883	Less difficult
3.7.	Interpreting the mean as a measure of central tendency.	123	157	124	44	2.20	.952	Less difficult
3.8.	Interpreting the median as a measure of central tendency.	113	172	121	42	2.21	.926	Less difficult
Category 2: Interpreting and determining measures of dispersion								
3.9.	Understanding the meaning of standard deviation.	111	133	131	73	2.37	1.03	Less difficult
3.10.	Understanding the meaning of variance.	94	146	126	82	2.44	1.02	Less difficult
3.11.	Interpreting standard deviation as a measure of dispersion.	71	121	129	127	2.70	1.05	Difficult
3.12.	Interpreting variance as a measure of dispersion.	61	153	125	109	2.63	.997	Difficult
3.13.	Calculating the lower quartile of grouped data.	253	127	43	25	1.64	.871	Less difficult
3.14.	Calculating the upper quartile of grouped data.	214	143	63	28	1.79	.908	Less difficult

Category 3: Representing data on graphs/plots								
3.1.	Drawing box-and-whisker diagrams on a number line.	278	83	49	38	1.66	.977	Less difficult
3.2.	Constructing a cumulative frequency table.	202	123	78	45	1.92	1.01	Less difficult
3.3.	Making a diagram of an ogive (cumulative frequency curve).	163	128	82	75	2.15	1.09	Less difficult
3.4.	Constructing a stem-and-leaf plot.	283	87	56	67	1.89	1.11	Less difficult
3.26	Representing bivariate numerical data as a scatter plot.	87	133	146	82	2.50	1.00	Difficult
3.27	Constructing a line of best fit on the scatter plot.	110	126	115	97	2.44	1.08	Less difficult
Category 4: Using graphs to predict results								
3.5.	Using a stem-and-leaf plot to determine the upper quartile.	220	103	87	38	1.87	1.00	Less difficult
3.6.	Using a stem-and-leaf plot to determine the lower quartile.	165	152	84	47	2.03	.988	Less difficult
3.7.	Using a diagram of ogive (cumulative frequency curve) to estimate the cumulative percentages (of aless than or more than nature).	168	154	66	60	2.04	1.03	Less difficult
3.8.	Using a diagram of ogive (cumulative frequency curve) to estimate the lower quartile from a set of grouped data.	67	145	100	136	2.68	1.06	Difficult
3.9.	Using a diagram of ogive (cumulative frequency curve) to estimate the middle quartile (median) from a set of grouped data.	81	130	114	123	2.62	1.07	Difficult
3.10.	Using a diagram of ogive (cumulative frequency curve) to estimate the upper quartile from a set of grouped data.	87	131	129	101	2.54	1.04	Difficult
3.28	Identifying a function that best fits the data.	72	140	128	108	2.52	.983	Difficult

The data in category 1 of table 13 shows that the learners had less difficulty with the following: Understanding the meaning of the mean as a measure of central tendency;

understanding the meaning of the median as a measure of central tendency; calculating the median of ungrouped data; calculating the mean of ungrouped data; calculating the median of grouped data; calculating the mean of grouped data; interpreting the median as a measure of central tendency; and interpreting the mean as a measure of central tendency.

The data in category 2 of table 13 also shows that the learners had less difficulty with interpreting and determining measures of dispersion difficult that is understands the meaning of standard deviation; understanding the meaning of variance; calculating the upper quartile of grouped data; calculating the lower quartile of grouped data. Meanwhile learners had difficulty with interpreting variance as a measure of dispersion; and interpreting standard deviation as a measure of dispersion.

Furthermore, the data in category 3 of table 13 shows that the learners had less difficulty with representing data on graphs/plots that is: drawing box-and-whisker diagrams on a number line Constructing a cumulative frequency table; making a diagram of an ogive (cumulative frequency curve); constructing a stem-and-leaf plot; representing bivariate numerical data as a scatter plot; and constructing a line of best fit on the scatter plot.

Lastly, the data in category 4 of table 13 shows that the learners had difficulty with using graphs to predict results that is: Using a diagram of ogive (cumulative frequency curve) to estimate the lower quartile from a set of grouped data; identifying a function that best fits the data; using a diagram of ogive (cumulative frequency curve) to estimate the upper quartile from a set of grouped data; and using a diagram of ogive (cumulative frequency curve) to estimate the middle quartile (median) from a set of grouped data. Meanwhile the learners experienced less difficulty with using a stem-and-leaf plot to determine the upper quartile; and using a stem-and-leaf plot to determine the lower quartile.

The findings presented in table 13 suggest that the learners experienced problems when using graphs to predict the results; interpreting and determining measures of dispersion; and representing data on graphs/plot; and when interpreting and determining measures of central tendency. However using a diagram of ogive (cumulative frequency curve) to estimate the lower quartile from a set of grouped data; identifying a function that best fits the data; using a diagram of ogive (cumulative frequency curve) to estimate the upper quartile from a set of grouped data; using a diagram of ogive (cumulative frequency curve) to estimate the middle quartile (median); interpreting variance as a measure of dispersion; and interpreting standard deviation as a measure of dispersion were the concepts which learners found them most difficult.

4.1.2.2 Problems with the learning of probability

Table 14 explores the problems encountered in the learning of probability. A four point scale was used where learners indicated their level of difficulties in learning aspects of probability such as (1) construction and interpretation of probability graphs and tables (items 4.1, 4.2, 4.3, 4.4, 4.5 and 4.6), and (2) understanding or interpretation of probability terminology (items 4.7, 4.8, 4.9, 4.10, 4.11, 4.12, 4.13, 4.14 and 4.15). Then overall means of learner responses were determined for each level of difficulty and the limit scale as indicated in section 7 determined whether the learner responses were not difficult, less difficult, difficult and highly difficult.

Table 14: Learners' problems in learning probability

Items	**Category 1: construction and interpretation of probability graphs and tables**	**Frequencies of learners**				**Mean**	**Standard deviation (SD)**	**Remark**
		Not difficult	Less difficult	Difficult	Highly difficult			
4.1.	Constructing a Venn diagram from a given word problem.	65	114	136	133	2.75	1.036	Difficult
4.2.	Constructing a two-way contingency table from a given word problem.	27	156	130	135	2.83	.930	Difficult
4.3.	Constructing a tree diagram from a given word problem.	80	117	102	149	2.71	1.109	Difficult
4.4.	Using Venn diagrams to solve probability problems.	79	113	116	144	2.73	1.084	Difficult
4.5.	Using two-way contingency tables to solve probability problems.	46	151	107	144	2.78	1.011	Difficult
4.6.	Using tree diagrams to solve probability problems.	76	108	123	141	2.73	1.080	Difficult
Category 2:understanding or interpretation of probability terminology								
4.7.	Identifying dependent events from Venn diagrams.	72	113	116	147	2.75	1.080	Difficult

4.8.	Identifying independent events from Venn diagrams.	94	106	98	150	2.68	1.145	Difficult
4.9.	Identifying independent events from two-way contingency tables.	89	134	95	130	2.59	1.105	Difficult
4.10.	Identifying dependent events from two-way contingency tables.	79	134	116	119	2.61	1.060	Difficult
4.11.	Using Venn diagrams to solve probability problems where events are not necessarily independent.	31	152	108	157	2.87	.976	Difficult
4.12.	Using tree diagrams to solve probability problems where events are not necessarily independent.	66	116	130	136	2.75	1.045	Difficult
4.13.	Calculating probability of two independent events by applying a product rule for independent events: P(A and B) = P(A). P (B).	106	85	103	154	2.68	1.175	Difficult
4.14.	Identifying mutually exclusive events from Venn diagrams.	55	95	125	173	2.93	1.042	Difficult
4.15.	Understanding the difference between independent and dependent events.	146	105	69	128	2.40	1.211	Less difficult

The data in category 1 of table 14 shows that the learners had difficulty with the construction and interpretation of probability diagrams and tables that is: constructing a tree diagram from a given word problem; constructing a two-way contingency from a given word problem; constructing a Venn diagram from a given word problem difficult; using Venn diagrams to solve probability problems, using two-way contingency tables to solve probability problems and using tree diagrams to solve probability problems.

The data in category 2 of table 14 also shows that the learners had difficulty understanding and interpreting probability terminology that is: identifying dependent events from two-way contingency tables; identifying independent events from Venn diagrams; identifying dependent events from Venn diagrams; using Venn diagrams to solve probability problems where events were not necessarily independent; using tree diagrams to solve probability problems where events were not necessarily independent; identifying independent events from two-way contingency tables; calculating probability of two independent events by applying a product rule for independent events; and identifying mutually exclusive events from Venn diagrams, and less difficulty with understanding the difference between independent and dependent events.

Clearly, the results in table 14 indicate that the learners encountered problems with the construction and interpretation of probability diagrams and tables, as well as with the interpretation of probability terminology.

4.1.2.3The cause of problems encountered in the learning of statistics

In order to gather information on the reasons behind learners' difficulties in statistics, they were asked to indicate the cause of their problems in the questionnaire, as shown in table 15. The study explored issues linked to: (1) teachers' content knowledge of statistics and methods of teaching (items 5.1, 5.2, 5.3, 5.4, 5.5, 5.10 and 5.18), (2) learning material (items 5.6, 5.7, 5.8 and 5.9), (3) learners' attitudes (items 5.11, 5.12, 5.16 and 5.17), and (4) the teaching of statistics topics in previous grades (items 5.13, 5.14 and 5.15).

Table 15: Reasons for learners' problems in the learning of statistics

The cause of the problems		Percentage of learners				
		Strongly disagree	Disagree	Neither Agree nor disagree	Agree	Strongly agree
Items	**Category 1: Teachers' content knowledge of statistics and methods of teaching**					
5.1.	I encounter problems because my teacher finds it difficult to explain concepts / does not explain clearly.	19.0	30.0	30.4	14.1	6.5
5.2.	I encounter problems because my teacher does not understand some of the topics in the syllabus.	29.7	44.4	10.7	12.5	2.7
5.3.	I encounter problems because my teacher teaches some topics but leaves others for us to do on our own.	23.0	24.6	14.5	25.4	12.5
5.4.	The teacher does not allow enough time to teach statistics.	22.8	32.1	16.7	21.0	7.4
5.5.	The teacher does not give learners enough exercises to practise statistics.	30.8	21.2	17.2	22.5	8.3
5.10	I encounter problems because I am learning statistics on my own; I do not have a teacher who teaches me data handling and probability.	58.3	22.0	9.2	8.3	2.2
5.18	I encounter problems because learners are not allowed to discuss the work during lessons.	36.8	23.0	14.1	18.3	7.8
TOTAL ROW % OF CATEGORY 1		**31.49**	**28.19**	**16.11**	**17.44**	**6.77**
	Category 2: Learning material					
5.6.	I encounter problems because I was never taught how to use the statistics function mode on the calculator.	32.6	32.8	16.5	11.4	6.7
5.7.	I encounter problems because I do not have a mathematics textbook.	63.5	23.9	4.5	5.6	2.5
5.8.	I encounter problems because I do not know how to use the statistics function mode on the calculator.	31.0	27.0	19.9	16.5	5.6
5.9.	I encounter problems because I do not have access to previous examination papers with solutions.	20.3	24.6	17.6	25.2	12.3
	TOTAL ROW % OF CATEGORY 2	**36.85**	**27.07**	**14.63**	**14.68**	**6.77**
	Category 3: Learners' attitudes					
5.11.	I encounter problems because I do not give myself enough time to practise data handling and probability problems.	18.1	11.3	18.1	31.7	20.8
5.12.	I encounter problems because statistics (data handling and probability) is too difficult for me.	21.8	32.6	24.6	16.1	4.9
5.16.	I do not see the importance of statistics in mathematics.	57.4	20.5	8.9	9.6	3.6

5.17.	I am not interested in statistics.	47.8	28.8	13.5	5.4	4.5
TOTAL ROW % OF CATEGORY 3		**36.27**	**23.3**	**16.28**	**15.7**	**8.45**
	Category 4: The teaching of statistics topics in previous grades					
5.13.	I encounter problems because certain sections of data handling and probability were not properly taught in lower grades.	15.8	22.1	23.2	21.7	17.2
5.14.	I encounter problems because certain sections of data handling and probability were never taught in previous grades.	25.2	27.3	21.4	12.5	13.6
5.15.	I encounter problems in data handling and probability because I did not pay enough attention when these topics were taught in previous classes.	22.3	34.2	22.5	10.5	10.5
	TOTAL ROW % OF CATEGORY 4	**21.1**	**27.86**	**22.36**	**14.9**	**13.78**

The data in category 1 of table 15 reveals that 24.21% (i.e. 17.44% of *agree* + 6.77% of *strongly agree*) of the learners believed that their problems stemmed from their teachers' inadequate content knowledge of statistics and poor methods of teaching. In this category, learners indicated that their difficulties arose because: their teachers experienced problems in explaining concepts to them (20.6% ((i.e. 14.1% of *agree* + 6.5% of *strongly agree*))); their teachers did not understand some of the topics in the syllabus (15.2% (i.e. 12.5% of *agree* + 2.7% of *strongly agree*)); their teachers taught some topics but left others for learners to do on their own (37.9% (i.e. 25.4% of *agree* + 12.5% of *strongly agree*)); their teachers did not allow enough time for the teaching of statistics (28.4% (i.e. 21.0% of *agree* + 7.4% of *strongly agree*)); their teachers did not give them enough exercises to practise statistics (30.8% (i.e. 22.5% of *agree* + 8.3% of *strongly agree*)); they were not allowed to discuss the work during a lesson (26.1% (i.e. 18.3% of *agree* + 7.8% of *strongly agree*)); they were studying statistics on their own and did not have someone to teach them data handling and probability (10.5% (i.e. 8.3% of *agree* + 2.2% of *strongly agree*)).

The data in category 2 of table 15 shows that 21.45% (i.e. 14.68% of *agree* + 6.77% of *strongly agree*) of the learners claimed that they encountered problems because of inadequate learning material. In this category, learners indicated that they experienced difficulties because: they did not have a mathematics textbook (8.1% (i.e. 5.6% of *agree* + 2.5% of *strongly agree*)); they were never taught how to use the statistics function mode on the calculator (18.1% (i.e. 11.4% of *agree* + 6.7% of *strongly agree*)); they did not know how to use the statistics function mode on the calculator (22.1% (i.e. 16.5% of *agree* 5.6% of *strongly agree*)); they did not have access to previous examination papers with solutions (37.5% (i.e. 25.2% of *agree* + 12.3% of *strongly agree*)).

In addition, the data in category 4 of table 15 reveals that 28.68% (i.e. 14.9% of *agree* + 13.78% of *strongly agree*) of the learners agreed that they encountered problems because of the inadequate teaching of statistics topics in the previous grades. In this category, they encountered problems because: certain sections of data handling and probability were not properly taught in previous grades (38.9% (i.e. 21.7% of *agree* + 17.2% of *strongly agree*)); certain sections of data handling and probability were never taught in previous grades (26.1% (i.e. 12.5% of *agree* + 13.6% of *strongly agree*)); they did not pay enough attention when these topics (data handling and probability) were taught in previous grades (21% (i.e. 10.5% of *agree* + 10.5% of *strongly agree*)).

Lastly, the data in category 3 of table 15 shows that 24.15% (i.e. 15.7% of *agree* + 8.45% of *strongly agree*) of the learners believed that they encountered difficulties because of their attitude towards statistics. In this category, learners indicated that they encountered problems because: they did not give themselves enough time to practise data handling and probability problems (52.5% (i.e. 31.7% of *agree* + 20.8% of *strongly agree*)); they found statistics (data handling and probability) very difficult (21% (i.e. 16.1% of *agree* + 4.9% of *strongly agree*)); they did not see the importance of statistics in mathematics (13.2% (i.e. 9.6% of *agree* + 3.6% of *strongly agree*)); they were not interested in statistics (9.9% (i.e. 5.4% of *agree* + 4.5% of *strongly agree*)).

The figures in table 15 suggest that the main causes of the problems encountered by learners in the learning of statistics are:

(1) The poor teaching of statistics in previous grades. Certain sections of data handling and probability were not properly taught or not taught at all in previous years.

(2) Learners did not give themselves enough time to practise data handling and probability problems.

(3) Teachers with inadequate content knowledge of statistics and poor methods of teaching. Learners indicated that their teachers taught some topics but left others for learners to do on their own. Also, their teachers experienced problems in explaining concepts to them.

(4) Learners' attitudes toward statistics affected their performance in statistics.

4.1.2.4 Summary of analysis of learner questionnaire

This section provides a brief summary of the findings from the analysis of the learner questionnaires related to the research questions.

4.1.2.4.1 Problems encountered by learners when dealing with certain topics in statistics

The results from the learner questionnaire showed that the learners had difficulties using graphs to predict the results (i.e. using a diagram of ogive to estimate the lower quartile, middle quartile and upper quartile from a set of grouped data, identifying functions that best fit the data); (2) the learners had difficulty interpreting and determining measures of dispersion (i.e. variance and standard deviation); (3) learners had problems representing data on graphs or plots (i.e. representing bivariate numerical data as a scatter plot, constructing a line of best fit on the scatter plot, and making a diagram of an ogive); (4) the learners had less difficulty interpreting and determining measures of central tendency (i.e. calculating the mean with grouped data, and interpreting the mean and the median); (5) the learners had difficulty with construction and interpretation of probability graphs and tables (i.e. constructing a Venn diagram, a two-way contingency table and a tree diagram from a given word problem; using Venn diagrams, tree diagrams and two-way contingency tables to solve problems); and (6) the learners found it difficult to understand and interpret probability terminology (i.e. identifying dependent and independent events from Venn diagrams, identifying dependent and independent events from two-way contingency tables, and identifying mutually exclusive events from Venn diagrams).

4.1.2.4.2 Causes of observed learners' problems when learning statistics.

The data from the learner questionnaires revealed that learners were encountering problems because of (1) the poor teaching of statistics topics in the previous grades; (2) teachers' lack of content knowledge of statistics and their teaching methods; (3) learners' attitudes toward statistics; and (4) learning materials.

4.2. Classroom observations

In order to find out more about problems besetting the teaching and learning of statistics (data handling implied) and their cause(s), classroom observations were conducted in four conveniently selected schools. A classroom observation schedule was used for this purpose (see Appendix 3). Probability lessons were not observed because this topic is treated as an optional assessment and it is also taught in the fourth semester, during which time researchers are not allowed to collect data at schools (see section 3.3.1.2). Eight classrooms taught by seven teachers over three weeks were observed by the researcher. In line with the *2011 Gauteng Province teacher work schedule for mathematics grade 11*, data handling was supposed to be taught over two weeks (see Appendix 8). However, it appeared that some schools struggled to comply with this work schedule. Only two schools managed to teach data handling in the required period. Two schools taught the topic over three weeks, thus stretching classroom observations to three weeks. The codes T1, T2, T3, and so on were used to identify teachers whose lessons were observed, with the teacher who was observed first referred to as T1 and so on. Therefore teachers' codes ran from T1 to T7.The duration of mathematics lessons in schools where T1, T2, T3 and T4 taught was 40 minutes; lessons taught by T5 and T6 were 30 minutes each; and T7's lesson was 45 minutes long.

In T1's class, the researcher noted that the teacher had already started teaching data handling, that is, before the time allocated to it in the work schedule. However, the other six teachers taught data handling during the required period and this posed a challenge to managing the classroom observations because some of the lessons clashed. Rearranging the periods was not possible, nor was it possible to teach the lessons after school hours since most learners lived some distance from their schools. Each aspect of data handling was observed by the researcher until it was agreed between the learners and teacher that it had been exhausted. Hence, only seven aspects of data handling (ogive, measure of central tendency, five number summary, stem-and-leaf plot, box-and-whisker plot, measure of dispersion, and scatter plot and line of best fit) were observed. The results of the classroom observations are displayed in Table 16 and are reported per aspect taught.

Table 16: Learners' and teachers' problems in the topics observed

Topics observed, total time spent in class (in minutes)&number of learners observed per teacher	**Classroom observation feedback**						
	T1	T2	T3	T4	T5	T6	T7
Topic: Ogive (cumulative frequency graph)	*	*	×	×	*	*	*
Total time for topic observed	160	40	×	×	70	35	45
Number of learners observed	91	21	×	×	31	37	32
Topic: Measure of central tendency (mean, mode, median)	×	*	*	*	*	*	*
Total time in minutes for topic observed	×	80	40	80	70	70	45
Number of learners observed per teacher	×	21	30	27	31	37	32
Topic: Five number summary (minimum number, Q1, Q2, Q3, maximum number)	*	*	*	*	*	*	*
Total time for topic observed	160	40	40	40	70	35	45
Number of learners observed	91	21	30	27	31	37	32
Topic: Stem-and-leaf plot	×	√	√	√	√	√	√
Total time for topic observed	×	80	40	40	35	70	45
Number of learners observed	×	21	30	27	31	37	32
Topic: Box-and-whisker plot	×	√	√	√	√	*	√
Total time for topic observed	×	40	40	40	35	35	45
Number of learners observed	×	21	30	27	31	37	32
Topic: Measure of dispersion (variance and standard deviations)	×	*	*	*	*	*	*
Total time for topic observed	×	40	40	40	35	70	45
Number of learners observed	×	21	30	27	31	37	32
Topic: Scatter plot and line of best fit	×	×	*	*	×	×	√
Total time for topic observed	×	×	80	40	×	×	90
Number of learners observed	×	×	30	27	×	×	32

√ Observed with no problems
* Observed and problems manifest
× Not observed

4.2.1 Ogive (Cumulative frequency graph)

Table 16 illustrates that there were problems experienced in five of the classrooms taught by T1, T2, T5, T6 and T7. These problems surfaced during lessons on the cumulative frequency graph. This concept was not observed in lessons taught by T3 and T4 because their lessons clashed with those taught by T1. The following problems were noted in the observed lessons:

4.2.1.1 Teachers' problems

(i) T1, T2, T5, T6 and T7 did not teach learners that the cumulative frequency graph can also be used to estimate cumulative percentages (of a less than or more than nature). Instead, these teachers taught learners to estimate these cumulative percentages using only the cumulative frequency table. T1and T2 used only one example with ungrouped numerical data to teach the ogive.

(ii) T1, T2, T5, T6 and T7 did not teach learners how to construct cumulative frequency graphs when the numerical data has been grouped into intervals, nor did they teach learners how to use a cumulative frequency curve to find an estimate of the median, lower quartile and upper quartile.

4.2.1.2 Learners' problems

(i) Learners in T1's classroom had difficulty drawing a cumulative frequency graph. For instance, none of the learners knew whether they should use the lower or the upper limit number of the interval to represent the x-axis when drawing this graph (ogive).

