INVESTIGATING POSSIBLE CAUSES OF SOME OF THE PROBLEMS EXPERIENCED IN THE TEACHING AND LEARNING OF GRADE 11 STATISTICS

Eva G. Makwakwa and David Mogari

Institute for Science and Technology Education, University of South Africa

This study investigated the possible causes of some of the problems experienced by teachers and learners when teaching and learning statistics. A convenient sample of 100 grade 11 mathematics teachers and 269 grade 11 mathematics learners in the Gauteng province schools of South Africa 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 classroom observation schedule, diagnostic test and learner interview. Results showed that possible causes of the teachers' and learners difficulties to be (1) teachers' 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 (4) learners' lack of conceptual knowledge of certain aspects of statistics; and (5) inadequate teaching of statistics topics in previous grades.

INTRODUCTION

Prior to the advent of democratic South Africa, data handling and probability were only taught as part of statistics in universities, and was not even taught in teacher training institutions. Thus, there is 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 (Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya & Makwakwa 2009; Makwakwa & Mogari, 2012). 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, non-governmental organisations such as the South African Statistical Association (SASA) and the Association for Mathematics Education of South Africa (AMESA) and the University of South Africa's (UNISA) Institute for Science and Technology Education (ISTE) organised numerous in-service training programmes to develop teachers' content and pedagogical knowledge in statistics (Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya, Dhlamini & Makwakwa 2010, 2011; Atagana et al. 2009; DoE, 2008; North and Zewotir, 2006). Nevertheless, these interventions has had little effect because teachers still have difficulties 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 (Atagana et al. 2010, 2011; Atagana et al. 2009; Makwakwa & Mogari, 2011, 2012; Wessels & Nieuwoudt

Makwakwa, E.G. & Mogari, D. (2013). Investigating possible causes of some of the problems experienced in the teaching and learning of grade 11 Statistics. In Z. Davis & S. Jaffer (Eds.), *Proceedings of the 19th Annual Congress of the Association for Mathematics Education of South Africa*, Vol. 1. (pp. 115 – 131). Cape Town: AMESA.

2011), and learners still have difficulties when interpreting, plotting and determining given statistical questions (Cassim 2012).

In this study, it is assumed that teacher knowledge influences the depth of teaching and, in turn, the quality of learning. The authors' point of view is that the knowledge of teachers can be upgraded through in-service education and training programmes, the use of quality textbooks and by elevating teacher background (qualifications, subject major, teaching experience). By understanding the possible causes of the problems in the teaching and learning of statistics, and addressing the possible causes of the problems would hopefully improve the statistics teachers' content knowlege and pedagogical knowledge, and statistics learners' content knowledge. Therefore, this paper seeks to answer the question: what are the possible cause (s) of the problems encountered in the teaching and learning of statistics in grade 11. The rationale for focusing on grade 11 was that: (1) there was no new content on data handling covered in grade 12 mathematics and probability topics were optional to all grade 12 mathematics learners during data collection of this study 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; (2) grade 11 would have learnt many topics in data handling (grade 10 and grade 11 data handling & probability topics) than grade 10 learner. It was on this basis that the study focused on grade 11. In this study, the term *statistics* is used as a collective term for data handling and probability and the word "problem" is defined as any obstacle that may hinder teachers' in creating an environment conducive to the learning and teaching of statistics. These obstacles include teachers' and learners' statistics content knowledge (i.e. learners' and teachers' difficulties in grasping concepts in statistics, learners' and teachers' difficulties to reason about the data, learners' and teachers' difficulties in using statistical formulae in calculations, learners' and teachers' difficulties to choose the correct formulae and learners' and teachers' difficulties in interpreting statistical results) and teachers' method of teaching statistics. 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.

LITERATURE REVIEW

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 & 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.

Further, there have also been instances where in-service education and training (inset) programmes are criticized for not first identifying the needs of teachers before organising them (Owusu-mensah 2008). These are inset programmes that do not adequately address the developmental needs of teachers. Such 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. It is against this background that the current study is pursued.

METHODOLOGY

Sample and Sampling technique

The study followed a descriptive survey research design. The design enabled the researchers 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. A convenient sample consisted of 269 grade 11 mathematics learners and 100 grade 11 mathematics teachers who were based in 10 schools in one of the districts of the Gauteng Province. 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. The authors selected participants because they were willing and available to be studied (Creswell, 2011). All 269 learners took part in classroom observations (see Appendix 1), 248 learners wrote a diagnostic test and 10 learners were interviewed. The group of 10 learners was chosen because they were willing and available to be interviewed. The Researcher conducted the interviews in separate room provided. In the case of the teachers, 100 completed a questionnaire, seven were observed teaching and six were interviewed. The group of seven teachers observed and six teachers interviewed were chosen because they were willing and available to be observed and interviewed. Classroom observations were conducted in eight classrooms at four of the 10 schools involved in the study. Also, Semi-structured interviews (both learners and teachers) were conducted from the four schools which were selected for classroom observations. 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.