(ii) Some learners in the classrooms of T1 and T2 had difficulties applying what they had learned in previous grades, such as writing the numbers in a tally format.

(iii) Learners in the classroom of T2 had difficulty understanding the phrases "at least" and "at most" in questions.

4.2.1.3 The cause(s) of these problems

(i) T1 and T2 taught the ogive using a textbook which did not have many examples.

(ii) Learners in the classrooms of T1, T2, T5, T6 and T7 were not taught to draw a cumulative frequency graph using numerical data that had been grouped into intervals.

4.2.2 Measure of central tendency

According to Table 16 there were significantly common problems in lessons on the measure of central tendency taught by T2, T3, T4, T5, T6 and T7. Teachers neither defined nor explained the mean concept (i.e. explaining its use, its disadvantages and when it could be used). T2, T3, T4, T5 and T7 could only explain how to compute the mean and how to define the formula for computing the mean for ungrouped data. The table also shows that lessons on measures of central tendency taught by T1 were not observed because this topic was taught before the time set in the work schedule. No problems were observed in the classroom of teacher T6 when teaching measures of central tendency. Learners in the classrooms of T2, T3, T4, T5 and T7 seemed to understand what they had been taught. The following problems were identified in the lessons:

4.2.2.1 Teachers' problems

(i) T2, T3, T4, T5, T6 and T7 could not define or explain the concept of mean. All but T6 were able only to explain how to compute the mean and to define the formula for computing the mean for ungrouped data.

(ii) T2, T3, T4, T5 and T7 did not teach how to compute the mean, mode or median with a set of numerical data when it was grouped into intervals.

(iii) T6 did not teach how to compute the mode and median with a set of numerical data when it was grouped into intervals.

(iv) T2, T3, T4 and T5 did not teach learners how to compute the mean using a calculator.

4.2.2.2 Learners' problems

The researcher observed that:

(i) Learners in the classrooms of teachers T2, T3 and T4 did not know how to use a calculator to compute the mean.

4.2.2.3 The cause(s) of the problems

(i) Teachers T2, T3 and T4 did not teach learners how to use a calculator to compute the mean.

4.2.3 Five number summary

Table 16 shows that all seven teachers encountered problems when teaching the five number summary concepts. The problems were as follows:

4.2.3.1 Teachers' problems

(i) The teachers T1, T2, T3, T5 and T6 used only the counting method to teach learners to determine the lower quartile (q1), median (q2) and upper quartile (q3).

(ii) T1, T2, T3, T5 and T6 did not teach learners how to determine (compute) the quartiles (q1, q2, q3) using the formulae of quartile positions.

(iii) T4 and T5 had difficulty computing the quartiles using the counting method when the total number of the data values (n) was even. The counting method worked for teachers and learners when the total number of data values was odd.

(iv) A problem observed among all teachers was that they understood the meaning of quartiles (q1, q2, q3) but had difficulty explaining or interpreting this for learners.

(v) None of the teachers taught learners how to compute the quartiles with numerical data when data had been grouped into intervals.

4.2.3.2 Learners' problems

Learners in the classrooms of T4 and T5 knew only how to determine the quartiles using the counting method and encountered problems when the total number of data values was even.

4.2.3.3 The cause(s) of the problems

T1, T2, T3, T4, T5 and T6 relied on only one textbook, a learner textbook, to teach quartiles and followed its style and sequence of presentation. For instance, the textbook explained quartiles with formulae of positions in the general discussion section and not under the main topic of five number summary, and this approach was also adopted by the teachers. In addition, the learner textbook did not contain examples of quartiles where the numerical data had been grouped into intervals; as a result, the number of examples provided by teachers during their lessons was limited to those in the learner textbook.

4.2.4 Stem-and-leaf plot

As is evident from Table 16 T2, T3, T4, T5, T6 and T7 did not have difficulties teaching stem-and-leaf plot concepts. Lessons on this topic taught by T1 were not observed because they clashed with lessons taught by T3 and T4.

4.2.5 Box-and-whisker plot

Table 16 indicates that no problems were observed in the classrooms of T2, T3, T4, T5 and T7 while they were teaching the topic of box-and-whisker plots. Again, lessons on this aspect taught by teacher T1were not observed because these clashed with the lessons taught by T3 and T4. The problem observed in T6's lesson concerned the teacher only; learners in this class did not encounter any problems. The problem is discussed below:

4.2.5.1 Teachers' problems

Some teachers, particularly T6, had difficulty explaining measures of skewness observed in box-and-whisker plots to learners. T6 explained incorrectly to learners that a box-and-whisker plot is negatively skewed when the median is close to the lower quartile and positively skewed when it is close to the upper quartile.

4.2.6 Measures of dispersion (variance and standard deviation)

In table 16 we can see that problems were identified in lessons on variance and standard deviation taught by T2, T3, T4, T5, T6 and T7. The table indicates that T1 was not observed teaching variance and standard deviation because he had already taught these as he did not follow the official work schedule. The following problems were noted in various lessons:

4.2.6.1 Teachers' problems

(i) T2, T3, T4, T5, T6 and T7 did not teach learners how to interpret variance or standard deviation. Teachers taught learners only how to compute variance and standard deviation with a set of ungrouped numerical data.

(ii) T2, T3, T4, T5 and T7 did not teach learners how to compute variance and standard deviation with a set of grouped data where the data had been grouped into intervals.

(iii) T2, T3, T4, T5 and T6 did not teach learners how to compute variance or standard deviation using a calculator.

(iv) T2, T3, T4, T5, and T6 taught learners to compute variance with the formula of variance (σ^2) for population data only.

4.2.6.2 Learners' problems

(i) Some of the learners in the classrooms of T2, T3, T4, T5 and T6 did not know how to use calculators to compute variance and standard deviation.

(ii) None of the learners in the classrooms of T2, T3, T4, T5, T6 and T7 knew how to interpret variance and standard deviation.

(iii) None of the learners in the classrooms of T2, T3, T4, T5, T6 and T7 knew how to compute variance and standard deviation with a set of grouped data.

(iv) Some of learners in the classrooms of T2, T3, T4, T5, and T6 confused the formula of variance (σ^2) for population data and variance (s^2) for sample data, meaning that they did not know when to use s^2 or σ^2.

4.2.6.3 The cause(s) of the problems

(i) T2, T3, T4, T5, and T6 followed the learner textbook throughout their teaching of measures of dispersion (variance and standard deviation). For instance, the learner textbook provided information about variance (σ^2) for population data only and did not provide an example of how to compute variance with a set of grouped data.

(ii) Learners were taught how to compute variance using the formula of variance (s^2) for sample data in grade 10. In grade 11, their teachers did not explain the difference between $s^2$and σ^2, or when to use these.

(iii) Learners in the classrooms of T2, T3, T4, T5 and T6 were not taught how to use calculators to compute variance or standard deviation.

4.2.7 Scatter plot and line of best fit

Only the lessons taught by T3 and T4 on the scatter plot revealed problems. T7's lessons were free of problems. T1 and T2 did not teach the scatter plot and line of best fit at all, with the result that there is no record of their lessons. Table 16 also shows that lessons on scatter plot and line of best fit were not observed in the classes taught by teachers T5 and T6 because these clashed with the lessons taught by T3 and T4. The problems are discussed below:

4.2.7.1 Teachers' problems

i) T3 and T4 did not explain the concept of the scatter plot, or why and when it should be used.

ii) T3 and T4 did not teach the correct methods (i.e. median-median line and least squares regression line) for constructing a line of best fit.

iii) T3 and T4 misled learners by telling them that to construct a line of best fit they should construct a positive slope between the points.

4.2.7.2 Learners' problems

i) Some learners in the classrooms of T3 and T4 had difficulty identifying the independent variable (x) and the dependent variable (y) from two numerical variables.

ii) Some learners in the classrooms of T3 and T4 had difficulties with scale measurement for graphing the plot.

iii) None of the learners in these classrooms were able to draw a correct line of best fit. Learners did not know the correct methods (i.e. median-median line and least squares regression line) for determining (drawing) a line of best fit. As a result of incorrect teaching, these learners believed that the slope of the line of best fit is always positive.

4.2.7.3 The cause(s) of the problems

(i) Learners in the classrooms of T3 and T4 were not taught the correct methods to determine a line of best fit.

(ii) Teachers (T1 and T2) taught all data handling topics using only one learner textbook, which did not cover the topic of scatter plots. That is the reason T1 and T2 did not teach learners this topic.

4.2.8 Summary of analysis of classroom observations

This section provides a brief summary of the analysis of classroom observations in relation to the research questions.

4.2.8.1 Problems encountered by teachers when teaching certain topics in statistics

The findings from the classroom observations revealed that T1, T2, T5, T6 and T7 encountered problems when teaching cumulative frequency graphs (ogive). It also appeared that T2, T3, T4, T5, T6 and T7 had difficulty teaching measures of central tendency (i.e. mean and median and mode with grouped data); T1, T2, T3, T4, T5, T6, and T7 encountered problems when teaching the five number summary (i.e. lower quartile, middle quartile and upper quartile; with grouped data and when the total number of data values (n) was even); T2, T3, T4, T5, T6 and T7 found teaching the measure of dispersion difficult (i.e. interpretation of variance and standard deviation, and calculating variance and standard deviation with grouped data); and T2 and T4 found teaching the construction of scatter plots and lines of best fit particularly difficult. This data suggests that there are many teachers who encounter problems with the teaching of data handling.

4.2.8.2 Problems encountered by learners when learning certain topics in statistics

The findings from the classroom observations indicated that learners in the classrooms of T1 and T2 encountered problems with the cumulative frequency curve (i.e. construction and interpretation of ogive). Most learners in the classrooms of T2, T3, T4, T5 and T6 did not know how to use a calculator to compute mean and variance. Furthermore, the findings showed that most of the learners in the classrooms of T4 and T5 experienced difficulties with the computation of quartiles (Q1, Q2, Q3) when the total number of data values was even. It also emerged that most of the learners in the classrooms of T2, T3, T4, T5, T6 and T7 had difficulty learning the measures of dispersion (i.e. interpretations of variance and standard deviation; and calculations of variance and standard deviation when data has been grouped into intervals). Lastly, the classroom observations found that most learners in the classrooms of T3 and T4 encountered difficulties with the construction of scatter plots and lines of best fit. These findings suggest that there are many learners who find data handling difficult to master.

4.2.8.3 Causes of observed teachers' problems.

Most of the teachers (T1, T2, T3, T4, T5 and T6) relied on only one prescribed learner textbook when teaching these topics. But this textbook did not provide adequate examples or explain how to interpret the mean, standard deviation or variance. Neither did it provide examples or explain how to calculate lower quartile, middle quartile and upper quartile with numerical data that has been grouped into intervals. The textbook did not cover the topics of scatter plots and lines of best fit; for this reason, some of the teachers (T3 and T4) ended up not teaching these topics at all. Furthermore, the observations found that several teachers encountered problems because they lacked both content knowledge and pedagogical knowledge in statistics. For instance, most of them (T2, T3, T4, T5, T6 and T7) had difficulty explaining statistical terms such as mean, standard deviation and quartiles.

4.2.8.4 Causes of observed learners' problems.

The classroom observations found that learners encountered difficulties in statistics (data handling implied) because of their teachers' lack of content knowledge in the topic. It was observed that teachers tended to skip the teaching of certain topics in statistics because of this. For instance, none of the teachers observed taught learners how to interpret standard deviation, variance, or the mean. In addition, none of them taught learners how to calculate standard deviation, variance, mean or quartiles with numerical data grouped into intervals; none of them showed learners how to use the ogive to estimate the lower quartile, upper quartile and middle quartile. Another cause of learners' problems was their textbook: it did not cover all the topics suggested by the curriculum and did not cover the whole syllabus. Some formulae were missing and inadequate examples were provided. For instance, it did not provide examples or explain how to interpret the mean, standard deviation or variance. There were no examples of how to calculate the lower quartile, middle quartile and upper quartile with numerical data grouped into intervals, nor did the book cover the topics of scatter plots and lines of best fit.

4.3 Diagnostic test

The diagnostic test used in this study consisted of four questions. Question1 tested whether learners could describe the appropriate central location and dispersion measures with the presence of outliers in the data set. Question 2 tested the computation of the quartiles (lower and upper) for grouped numerical data. Question 3 tested learners on: (i) construction of box-and-whisker plots, (ii) skewness in box-and-whisker plots, (iii) computation of the standard deviation and the mean for grouped numerical data, and (iv) the ogive. Question 4 tested learners on scatter plots and the line of best fit (see also Appendix 4).The purpose of using a diagnostic test was to identify any difficulties learners might have in the mastery of data handling (Johnson & Christensen, 2004). A group of 248 learners drawn from the seven classrooms of the four schools in which classroom observations had been conducted wrote this test. It was administered to learners once the teaching of data handling had been completed by all teachers in the sample. The test was an hour in duration. The researcher administered it herself to ensure that it was written according to the requirements. All four schools wrote the test on the same day.

The researcher scored the answer scripts using a marking rubric. Table 17 reflects the analysis of the learners' performance in each sub-question. Frequencies of completely correct answers, partially correct answers, completely wrong answers, and no answers provided by each learner are listed in the table. A partially correct answer was one where the learner failed to get the answer completely correct, perhaps because the choice of formula was incorrect, or the arithmetical computation was incorrectly executed or because the learner provided a faulty explanation. Partially correct answers were categorised as arithmetical problems or conceptual problems. Arithmetical problems occurred when the correct concept and procedure was followed in problem solving but somewhere along the line a miscalculation occurred, or the learner made a mistake in the arithmetical or computational operation such as the application of the wrong value to a variable. Conceptual problems arose when a learner did not understand the statistical concept or concepts embodied in the task, that is, the learner did not understand the properties that were required in the task. The findings of the analysis of the diagnostic test are presented in table 17 in the order in which the questions occurred in the test.

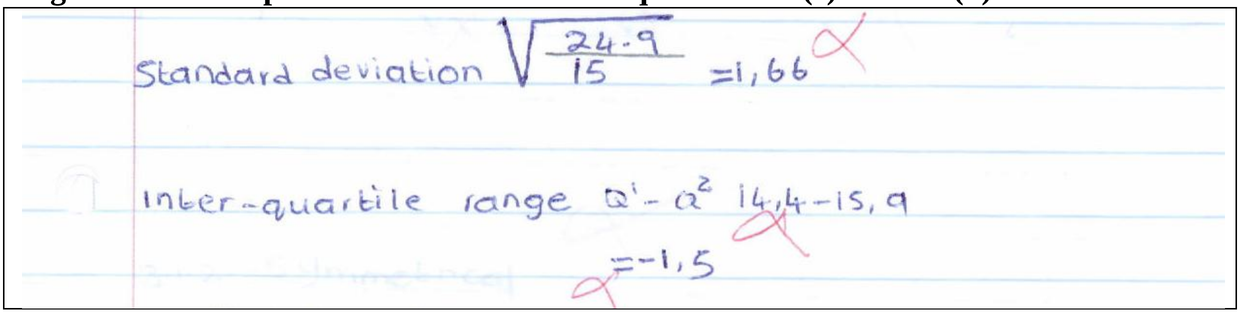
Table 17: Analysis of learners' performance on diagnostic test (n=248)

Question	Sub-questions	Performance of learners on each test item (sub-question)					Number of learners
			Partially correct				
		Correct answer	Arith-metical problem	Conceptual problem	Wrong answer	No answer	
Question 1	1.1 (a)	162	7	0	79	0	248
	1.1 (b)	189	0	0	56	3	248
	1.1 (c)	31	3	0	184	30	248
	1.1 (d)	65	0	1	145	37	248
	1.2	3	20	76	57	92	248
	1.3	0	0	158	77	13	248
	1.4	0	0	80	124	44	248
Question 2	2.1	0	0	0	209	39	248
	2.2	0	0	0	198	50	248
Question 3	3.1.1	42	51	2	135	18	248
	3.1.2	21	0	71	111	45	248
	3.2.1	2	0	0	228	18	248
	3.2.2	1	0	0	195	52	248
	3.3.1	76	0	0	159	13	248
	3.3.2	4	0	0	230	14	248
	3.3.3	8	0	0	204	36	248
Question 4	4.1	88	0	0	108	52	248
	4.2	100	0	51	84	13	248
	4.3	85	0	0	92	71	248
	4.4	9	0	0	203	36	248

Question 1

In table 17 we observe that most learners performed badly in question 1, which covered aspects of standard deviation, inter-quartile range, outliers, better measure of central location and better measure of dispersion. For instance, in question 1.1 (c), 184 (74%) learners performed badly by calculating the standard deviation incorrectly: it should have been calculated for ungrouped numerical data. In question 1.1 (d), 145 (58%) performed badly when computing the inter-quartile range for ungrouped numerical data. Also, in question 1.4, 124 (50%) of the group did not perform well when choosing the better measure of spread for the data between the standard deviation and the inter-quartile range when the data contained outliers. See the vignettes of learner A and B's answers to questions 1.1(c), 1.1(d) and 1.4.

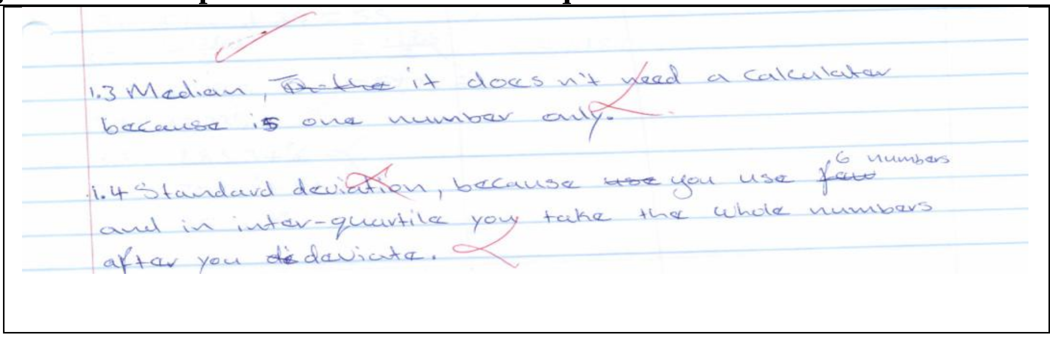
Vignette A: Example of learner's answers to question 1.1 (c) and 1.1 (d)



Vignette A shows that the learner was completely unable to calculate the standard deviation for ungrouped numerical data. The learner was also unable to calculate the inter-quartile range for ungrouped numerical data correctly. The learner revealed a lack of conceptual knowledge of the inter-quartile range and as a result failed to provide the correct formula for this calculation.

Table 17 shows that most of the learners in question 1.2 had difficulty determining outliers. The majority did not know how to identify these in the data or to describe the better measure of central tendency (measure of central location) or the better measure of dispersion when outliers are present in data. For instance, in question 1.2, only three (0.01%) learners' responses were rated as "completely correct". The performance in question 1.2 of the other 86 (35%) learners was classified as "partially correct" as they were either able to guess whether the data contained outliers but failed to describe them or could describe only a few. There are formulae for determining whether data contains outliers but none of the learners who performed "completely correctly" and "partially correctly" could provide these. This suggests that learners must have simply guessed the correct solutions.

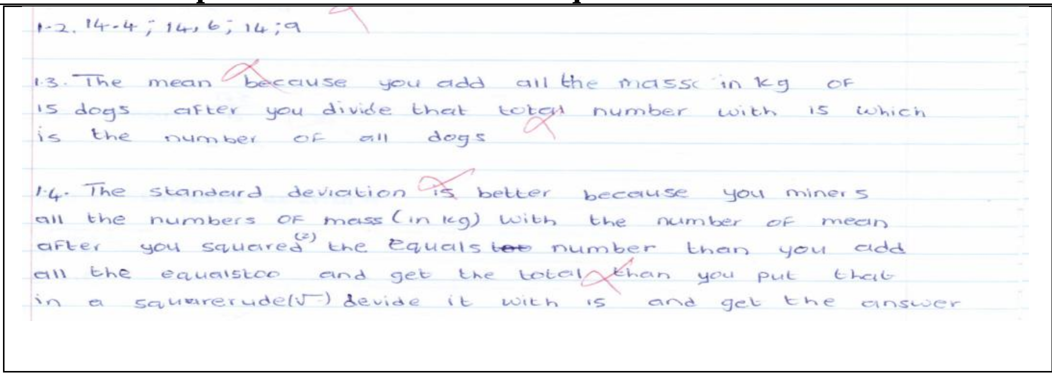
Vignette B: Example of learner's answer to question 1.4



1.3 Median, it does n't need a calculator because is one number only.

1.4 Standard deviation, because you use 6 numbers and in inter-quartile you take the whole numbers after you deviate.

Vignette C: Example of learner's answers to question 1.2 and 1.4



1.2. 14-4 ; 14,6 ; 14 ; 9

1.3. The mean because you add all the mass in kg of 15 dogs after you divide that total number with 15 which is the number of all dogs

1.4. The standard deviation is better because you miner s all the numbers of mass (in kg) with the number of mean after you squared(2) the equals number than you add all the equalstoo and get the total than you put that in a squarerude(√‾) devide it with 15 and get the answer

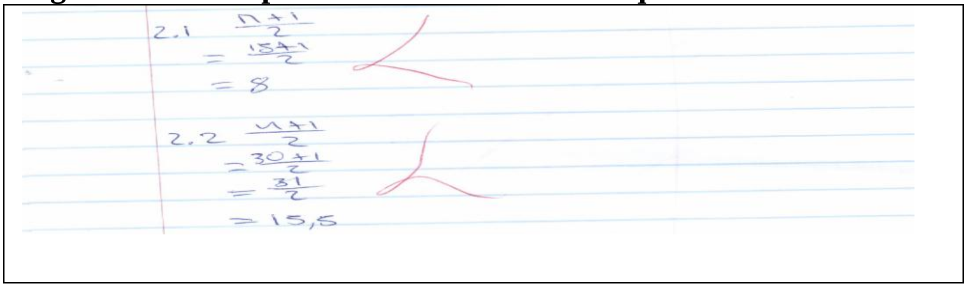
Vignette C shows that the learner was completely unable to describe the outliers in the provided data (see question 1.2). The learner could neither tell whether the data contained outliers nor describe the correct outliers from the data. Also, Vignettes B and C indicate that neither learner was able to answer question 1.4: they could not explain the better measure of dispersion when outliers are present in the data.

Most learners performed better in questions 1.1 (a) and 1.1 (b), which covered aspects of the mean and the median for ungrouped numerical data. For instance, in question 1.1 (a), 162 (65%) learners calculated the mean of ungrouped numerical data correctly and in question 1.1 (b), 189 (76%) could calculate the median correctly.

Question 2

Most learners performed badly in question 2, which dealt with aspects of the lower quartile and upper quartile. For instance, in question 2.1, the majority (209 or 84%) calculated the lower quartile for grouped numerical data incorrectly. In question 2.2, 198 (80%) learners were unable to compute the upper quartile for grouped numerical data. See vignette E of learner's answer to question 1.2.

Vignette E: Example of learner's answers to question 2.1 and 2.2

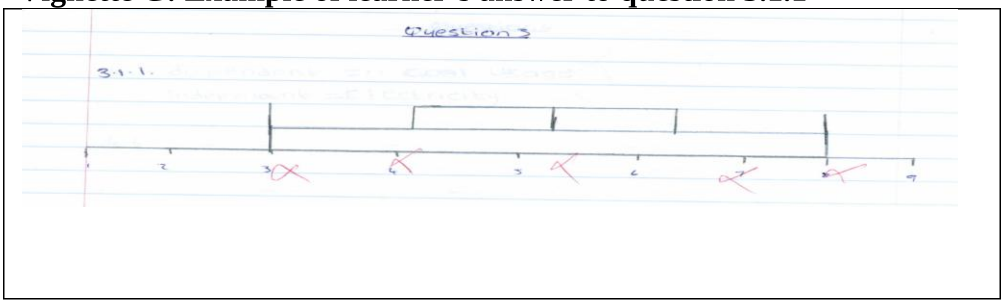


Vignette E shows that the learner provided completely wrong answers to questions 2.1 and 2.2. The learner failed to provide the correct formulae for lower quartile and upper quartile for grouped numerical data, which made the entire solution incorrect.

Question 3

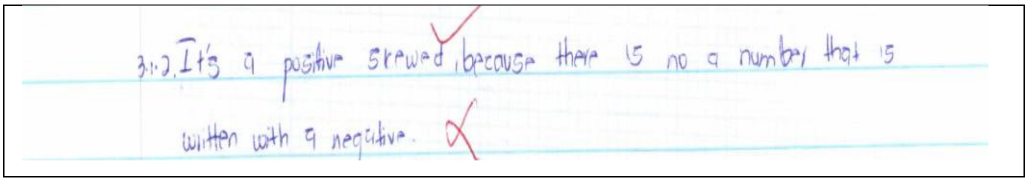
Table 17 indicates that most learners performed badly in question 3, which dealt with aspects of the box-and-whisker plot. For instance, in question 3.1.1, 135 (54%) of the learners were unable to draw a box-and-whisker plot. In question 3.1.1, most of the learners had difficulty determining the five number summary since box-and-whisker plots are constructed from this (that is, the minimum data value, lower quartile, median, upper quartile and maximum data value). Also, in question 3.1.2, 111 (45%) learners failed to describe and explain the skewness of the box-and-whisker plot. In this question, 71 (29%) of 248 learners answered partially correctly in that they could determine the type of skewness they observed in the box-and-whisker plot but were unable to explain this skewness. See vignettes G and H of learners' scripts for question 3.1.1 and 3.1.2 respectively.

Vignette G: Example of learner's answer to question 3.1.1



Vignette G shows that the learner was unable to answer question 3.1.1. The learner had an understanding of the shape of the box-and whisker plot, but had difficulty determining the five number summary since box-and-whisker plots are constructed from the five number summary.

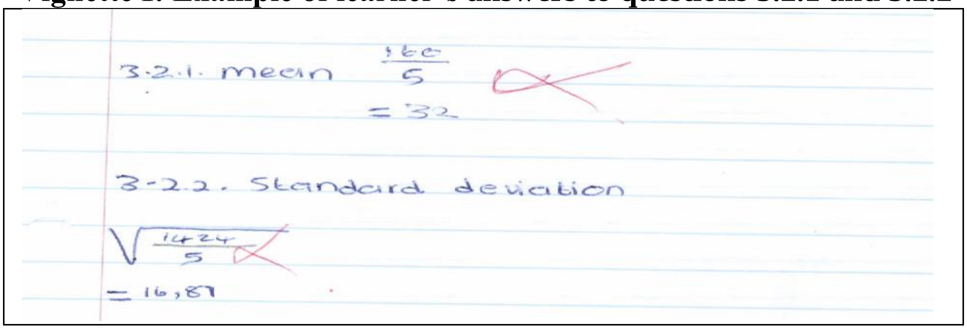
Vignette H: Example of learner's answer to question 3.1.2



Vignette H shows that the learner performed partially correctly in question 3.1.2 because he/she could tell the type of skewness he/she observed in the box-and-whisker plot but could not explain the skewness. For example, the learner explained that the plot was positively skewed because "...there is no number that is written with a negative". The explanation of the skewness of the plot or data indicated that the learner had difficulty understanding the concept of the distribution of data.

Most of the learners performed badly in question 3 which covered the mean and standard deviations for grouped numerical data. For instance, in question 3.2.1, 228 (92%) of the learners calculated the mean for grouped numerical data incorrectly. Again, in question 3.2.2, 195 (79%) learners were unable to calculate the standard deviation for grouped numerical data (see vignette I of learner's script for question 3.2.1 and 3.2.2).

Vignette I: Example of learner's answers to questions 3.2.1 and 3.2.2



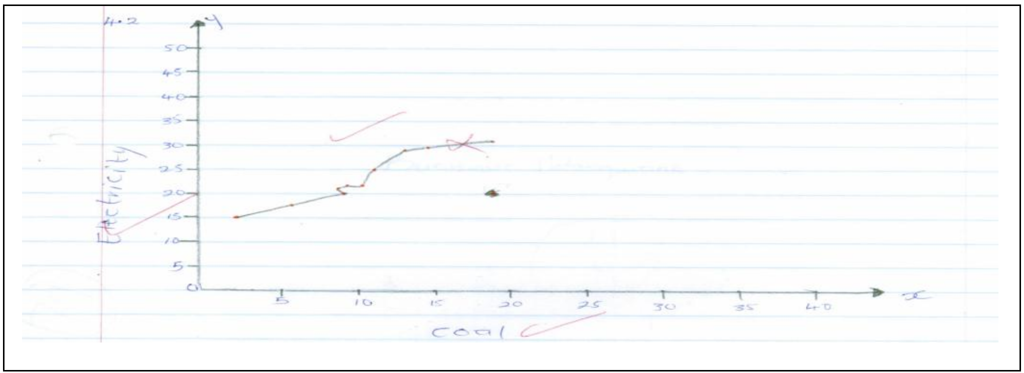
Vignette I shows that the learner answered completely incorrectly in questions 3.2.1 and 3.2.2. The learner used the wrong formulae for calculating the mean and the standard deviation for grouped numerical data.

Furthermore, most learners (159 (64%), 230 (93%), and 204 (83%)) performed badly in questions 3.3.1, 3.3.2, and 3.3.3 respectively, which covered aspects of the ogive. Learners calculated and interpreted the graph incorrectly when they were required to calculate the percentages of less than or more than type questions using the ogive.

Question 4

The majority of learners performed badly in question 4, which covered the aspects of the scatter plots and lines of best fit. For instance, in question 4.1, 108 (44%) learners had difficulty identifying the independent and dependent variables from the two numerical variables. In question 4.4, 203 (82%) learners were unable to draw a line of best fit on the scatter plot. See vignette J of learner's script for question 4.4 below.

Vignette J: Example of learner's answer to question 4.4



Vignette J indicates that the learner demonstrated the ability to plot a scatter plot, but could not draw a line of best fit on this plot, and was unable to use the correct methods of constructing a line of best fit (see section 4.2.7.1)

However, we observe from table 17 that most of the learners performed better on question 4.2. (see vignette J). For instance, 100 (40%) of the 248 learners demonstrated the ability to draw a scatter plot.

4.3.1 Summary of analysis of the diagnostic test

This section gives a brief summary of the analysis of the diagnostic test in relation to the research questions.

4.3.1.1 Problems encountered by learners in certain topics in statistics

The data from the diagnostic test showed that most of the learners encountered problems when computing the mean (92%) and standard deviation (79%) for grouped numerical data; computing the lower quartile (84%) and upper quartile (80%) for grouped data; constructing a box-and-whisker plot (54%); constructing a line of best fit on a scatter plot (82%). The data from the diagnostic test also showed that half the learners (50%) had insufficient knowledge about outliers and their effect on the data. This suggests that most of learners encountered difficulties with data handling.

4.3.1.2 Causes of observed learners' problems when learning statistics

The data from the diagnostic test indicated that learners encountered problems because (1) certain sections of data handling were not properly taught in the lower grades (i.e. outliers). Most learners (50%) did not know how to identify the outliers in the data or to describe the better measure of central tendency (measure of central location) and the better measure of dispersion when outliers are present in the data. The teaching of outliers started in the previous grades (i.e. grade 10), but grade 11 learners also encountered problems with the concept of outliers.(2) Learners were not taught certain topics in data handling. The majority of the learners had no idea of how to calculate the lower quartile (84%), upper quartile (80%) or standard deviation (79%) with numerical data that has been grouped into intervals (see section 4.2.3.1 and 4.2.6.1). (3) Most of the learners lacked conceptual knowledge of certain concepts in statistics (data handling implied). (4) Learners were not taught how to use the statistical function mode on the calculator to make statistical calculations. Most of the learners (74%) calculated standard deviation for ungrouped numerical data incorrectly because they used long formulae.

4.4 Interviews with learners and teachers

The objectives of the interviews were to gain further insight, understanding, meaning, constructions and perspectives of the interviewee's own experiences or knowledge on various issues (Denzil & Lincoln, 2005). Semi-structured interviews were conducted with a group of six teachers and ten learners from the four schools which were selected for classroom observations. The semi-structured interviews for both teachers and learners took place immediately after the diagnostic test had been administered; all the interviews were conducted by the researcher. All questions for learners and teachers were in English and posed to all respondents in the same sequence as they appeared in the interview schedules, in an attempt to reduce bias as much as possible. All learners and teachers were individually interviewed. The outcomes of these interviews are presented in the sub-sections below.

4.4.1 Learner interviews

Learners were interviewed to gather more information on problems they encountered with data handling and probability in grade 11, as well as to investigate the causes of these problems. The following themes emerged from the interviews: (1) learners' attitudes to the learning of data handling and probability, (2) learners' problems with data handling and their views on the possible causes of these problems. (3) learners' opinions about the teaching of data handling, (4) learners' problems with probability, (5) learners' views on the problems they had with probability, and (6) learners' opinions about teachers' problems in teaching probability. The interviewees were identified as L1, L2, L3 and so on, with the letter "L" standing for learner and the number indicating the order in which they were interviewed.

4.4.1.1 Learners' attitudes to the learning of data handling and probability

This question aimed to explore learners' attitudes toward learning statistics (data handling and probability). Learners were asked whether they liked learning data handling and probability, and also to provide the reasons for their responses. All the learners indicated that they enjoyed data handling because they found it the easiest section in mathematics. They said data handling was a practical topic and could easily be related to real-life experience. Almost all the learners, seven of the ten, indicated that they could not say whether they liked learning about probability since they had last learned the basics in previous grades (grade 9 and 10). Only learners L8, L9 and L10 indicated that they enjoyed learning about probability. Some responses were as follows (all responses are linked to the researcher's question provided below):

Researcher: *Do you like learning data handling and probability? Why?*

L1*: I enjoy data handling because it is practical. I learned probability in grade 9.*

L2*: I enjoy data handling, it's not difficult. I learned probability in grade 9.*

L3*: I enjoy data handling; it's an easiest topic in mathematics. I learned probability in grade 10.*

L4*: I enjoy data handling, it's practical. I do not know what probability is.*

***L5**: I enjoy data handling because so far in mathematics it is the only topic that I understand although there are topics in data handling that I encounter problems with. I think I encounter problems with probability since I have not learned probability in the previous grade.*
***L6**: I enjoy data handling: it is an easiest chapter in mathematics. I am not sure if I enjoy the learning of probability because we learned little about probability in grade 9 and we did not learn probability in grade 10 so I do not understand anything about probability.*
***L7**: I enjoy data handling and probability; it is practical*
***L8**: I enjoy data handling and probability; it is an easy chapter as compared to other chapters in mathematics.*
***L9**: I enjoy data handling and probability; it is easy, practical and can be related to many things.*
***L10**: I enjoy data handling and probability; it is easy, practical and can be related to many things.*

4.4.1.2 Learners' problems with data handling

The following questions were asked about problems learners had with data handling: *Do you encounter problems with the learning of data handling? What areas of data handling do you encounter problems to learn?* The questions were meant to determine whether or not learners encountered problems with data handling. The learners indicated that they had problems with quartiles (lower quartile (Q1), middle quartile (Q2), upper quartile (Q3) with grouped data; inter-quartile range; box-and-whisker plots; scatter plots; drawing and interpretation of ogive; calculation of mean and standard deviation with grouped data; and outliers. For instance, L1 indicated that she had great difficulties with box-and-whisker plots. She responded: "*...to draw a box-and-whisker plot you need to have Q1, Q2, Q3, Q4 and Q5 so I do not know to arrange them on the plot*". The box-and-whisker plot has a minimum number, Q1, Q2, Q3, and a maximum number not Q4 and Q5. Other learners' responses were as follows:

Researcher: *Do you encounter problems with the learning of data handling? What areas of data handling do you encounter problems to learn?*
***L1**: Box-and-whisker plot. To draw a box-and-whisker plot you need to have Q1, Q2, Q3, Q4, Q5 so I do not know how to arrange them on the plot.*
***L2**: Quartiles (Q1,Q2,Q3) with grouped data and scatter plot. I saw the word scatter plot for first time when we were writing the test.*
***L3**: Quartiles (Q1,Q2,Q3); and drawing and interpreting ogive.*
***L4**: Quartiles (Q1,Q2,Q3) with grouped data; mean and standard deviation with grouped data.*

L5*: Inter-quartile range, quartiles (Q1,Q2,Q3) with grouped data; mean and standard deviation with grouped data.*
L6*: Mean and standard deviation with grouped data; outliers. I was never taught outliers I just saw it in the textbooks*
L7*: Box-and-whisker plot; drawing and interpreting ogive, standard deviation, variance, inter-quartile range.*
L8*: Quartiles (Q1, Q2, Q3); drawing and interpreting ogive, standard deviation and mean of grouped data. Concepts like outliers. I was not taught about the outliers, I just saw the term outliers in the test. It would have been better if we had a mathematics dictionary that explains statistics concepts.*
L9*: Quartiles (Q1, Q2, Q3); drawing and interpreting ogive, standard deviation and mean of grouped data. Scale is a problem when drawing ogive diagram sometimes we do not get an s-shape.*
L10*: Quartiles (Q1,Q2,Q3) with grouped data; standard deviation and mean of grouped data; interpreting ogives, scatter plot.*

4.4.1.3 Learners' views about possible causes of problems they had with data handling

The learners were asked to indicate what they thought were the possible cause(s) of their problems, if any, when learning about data handling. LearnersL1, L2, L3, L4, L9and L10,for example, cited their teachers as the source of their problems. They claimed their teachers skipped difficult topics and only taught the easy ones. Learners' responses were as follows:

Researcher*: What are the cause(s) of the problems you encounter in the learning of data handling, do you think?*
L1*: The teacher does not teach all the topics in data handling and when the teacher teaches some of the topics he does not explain thoroughly.*
L2: *My teacher does not teach some of the topics in data handling.*
L3: *My teacher does not teach us some of the topics, he only teaches us simplest things.*
L4*: My teacher does not teach us some of the topics. My teacher is impatient, does not do the follow-ups to see if all the learners understand. My teacher taught us quartiles, mean and standard deviation with ungrouped data. Only teaches us simplest things.*
L9*: The teacher, she skips some of the topics. For example, I was not taught how to determine mean and standard deviation with a grouped data.*
L10*: The teacher, she skips some of the topics. For example, I was not taught how to determine mean and standard deviation with a grouped data and scatter plot.*

However, L5, L6 and L8 attributed the cause of their problems to the way data handling was taught and to their lack of commitment because, as L8 responded: " ...*The teacher do not teach some of the topics and learners do not practise data handling since we think data* *handling is simple*". L7 thought the time of day statistics was taught was the problem. Statistics was taught after lunch and by that time most learners were tired. The following are the learners' responses:

> ***L5***: *Myself and the teacher.*
> ***L6***: *Learners are making noise during the lessons.*
> ***L7***: *The timing of mathematics period. Mathematics is taught after lunch and by that time we are tired.*
> ***L8***: *Both teacher and learners. The teacher does not teach some of the topics and learners do not practise data handling since we think data handling is simple.*

4.4.1.4 Learners' opinions about the teaching of data handling

When learners were asked about possible problems experienced by their teachers in the teaching of some aspects of data handling they conceded that their teachers did seem to have problems. For instance, L4, L5 and L7 felt their teachers battled teaching quartiles (Q1, Q2, Q3) with grouped data; mean and standard deviation with grouped data; the use of a calculator in teaching statistics and the scatter plot. The learners responded as follows:

Researcher: *Which aspects of data handling do you think your teacher finds difficult to teach?*

> ***L4***: *Quartiles (Q1, Q2, Q3) with grouped data; mean and standard deviation with grouped data.*
> ***L5:*** *Quartiles (Q1, Q2, Q3) with grouped data; mean and standard deviation with grouped data. Topics that she does not teach us. The use of a calculator to calculate the mean, standard deviation and variance.*
> ***L7***: *Standard deviation with grouped data, the use of the calculator, scatter plot.*
> ***L8***: *The teacher is good in what she is teaching, the problem is she skips some of the topics and if she was teaching all the topics we would not encounter problems with data handling.*

However, L1, L2 and L3 felt that their teachers did not have difficulty teaching data handling. Their responses were based on the view that their teachers had been teaching data handling for a long time.

> *Three learners,* L1, L2 *and* L3*provided a "none" response, suggesting that their respective teachers did not encounter problems with the teaching of data handling.*
> ***L6***: *No, I do not think my teacher encounter problems with the teaching of data handling.*

__L9__: My teacher is experienced with the teaching of data handling and has been teaching it for long time.
__L10:__ My teacher is experienced with the teaching of data handling and has been teaching it for long time.

4.4.1.5 Learners' problems with probability

When asked about the problems they encountered when learning probability, Learners L3, L7, and L8 indicated that they had problems with the construction of tree and Venn diagrams; mutually exclusive and independent events. Other learners such as L1, L2, L4, L5 and L6 indicated that they encountered problems with probability but when they were asked which aspects they did not mention the problems because they had learned probability in previous grades. Only L9 and L10 indicated that they had not encountered any problems with the learning of probability. Learners responded below as follows:

__Researcher__: Do you encounter problems with the learning of probability? What aspect(s) of probability do you find difficult to learn?
__L1__: I do not know because I have not learned probability in grade 11, since I studied probability in grade 9, and it was not that deep.
__L2__: I do not know because I have not learned probability in grade 10 and 11, but I had some problems with the learning of probability in grade 9.
__L3__: I have problems with the learning of tree and Venn diagrams; mutual exclusive and independent events.
__L4__: Yes, I do not even know what probability is.
__L5__: Yes, I do not even know what is all about probability.
__L6__: Yes, I was not taught probability for two years now.
__L7__: I have problems with the learning of probability. Tree and Venn diagrams; and mutual exclusive and independent events.
__L8__: Yes. Tree diagrams, Venn diagrams, mutual exclusive events and independent events.
__L9__:No.
__L10__: No.

4.4.1.6 Learners' views about possible causes of problems they experienced when learning about probability

When asked what they thought the reasons for their with the learning of probability were, L3, L5, L6 and L7 indicated that their teachers caused them problems because they did not teach them probability, and those who had taught probability were not as confident as they were when they taught other topics. Learner L8 indicated that he experienced difficulties because "….we do not have good foundation of probability, in grade 9 they taught us the simplest things and in grade 10 we learned few things, like tossing a dice. Then we get to grade 11, the teacher assumes that we have learned everything in the previous grades and we were not taught some of the topics. Poor background in mathematics, some of the topics we were not taught in the previous grades".

***Researcher**: what are the cause(s) of the problems that you are experiencing with the learning of probability, do you think?*

> ***L3**: My teacher, he is not confident when teaching probability like other topics. my teacher does not give more examples in probability.*
> ***L5**: My teacher because he is not teaching the chapter of probability, my teacher always skips the chapter of probability since from grade 10.*
> ***L7**: My teacher, she is not clear when teaching probability. She does not explain thoroughly to make learners understand.*
> ***L8**: We do not have good foundation of probability, in grade 9 they taught us the simplest things and in grade 10 we learned few things, like tossing a dice. Then we got to grade 11, the teacher assumes that we have learned everything in the previous grades and we were not taught some of the topics. Poor background in mathematics, some of the topics we were not taught in the previous grades.*

Some of the learners, such as L1, L2, L4, and L6, indicated that they did not know whether they had any problems with probability because they had not been taught the topic yet.

> ***L1**: I do not know because I studied it in grade 9.*
> ***L2**: I do not know because I have not learned probability in grade 10 and 11.*
> ***L4**: I was not taught probability for about two years now, so I think is one of the problems.*
> ***L6**: I am not sure because I was not taught probability for two years now. I think is the teacher for not teaching us.*

4.4.1.7 Learners' opinions about teachers' problems in teaching probability

This question allowed the researcher to probe whether learners felt that their teachers encountered problems with teaching probability. All but L4, L9 and L10 indicated that they did not know whether their teachers encountered any problems with the teaching probability, since they had not had the experience of observing their teachers teaching this topic. Only L9 and L10 indicated that they did not think that their teachers encountered any problems with the teaching of probability. Learners' responses are indicated below:

***Researcher**: Do you think your teacher is encountering problems with teaching probability? Which aspect(s) of probability do you think your teacher finds difficult to teach?*

> ***L1**: I do not know because I have never seen my teacher teaching probability.*

***L2**: I do not know because I have never seen my teacher teaching probability.*
***L3**: I do not know. All I know is he is not confident like other chapters and that he does not provide more examples and explain thoroughly.*
***L4**: Yes, I think my teacher encounter problems with probability. I do not know which aspects of probability because he did not teach us probability even in grade 10.*
***L5**: I do not know because he did not teach us probability even in grade 10.*
***L6**: I do not know because I have not seen my teacher teaching probability.*
***L7**: I do not know.*
***L8**: I do not know.*
***L9**: No.*
***L10**: No.*

4.4.2 Teacher interviews

The purpose of the teachers' interviews was to investigate any problems they might have experienced with statistics (data handling and probability) and to investigate the possible causes of these problems and teachers' suggestions on how to address them. Table 6 presents teachers' demographic data (i.e. qualifications, subject specialisation, grades to which they were currently teaching mathematics, highest qualification in statistics and years of experience in teaching statistics). A total of six teachers were interviewed. Teachers T2, T3, T4 and T6 were observed teaching while T8 and T9 were not. Teachers T1, T5 and T7 were not interviewed because they did not want their voices to be tape-recorded.