INSTRUMENTS

This study used a teacher questionnaire, a diagnostic test, a classroom observation schedule and semi-structured interviews (teacher interview and learner interview) to collect the data. All the instruments were designed by the authors (see appendices).

Validity and reliability of instrument

Convergent validity was used to test the validity of an interviews (teacher and learner) instrument and classroom observation instrument. This was done by comparing interview instrument and classroom observation instrument (Cohen, Manion, & Morrison, 2007). Then content validity of the diagnostic test was ascertained by the same five experts in statistics education and statistics field. The teacher questionnaire was validated using content validity. Content validity was carried out by experts is statistics and statistics education field.

The reliability of an interviews and classrooms observation instruments was determined by using test-retest reliability. This was determined through a process of repeated usage of the interview and classroom observation schedule (Gay & Airasian, 2003). The reliability of teacher questionnaire was determined by computing Cronbach's alpha (the value of alpha was 0.93). Then, the Stability reliability was used to determine the reliability of the diagnostic test. 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 test showed significant correlation of 0.782 obtained at 99% confidence interval. See the appendices for example of the content of the instruments

DATA ANALYSIS

Teacher questionnaire and teacher interviews

Number of days spent in in-service education and training (inset) programmes

Table 1 shows the number of days spent by teachers in inset programmes over the last 24 months.

Table 1: Time spent on 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 22 days	16.0
Twenty-two to 31 days	5.0
More than 31 days	6.0

Table 1 also 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. Further, the findings imply that most (32%) of the teachers did not attend inset programmes regularly to improve their statistics knowledge.

The extent to which probability and data handling were covered in the workshops

Table 2 displays 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 2: Emphasis on probability and data handling in inset programmes

	Percentage of teachers' responses to coverage of probability and data handling content				
	Did not attend	Not at all	Slightly	Moderately	Largely
Probability	30.0	23.0	7.0	23.0	17.0
Data handling	30.0	16.0	11.0	21.0	22.0

Table 2 also 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; the inset programmes did not deal will all the aspects of probability required in the curriculum.

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. The findings from table 2 implies that almost half (30%) of the teachers who had attended programmes during this time are not satisfied with how the probability topic was facilitated in inset programmes.

Teachers' interview responses on the causes of the problems encountered in the teaching of data handling and probability

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.

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.

The improvement of teachers' subject and pedagogical knowledge after inset

Table 3 shows 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 3: Improvement of teachers' subject matter and pedagogical knowledge.

		Percentage of teachers				
Forms of Knowledge	Topics	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 3 also 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 Not very well) believed that their subject matter knowledge had not improved. Only a few 18% (i.e. 9% of not at all + 9% of Not very well) of those who had attended these inset programmes on data handling believed that their subject matter knowledge had not improved.

Again, table 3 shows that 38% (i.e. 21% of not at all + 17% of Not very well) 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 Not very well) 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 3 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.

Classroom observation

Teachers

Eight classrooms taught by seven teachers over three weeks were observed by the researchers. 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 researchers 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 (see Appendix 1).

The results from classroom observations showed that teachers (i.e. T1, T2, T3, T4, T5 and T6) relied on only one prescribed learner textbook when teaching mean, standard deviation or variance. 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 there were teachers who encountered problems because they lacked both content knowledge and pedagogical knowledge in statistics. For instance, some of them (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. Further example, one of the teachers, 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 (Makwakwa & Mogari, 2012). Makwakwa and Mogari reported on the problems encountered in the teaching and learning of statistics which were discovered through the use of classroom observation instrument. Their paper shed light on the teaching of some of the aspects of data handling and kind of tasks teachers gave to learners.

Learners

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.

Learners' interview responses about the cause(s) of the problems encountered in the learning of 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.

Diagnostic test

The diagnostic test used in the 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 appendix 2). 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. Time allowed for the test was an hour. The researchers scored the answer scripts using a marking rubric. Table 4 illustrates 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 data from the diagnostic test (also see Table 4) and learner interviews indicated that learners encountered problems because (1) certain sections of data handling were not adequately taught in the lower grades (i.e. outliers). Most learners (124 or 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 grade 10, but learners in grade 11 also still had problems with the concept of outliers. (2) Learners were not taught certain topics in data handling. The majority of the learners could not calculate the lower quartile (209 or 84%), upper quartile (198 or 80%) or standard deviation (195 or 79%) with numerical data that has been grouped into intervals (see the learner interviews and teacher interviews below). (3) Most of the learners lacked conceptual knowledge of certain concepts in statistics (data handling implied) (see appendix 4). (4) Learners were not taught how to use the statistical function mode on the calculator to make statistical calculations. Most of the learners (184 or 74%) calculated standard deviation for ungrouped numerical data incorrectly because they used long formulae.