The themes that emerged from the semi-structured interviews with teachers were: teachers' problems with data handling, their views on possible causes of problems they had with data handling, their problems with probability, views on possible causes of problems they had with probability, and their suggestions on how to address problems in the teaching of data handling and probability. The themes are discussed in detail below.

Table 18: Teachers Background

	Teacher eight(T8)	Teacher four (T4)	Teacher nine(T9)	Teacher two(T2)	Teacher six(T6)	Teacher three(T3)
Teaching qualification	B. Ed. Management Sciences	B. Ed. FET specialisation	B.Tech. Education Management	STD, ACE and B.Tech.	B. Ed.	ACE
Subject of specialisation	Mathematics& Computer Application Technology	Mathematics &Physical Science	Physical Science &Mathematics	Mathematics	Mathematics	Mathematics
Grades currently being taught Mathematics	Grades 10 and 11	Grades 10, 11 and 12	Grades 10, 11 and 12	Grades 11 and 12	Grade 11	Grades 10 and 11
Studied statistics at tertiary level	No	Studied statistics as a topic in Mathematics, not as a course	No	Studied statistics as module when I was studying B.Tech.	Yes	Yes
Highest qualification in statistics	Never studied statistics	Studied it in level three (third level)	Never studied statistics	Studied statistics as a module	Statistics at first year	Level three
Years of experience of teaching statistics	Six years	First year of teaching	About six years –since the introduction of National Curriculum Statement	About six years	About six years	Four years

4.4.2.1 Teachers' problems with data handling

When asked about the problems they encountered when teaching data handling, T4, T8 and T9 cknowledged that they found data representation difficult: histograms with grouped data, ogive, box-and-whisker plots, calculation of quartiles (Q1 & Q3), variance and standard deviations with grouped data, median when the total number of observations was even, and scatter plots. T2, T3and T6 indicated that they had no difficulties teaching data handling. The responses are indicated below.

Researcher*: What problems do you experience when teaching data handling?*

> ***T8****: Data representations, that is, histogram (grouped data), ogive (cumulative frequency graphs) with grouped data and box-and-whisker plot with grouped data.*
>
> ***T4****: Calculation of quartiles (Q1 and Q3). The textbooks, memorandum and workshops are not saying the same thing. National memorandum and textbooks*

sometimes use (n+1)/4 or 3(n+1)/4 and at other times they list and count the values.
***T9**: Calculating variance and standard deviation with grouped data, median when the total number of observations (n) is even, and scatter plot.*
***T2**: I do not experience any problems in the teaching of data handling.*
***T6**: I do not experience any problems in the teaching of data handling.*
***T3**: I do not experience any problems in the teaching of data handling.*

However, what T2, T3 and T6 said in the interviews contradicted the researcher's classroom observations. It was observed that T2 experienced problems when teaching cumulative frequency graphs (see section 4.2.1), measures of central tendency (see section 4.2.2) and measures of dispersion (see section 4.2.6); T3 encountered problems when teaching measures of central tendency (see section 4.2.2), the five number summary (see section 4.2.3), measures of dispersion (see section 4.2.6) and scatter plots (see section 4.2.7); and T6 had problems teaching cumulative frequency graphs (see section 4.2.1), measures of central tendency (see section 4.2.2), the five number summary (see section 4.2.3), box-and-whisker plots (see section 4.2.5) and measures of dispersion (see section 4.2.6). It is clear, as Robson (2002:310) noted, that observation "provides a reality check", because "what people do may differ from what they say". Owusu-Mensah (2008) suggests that, in addition to self-evaluation by teachers, classroom observations should be conducted to determine further in-service needs as it is not easy for some teachers to acknowledge their teaching difficulties.

4.4.2.2 Teachers' views on possible causes of problems they may have had with data handling

Furthermore, when teachers were asked to indicate what they thought the reasons for their problems when teaching data handling were, T4, T8 and T9 indicated that inadequate content knowledge of statistics was the chief factor. One of the teachers, T4, also added that the content of textbooks, memoranda and workshops did not correspond. This teacher responded, "*...textbooks, memoranda and workshops are not saying the same thing about calculating lower quartile (Q1) and upper quartiles (Q3)*". Also, T9 indicated that a cause of these problems was the failure to attend workshops.

***Researcher**: What are the causes of the problem(s) you experience in the teaching of data handling, do you think?*

__T8__: I lack statistics content knowledge. I did not study statistics during my pre-service training.
__T4__: I lack statistics content knowledge. Textbooks, memoranda and workshops are not saying the same thing about the calculating lower quartile (Q1) and upper quartiles (Q3).
__T9__: I did not study statistics during my pre-service training. I did not attend any in-service education and training in the past.

4.4.2.3 Teachers' problems with teaching probability

When asked about the problems encountered when teaching probability, T4, T6 and T8 indicated that they had difficulties constructing probability diagrams, tree and Venn diagrams, two-way contingency tables, and difficulty understanding probability terminology and mutually exclusive events and independent events. T2 and T9 said that they did not know whether they encountered any problems teaching probability since they had not taught it yet. Only T3 indicated that the teaching of probability presented no difficulties.

***Researcher**: What problems do you experience when teaching probability?*
__T8__: Probability diagrams, for an example tree and Venn diagrams
__T4__: Mutually exclusive events and independent events confuse me. I do not understand them. Again I have problems with constructing tree diagrams where events are dependent. Still the terms independent events and dependent events give me a hard time.
__T9__: I have not yet taught probability so I do not know if there are problems that I encounter in the teaching of probability.
__T2__: I am not sure if I encounter problems with the teaching of probability, most of the times we do not teach probability because we do not have learners who write optional paper. Our learners are weak so we do not teach probability as it will extend the syllabus whereas learners struggle with some topics which are compulsory.
__T6__: Constructing tree and Venn diagrams; constructing two way contingency tables.
__T3__: I do not experience any problems with the teaching of probability.

4.4.2.4 Teachers' views on possible causes of problems they may have had with probability

When asked what they thought the causes of their problems with the teaching of data handling might be, T4 and T8 admitted that they encountered problems because of a lack of content knowledge. T6 indicated that he/she found the teaching of probability problematic because workshops focused more on data handling than on probability. The teachers responded as follows:

Researcher*: What are the causes of your problem(s) in the teaching of probability, do you think?*

> ***T8****: I lack statistics content knowledge. I did not study probability during my pre-service training. Probability is always not taken seriously even by the facilitators during the workshops.*
> ***T4****: I lack statistics content knowledge. I was not taught probability during my pre-service training.*
> ***T6****: Most of the workshops focus more on data handling than probability.*

4.4.2.5 Teachers' suggestions on how to address problems in the teaching of data handling and probability

When asked what should be done to address problems in the teaching of data handling and probability,T2, T3, T6, T8 and T9 suggested that teachers should attend more workshops and that these should run for at least five days; T4 suggested that teachers should have discussion groups at their schools to assist each other. T6 thought there should be lengthy workshops on probability as this had been neglected in past workshops. The teachers responded as follows:

Researcher*: What do you think should be done to address the problems experienced by teachers in data handling and probability?*

> ***T8****: Teachers should attend many workshops, for example five-day workshops.*
> ***T4****: Teachers should start discussions groups at their schools to assist each other about the problems they encounter with the teaching of statistics, because workshops are not found at anytime teachers need them.*
> ***T9****: Teachers should attend as many workshops as they can.*
> ***T2****: Stakeholders, organisations and departments of education should organise the workshops. Before they organise the workshops they should find out from the teachers about the topics they encounter problems with.*
> ***T6****: There should be workshops on probability topic with long hours as it has been neglected with the past workshops. Many workshops I have attended have been concentrating more on data handling than probability.*
> ***T3****: There is a need for lesson-study type of workshops in order to integrate probability and statistics content knowledge and pedagogical content knowledge.*

4.4.2.6 Summary of analysis of interviews

This section provides a brief summary of the analysis of the teacher and learner interviews in relation to the research questions.

4.4.2.6.1 Teacher interviews

4.4.2.6.1.1 Problems encountered by teachers when teaching certain topics in statistics

The data from teacher interviews showed that some teachers (T8, T4 and T9) encountered problems with the construction of a histogram, ogive, box-and-whisker plot with grouped data, calculations of the quartiles (Q1 andQ3), variance and standard deviation with numerical data that had been grouped into intervals and the median (Q2) when the total number of observations were even. Lastly, the results revealed that some of the teachers (T4, T6 and T8) encountered difficulties when constructing probability diagrams, tree and Venn diagrams and two-way contingency tables. They also found it difficult to understand probability terminology (i.e. mutually exclusive events and independent events).

4.4.2.6.1.2 Causes of teachers' problems observed when teaching statistics

The data from the teacher interviews indicated that T4, T6 and T8 encountered problems when teaching because (1) they lacked statistics content knowledge; (2) in-service teacher workshops focused more on data handling than on probability; (3) textbooks, memoranda and workshops were not consistent regarding certain topics in data handling (i.e. quartiles); and (4) some of the teachers indicated that they encountered problems because they had not attended in-service teacher workshops.

4.4.2.6.1.3 Possible ways to address the problems encountered by teachers in their teaching of statistics

The data from the teacher interviews suggested that possible solutions to the problems encountered by teachers in their teaching of statistics might be addressed by: (1) in-service teacher programmes meeting the needs of teachers by offering the topics they find difficult in their teaching; (2) offering longer and more frequent inset programmes on probability, preferably five-day workshops.

4.4.2.6.2 Learner interviews

4.4.2.6.2.1 Problems encountered by learners when learning certain topics in statistics

The findings from the learner interviews were that more than a quarter of the learners interviewed (L4, L5 and L7) encountered difficulties with: the computation of quartiles (Q1, Q2, Q3) with grouped data; computation of inter-quartile range; construction of a box-and-whisker plot; construction of a scatter plot; construction and interpretation of the ogive; calculation of mean and standard deviation with grouped data; and outliers. In addition, most of the learners, L4, L5, L7 and L8, mentioned that they encountered problems with the construction of tree and Venn diagrams and with understanding and calculating probability of mutually exclusive and independent events.

4.4.2.6.2.2 Causes of observed learners' problems when learning statistics

The interviews revealed that the majority of the learners, L3, L5, L7 and L8, encountered problems because(1) certain sections of data handling and probability had not been properly taught in the lower grades; (2) teachers taught only some and not all topics; (3) teachers did not explain some of the concepts in statistics thoroughly.

4.5 Answers to research questions

The results of the data analysis presented above were used to address the research questions posed in this study.

4.5.1 Research question one

The first research question was:

- What problems, if any, are encountered by teachers in the teaching of statistics (data handling and probability) in grade 11?

The findings were that the teachers encountered various problems with the construction and interpretation of probability diagrams and tables (i.e. Venn and tree diagrams, two-way contingency tables) (see section 4.1.1.3), and understanding or interpreting probability terminology difficult (i.e. mutually exclusive events, independent and dependent events etc.) (see section 4.1.1.3). The findings were that the some teachers encountered various problems with interpretation and determination or calculation of measures of dispersion (i.e. interpretation of variance and standard deviation, determining the five number summary (i.e. lower quartile, middle quartile and upper quartile with grouped data and when the total number of the data values (n) is even), and calculating variance and standard deviation with grouped data) (see section 4.2.6.1, 4.2.3.1). Further, the findings were that the some teachers experienced difficulties with the representation and interpretation of data on graphs or plots (i.e. cumulative frequency graph (ogive), box-and-whisker plot, scatter plot and line of best fit) (see section 4.4.2.1, 4.2.7, 4.2.5.1). However, the findings from the teacher questionnaire indicated that teachers had no problems with interpretation and determination or calculation of measures of dispersion and also with representation and interpretation of data on graphs or plots (see section 4.1.1.2). This could have been expected since it was a self-evaluation by teachers (Owusu-Mensah, 2008).

4.5.2 Research question two

The second research question was:

- What problems, if any, are encountered by learners in the learning of statistics (data handling and probability) in grade 11?

The findings of the analysis were that the learners had difficulty using graphs to predict results (i.e. using a diagram of ogive to: estimate the cumulative percentages; estimate the middle quartile from a set of grouped data; to estimate the lower quartile). They also had difficulty identifying functions that best fit the data and answering less than or more than type questions with the ogive graph (see section 4.1.2.1, 4.2.1.2, 4.3.1.1); learners had problems interpreting and determining measures of dispersion (i.e. using a calculator to compute variance and standard deviation; interpretations of variance and standard deviation; calculations of variance and standard deviation when data has been grouped into intervals); and they experienced difficulty with the computation of quartiles (Q1, Q2, Q3) when the total number of data values was even (see section 4.1.2.1, 4.3.1.1, 4.4.1.2, 4.2.6.2); learners found representing data on graphs or plots difficult (i.e. constructing a line of best fit on the scatter plot, constructing an ogive, and representing bivariate numerical data as a scatter plot) (see section 4.1.2.1, 4.2.7.2, 4.3.1.1); Learners had difficulty interpreting and determining measures of central tendency (i.e. using a calculator to compute the mean with ungrouped data, calculating the mean and mode with grouped data, interpreting the mean and median; in sufficient knowledge about outliers and their effect in the data) (see section 4.1.2.1, 4.3.1.1); learners had problems with the construction and interpretation of probability graphs and tables (see section 4.1.2.2); and further learners had problems to understand or interpret probability terminology (i.e. construction of tree and Venn diagrams, and understanding and computing probability of mutually exclusive and independent events) (see section 4.1.2.2).

4.5.3 Research question three

The third research question was:

- What are the cause(s) of the problems encountered in the teaching and learning of statistics in grade 11?

The causes of the problems in the teaching of statistics in grade 11

According to the findings, possible causes of these problems were: (1) teachers lacked content knowledge of statistics because they had not studied statistics; those who had, had not studied further (see section 4.1.1.4); (2) textbooks were not well written or thoroughly explained (i.e. formulae were missing and the books did not provide adequate examples) (see section 4.1.1.4, 4.2.3.3, 4.2.7.3); (3) teachers were not given the financial support to attend in-service education and training (inset) programmes (see section 4.1.1.4); and (4) most of the in-service teacher programmes teachers had attended did not cover statistics topics (i.e. most programmes did not cover the topic of probability [see section 4.4.2.4]).

The causes of the problems in the learning of statistics in grade 11

The following possible causes of problems in the learning of statistics were found: (1) inadequate teaching of statistics topics in lower grades (i.e. learners indicated that they encountered problems because certain sections of data handling and probability had not been properly taught in previous grades or had never been taught them (see section 4.1.2.3, 4.1.1.7); (2) teachers' content knowledge of statistics and their methods of teaching (i.e. learners indicated that they experienced difficulties because their teachers taught some topics but left others for them to do on their own (see section 4.1.2.3, 4.2.8.4, 4.4.1.3); their teachers did not allow enough time to teach statistics (section 4.1.2.3); their teachers did not give them enough exercises to practise statistics (section 4.1.2.3); and teachers had difficulty explaining concepts to learners (see section 4.1.2.3, 4.2.8.4, 4.1.1.7); (3) learning material (i.e. learners claimed their problems stemmed from the fact that they did not know how to use the statistics function mode on their calculators (see section 4.1.2.3, 4.3.1.2, 4.1.1.7); their prescribed textbook did not cover all the topics in the curriculum nor did it provide thorough explanations (see section 4.3.1.2); learners did not have access to previous examination papers with solutions (see section 4.1.2.3); (4) learners' attitudes to the learning of statistics (i.e. learners indicated that they encountered problems because they did not give themselves enough time to practise data handling and probability problems (see section 4.1.2.3, 4.1.1.7).

4.5.4 Research question four

The fourth research question was:

- What are possible solutions to the problems encountered in the teaching and learning of statistics in grade 11?

In an effort to answer the fourth research question, data from teacher questionnaires and teacher interviews was analysed. Findings suggested the following ways in which problems in the teaching and learning of statistics in grade 11 could be addressed: (1) teachers should receive financial support from their schools/districts to attend in-service education and training programmes(see section 4.1.1.4); (2) textbooks should be well written (provide thorough explanations) and contain all the information necessary to teach data handling and probability (i.e. formulae, more examples) (see section 4.1.1.4, 4.2.3.3, 4.2.7.3); (3) in-service teacher programmes should meet the needs of the teachers by offering topics that teachers find difficult to teach(see section 4.4.2.5); (4) more and longer inset programmes on probability, preferably five-day workshops, should be arranged (see section 4.4.2.5).

CHAPTER 5

SUMMARY OF THE STUDY, DISCUSSION OF FINDINGS, CONCLUSION AND RECOMMENDATIONS

5.1 Summary of the study

The purpose of the study was to explore problems encountered in the teaching and learning of statistics in grade 11 and to investigate ways of addressing these problems. Data on teachers' problems and their causes were collected through a questionnaire for teachers, classroom observations, teacher interviews and learner interviews, while data about learners' problems and their causes were collected from a questionnaire for learners, classroom observations, a diagnostic test, a teacher questionnaire and learner interviews. The data was analysed qualitatively using descriptive statistics (frequencies). It was found that teachers and learners did indeed encounter difficulties when teaching and learning statistics.

Teachers' problems lay particularly with the interpretation and calculation of measures of dispersion (i.e. interpretation of variance and standard deviation, and calculating variance and standard deviation with grouped data) as well as with the representation and interpretation of data on graphs or plots (i.e. cumulative frequency graph (ogive) with grouped data, scatter plot and line of best fit.). They had difficulty determining the five number summary (i.e. lower quartile, middle quartile and upper quartile with grouped data and when the total number of data values (n) was even). In addition, the findings were that teachers had difficulty with the construction and interpretation of probability diagrams and tables (i.e. Venn and tree diagrams, two-way contingency tables) and with understanding and interpreting probability terminology (i.e. mutually exclusive events, independent and dependent events etc.).

The findings also revealed that learners encountered problems when using graphs to predict results (i.e. using a diagram of ogive to estimate the cumulative percentages, to estimate the middle quartile from a set of grouped data and to estimate the lower quartile, in identifying functions that best fit the data and answering less than or more than type questions with ogive graphs). Learners were also revealed to experience difficulties with interpreting and determining measures of dispersion (i.e. using a calculator to compute variance and standard deviation, calculations of variance and standard deviation when data has been grouped into intervals,with the computation of quartiles (Q1, Q2, Q3) when the total number of data values is even). It appeared that learners also had problems with representing data on graphs or plots (i.e. constructing a line of best fit on the scatter plot, constructing an ogive, and representing bivariate numerical data as a scatter plot) and with interpreting and determining measures of central tendency (i.e. using a calculator to compute the mean with ungrouped data, calculating the mean and mode with grouped data, interpreting the mean and median; they also had insufficient knowledge of outliers and their effect in the data). Lastly, the majority of learners seemed to find the construction and interpretation of probability graphs and tables very difficult, as well as understanding or interpreting probability terminology (i.e. construction of tree and Venn diagrams, and understanding and computing the probability of mutually exclusive and independent events).

Possible reasons for teachers' problems were revealed as (1) their lack of content knowledge in statistics; (2) textbooks lacking in explicit explanations, with some formulae missing and within adequate examples; (3) most of the in-service teacher programmes they attended did not cover statistics topics, and those that did, did not pay adequate attention to probability, with the result that the programmes had no significant effect in improving the teachers' content knowledge; (4) teachers did not attend in-service teacher workshops.

Possible causes of learners' problems emerged as (1) the inadequate teaching of statistics topics in lower grades (i.e. certain sections of data handling and probability had not been taught, or had not been properly taught in previous years); (2) teachers' lack of statistics content knowledge and poor methods of teaching (i.e. teachers experienced difficulties in explaining concepts to learners); (3) learning material (i.e. learners did not know how to use the statistics function mode on a calculator when doing statistical calculations (i.e. the mean and variance or standard deviation), the prescribed learner textbook did not cover all the topics suggested by the curriculum nor did it provide explicit examples and explanations. Some formulae were missing and not enough examples were provided); (4) learners lacked conceptual knowledge of certain concepts in statistics.

In order to address these problems in the teaching and learning of statistics, the findings revealed that the following should be taken in to consideration: (1) in-service teacher programmes should meet the needs of the teachers by offering those topics that teachers find problematic during instruction. There should also be more and longer inset programmes on probability, preferably five-day workshops; (2) textbooks should be better written, should provide explicit explanations and contain all information necessary to the teaching and learning of data handling and probability.

5.2 Discussion of the findings

The findings are discussed below in the light of the research questions.

5.2.1 Problems experienced in the teaching of statistics

The findings of the analysis of the data were that some teachers in this study experienced difficulty with the interpretation and calculation of measures of dispersion such as variance and standard deviation. These findings are consistent with those of Cardoso(2007) and Da Silva and De Queiroz e Silva Coutinho (2006). In contrast, Sánchez and García (2008) found that teachers could compute measures of dispersion but could not analyse the values of the measures of dispersion. In this study, the teachers had problems with the representation and interpretation of plot.This finding can be compared to those of a study by Bruno and Espinel (2009), which found that a large percentage of primary school teachers had difficulties constructing histograms and making evaluations of graphs. Even though the study by Bruno and Espinel (2009) was not on ogive and scatter plot or line of best fit, per se, what is significant about it in the current discussion is the fact that the teachers battle with some form of graphs. Some teachers had difficulties with five number summary. Little research seems to have been conducted on problems experienced by teachers when teaching quartiles (lower quartile, middle quartile and upper quartile with grouped data and when total number of the data values (n) is even). It is hardly surprising that teachers in the present study had problems with these aspects, given that their statistical knowledge was not well grounded (see Atagana et al., 2009; Atagana et al., 2010; Wessels & Nieuwoudt, 2011). This study also found that teachers had difficulty constructing and interpreting probability diagrams and tables; and with the interpretation of probability terminology. With regard to the former, Groth (2007) also found that teachers had problems with the interpretation of data but not specifically with data presented in diagrams and tables. On the other hand, Bruno and Espinel (2009) found that teachers were having difficulty interpreting histograms and frequency polygons. Evidently, teachers have a track record of struggling with the interpretation of some aspects of statistics and this is certainly backed up by the findings of the current study. Regarding the teachers' difficulties with terminology, the argument is made that teachers with deficient statistical knowledge (see Atagana et al., 2009; Atagana et al., 2010; Wessels & Nieuwoudt, 2011) are most likely to have problems with the terminology of those particular aspects of statistics.