The authors asked learners the following questions about problems they encountered with the learning of 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.

Table 4: Analysis of learners' performance on diagnostic test (n=248)

		Performance of learners on each test item (sub-question)					
			Partially correct				
Question	Sub-questions	Correct answer	Arithmetical problem	Conceptual problem	Wrong answer	No answer	Number of learners
Q1	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
Q2	2.1	0	0	0	209	39	248
	2.2	0	0	0	198	50	248
Q3	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
Q4	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

Learners' interview responses on the problems encountered in the learning of 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.

Teachers' interview responses on the problems encountered in the teaching of data handling

When authors asked teachers about the problems they encountered when teaching data handling, T4, T8 and T9 acknowledged 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.

Learner interviews

Learners were interviewed to gather more information on the possible causes of the problems encountered with data handling and probability in grade 11. 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.

The interviews revealed that learners L3, L5, L7 and L8, encountered problems because (1) certain sections of data handling and probability were not adequately taught in the lower grades; (2) teachers taught only some topics; (3) some concepts were not thoroughly taught.

For example, when the learners were asked to indicate what they thought were the possible cause(s) of their problems, if any, when learning about data handling. Learners L1, L2, L3, L4, L9 and 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.

Further, when learners were asked what they thought the reasons for their problems 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.

DISCUSSION: CAUSES OF TEACHING PROBLEMS IN STATISTICS

The findings in this study show that there were teachers who experienced difficulties teaching statistics because of poor content knowledge. The finding is consistent with those by 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). For example, 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.

The analysis also revealed that the teachers encountered problems because the textbook they were using were not explicit enough. Certain formulae were missing and not enough examples were provided. These results are consistent with the findings by Lue (1998), who found that teachers experienced difficulties because some statistical concepts and some sections in the textbook were unclear. Textbooks are expected to be explicit and to provide a framework for what is taught, how it may be taught, in the sequence in which it should be taught (Lemmer, Edwards & Rapule 2008). Furthermore, the findings seem to attribute the teachers' difficulties to the quality and relevance of the inset programmes offered. The inset programmes do not seem to cover essential statistics topics. It implies that the inset programmes are not normally based on the problems experienced by teachers in their teaching. It is our contention that there should be inset programmes offered based on the needs of the teachers. Moeini (2008) warns that teacher training that ignores the 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 it dampens the motivation and enthusiasm of the teachers.

Causes of learning problems in statistics

The study found that the learners had problems because of their teachers' lack of statistics content knowledge and relevant pedagogical knowledge. Learners indicated that their teachers were finding it difficult to explain concepts in class and also left difficult topics for learners to study on their own. This adversely affected the learners' learning of statistics. As a result, learners' conceptual knowledge of statistics was somehow compromised. This finding is unusual; given that there are teachers who struggle with some aspects of statistics as a result they do not teach those aspects (cf. Cardoso, 2007; Da Silva & De Queiroz e Silva Coutinho, 2008; Groth, 2009; Sharma, 2006). Also, the study showed that the learners experienced difficulties because of how they were taught statistics in lower grades. Certain aspects of data handling and probability were treated superficially by teachers or not taught at all. This finding is consistent with observation made by Da Silva and De Queiroz e Silva Coutinho (2008). Lastly, the study found that learners encountered problems with statistics because of either they could not use the learning resources (e.g. learners indicated that they did not know how to use the statistics function mode in their calculators when doing statistical calculations) or the resources were of poor quality (e.g. learners mentioned that the prescribed textbook did not cover all the topics and provided few examples). The results are consistent with findings by Lue (1998), who found that the major factors that constrain learners' learning of descriptive statistics (data handling) are the lack of clarity on some statistical concepts in some sections in the textbook.

In sum the study has showed that there are problems in the teaching and learning of statistics as a result of (1) teachers' 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; (4) teachers failure to attend in-service teacher workshops; (5) learners' lack of conceptual knowledge of certain aspects of statistics; and (6) inadequate teaching of statistics topics in previous grades.

REFERENCES

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.

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

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

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

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

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.

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

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.

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

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.

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*.

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

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

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.

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

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

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

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. *Icots 5* (pp. 580-585). Singapore: IASE. http://www.stat.auckland.ac.nz/~iase/publication.php.

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).

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

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.

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).

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

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.

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). *Dynamics of Effective Secondary Teaching*. Boston, Massachusetts: Allyn & Baco.