5.2.2 Problems experienced in the learning of statistics

This study found that learners had difficulties using graphs to predict results, such as using a diagram of ogive to estimate cumulative percentages or to estimate the middle quartile, lower quartile and upper quartile from a set of grouped data, identifying functions that best fit the data, and answering less than or more than type questions with the ogive graph. This finding can be compared to those of a study by Meletiou-Mavrotheris and Lee (2010), which found that learners have difficulties in tackling tasks involving group comparison and critical information presented graphically, especially simple reading and interpretation tasks. In the present study, learners had difficulty representing data on graphs or plots (i.e. constructing a line of best fit on a scatter plot and constructing an ogive). This finding supports those of Meletiou-Mavrotheris and Lee's (2005) study, in which a noticeable proportion of students had dificulties understanding graphical representations and interpretation of graphs such as the histogram, bar graph and line plot. Even though the study by Meletiou-Mavrotheris and Lee (2005) was not concerned with the line of best fit, the scatter plot, ogive or box-and-whisker plot, per se, what is significant about it in the current discussion is the fact that the learners struggled with any form of graph, either using transformed of data or raw data. Some of the learners had problems interpreting and determining measures of dispersion (i.e. interpretations of variance and standard deviation, using a calculator to compute variance and standard deviation, calculations of variance and standard deviation when data had been grouped into intervals). These findings are similar to those of a study by Slauson (2008), which found that students have difficulties reasoning with the measure of dispersion. In a study by Ghinis, Korres and Bersimis (2009) learners also faced difficulties, mainly in the procedure of solving problems, and in applying known statistical methodology to unfamiliar, real-life situations. Further, the learners in this study had problems with interpreting and determining measures of central tendency (i.e. interpreting the mean and median, using a calculator to compute the mean with ungrouped data, calculating the mean and mode with grouped data) and they revealed inadequate knowledge about outliers and their effect in the data. This is not surprising as studies by Da Silva and De Queiroz e Silva (2008) and Cardoso (2007) found that teachers taught learners how to calculate the mean, median and the standard deviation but not how these concepts could be related. Lastly, the findings in the present study show that the learners had problems with the construction and interpretation of probability graphs and tables, and also had difficulties understanding or interpreting probability terminology (i.e. construction of tree and Venn diagrams, and understanding and computing probability of mutually exclusive and independent events). It seems to be a universal problem, then, that learners have trouble interpreting aspects of statistics. As far as their problems with terminology are concerned, this study suggests that learners' ability to answer the questions is a reflection of their teachers' ability to answer similar questions and the knowledge and errors that learners display are the legacy of the lack of knowledge and the errors of their teachers. Learners taught by teachers with inadequate statistical knowledge are most likely to have difficulties with the terminology of aspects of statistics, as Cobb and Moore (1997) suggest: probability is more confusing and confusion persists even among those who can do textbook exercises. The findings of the present study support the researcher's earlier observations that teacher knowledge influences the depth of teaching and, in turn, the quality of learning. Most of the problems encountered by learners in the current study were also encountered by their teachers.

5.2.3 Possible causes of these problems

5.2.3.1 The causes of problems in the teaching of statistics in grade 11

The findings in this study were that almost all the teachers experienced difficulties because they lacked content knowledge of statistics. These findings are consistent with those of Davis and Simmt (2006), Koehler and Mishra (2009) and Shulman (1986). Teachers who do not have comprehensive base content knowledge cannot interpret idiosyncratic learner responses, prompt multiple interpretations, trace misconceptions, or plan rich learning experiences for learners; instead, these teachers may misrepresent these subjects to their learners, giving them, for example, incorrect information and developing misconceptions about the content area (Shulman,1986). The analysis also revealed that the teachers encountered problems because the textbooks they were using were not explicit enough. Certain formulae were missing and not enough examples were provided. These results are consistent with the findings of Lue (1998), who found that teachers experienced difficulties because some statistical concepts and some sections in the textbook were unclear and the content was boring. Textbooks are expected to provide a framework for what is taught, how it may be taught, in what sequence it can be taught, and to be explicit (Lemmer, Edwards& Rapule, 2008). Furthermore, the findings were that the teachers had difficulties because most of the inset programmes they attended did not cover all the statistics topics they required. This finding implies that the organisers of these programmes may not have conducted a needs analysis of the problems experienced by teachers during their classroom instruction. Moeini (2008) warns that teacher training that ignores a needs analysis as the first critical step in the development of a training programme leads to a waste of time, human resources and money, while at the same time dampening the motivation and enthusiasm of the majority of people involved in these programmes. Lastly, this study revealed that teachers encountered problems because they were not attending inset programmes. This revelation is crucial as studies by Jamil, Atta, Ali, Balock and Ayaz (2011) and Farroq and Shahzadi (2006) found that in-service trained teachers produce better results than untrained teachers. Teachers should attend these inset programmes in an effort to develop their content knowledge, their pedagogy, their own practice and skills related to curriculum changes, given that there is evidence that teachers with inadequate statistical knowledge (see, for example, Atagana et al., 2009; Atagana et al., 2010; Wessels & Nieuwoudt, 2011).

5.2.3.2 The causes of the problems in the learning of statistics in grade 11

The study found that the learners had problems because of their teachers' lack of statistics content knowledge and their methods of teaching (pedagogy). Learners indicated that their teachers found it difficult to explain concepts, and that they sometimes left learners to study sections of the work on their own. As a result, learners also lacked conceptual knowledge of certain concepts in statistics. This finding is to be expected, given that teachers commonly struggle with some aspects of statistics. Teachers do not teach the meaning and interpretation of the values of the measure of central tendency and measure of dispersion; they focus instead on algorithms and mechanical approaches (Cardoso, 2007; Da Silva & De Queiroz e

Silva Coutinho, 2008; Groth, 2009; Sharma, 2006). Also, this study showed that the learners experienced difficulties because of how they were taught statistics in the lower grades; certain sections of data handling and probability were not properly or never taught. This finding is also not surprising and is supported by findings of a study by Groth (2009) that many teachers believed that learners should not be taught certain aspects of statistics. This study also revealed that the learners had problems because of their attitude toward statistics. Learners indicated they did not give themselves enough time to practise data handling. This finding was to be expected as learners are only taught the simplest concepts of statistics, as a result of which they are inclined not to practise statistics, thinking that it is easy (Ghinis, Korres & Bersimis, 2009). Lastly, the study found that the learners encountered problems because of the learning materials they were using. They indicated that they had no access to previous examination papers with solutions and that they did not know how to use the statistics function mode on their calculator in statistical calculations. It also appeared that a prescribed textbook did not cover all the topics in the curriculum and that was inadequate, providing only a few examples. The results are consistent with findings by Lue (1998), who found that the major factors that constrain learners' learning of descriptive statistics are the lack of clarity on some statistical concepts in some sections in the textbook, and the lack of the usage of the calculators which makes calculations complicated and difficult.

5.2.1 Ways to address these problems

This study suggests that one way to address the problems encountered in the teaching and learning of statistics in grade 11 could be the improvement of in-service teacher programmes. These should meet the needs of teachers by offering topics that they find difficult to teach. There should be more inset programmes on probability and these should be fairly lengthy. These findings are consistent with those of Garet, Porter, Desimone, Birman and Yoon (2001), Gaible and Burns (2005), Lieberman (1994) and Yigit (2008), who found that in organising workshops for teachers the following factors need to be considered: (1) prior to the workshop an analysis must be made to understand teachers' needs; (2) the workshop should be sustained over an extended time period involving a substantial number of hours that can lead to the improvement of teachers' knowledge; and, (3) if steps (1) and (2) have been considered, then the workshop may help teachers to improve their own knowledge and their learners' performance.

The analysis of the data in this study also suggested that a solution to the problems encountered in the teaching and learning of statistics in grade 11could lie in the textbooks in use. These should be better written, should provide explicit explanations and contain all the necessary information to teach data handling and probability. These findings support the ideas of Lemmer, Edwards and Rapule (2008) and Lue (1998), who found that one of the most cost-effective ways of improving academic performance of both learners and teachers was to improve the quality of textbooks.

5.3 Implications of the findings

5.3.1 Implications for learning

It was discovered in this study that learners encounter problems when using graphs to predict results; when calculating measures of dispersion and measures of central tendency, more especially with numerical data that has been grouped into intervals; interpreting measures of dispersion; using graphs to predict the results; constructing and interpreting probability diagrams and tables; and understanding or interpreting probability terminology. This implies that teachers need to change their methods of teaching to those that will help learners understand the topics better, in this way helping learners to improve their performance in statistics.

The fact that certain sections of statistics (data handling and probability) are not always taught in the lower grades gives rise to learner difficulties. One reason for this is that some teachers find certain sections of data handling and probability difficult to explain. The obvious implication of these findings is that teachers are skipping certain areas of statistics because they lack the content knowledge to teach these topics. A common area of difficulty in this study was that learners were unable to use the statistics function mode on their calculators. This problem is a simple enough one to solve but it underlines the point that teachers need to integrate information and communication technology (ICT) in their teaching of statistics. In addition, instructional supervision by the school inspection and supervision division of the Department of Education should focus on making sure that all topics in statistics are mastered by learners.

The findings of this study exposed the limitations of the textbook learners were using. This did not cover all the topics in the curriculum and the ones it did were not adequately explained. This implies that the Ministry of Education and school advisors/inspectors need to look more closely at the selection of textbooks in order to recommend those which will improve academic performance in statistics in grade 11. Authors of textbooks must ensure that they cover all topics included in the curriculum in appropriate depth and provide correct content and instructional support.

5.3.2 Implications for teaching

This study established that teachers encounter problems with statistical calculations and interpretations of: measures of central tendency and measures of dispersion with grouped data, probability diagrams and tables and probability terminology. This suggests that teachers require more training in these aspects, which could be provided by in-service training programmes.

Teachers complained that those programmes they were able to attend often failed to address statistics topics, and when they did they did not pay adequate attention to probability. The result was that these programmes had little effect in improving teachers' knowledge. Again, this suggests that the Ministry of Education or school advisors/inspectors should look critically at the content covered in the in-service teacher workshops. Some of the teachers in the study had been attending in-service training programmes since the introduction of statistics into the school mathematics curriculum but they were still finding it difficult to teach the topic. As observed by Atagana et al. (2009) and Atagana et al. (2010), teachers need regular in-service training programmes on the content of some topics (including data handling and probability) in order to improve their teaching.

5.4 Conclusions

This study has shown that there are problems associated with the teaching and learning of statistics in grade 11. Teachers still do not have a complete grasp of some aspects of statistics which they are expected to teach, and this in turn has presented learners with difficulties in learning those aspects. The study therefore confirms that problems encountered in teaching tend to affect learning as well. In other words, lack of effective teaching results in lack of meaningful learning. It is also evident from the findings that when the NCS Grades 10-12 curriculum was introduced the issue of teacher preparedness and readiness to implement the curriculum was not taken proper note of, thus teachers still have problems teaching new aspects of the curriculum. Furthermore, it is evident that there is lack of support for newly introduced topics and there are also inadequacies in the available teaching materials. This adds to teachers' problems with statistics.

5.5 Recommendations

Based on the findings of this study the following recommendations are made.

5.5.1 Capacitating teachers in statistics content and pedagogical knowledge

Teacher in-service workshops should focus more on content knowledge. The current study has shown that teachers lack content knowledge of statistics, more especially knowledge of probability topics. This is important considering that most teachers interviewed had not studied statistics during their pre-service training. It is suggested that in-service workshop organisers should consider the results of this study when conducting workshops. Owusu-mensah (2008) recommends that in addition to self-evaluation methods of identifying the needs and weaknesses of a teacher, classroom observations should be used to determine further in-service needs. The observations in the current study revealed many teacher and learner problems.

The recommendation is also made that these in-service teacher training programmes should be long enough to allow thorough coverage of the content if they are to be of any benefit to teachers. One- or two-day workshops do not offer much help to teachers as too many topics are covered in too short a period of time (Ogbonnaya, 2011).

South African teachers are underpaid compared to those in other countries (Nesane, 2009). Even though most teachers may wish to improve their content and pedagogical knowledge through in-service workshops, money will always be an obstacle that hinders them from attending these workshops. Therefore this study recommends that the Department of Education provides funds for teachers to engage in in-service teacher training.

5.5.2 Effective teaching and learning resources

This study has demonstrated that the prescribed textbook used by learners and teachers is inadequate: it does not cover all the topics in the curriculum, does not provide clear explanations, is missing some important formulae and does not provide enough examples.

Textbooks should be well written, should provide explicit explanations and be comprehensive on the topics of data handling and probability. Taking into account the fact that many serving mathematics teachers did not study statistics during their pre-service training, the study recommends that textbook authors ensure that learner activities are challenging and contribute to the learning of the content, scientific skills and processes, covering core knowledge in appropriate depth and with scientific accuracy (Lemmer, Edwards & Rapule, 2008). This would help to prevent teachers and learners from struggling with statistics, particularly if textbooks were to explain the concepts of statistics and provide more examples. The literature indicates that most inexperienced and un(der)qualified teachers regard the textbook as absolutely correct, sometimes even as the only source of information, and follow it rigidly.

5.5.3 Integration of Information and Communication Technology (ICT) in the teaching and learning of statistics

This study has also demonstrated that learners encounter problems because they do not know how to use the statistics function mode on their calculators when doing statistical calculations (i.e. the mean and variance or standard deviation). Learners make mistakes when making some calculations (i.e. variance and standard deviation) and these are time consuming when using formulae to calculate them. Besides calculators, teachers and learners could use computers (and programs such as Excel) to compute statistical calculations and to construct statistical graphs/plots. The use of Information and Communication Technology (ICT) could help learners to verify their solutions and save them time when making statistical calculations and constructing statistical graphs or plots because learners only need to enter the data. This would allow teachers and learners to put more effort into teaching and learning conceptual understanding of certain statistical concepts, something which this study revealed to be lacking among many learners and teachers. Research studies have shown that the use of ICT improves the quality of instruction, motivates the learning process, encourages students' active learning at their own pace in the form of participation and feedback, and provides students with the psychological incentives they need to work hard (Garfield, 1995; Higazi, 2002; Ogbonnaya, 2010). Higazi (2002) found that when he taught topics such as graphical representations using both a traditional approach where he gave the purpose, assumptions, meanings, formulae and computation; and when integrating ICT by displaying a Minitab or Excel output and explaining the meaning of the output obtained in relation to what had been introduced to, students were impressed to hear that it took only "data entry time" to get all that output, and then asked for special sessions to learn "how". They became really motivated by the ease of use; self-learning was stimulated and that opened doors for students to specialise in statistics (Higazi, 2002). Therefore, the current study recommends the integration of ICT into the teaching and learning of statistics.

5.6 Possible further research

This study did not investigate the issue of teachers' and learners' beliefs about the teaching and the learning of statistics. It is therefore suggested that it may be of interest to focus on beliefs of teachers and learners to see their roles in the teaching and learning of statistics.

5.7 Limitations of the study

The limitation of the current study is that data about problems encountered in the teaching and learning of probability topics were collected using questionnaires (teacher and learner questionnaires) and interviews (learner and teacher interviews). This was not completely adequate. A more balanced technique would have been to use classroom observation and a diagnostic test for learners and classroom observation of teachers and learners, as was done in the case of data handling. Classroom observation and a diagnostic test would have given the researcher a deeper insight into the problems encountered by learners and teachers in the teaching of probability topics.

REFERENCES

Adler, J. (1994). Mathematics Teachers in the South African Transition. *Mathematics Education Research Journal, 6* (2), 101-112.

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Could, R. and Lock, R. (2010). *Guidelines for Assessment and Instruction in Statistics Education: College Report.*California: American Statistical Association. Retrieved from http://it.stlawu.edu/~rlock/gaise/GAISECollege.pdf

Angrist, J. D.& Lavy, V. (2001). Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools. *Journal of Labor Economics, 19* (2), 343-369. Retrieved from http://www.jstor.org/stable/10.1086/319564

Armstrong, P. (2009). *The impact of teacher characteristics on student performance: An analysis using hierarchical linear modelling*. Stellenbosch Economic Working Papers No. 07/09.

Arriassecq, I. & Greca, I. M. (2007).Approaches to the teaching of special relativity theory in high school and university textbooks of Argentina, *Science &Education,* 16, 65–86.

Atagana, H. I., Mogari, L. D., Kriek, J., Ochonogor, E. C., Ogbonnaya, U. I., Dhlamini J.J. & Makwakwa, E. G. (2010). *An intervention into educators' and learners' difficulties in some topics in mathematics, science and technology: A report of the ISTE 2010 Winter School.* ISBN: 978-1-86888-684-5.

Atagana, H. I., Mogari, L. D., Kriek, J., Ochonogor, E. C., Ogbonnaya, U. I. & Makwakwa, E. G. (2009). *An analysis of educators' and learners' perceived difficult topics in mathematics, physical science and life science at the Further Education and Training (FET) band in Gauteng province: A Report of the ISTE 2009 Winter School.* ISBN: 978-1-86888-612-8.

Atagana, H. I., Mogari, L. D., Kriek, J., Ochonogor, E. C., Ogbonnaya, U. I., Dhlamini J.J. & Makwakwa, E. G. (2011). *Aligning pedagogy with content knowledge: An improvement strategy for MST Educators: ISTE 2011 Winter School Project Report.* ISBN: 978-1-86888-690-6.

Australian Curriculum, Assessment and Reporting Authority (2013, March 26). *Australian Curriculum, Assessment and Reporting Authority (ACARA)*. Retrieved March 26, 2013, from The Australian curriculum v4.2 Senior Scondary Curriculum: http://www.australiancurriculum.edu.au/SeniorSecondary/Mathematics/MathematicalMethods/curriculum/SeniorSecondary.

Baloyi, H. G. (2011). Learner performance disparities between former white and former black schools in Gauteng province of South Africa after more than a decade of democracy (PhD thesis, University of Witwatersrand, South Africa). Retrieved from http://wiredspace.wits.ac.za/handle/10539/9829.

Batanero, C., Burrill, G. & Reading, C. (eds.) (2011). Teaching statistics in school mathematics – Challenges for teaching and teacher education. Springer.

Batanero, C.& Diaz, C. (2010). Training teachers to teach statistics: What can we learn from research? *Statistique et Enseignement, 1* (1), 5-20.

Batanero, C., Godino, J. D.& Roa, R. (2004). Training teachers to teach probability. *Journal of Statistics Education, 12* (1). Retrieved from http://www.amstat.org/publications/jse/v12n1/batanero.html

Creswell, J. W. (2011). *Educational Research: planning, conducting, and evaluating qualitative and quantitative research (4th edition)*. Boston: Pearson Education, Inc.

Behr, M., Lesh, R., Post, T. & Silver, E. (1983). Rational number concepts. In R. Lesh (Ed.), *Acquisition of mathematical concepts and processes* (pp. 91-126).New York: Academic Press.

Ben-Zvi, D. & Garfield, J. (2004). *The challenge of developing statistical literacy, reasoning and thinking.* The Netherlands: Kluwer Academic Publishers.

Betts, J. R., Zau, A. C. & Rice, L. A. (2003).*Determinants of student achievement: New Evidence from San Diego*. San Francisco, CA: Public Policy Institute of California.

Bogdan, R. C. & Biklen, S. K. (1992).*Qualitative research for education: An introduction to the theory and methods (2nd edition)*. Boston: Allyn & Bacon.

Boyd, D. J., Lankford, H., Loeb, S., Rockoff, J. E. & Wyckoff, J. H. (2007). The Narrowing Gap in New York City Teacher Qualifications and Its Implications for Student Achievement in High-Poverty Schools. *CALDER Working Paper 10*. Washington, DC: The Urban Institute.

Bruno, A. & Espinel, M. C. (2009). Construction and evaluation of histograms in teacher training. *International Journal of Mathematical Education in Science and Technology, 40* (4), 473–493.

Burton, D. M. (2007). *The History of Mathematics. An introduction.* New York: The McGraw-Hill Companies, Inc.

Cardoso, R. (2007). *The mathematics teachers and the exploratory data analysis in high school.* published Master's thesis. PUC-SP.Retrived form http:// www.pucsp.br/pos/edmat.

Carnoy, M. & Chisholm, L. (2008). *Towards an understanding of student academic performance in South Africa: A pilot study of grade 6 mathematics lessons in South Africa.* Report prepared for the Spencer Foundation. Pretoria: HSRC.

Cassim, I. (2012, February 9). Mathematics Roadshow: Mathematics Paper 2 (2011) grade 12. Gauteng Department of Education. South Africa.

Clotfelter, C. T., Ladd, H. F. & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, 26(6), 673-682.

Cobb, G. W. & Moore, D. S. (1997). Mathematics, Statistics, and Teaching. *JSTOR, 104* (9), 801-823.

Cohen, L., Manion, L. & Morrison, K. R. (2007). *Research Methods in Education,* Sixth Edition. USA: Taylor & Francis e-library.

Community College Survey of Student Engagement (2006).Classroom observation form. http://www.ccac.edu/files/PDF_Document/b1578a72911d40f3b0180b0c9f822ec4.pdf

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Croninger, R., Rice, J., Rathbun, A. & Nishio, M. (2007). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement. *Economics of Education Review*, 26, 312-324.

Da Silva, C. B.& De Queiroz e Countinho, C. (2006). The Variation Concept: A Study With Secondary School Mathematics Teachers. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics (*ICOTS 7), Salvador, Bahia, Brazil. http://www.stat.auckland.ac.nz/~iase/publications/17/C216.pdf

Da Silva, C. B. & De Queiroz e Countinho, C. (2008). Reasoning about variation of a univariate distribution: A study with secondary mathematics teachers. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.),Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. *Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.*ICMI/IASE.

Davis, B.& Simmt, E. (2006). Mathematics-for-teaching: An ongoing investigation of the mathematics that teachers (need to) know. *Educational Studies in Mathematics, 61* (3), 293-319.

DeCaro, S. A. (2003). *A student's guide to the conceptual side of inferential statistics.* Retrieved September 2, 2010, from http://psychology.sdecnet.com/stathelp.htm.

Department of Basic Education. (2011, July 05). Curriculum And Assessment Policy Statement(Caps): Mathematics Fet Phase Final Draft. Pretoria, Gauteng, South Africa.

Department of Education. (2002). *Revised National Curriculum Statement Grades R-9 (Schools).* Pretoria: The Department of Education of South Africa.

Department of Education. (2003). *National Curriculum Statement Grades 10-12 (General): Mathematics.* Pretoria: Department of Education.

Department of Education. (2004). *Revised National Curriculum Statement Grades 7-9 (Schools) Senior Phase.* Pretoria: Department of Education.

Department of Education. (2008). *National curriculum Statement Grades 10-12 (General) Subject Assessment Guidelines: Mathematics.* Pretoria: Department of

Education.Erduran, S., Bravo, A. A.&Naaman, R. M. (2007). Developing epistemologically empowered teachers: examining the role of Philosophy of Chemistry in teacher education, *Science & Education,* 16, 975-989.

Fan, L.,Chen, J., Zhu, Y. Qiu, X. & Hu, J. (2004). Textbook use within and beyond Mathematics Classrooms: A Study of 12 Secondary Schoolsin Kunming and Fuzhou of China. In L. Fan, N. Y. Wong, J. Cai, & S. Li (Eds.), *How Chinese learn mathematics: Perspectives frominsiders*(pp.228-262). Singapore: World Scientific.

Farooq, M.S. & Shahzadi, N. (2006). Effects of teachers' professional education on students' achievement in mathematics.*Bulletin of Education & Research*, 28(1), 47-55.

Flores, F., Lopez, A., Gallegos, L. & Barojas, J. (2000). Transforming Science and learning concepts of Physics teachers. *International Journal of Science Education,*22(2), 197-208.

Franklin, C., Kader, G., Mewborn, D. S., Moreno, J., Peck, R., Perry, M. & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework.* Alexandria: American Statistical Association. Retrieved from http://www.amstat.org/education/gaise/GAISEPreK12_Intro.pdf

Gachutha, C. W. (2009). *The role of supervision in the management of counsellor burnout.*Doctoral thesis, University of South Africa. Pretoria, South Africa: http://hdl.handle.net/10500/1876 .

Gaible, E. & Burns, M. (2005).*Using Technology to Train Teachers: Appropriate Uses of ICT forTeacher Professional Development in Developing Countries*. Washington, DC: infoDev / World Bank. Available at: http://www.infodev.org/en/Publication.13.html

Gal, I. (2002). Adults' Statistical Literacy: Meanings, Components, Responsibilities. *International Statistical Review, 70* (1), 1-51.

Garegae, K. G. (2008). Training statistics teachers for Botswana senior schools: is it necessary? In C. Batanero, G, Burril, C. Reading & A. Rossman (Eds.). *Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* ICMI/IASE.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F. & Yoon, K. S. (2001). What Makes Professional Development Effective? Results from a National Sample of Teachers. *American Educational Research Journal, 38* (4), 915-945.

Garfield, J. (1995). How Students Learn Statistics. *International Statistics Review, 63* (1), 25-34.

Garfield, J. & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: implications for research. *Journal for Research in Mathematics Education, 19* (1), 44-69.

Garfield, J. & Everson, M. (2009). Preparing teachers of statistics: A graduate course for future teachers. *Journal of Statistics Education, 17* (2).

Gay, L. R., & Airasian, P. (2003). *Educational research: Competencies for analysis and Applications. Seventh Edition.* Upper Saddle River, New Jersey: Pearson Education, Inc.

Gay, L. R., Mill, E.& Airasian P. (2011). *Educational research: Competencies for analysis and application.* Tenth Edition. Upper Saddle River, New Jersey: Pearson Education, Inc.

Ghinis, D., Korres, K. & Bersimis, S. (2009). Difficulties Greek Senior High School Students identify in the Learning and the Teaching of Statistics: The Case of Experimental and Private High Schools. *Journal of Statistics Education, 17* (3).

Giliomee, H. (2009). A Note On Bantu Education, 1953 To 1970. *South African Journal of Economics*, *77* (1), 190-198.

Godino, J. D., Ortiz, J. J., Roa, R. & Wilhelmi, M. R. (2011). *Models for statistical pedagogical knowledge.* In C. Batanero, G. Burrill, & C. Reading (Eds.), Teaching Statistics in School-Mathematics-Challenges for Teaching and Teacher Educ ation: A Joint ICMI/IASE Study (pp. 271-282). New York: Springer. http://www.springerlink.com/

Goldhaber, D. D. & Brewer, D. J. (1997).Evaluating the effect of teacher degree level on educational performance. In W. J. Fouler (Ed.), *Development in school finance, 1996* (pp. 197–210). Washington, DC: National Center for Education Statistics, US Department of Education.

Goldhaber, D. D. & Brewer, D. J. (1998). When should we reward degrees for teachers? *Phi Delta Kappan, 80*(2), 134-138.

Gould, R. & Peck, R. (2004). Preparing Secondary Mathematics Educators to Teach Statistics. *Curriculum Development in Statistics Education*, 244-255.

Gravetter, F. J.& Wallnau, L. B. (2002). *Essentials of statistics for the Behavioral sciences.* Fourth Edition. Canada: Wadsworth Group.

Gravetter, F. J. & Wallnau, L. B. (2008). *Essentials of Statistics For the Behavioral Sciences,* Sixth Edition.Thomson Learning, Inc.

Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education*,38, 427-437.

Groth, R.E. (2009). Characteristics of teachers' conversations about teaching mean, median, and mode. *Teaching and Teacher Education, 25,* 707-716.

Haggarty, L. & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German classrooms: Who gets an opportunity to learn what? *British Educational Research Journal, 28(4)*, 567-590.

Hald, A. (2003). *History of Probability and Statistics and their applications before 1750.* Hoboken, New Jersey: John Wiley & Sons Inc.

Harrison, A. G. (2001). How do teachers and textbook writers model scientific ideas for students?*Research in Science Education*,31, 401-435.

Higazi, S. M. F. (2002). Teaching statistics using technology. *The Sixth International Conference on Teaching Statistics. Developing a statistically literate society.* Cape Town, South Africa: http://www.stat.auckland.ac.nz/~iase/publications/1/5f4_soha.pdf.

Hollinger, J. D. (n.d.). *How to improve teacher morale in a school* . Retrieved Ju1y 15, 2012, from Ezinearticles: http://ezinearticles.com/?Why-and-How-to-Improve-Teacher-Morale-in-a-School&id=5352075

Howie, S. J. & Pietersen, J. J. (2001). Mathematics literacy of final year students: South African realities. *Studies in Educational Evaluation,* 27, 7-25.

Hubbard, M. R. (2003). *Statistical Quality Control for the Food Industry,* Third Edition. New York: Kluwer Academic/Plenum Publishers.

Hubisz, J. (2003). Middle-school texts don't make the grade. *Physics today*, 56 (5), 50-54.

Izsak, A. & Sherin, M. G. (2003). Exploring the use of new representations as a resource for teacher learning, *School Science and Mathematics*, 103(1), 18-27.

Jacobbe, T. (2008). *Elementary school teachers' understanding of the mean and median.* Monterey, Mexico: Proceedings of the Joint ICMI/IASE Study Statistics Education in School Mathematics.

Jamil, A., Attar, M. A., Ali, U., Balock, J.&Ayaz, M. (2011).Effects of In-service training in ameliorating teachers' performance at secondary level. *International Journal of Academic Research*, 3(2), 624-628.retrieved from http://www.Ijar.lit.az

Jafri, H. & Shahzadi, S. (2002). Comparison between untrained teachers and achievement of students at secondary level, Lahore: Unpublished Master's dissertation, University of Punjab.

Johnson, N. C. (1985). Role of dynamic microcomputer graphics in teaching the meaning of random sampling and the central limit theorem. Doctoral thesis, University of Minnesota. Dissertation Abstracts International, 47, 455A.

Johnson, B. & Christensen, L. (2004). *Educational Research: Quantitative, qualitative, and mixed approaches.*Second edition. Boston, MA: Pearson Education Inc.

Joyce, B.R.& Weil, M. (1980).*Models of Teaching.*Englewood Cliffs NJ: Prentice-Hall.

Julian, P. R.& Murphy, A. H. (1972). Probability and Statistics in meteorology: a review of some recent developments. *American Meteorological Society, 53* (10), 957-965.

Khurshid, K. (2008). A Study of the Relationship between the professional Qualifications of the Teachers and Academic Performance of Their Students at Secondary School Level.*World Academy of Science, Engineering and Technology,* 38, 445-451.

Khuzwayo, H. B. (2005). A history of mathematics education research in South Africa: The apartheid years. In: Vithal, R., Adler, J., and, Keitel, C. (eds). *Mathematics education research in South Africa: Challenges and possibilities.*

King, R. S.& Julstrom, B. (1982). *Applied statistics using the computer*. United States of America: Alfred Publishing Co. Inc.

Koehler, M. J.& Mishra, P. (2009). What is Technological Pedagogical Content Knowledge? . *Contemporary Issues in Technology and Teacher Education, 9* (1).

Ladd, H. F. (2008). Value-Added Modeling of Teacher Credentials: Policy Implications. Paper presented at the second annual CALDER research conference, "The Ins and Outs of Value-Added Measures in Education: What Research Says", Washington, D.C., November 21. http://www. caldercenter.org/upload/Sunny_Ladd_presentation.pdf.

Laridon, P., Barnes, H., Jawurek, A., Kitto, A., Myburgh, M., Pike, M., Myburgh, M., Rhodes-Houghton, R., Scheiber, J., Sigabi, M.& Wilson, H. (2004). *Classroom Mathematics. Grade 10 learners book.* Sandton: Heinemann Publishers (Pty) Ltd.

Lee, C. (2003). Some difficulties inlearning histograms in introductory.*Joint Statistical Meetings - Section on Statistical Education*, 2326-2333.

Lemmer, M., Edwards, J.& Rapule, S. (2008). Educators' selection and evaluation of natural sciences textbooks. *South African Journal of Education* , 28(2), 175-187.

Liberman, A. (1994). *The work of Restructuring Schools: building from the ground up*. New York: Teacher College Press.

Liston, D.P. & Zeichner, K.M. (1990).*Teacher Education and the Social Conditions of Schooling*. New York: Routledge.

Lue, Y. (1998). Evaluation on a descriptive statistics curriculum unit in the high school mathematics curriculum. *Icots 5* (pp. 580-585). Singapore: IASE. http://www.stat.auckland.ac.nz/~iase/publication.php.

Macrae, M. (1994). A legacy of apartheid: The case of mathematics education in South Africa. *International Journal of Educational Development,* 14 (3), 271-287.

Mahlobo, R. K. (2009). A model for an open-ended task-based approach in grade 11 mathematics classes.PhD thesis,University of North West, Potchefstroom, South Africa. Retrieved from http://hdl.handle.net/10394/5080

Makwakwa, E.G. & Mogari, D. (2011). Investigating problems experienced by grade 11 mathematics teachers in the teaching of probability. In D. Mogari, A. Mji, & U.I. Ogbonnaya (Eds), Towards Effective Teaching and Meaningful Learning in Mathematics, Science and Technology. *Proceedings of Institute of Science and Technology Education International Conference* held on October 17-20, 2011 (pp. 122-128). Pretoria, South Africa: Unisa Press

Makwakwa, E. G. & Mogari, D. (2012). Teachers' claims about their competence in teaching data handling in Grade 11. In S. Nieuwoudt, D. Laubscher & H. Dreyer (Eds). Mathematics as an Educational Task. *Proceedings of long and short papers: 18th National Congress* held on June 24 – 28, 2012 (pp. 249-259). Johannesburg, South Africa: Association for Mathematics Education of South Africa (AMESA).

Mann, P. S. (1998). *Introductory statistics.* Third Edition. John Wiley & Sons, Inc.

McMillan, J. H. & Schumacher, S. (2001). *Research in Education: A conceptual introduction.* Fifth Edition. New York: Addison Wesley Longman, Inc

Meletiou-Mavrotheris, M. & Lee, C. (2010). Investigating College-Level Introductory Statistics Students' Prior Knowledge of Graphing. *Canadian Journal Of Science, Mathematics And Technology Education,* 10 (4), 339-355.

Ministry of Education Malaysia. (2004). Integrated Curriculum for Secondary Schools Syllabus Mathematics. Malaysia: Curricululm Development Centre. Retrieved March 31, 2011, from http://smkamil.com/panitia/sp_math_kbsm.pdf

Mishra, P.& Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A framework for Teacher Knowledge. *Teachers College Record*, 108 (6), 1017-1054.

Moeini, H. (2008). Identifying Needs: A Missing Part in Teacher Training. *International Journal of Media, Technology and Lifelong Learning,* 4(1), 1-12.Retrieved from http://www.seminar.net/images/stories/vol4-issue1/moini-identifyingneeds.pdf

Moore, S. (1988). Should Mathematicians Teach Statistics? *The College Mathematics Journal,* 19 (1), 3-7.

Moore, D. S., McCabe, G. P., Duckworth, W. M.& Alwan, L. C. (2009). *The Practice of Business Statistics using data for Decisions.* Second Edition. New York: W. H. Freeman and Company.

Motshekga, A. (2010, July 06). Statement by The Minister Of Basic Education, Mrs Angie Motshekga, MP on the progress of the review of the National Curriculum Statement. Pretoria, Gauteng, South Africa: http://www.education.gov.za/dynamic.aspx?pageid=310&id=10245.

Naoreen, B., Aslam, S., Arshad, M. & Nausheen, R. (2011). In-service Teacher Training On Students' Learning Achievement in Mathematics. *International Conference on Social Science and Humanity,* 168-174.

National Council for Teacher Education (n.d.).Curriculum Framework for Quality Teacher Education. Retrieved from http://www.ncte-india.org/pub/curr/curr_0.htm

Nesane, M. A. (2009). Selected school-related reasons why teachers abandon the teaching profession: an educational management perspective.Master's dissertation, University of South Africa, Pretoria, South Africa.: http://hdl.handle.net/10500/2312.

North, D., Scheiber, J.& Ottaviani, M. G. (2010). Training teachers to teach statistics in South Africa: Realities and attitudes. In C. Reading (Ed.), Data and conttext in statistics education:Towards an evidence-based society. *Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8).* Ljubljana, Slovenia. www.stat,auckland.ac.nz/~iase/publications.php: IASE.

North, D. & Zewotir, T. (2006). Introducing statistics at school level in South Africa. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil: International Statistical Institute and International Association for Statistical Education. online: http://www.stat.auckland.ac.nz/~iase/publications.

Newton, D. P.& Newton, L. D. (2006). Could elementary Mathematics textbooks help give attention to reasons in the classroom? *Educational Studies in Mathematics,* 64, 69-84.

Nicholson, J. & Darnton, C. (2003). Mathematics teachers teaching statistics: What are the challenges for the classroom teacher? *International Association for Statistical Education..*

Nicholson, J., Ridgway, J.& McCusker, S. (2006). Reasoning with Data - Time for a Rethink? *Teaching Statistics, 28* (1), 2-9.

Ogan-Bekiroglu, F. (2007). To what degree do the currently used Physics textbooks meet the expectations? *Journal of Science Teacher Education,* 18, 599-628.

Ogbonnaya, U. I. (2007). The influence of teachers' background, professional development and teaching practices on students' achievement in Mathematics in Lesotho.Master's dissertation, University of South Africa, Pretoria, South Africa. Retrieved from http://uir.unisa.ac.za/bitstream/handle/10500/2257/dissertation.pdf?sequence=1

Ogbonnaya, U. I. (2010). Improving the Teaching and Learning of Parabolic Functions by the use of Information and Communication Technology (ICT). *African Journal of Research in MST Education,* 14 (1), 49-60.

Ogbonnaya, U. I. (2011). Exploring the relationship between Mathematics teachers' subject matter knowledge and their teaching effectiveness.PhD thesis, University of South Africa, Pretoria, South Africa: http://hdl.handle.net/10500/5240.

Ortiz, J. J., Cañizares, M. J., Batanero, C.& Serran, L. (2002). *An experimental study of probabilistic language in secondary school textbooks.* Cape Town: International Conference On Teaching Statistics 6 (ICOTS 6).

Ost, B. (2009). How do teachers improve? The relative importance of specific and general human capital. Retrieved [2012 February 11], from Cornell University, School of Industrial and Labor Relations site: http://digitalcommons.ilr.cornell.edu/workingpapers/125/

Owusu-Mensah, J. (2008). A critical investigation into the nature and quality of in-service Education and Training programmes for Further Education and Training (FET)Mathematics teachers in the Greater Taung Area.Master’s dissertation, University of South Africa,Pretoria, South Africa. Retrieved from http://uir.unisa.ac.za/bitstream/handle/10500/2422/dissertation.pdf?sequence=1

Papaieronymou, I. (2010). Recommended knowledge of probability for secondary mathematics teachers. *Cerme 6* (pp. 358-367). France: Institut National de Recherche Pédagogique.

Parsad, B., Lewis, L., Farris. E.& Greene, B. (2001).*Teacher Preparation and Professional Development: 2000.* Washington, DC: U.S. Department of Education, National Center for Education Statistics (NCES 2001–088). Retrieved from http://nces.ed.gov/pubs2001/2001088.pdf.

Pearson, K. (1978). *The history of statistics in the 17th and 18th centuries against the changing background of intellectual, scientific, and religious thought: lectures by Karl Pearson given at University College, London, during the academic sessions, 1921-1933*. London: Griffin & Co Ltd.

Pepin, B. and Haggarty, L. (2003) 'Mathematics textbooks and their use by teachers: a window into the education world of particular countries?, in Van den Akker, J. , Kuiper, W. and Hameyer, U. (eds) *Curriculum Landscapes and Trends,* Dordrecht: Kluwer (pp. 73-100).

Pollatsek, A., Lima, S.& Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics,* 12, 191-204.

RangaRao, R. & BhaskaraRao, D. (2004).*Methods of Teacher Training*. India: Discovery Publishing House.

Remillard, J. T.& Bryans, M. B. (2004). Teachers' orientations toward mathematics curriculum materials: Implications for teacher learning. *Journal for Research in Mathematics Eduaction,* 35 (5), 352-388.

Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes.*Washington,DC: Economic Policy Institute.

Rice, J. K. (2010). The impact of Teacher Experience Examining the Evidence and Policy implications. National Center for Analysis of Longitudinal Data in Education Research. Washington, DC: Urban Institute.

Robson, C. (2002). *Real World Research: A resource for social scientists and practitioner researchers.* Second Edition. Oxford: Blackwell.

Rockoff, J.E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, 94 (2), 247-252.

Sánchez, E. & García, J. (2008). Acquisition of notions of statistical variation by in-service teachers. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.),. Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. *Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*.

Schmidt, W. H., McKnight, C.C.& Raizen, S. A. (2002).*A splintered vision: An investigation of U.S. science and mathematics education.* New York, Boston, Dordrecht, London, Moscow: Kluwer Academic Publishers.

Serradó, A., Cardeñoso, J. M.& Azcárate, P. (2005). Los obstáculos en el aprendizaje del conocimiento probabilístico: Su incidencia desde los libros de texto (Obstacles in the learning of probabilistic knowledge: Influence from the textbooks). *Statistics Education Research Journal*, 4 (2), 59-81.

Sharma, S. V. (2006). High school students interpreting tables and graphs. *International Journal of Science and Mathematics Education*, 4, 241-268.

Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15 (2), 4-14.

Slauson, L. V. (2008). Students' conceptual undertanding of variability.PhD thesis, The Ohio State University, Columbus, United States.Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/dissertations/08.Slauson.Dissertation.pdf.

Steel, R. G.& Torrie, J. H. (1980). *Principles and procedures of statistics: A Biometric approach.* Second Edition.McGraw-Hill International Editions.

Swanepoel, S. (2010). The assessment of the quality of science education textbooks: Conceptual framework and instruments for analysis. Doctoral thesis, University of South Africa, Pretoria, South Africa. Retrieved from http://uir.unisa.ac.za/bitstream/handle/10500/4041/thesis_swanepoel_s.pdf?sequence=1

Tarr, J. E., Ch´avez, O., Reys, R. E.&Reys, B. J. (2006). From the written to the enacted curriculum: The intermediary role of Middle School Mathematics teachers in shaping students' opportunity to learn, *School Science & Mathematics*, 106(4), 191-201.

Tesch, R. (1990). *Qualitative research: analysis types and software tools*. New York: Falmer.

Watson, J. M. (1998). Professional Development for Teachers of Probability and Statistics: Into an Era of Technology. *lnternational Statistical Review,* 66 (3), 271-289.

Watson, A., Burrill, G., Landwehr, J.& Sheaffer, R. (1992). Remedial statistics? The implications for colleges of the changing secondary school curriculum. In K. G. Garegae (Eds.). *Proceedings of the ICMI study 18 and 2008 IASE Round Table conference*. Botswana: University of Botswana.

Wegner, T. (2007). *Applied Business Statistics: Methods and Excel-based Applications.*Second Edition. Cape Town, South Africa: Juta & Co. Ltd.

Wessels, H. & Nieuwoudt, H. (2011).Teachers' professional development needs in data handling and probability. *Pythagoras*, 32(1), 1-9.

Wilen, W., Hutchinson, J., Bosse, M. I.&Kindsvatter, R.(2004). *Dynamicsof Effective Secondary Teaching*. Boston, Massachusetts: Allyn& Bacon.

Williams, T. A., Sweeney, D. J.& Anderson, D. R. (2006). *Contemporary Business Statistics with Microsoft Excel.* China: Thomson South-Western.

Woodward, M.& Francis, L. M. (1988). *Statistics for Health Management and Research.*Edward Arnold, Hodder & Stoughton Limited.

Yadav, S. (February, 2010).Speech as Guest of honor at a two day seminar on " Impact of statistics on science & society" at M.G. Science Institute on 25th February 2010. Retrieved June 2, 2012, from http://shashiranjanyadav.blogspot.com/2010/02/speech-as-guest-of-honor-at-two-day.html

Yigit, N. (2008). A study on Evaluation of Effectiveness of an In-service Training (Inset) Course about th Use of Instructional Technologies and Material Development. *World Applied Sciences Journal*, 4 (1), 69-75. Retrieved from http://www.idosi.org/wasj/wasj4(s1)/12.pdf

APPENDICES

Appendix 1

TEACHER QUESTIONNAIRE

Purpose of the questionnaire: This questionnaire is designed to obtain the problems teachers encounter in the teaching of statistics (data handling and probability). This is in order to proffer useful solutions to improve teachers' performance in teaching and also to help learners to perform better in Statistics.

The response will be anonymous and used for research purposes only.

Instructions:

1. Please tick (✓) in the provided boxes what is appropriate to you.
2. There are no correct or wrong answers, if you make a mistake cross out and tick another opinion.
3. Do not write your name on the questionnaire.

Section A:	**Demographic information**

1. What is your gender? Please tick (✓) one choice.

Male ☐1 Female ☐2

2. Where is the school you teach located?

Township ☐1 Urban/town ☐2

3. What is your highest qualification? Please tick one choice.

Highest qualification	**Tick (✓)**
Grade 12 Certificate	
Diploma	
Bachelors Degree	
Masters Degree	
Doctorate	

4. What is your highest mathematics qualification? Please tick one choice.

Highest mathematics qualification	**Tick (✓)**
Grade 12 or lower	
Mathematics diploma	
Mathematics I	
Mathematics II	
Mathematics III	
Postgraduate	

5. Which grade(s) are you currently teaching statistics (data handling and probability)? Please tick one choice.

Grades in which I am currently teaching	**Tick (✓)**
Grade 10, grade 11 & grade 12	
Grade 10 &grade 12 only	
Grade 10 & grade 11	
Grade 11 & grade 12	
Grade 10 only	
Grade 11 only	
Grade 12 only	

6. What is your highest statistics qualification? Please tick one choice.

Highest statistics qualification	**Tick (✓)**
Never studied statistics	
Statistics I	
Statistics II	
Statistics III	
Postgraduate	

7. For how many years have you been teaching statistics (data handling) and Probability at the Further Education and Training (FET) level? Please tick (✓) in the appropriate box below.

	Less than a year	**1 to 2 years**	**3 to 4 years**	**More than 5 years**
Probability				
Data handling				

Section B:	Problems encountered in the teaching of data handling and probability topics

8. Indicate how competent you feel when doing the following teaching activities in data handling? Please choose in one of the response below by placing a tick (✓) in the appropriate box.

	Very competent	Competent	Slightly competent	Not at all Competent
8.1 Explaining the meaning of measures of central tendency (mean, mode, median) to learners.				
8.2 Teaching learners to do the calculations of measures of central tendency (mean, mode, median) with **ungrouped data**.				
8.3 Teaching learners to do the calculations of measures of central tendency (mean, mode, median) with the **grouped data**.				
8.4 Interpreting the measures of central tendency (mean, mode, median) for learners.				
8.5 Explaining the meaning of measures of dispersion (standard deviation, range, and variance) to learners.				
8.6 Teaching learners to do the calculations of measures for dispersion (standard deviation, variance, range).				
8.7 Interpreting the measures of dispersion (standard deviation, variance, range) for learners.				
8.8 Teaching learners the calculations of contents of five number summary (lower quartile, middle quartile, upper quartile).				
8.9 Teaching learners to make drawings of box-and-whisker diagrams on a number line.				
8.10 Teaching learners to construct stem-and-leaf plot.				
8.11 Teaching learners to use stem-and-leaf plot to determine the quartiles (lower quartile, median and upper quartile).				
8.12 Constructing cumulative frequency tables for learners.				
8.13 Making a drawing of an Ogive (cumulative frequency curves) for learners.				
8.14 Interpreting an Ogive (cumulative frequency curves) for learners.				
8.15 Teaching learners to represent bivariate numerical data as a scatter plot.				
8.16 Determining the line of best fit on a scatter plot for learners.				
8.17 Teaching learners to select a function that best fit the data.				

9. Indicate how competent you feel when doing the following teaching activities in probability? Please choose in one of the response below by placing a tick (✓) in the the appropriate box.

	Very competent	**Competent**	**Slightly competent**	**Not at all Competent**
9.1 Teaching learners to **construct Venn diagram** from a given word problem.				
9.2 Teaching learners to **construct two –way contingency** tables from a given word problem.				
9.3 Teaching learners to **construct tree diagrams** from a given word problem.				
9.4 Teaching learners to use two –way contingency tables for problem solving.				
9.5 Teaching learners to use Venn diagrams for problem solving.				
9.6 Teaching learners to use tree diagrams for problem solving.				
9.7 Teaching learners to identify dependent and independent events from Venn diagrams.				
9.8 Teaching learners to Identify dependent and independent events from two-way contingency tables.				
9.9 Teaching learners to use Venn diagram to solve probability problems where events are not necessarily independent.				
9.10 Teaching learners to use tree diagrams to solve probability problems where events are not necessarily independent.				
9.11 Teaching learners to calculate the probability of two independent events by applying product rule for independent events: P(A and B) = P(A). P(B)				
9.12 Teaching learners to identify mutually exclusive events from Venn diagrams.				
9.13 Teaching learners to differentiate between independent and dependent events.				

10. The following reasons might be some of the causes for the problems you are experiencing in the teaching of Statistics (data handling and probability). Please rate how strongly you agree or disagree with each of the statements by placing a tick (✓) in the appropriate box.

	Strongly disagree	**Disagree**	**Neither disagree nor agree**	**Agree**	**Strongly agree**
10.1 I do not have statistics content knowledge/ I did not study statistics.					
10.2 I do not like the teaching of statistics.					
10.3 I do not see the importance of statistics (data handling and probability) in the syllabus.					
10.4 I do not have enough experience in the teaching of statistics.					

10.5 There are different types of problems in statistics and I lack the problem solving skills to deal with them.					
10.6 I cannot give myself enough time to prepare properly for my teaching of statistics because I am committed in other learning areas.					
10.7 Most of the in-service training programmes that I have attended did not cover statistics topics.					
10.8 My school does not support my attendance of teacher development programmes to improve my statistics knowledge.					
10.9 The text books do not explain thoroughly (i.e. formulae are missing) and do not have enough examples.					
10.10 I do not seek assistance from other teachers when I experience problems.					
10.11 I do not have sufficient teaching resources to teach statistics.					
10.12 My learners do not pay enough attention when I am teaching statistics.					
10.13 I am still using a teacher-dominated method to teach as opposed to the recommended learner-centred approach.					

SECTION C: Problems encountered by learners in the learning of data handling and probability

***Section C** asks you about the problems experienced by learners in the learning of statistics (data- handling and Probability).*

11. The following items seek your opinion about the level of difficulty learners have with learning some aspects of data handling. Please put your response in a rating format by placing a tick (✓) in the appropriate box below.

	Not difficult	**Less difficult**	**Difficult**	**Highly difficult**
11.1 Understanding the meaning of measures of central tendency (mean, mode, median).				
11.2 Calculating measures of central tendency (mean, mode, median) with ungrouped data.				
11.3 Calculating measures of central tendency (mean, mode, median) with grouped data.				

11.4 Interpreting measures of central tendency (mean, mode, median).				
11.5 Understanding the meaning of measures of dispersion (standard deviation, range, and variance).				
11.6 Calculating the measures of dispersion (standard deviation, variance, range).				
11.7 Interpreting measures of dispersion (standard deviation, variance, range).				
11.8 Understanding the content and doing calculations for five number summary (lower quartile, middle quartile, upper quartile).				
11.9 Drawing box-and-whisker diagrams on a number line.				
11.10 Constructing cumulative frequency table.				
11.11 Making a diagram of an Ogive (cumulative frequency curves).				
11.12 Representing bivariate numerical data as a scatter plot.				
11.13 Selecting a function that best fit the data for linear, quadratic and exponential.				

12. The following items seek your opinion about the level of difficulty learners have with learning some aspects of probability. Please put your response in a rating format by placing a tick (✓) in the appropriate box below.

	Not difficult	Less difficult	Difficult	Highly difficult
12.1 Using and constructing Venn diagrams from a given word problem.				
12.2 Using and constructing two –way contingency tables from a given word. problem.				
12.3 Using and constructing tree diagrams from a given word problem.				
12.4 Identifying dependent and independent events from Venn diagrams.				
12.5 Identifying dependent and independent events from **t**wo-way contingency tables.				
12.6 Using Venn diagram to solve probability problems where events are not necessarily independent.				
12.7 Using tree diagrams to solve probability problems where events are not necessarily independent.				
12.8 Calculating probability of two independent event by applying a product rule for independent events: P(A and B) = P(A). P(B).				
12.9 Understanding the difference between independent and dependent events.				

13. The following statements are considered to be some of the reasons for learners not to do well in statistics (data handling and Probability).Please rate how strongly you agree or disagree with each of the statements by placing a tick (✓) in the appropriate box.

	Strongly disagree	Disagree	Neither disagree nor agree	Agree	Strongly agree
13.1 Learners encounter problems because I experiance problems in explaining concepts to learners.					
13.2 Learners encounter problems because I do not understand some of the topics in the syllabus.					
13.3 Learners encounter problems because I do not know which method of instruction I should use to teach statistics (data handling and probability).					
13.4 Learners encounter problems because I did not study statistics at tertiary level.					
13.5 Learners encounter problems because I do not have enough teaching experience.					
13.6 Learners encounter problems because I do not have enough class time to assist them individually with their problems.					
13.7 Learners encounter problems because they do not have a mathematics textbook.					
13.8 Learners encounter problems because they do not know how to use statistics function mode on a calculator.					
13.9 Learners encounter problems because they do not give themselves enough time to practice data handling and probability problems.					
13.10 Learners encounter problems because statistics (data handling and probability) is too difficult for them.					
13.11 Learners encounter problems because certain sections of data handling and probability were not properly taught in previous years.					
13.12 Learners encounter problems because certain sections of data handling and probability were never taught in the lower classes.					

Section D: In-service Education and Training(INSET) programmes

*In this research, **INSET Programmes** are activities undertaken by teachers to improve their skills, competencies and knowledge and also help to equip teachers to deal with curriculum and other changes, for example workshops, seminars, etc.*

14. How many days you have spent on INSET programmes for mathematics in the last 24 months? (Include attendance of workshops, and seminars, but do not include formal courses for which you received college credit or time you spend providing professional development for other teachers). Tick in the appropriate box below.

None☐$_1$ Fewer than three days☐$_2$ 3 to 7 days☐$_3$ 8 to 14 days☐$_4$

15 to 21 days☐$_5$ 22 to 31 days☐$_6$ More than 31 days☐$_7$

15. For the inset programmes which you have attended for mathematics in the last 24 months, to what extent was each of the following topics in statistics emphasized (facilitated)?

	Did not attend	Not at all	Slightly	Moderately	Largely
15.1 Probability					
15.2 Data handling					

16. For the inset programmes which you have attended for mathematics in the last 24 months, to what extent was your Subject matter knowledge and method of teaching improved in each of the following topics?

		Did not attend	Not at all	Not very well	Pretty well	Completely
16.1 Subject matter (content knowledge)	16.1.1 Probability					
	16.1.2 Data handling					
16.2 Method of teaching (pedagogical knowledge)	16.2.1 Probability					
	16.2.2 Data handling					

17 The following reasons might best explain what prevented you from participating in most inset programmes than you did in the last 24 months. Please rate how strongly you agree or disagree with each of the reasons below by placing a tick (✓) in the appropriate box.

	Strongly disagree	Disagree	Neither disagree nor agree	Agree	Strongly agree
17.1 The inset programmes were too expensive; I could not afford them as I was expected to pay for myself.					
17.2 There was a lack of financial support from our school.					
17.3 The inset programmes were arranged during teaching time.					
17.4 Inset programmes were arranged over weekends.					

17.5 I did not have time because of family responsibilities.					
17.6 There were no suitable inset programmes offered for me to attend.					
17.7 The inset programmes were offered by incompetent facilitators and were not productive.					
17.8 The inset programmes were offered very far from where I teach.					

Section E: Ways to improve the teaching of statistics

18. The following statements are some of the suggestions on how to address the problems encountered in the teaching of statistics. Please rate how strongly you agree or disagree with each of the statements below by placing a tick (✓) in the appropriate box.

	Strongly disagree	**Disagree**	**Neither disagree nor agree**	**Agree**	**Strongly agree**
18.1 Teachers should attend inset programmes regularly prior to the teaching of the topics in which they encounter problems.					
18.2 Teachers should get financial support from their schools/districts to attend inset programmes.					
18.3 Schools should plan ahead to allow their teachers to attend inset programmes.					
18.4 Inset programmes should be arranged and organised by reputable and competent service providers.					
18.5 Inset programmes should meet the needs of the teachers by offering the topics in which they encounter difficulties.					
18.6 Inset programmes should be organised in the same circuit /district as the teachers' place of work.					
18.7 It is recommended that teachers should take a formal tertiary course on statistics to improve their content knowledge and teaching.					
18.8 Textbooks should be well written (explained thoroughly) and contain all necessary information needed to teach data handling and probability (i.e. formulae, more examples, etc.)					

This is the end of the questionnaire. Thank you very much!

Appendix 2

LEARNER QUESTIONNAIRE

Purpose of the questionnaire: This questionnaire is designed to obtain the problems learners encounter in the learning of statistics (Data handling and probability). The information gathered will help to improve the learning of statistics (data handling and probability) and also help learners to perform better in Mathematics.

The response will be anonymous and used for research purposes only.

Instructions:

1. Please give an answer for every question.
2. Tick (✓) in the provided boxes what is appropriate to you.
3. There are no correct or wrong answers, if you make a mistake cross out and tick another opinion.
4. Do not write your name on the questionnaire.

	Section A: Demographic information

1. **What is your gender? Please tick (✓) one choice.**
   Male ☐$_1$ Female ☐$_2$

2. **Where is the school you attend located?**
   Township ☐$_1$ Urban/town ☐$_2$

Section B: Data handling

3. The following items seek your opinion about the level of difficulty you have with learning some aspects of data handling. Please put your response in a rating format by placing a tick (✓) in the appropriate box below.

		Not difficult	**Less difficult**	**Difficult**	**Highly difficult**
3.1.	Understanding the meaning of the mean as a measure of central tendency.				
3.2.	Understanding the meaning of the median as a measure of central tendency Median.				
3.3.	Calculating the median of ungrouped data.				
3.4.	Calculating the **mean** of ungrouped data.				
3.5.	Calculating the **median** of **grouped data.**				
3.6.	Calculating the **mean** of **grouped data.**				
3.7.	Interpreting the mean as a measure of central tendency.				
3.8.	Interpreting the median as a measure of central tendency.				
3.9.	Understanding the meaning of standard deviation.				
3.10.	Understanding the meaning of variance.				
3.11.	Interpreting standard deviation as a measure of dispersion.				
3.12.	Interpreting variance as a measure of dispersion variance.				
3.13.	Calculating the lower quartile of grouped data.				
3.14.	Calculating upper quartile of grouped data.				
3.15.	Drawing box-and-whisker diagrams on a number line.				
3.16.	Constructing cumulative frequency table.				
3.17.	Making a diagram of an ogive (cumulative frequency curves).				
3.18.	Constructing a stem-and-leaf plot.				
3.19.	Using a stem and leaf plot to determine the upper quartile.				
3.20.	Using a stem and leaf plot to determine the lower quartile.				
3.21.	Using a diagram of ogive (cumulative frequency curve) to estimate the cumulative percentages (of a less than or more nature).				
3.22.	Using a diagram of ogive (cumulative frequency curve) to estimate the lower quartile from a set of grouped data.				
3.23.	Using a diagram of ogive (cumulative frequency curve) to estimate the middle quartile (median) from a set of				

	grouped data.				
3.24.	Using a diagram of ogive (cumulative frequency curve) to estimate the upper quartile from a set of grouped data.				
3.25.	Representing bivariate numerical data as a scatter plot.				
3.26.	Constructing a line of best fit on the scatter plot.				
3.27.	Identifying a function that best fit the data.				

Section C: Probability

4. The following items seek your opinion about the level of difficulty you have with learning some aspects of probability. Please put your response in a rating format by placing a tick (✓) in the appropriate box below.

		Not difficult	**Less difficult**	**Difficult**	**Highly difficult**
4.1	Constructing a Venn diagram from a given word problem.				
4.2	Constructing a two –way contingency table from a given word problem.				
4.5	Constructing a tree diagram from a given word problem.				
4.4	Using Venn diagrams to solve probability problems.				
4.5	Using two –way contingency tables to solve probability problems.				
4.6	Using tree diagrams to solve probability problems.				
4.7	Identifying dependent events from Venn diagrams.				
4.8	Identifying independent events from Venn diagrams.				
4.9	Identifying independent events from **t**wo-way contingency tables.				
4.10	Identifying dependent events from **t**wo-way contingency tables.				
4.11	Using Venn diagram to solve probability problems where events are not necessarily independent.				

4.12	Using tree diagrams to solve probability problems where events are not necessarily independent.				
4.13	Calculating probability of two independent events by applying a product rule for independent events: P(A and B) = P(A). P(B).				
4.14	Identifying mutually exclusive events from Venn diagrams.				
4.15	Understanding the difference between independent and dependent events.				

Section D: Possible cause of the problems

5. The following reasons might best explain the causes of the problem(s) you encountered in the learning of statistics (data handling and probability). Please rate how strongly you agree or disagree with each of the reasons below by placing a tick (✓) in the appropriate box.

	Strongly disagree	**Disagree**	**Neither agree nor disagree**	**Agree**	**Strongly agree**
5.1 I encounter problems because my teacher finds it difficult to explain concepts/ does not explain clearly.					
5.2 I encounter problems because my teacher does not understand some of the topics in the syllabus.					
5.3 I encounter problems because my teacher teaches some topics but leaves others for us to do on our own.					
5.4 The teacher does not allow enough time to teach statistics.					
5.5 The teacher does not give learners enough exercises to practice statistics.					
5.6 I encounter problems because I was never taught how to use statistics function mode on the calculator.					
5.7 I encounter problems because I do not have a mathematics textbook.					
5.8 I encounter problems because I do not know how to use statistics function mode on the calculator.					
5.9 I encounter problems because I do not have access to previous examination papers with solutions.					
5.10 I encounter problems because I am learning statistics on my own, I do not have a teacher who teaches me data handling and probability.					
5.11 I encounter problems because I do not give myself enough time to practice data handling and probability problems.					

5.12 I encounter problems because statistics (data handling and probability) is too difficult for me.					
5.13 I encounter problems because certain sections of data handling and probability were not properly taught in lower grades.					
5.14 I encounter problems because certain sections of data handling and probability were never taught in previous grades.					
5.15 I encounter problems in data handling and probability because I did not pay enough attention when these topics were taught in previous classes.					
5.16 I do not see the importance of statistics in mathematics.					
5.17 I am not interested in statistics.					
5.18 I encounter problems because learners are not allowed to discuss the work during lessons.					

This is the end of the questionnaire.
Thank you very much!

Appendix 3

CLASSROOM OBSERVATION SCHEDULE

Subject: Mathematics **Topic: Data handling**

Grade level: Grade 11

Gender of teacher observed:____________________ Date:_______________________
Number of learners observed:_________________ Time:______________________

Sub-topic observed:__________________________
School:__________________

***ALL items marked "Not observable" must be explained in comments.**

Teacher's method of instruction	observed	Not observable
Teacher provides well-designed materials for the lesson and use different textbooks.		
Teacher employs learner –centred instruction (i.e. learners work in pairs and in groups, learners interact with teacher).		
Teacher discusses or uses strategies for engaging learners.		
Teacher make follow-up to check whether learners understand.		

Comments:

Teacher's content knowledge	Observed	Not observable
Teacher can explain concepts of statistics clearly to learners.		
Teacher addresses learner difficulties and misconceptions		
Teacher uses lot of data during the teaching of statistics		
Teacher provides examples during instruction.		
Teacher teaches learners to use different formulae to solve problems where necessary.		
Teacher interprets statistical results to learners.		
Teacher presents lesson according to curriculum expectations.		

Comments:

Learners	Observed	Not observable
Learners pay attention during lesson presentation.		
Learners ask questions where they do not understand.		
Learners understand statistical concepts.		
Learners understand statistical formulas.		
Learners apply statistical formulae correctly		
Learners know how to use scientific calculators to calculate mean and standard deviation/ variance.		
Learners interpret the results correctly.		
Learners discuss with each other.		
Learners do class-activity correctly.		
Learners have study materials (i.e. text books and scientific calculators).		

Comments:

Appendix 4

DIAGNOSTIC TEST

Subject: Mathematics
Grade level: Grade 11
Topic: Data Handling
Time: 80 minutes
Total: 50 MARKS

INSTRUCTIONS AND INFORMATION

Read the following instructions carefully before answering the questions.

1. This question paper consists of 4 questions.
2. Answer ALL the questions.
3. Clearly show ALL calculations, diagrams, graphs, etc., that you have used in determining your solutions.
4. Write all steps because answers only will not necessarily be awarded full marks.
5. You may use an approved scientific calculator (non-programmable).
6. If necessary, round answers off to TWO decimal places, unless stated otherwise.
7. An information sheet, with formulae, is included at the back of the question paper.
8. Number questions according to the numbering system used in this question paper.
9. A diagram sheet for answering QUESTION 4.2 is attached at the back of this question paper. Write your **NAME** on this sheet in the space provided and insert them in the back of the cover of your ANSWER BOOK.
10. **DO NOT WRITE ON THE QUESTION PAPER.**
11. Write legibly and present your work neatly.

QUESTION 1

The mass (in kg) of 15 dogs were recorded as follows:

15.9, 14.1, 14.4, 14.4, 14.4, 14.5, 14.5, 14.6, 14.7, 14.7, 14.7, 14.9, 15.1, 10.2, 16.4

1.1 Calculate the mean, median, standard deviation, and inter-quartile range for the above data. (6)

1.2 Do these data contain any outliers? If so, describe them. (4)

1.3 Which is the better measure of the centre for these data, the mean or the median? Explain your answer. (2)

1.4 Which is the better measure of spread for the data, the standard deviation or the inter-quartile range? Explain your answer. (2)

QUESTION 2

Consider the time (in minutes) it takes a courier service to deliver parcels from its depot in Umbilo to its customers in Durban. A sample of delivery times were taken in the last month. The frequency counts for delivery times are given in the table below:

Table 2.1: Frequency count table and ogive for the courier delivery times

Time, x, (in minutes)	Frequency	Cumulative frequency
5-<10	3	3
10-<15	5	8
15-<20	9	17
20-<25	7	24
25-<30	6	30
Total	30	

2.1 Calculate the lower quartile of the delivery time. (4)

2.2 Calculate the upper quartile of the delivery time (4)

QUESTION 3

3.1 The stem and leaf plot below represents examination marks obtained in mathematics by 25 students.

Table 3.1: Stem and leaf plot for Mathematics Examination Marks

Stem	Leaves
3	4 7 9
4	1 2 5 8
5	0 0 1 2 2 3 5 7 7
6	1 5 6 8 9
7	1 2 8
8	7

3.1.1 Draw a box and whisker diagram of the data. (5)

3.1.2 Is the box and whisker plot symmetrical, negatively skewed, or positively skewed? Explain your answer. (2)

3.2 The following is the distribution of the amount spent on cell phone calls per month by Grade11 learners.

Table 3.2: Frequency table of the amount spend on cell phone per month.

Rand	Frequency
1-20	19
21-40	46
41-60	55
61-80	31
81-100	9

3.2.1 Calculate the mean (correct to 1 decimal digit) of the amount spend on cell phone per month. (3)

3.2.2 Calculate standard deviation (correct to 1 decimal digit) of the amount spend on cell phone per month. (5)

3.3 The figure 3.1, below, is a Cumulative frequency curve (Ogive) for the amount spent on groceries last month by 30 shoppers.

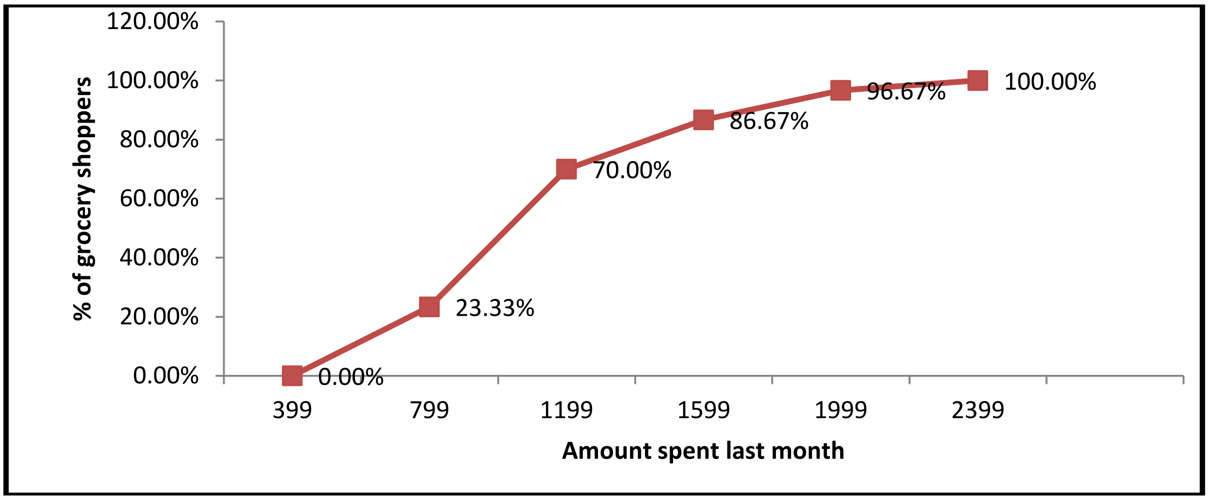


Figure 3.1: Cumulative frequency curve (Ogive) of amount spent

3.3.1 What percentage of shoppers spent less than R 1 199 last month? (1)

3.3.2 What percentage of shoppers spent more than R 1 599 last month? (2)

3.3.3 What proportion of shoppers spent between R800 and R1 600 last month? (2)

QUESTION 4

An experiment was done to determine how much coal is used to generate electricity. The results were as follows:

Coal usage (in tons)	3	13	9	10	6	14	17	7	11	8
Electricity(in kilowatt hours)	15	28	23	23	17	29	31	20	25	21

4.1 Determine the dependent and independent variable. (2)

4.2 Draw a scatter plot of the data (coal usage on the x-axis) on DIAGRAM SHEET. (4)

4.3 Suggest whether a linear, quadratic or exponential function would best fit the data. (1)

4.4 Draw a line of good fit on the graph. (1)

FORMULA SHEET

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

$$\bar{x} = \frac{\sum_i^n x_i}{n}$$

$$\sigma^2 = \frac{\sum f(x-\bar{x})^2}{\sum f}$$

$$\bar{x} = \frac{\sum fx}{n}$$

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$$s^2 = \frac{\sum f(x-\bar{x})^2}{\sum f-1}$$

Appendix 5

INTERVIEW SCHEDULE FOR TEACHERS

PROBLEMS ENCOUNTERED BY GRADE 11 MATHEMATICS TEACHERS IN THE TEACHING OF STATISTICS

Teacher background

1. What is/are your teaching qualification(s)?
2. What is your subject or area of Specialisation?
3. In which grade(s) are you currently teaching mathematics?
4. As a grade 11 mathematics teacher do you also teach statistics as a topic?
5. Have you ever studied statistics as a course (subject) while you were training as a teacher or at the tertiary institution?
6. If yes, in previous question, until what level have you studied statistics?
7. How many years have you been teaching statistics (data handling and Probability) at your schools?

Teacher knowledge

8. Do you think you have sufficient content knowledge to teach statistics at grade 11?
9. Do you think there is a need to upgrade your subject knowledge in statistics?
10. Have you done anything to improve your knowledge in data handling and Probability?

Teachers' problems and the cause(s) in the teaching of data handling

11. What problems do you experience when teaching data handling?
12. What are the causes of the problem(s) you experience in the teaching of data handling, do you think?

Teachers' problems and the cause(s) in the teaching of probability

13. What problems do you experience when teaching probability?
14. What are the causes of your problem(s) in the teaching of probability, do you think?

15. Do you think other teachers are also experiencing problems with the teaching of probability in statistics? Why?

Suggestions on how to address the problems in the teaching of data handling and probability

16. What do you think should be done to address the problems experienced by teachers in data handling and probability?

Thank you very much. Your participation is highly appreciated

Appendix 6

INTERVIEW SCHEDULE FOR LEARNERS

PROBLEMS ENCOUNTERED BY GRADE 11 MATHEMATICS LEARNERS IN THE LEARNING OF STATISTICS

1. Do you like learning data handling and probability? Why?
2. Do you encounter problems with the learning of data handling?
3. What area(s) of data handling do you encounter problems to learn?
4. What are the cause(s) of the problems you encounter in the learning of data handling, do you think?
5. Which aspects of data handling do you think your teacher finds difficult to teach?
6. Do you encounter problems with the learning of probability?
7. What aspect(s) of probability do you find difficult to learn?
8. Why do you have problems with the aspect(s) of probability that you have mentioned in question 11?
9. What are the cause(s) of the problems that you are experiencing with the learning of probability, do you think?
10. Do you think your teacher is encountering problems with the teaching of probability?
11. Which aspects of probability do you think your teacher find difficult to teach?
12. What makes you think that your teacher experiences problems with the teaching of the aspects of probability that you have mentioned in the previous question?

Thank you very much. Your participation is highly appreciated

Appendix 7

A SAMPLE OF A VALIDITY FORM BACK FROM SUBJECT EXPERT

VALIDATION FORM for Teacher questionnaire

Researcher	Validator
Name: Ms Eva Makwakwa **Institution:** University of South Africa	**Name:**_BLKMofolo-Mbokane__ **Title** _Mrs **Institution/ organisation**: University of Pretoria_ **Occupation**:**Teacher/Subjectadvisors/Lecturer__**

The following questions are based on a questionnaire to be administered to a group of grade 11 mathematics learners, in township and urban schools. The purpose of the questionnaire is to investigate the problems/difficulties encountered in the learning of statistics (data handling) and probability at Further Education and Training (FET) schools.

You are kindly requested to provide feedback on the validity of the questionnaire by answering the questions below. You can provide your feedback by inserting a cross (X) in appropriate spaces. Your feedback to the questions will be highly valued for the success of this research.

Question	YES	NO
1. Are the items in a questionnaire representative of the topics covered in statistics (data handling) and probability in the grade 11 mathematics curriculum?	x	
2. Are the items in a questionnaire at the level of understanding of the teachers in the grade 11 mathematics class?	x	
3. Will a questionnaire be able to gather information on problems encountered by learners in the learning of statistics (data handling) and probability?	x	
4. Is a questionnaire valid?	x	

Please provide comments, if necessary, on the strengths and weaknesses of a questionnaire.

The questionnaire covers all aspects necessary for the teaching and learning of statistics at Grade11level.______

Signature: 

Date: 22 January 2011

Official Stamp:

Appendix 8

2011 GAUTENG PROVINCE TEACHER WORK SCHEDULE FOR MATHEMATICS
GRADE 11

GAUTENG PROVINCE
MATHEMATICS – WORK SCHEDULE – GRADE II
2011

DATE	TOPIC	CONTENT	LO/AS	ASSESSMENT	DATE COMPLED
TERM 1					
12/1 - 14/1 (3 days)	Real N's / Exponents / Surds / Error margins	• Real and non real numbers • Add, subtract, multiply and divide surds • Error margins	11.1.1/ 11.1.2		
17/1 - 21/1	Rational exponents Number patterns	• Rational exponents • Number patterns (constant second difference between consecutive terms)	11.1.2 11.1.3		
24/1 - 28/1	Number patterns	• Number patterns (general term is quadratic)	11.1.3		
31/1 - 04/2	Manipulation of algebraic expressions	• Completion of square • Fractions with binomial denominators	11.2.4		
07/2 - 11/2	Quadratic equations / Quadratic Formula	• Quadratic equations (factorisation, completing the square and formula)	11.2.5		
14/2 - 18/2	Inequalities/ Simultaneous equations	• Quadratic inequalities with one variable • Simultaneous equations in two variables, one linear and one quadratic	11.2.5		
21/2 - 25/2	Transformations	• Rotation around the origin through an angle of 90^0 or 180^0 • The enlargement of a polygon, through the origin, by a factor of k.	11.3.4	Investigation	
28/2 - 04/3	Transformations Analytical geometry	• The enlargement of a polygon, through the origin, by a factor of k. • The equation of a line through two points	11.3.4 11.3.3		
07/3 - 11/3	Analytical Geometry	• The equation of a line through one point and parallel or perpendicular to a given line • The inclination of a given line	11.3.3		
14/3 - 18/3	Modelling	• Mathematical models to investigate real-life problems	11.2.6		
22/3 - 25/3 (4 days)	Test/Exam			Term Test	
TERM 2					
11/4 - 15/4	Graphs-parabola	• Relationships and conversions between variables: numerical, graphical, verbal and symbolical • $y = (x + p)^2 + q$	11.2.1/ 11.2.2		
18/4 - 21/4 (4 days)	Graphs-hyperbola and exponential	• $y = \frac{a}{x+p} + q$ and $y = ab^{x+p} + q;\ b>0$	11.2.1/ 11.2.2	Investigation	
28/4 - 29/4 (2 days) 03/5 - 06/5 (4 days)	Graphs-trigonometry	• Trigonometry graphs (sin kx, cos kx, tan kx, sin$(x + p)$, cos$(x + p)$, tan$(x + p)$)	11.2.1/ 11.2.2		
09/5 - 13/5	Graphs-applications	• Practical problems	11.2.3		
16/5 - 20/5	Gradient/Ave rate of change	• Average gradient and gradient of a curve at a point	11.2.7		
23/5 - 27/5	Special angles	• Function values of the special angles 30^0, 45^0 and 60^0 (in surd form where applicable)	11.3.5		
30/5 - 03/6	Identities/Reduction formula	• Derivation and use of the identities $\tan\theta = \frac{\sin\theta}{\cos\theta}$ and $\sin^2\theta + \cos^2\theta = 1$ • Derivation and use of reduction formulae for $\sin(90^0 \pm \theta)$, $\cos(90^0 \pm \theta)$, $\sin(180^0 \pm \theta)$, $\cos(180^0 \pm \theta)$, $\tan(180^0 \pm \theta)$, $\sin(360^0 \pm \theta)$, $\cos(360^0 \pm \theta)$, $\tan(360^0 \pm \theta)$, $\sin(-\theta)$, $\cos(-\theta)$, $\tan(-\theta)$	11.3.5		
06/6 - 10/6	JUNE EXAMS				
13/6 - 15/6 (3 days)	JUNE EXAMS			June Exams	
20/6 – 24/06	JUNE EXAMS				

TERM 3					
18/7 - 22/7	Revision	• June Exams Papers			
25/7 - 29/7	Data Handling: Dependant & Independent	• Measures of central tendency and dispersion in univariate numerical data by: ○ Five number summary (maximum, minimum and quartiles) ○ Box and whisker diagrams ○ Ogives ○ Calculating the variance and standard deviation sets of data manually (for small sets of data) and using available technology (for larger sets of data) and representing results graphically using histograms and frequency polygons • Scatter plot of bivariate data and intuitive choice of function of best fit	11.4.1		
01/8 - 05/8	Misuse of Statistics / Misleading graphs **(optional)**	• Symmetric and skewed data • Potential use and misuses of statistics and charts	11.4.3	**Project**	
10/8 - 12/8 (3 days)	Financial Maths	• Simple and compound decay formulae	11.1.4		
15/8 - 19/8	Financial Maths	• Different periods of compound growth and decay, Effective and nominal interest rates	11.1.5/ 11.1.6	Assignment **(optional)**	
22/8 - 26/8	Geometry **(optional)**	• Conditions for polygons Prove and use : • Midpoint Theorem • Equiangular triangles are similar	11.3.2	**(expanded opportunity)**	
29/8 - 02/9	Geometry **(optional)**	• Triangles with sides in proportion are similar • Pythagorean Theorem by similar triangles	11.3.2		
05/9 - 09/9	General solutions	• The general solution of trigonometric equations	11.3.5		
12/9 - 16/9	Sin/Area/Cos rule	• Establish and apply Sine, Cosine and Area Rules	11.3.5		
19/9 - 23/9	Sine/Area/Cos rules 2-D problems	• Application to problems in two dimensions, of the Sine, Cosine and Area Rules • Constructing and interpreting geometric and trigonometric models	11.3.6		
26/9 - 30/9	Test			**Term Test**	
TERM 4					
10/10 - 14/10	Volume & Surface area	• Volume and surface area for right pyramids, right cones and spheres	11.3.1	**Assignment**	
17/10 - 21/10	Linear Programming	• Linear programming by numerical search	11.2.8		
24/10 - 28/10	Linear Programming	• Linear programming by system of linear equations	11.2.8		
31/10 - 04/11	Probability **(optional)**	• Dependent and independent events • Two-way contingency tables • the product rule for independent events: P(A and B) = P(A).P(B)	11.4.2		
07/11 - 11/11	Probability **(optional)**	• Venn diagrams and other techniques to solve probability problems (where events are not necessarily independent	11.4.2		
14/11 - 18/11	Revision and Final Exams				
21/11 - 25/11	FINAL EXAMINATIONS				
28/11 - 02/12	FINAL EXAMINATIONS				
05/12 - 09/12	FINAL EXAMINATIONS				

Appendix 9

APPROVAL LETTER FROM HEAD OFFICE DEPARTMENT OF EDUCATION IN GAUTENG PROVINCE

2011-06-21 12:56 >> 0865941781 P 1/2

education
Department: Education
GAUTENG PROVINCE

For administrative use:
Reference no. D2012/87

GDE RESEARCH APPROVAL LETTER

Date:	**21 June 2011**
Name of Researcher:	**Makwakwa E.G.**
Address of Researcher:	**8003 Ceres South**
	229 Jacob Mare Street
	Pretoria
	0002
Telephone Number:	**012 429 6175 / 084 806 8727**
Fax Number:	**012 429 8690**
Email address:	**makwaeg@unisa.ac.za**
Research Topic:	**Exploring problems encountered in the teaching and learning of statistics in Grade 11**
Number and type of schools:	**Twenty Secondary Schools**
District/s/HO	**Tshwane South**

Re: Approval in Respect of Request to Conduct Research

This letter serves to indicate that approval is hereby granted to the above-mentioned researcher to proceed with research in respect of the study indicated above. The onus rests with the researcher to negotiate appropriate and relevant time schedules with the school/s and/or offices involved to conduct the research. A separate copy of this letter must be presented to both the School (both Principal and SGB) and the District/Head Office Senior Manager confirming that permission has been granted for the research to be conducted.

The following conditions apply to GDE research. The researcher may proceed with the above study subject to the conditions listed below being met. Approval may be withdrawn should any of the conditions listed below be flouted:

1. *The District/Head Office Senior Manager/s concerned must be presented with a copy of this letter that would indicate that the said researcher/s has/have been granted permission from the Gauteng Department of Education to conduct the research study.*
2. *The District/Head Office Senior Manager/s must be approached separately, and in writing, for permission to involve District/Head Office Officials in the project.*
3. *A copy of this letter must be forwarded to the school principal and the chairperson of the School Governing Body (SGB) that would indicate that the researcher/s have been granted permission from the Gauteng Department of Education to conduct the research study.*

1

Making education a societal priority

Office of the Director: Knowledge Management and Research
9th Floor, 111 Commissioner Street, Johannesburg, 2001
P.O. Box 7710, Johannesburg, 2000 Tel: (011) 355 0506
Email: David.Makhado@gauteng.gov.za
Website: www.education.gpg.gov.za

2011-06-21 12:57 >> 0865941781 P 2/2

4. A letter / document that outlines the purpose of the research and the anticipated outcomes of such research must be made available to the principals, SGBs and District/Head Office Senior Managers of the schools and districts/offices concerned, respectively.
5. The Researcher will make every effort obtain the goodwill and co-operation of all the GDE officials, principals, and chairpersons of the SGBs, teachers and learners involved. Persons who offer their co-operation will not receive additional remuneration from the Department while those that opt not to participate will not be penalised in any way.
6. Research may only be conducted after school hours so that the normal school programme is not interrupted. The Principal (if at a school) and/or Director (if at a district/head office) must be consulted about an appropriate time when the researcher/s may carry out their research at the sites that they manage.
7. Research may only commence from the second week of February and must be concluded before the beginning of the last quarter of the academic year.
8. Items 6 and 7 will not apply to any research effort being undertaken on behalf of the GDE. Such research will have been commissioned and be paid for by the Gauteng Department of Education.
9. It is the researcher's responsibility to obtain written parental consent of all learners that are expected to participate in the study.
10. The researcher is responsible for supplying and utilising his/her own research resources, such as stationery, photocopies, transport, faxes and telephones and should not depend on the goodwill of the institutions and/or the offices visited for supplying such resources.
11. The names of the GDE officials, schools, principals, parents, teachers and learners that participate in the study may not appear in the research report without the written consent of each of these individuals and/or organisations.
12. On completion of the study the researcher must supply the Director: Knowledge Management & Research with one Hard Cover bound and an electronic copy of the research.
13. The researcher may be expected to provide short presentations on the purpose, findings and recommendations of his/her research to both GDE officials and the schools concerned.
14. Should the researcher have been involved with research at a school and/or a district/head office level, the Director concerned must also be supplied with a brief summary of the purpose, findings and recommendations of the research study.

The Gauteng Department of Education wishes you well in this important undertaking and looks forward to examining the findings of your research study.

Kind regards

21 June 2011

Shadrack Phele MIRMSA
[Member of the Institute of Risk Management South Africa]
CHIEF EDUCATION SPECIALIST: RESEARCH COORDINATION

2

Making education a societal priority

Office of the Director: Knowledge Management and Research
9th Floor, 111 Commissioner Street, Johannesburg, 2001
P.O. Box 7710, Johannesburg, 2000 Tel: (011) 355 0506
Email: David.Makhado@gauteng.gov.za
Website: www.education.gpg.gov.za

Appendix 10

RESULTS OF RELIABILITY-TEACHER QUESTIONNAIRE

Reliability Statistics

Cronbach's Alpha	N of Items
.723	108

Appendix 11

RESULTS OF RELIABILITY-LEARNER QUESTIONNAIRE

Reliability Statistics

Cronbach's Alpha	N of Items
.938	64

Appendix 12

CORRELATION OF SCORES OF THE TWO EQUIVALENT GROUP IN PILOT TEST

Correlations

		PRETEST	POSTTEST
PRETEST	Pearson Correlation	1	.782**
	Sig. (2-tailed)		.000
	N	40	40
POSTTEST	Pearson Correlation	.782**	1
	Sig. (2-tailed)	.000	
	N	40	40

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix 13

EXAMPLE OF A LETTER TO PRINCIPALS

....June 2011

The Principal

………………………………………..

Dear Mr…………:

RE: Ms EVA MAKWAKWA APPLY FOR PERMISSION TO CONDUCT A RESEARCH

I am Eva Makwakwa, a Masters student at the Institute for Science and Technology Education (ISTE) at the University of South Africa. I am interested in exploring problems encountered in the teaching and learning of statistics (Data handling and probability) in grade 11. In order to complete the research, I need to collect data using learner questionnaire, teacher questionnaire, interviews, diagnostic test and classroom observation schedule instruments. Data collection will take only two weeks during the teaching periods of data handling and probability and only classroom observation will take place during normal working hours (classroom will not be interrupted). As for Interviews and questionnaires will take place during lunch/ breaks and after schools.

I have requested a permission from Gauteng Department of Education Head Office and Tshwane South District Office to conduct a research in Tshwane South District schools and I was granted the permission to conduct the research. Please find attached permission letters from Head office and District. I therefore request that I be allowed to conduct the research in your school. Kindly note that the information gathered will be used for research purposes only, and the name of school and participants will not be revealed in the report. Your assistance will be greatly appreciated.

Thank you in anticipation

Yours Sincerely

Ms EG Makwakwa

Tel: 012 429 6175/6473

Fax: 012 429 8690

Appendix 14

PERMISSION LETTER FROM TSHWANE SOUTH DISTRICT GAUTENG PROVINCE TO CONDUCT RESEARCH

29.JUN.2011 08:17 0123410054 POL&PLAND #7654 P.001 /001

education
Department: Education
GAUTENG PROVINCE

Reference : Policy and Planning: Partnerships
Enquiries : Sello George Ngwenya
Telephone : 012 401 6322
Fax : 012 401 6323
Fax-2-email: 0865 674 276
E-mail : Sello.Ngwenya@gauteng.gov.za
23 June 2011

Ms Eva Gavhaza Makwakwa
8003 Ceres South
229 Jacob Mare Street
Pretoria, 0002
012 429 6175(T); 012 429 8690 (F); 084 806 8727 (Mobile)
E- mail: makwaeg@unisa.ac.za

Cc: The Principal and SGB (Selected schools)

Dear Eva Gavhaza Makwakwa

PERMISSION TO CONDUCT RESEARCH: EVA MAKWAKWA

Your research application has been approved by Head Office and a copy of the letter of approval has been received by the district. **You are kindly advised to adhere strictly to the terms and conditions given by Head Office.** You are also advised to communicate with the school principal/s and/or SGB/s regarding your research and time schedule.

Our commitment of support may be rescinded if any form of irregularity/ no compliance to the terms in this letter/ Head Office research unit's terms or any other departmental directive/ if any risk to any person/s or property or our reputation is realised, observed or reported.

Terms and conditions

1. The safety of all the learners and staff at the school must be ensured at all times.
2. All safety precautions must be taken by the researcher and the school. The Department of Education may not be held accountable for any injury or damage to property or any person/s resulting from this process. The school/s must ensure that sound measures are put in place to protect the wellness of the researcher and her property.

NB Kindly submit your report including findings and recommendations to the District at least two weeks after conclusion of the research. You may be requested to participate in the Department of Education's mini-research conference to discuss your findings and recommendations with departmental officials and other researchers.

The District wishes you well.

Yours sincerely

Mrs H.E Kekana
Director: Tshwane South District

1/1

Making education a societal priority

Office of the District Director: Tshwane South District
(Mamelodi/Eersterust/ Pretoria East/Pretoria South/Atteridgeville/Laudium)
265 Pretorius Street, Pretoria 0001
Private Bag X 27825, Sunnysie, 1322 .Tel: (012) 401 6317. Fax: (012) 401 6318
Website: www.education.gpg.gov.za