

Some of the problems experienced by teachers in the teaching of data-handling in grade11

Eva G. Makwakwa & David Mogari

Institute for Science and Technology Education, University of South Africa

Email: makwaeg@unisa.ac.za, mogarld@unisa.ac.za

Abstract

The study investigated problems experienced by teachers when teaching data-handling in grade 11 mathematics classes. Data-handling is a relatively new topic in school mathematics in South Africa. Through a descriptive survey design entailing the use of a semi-structured interviews and classroom observations problems experienced by teachers when teaching data-handling were investigated. A purposive sample of nine grade 11 mathematics teachers based in the Gauteng province schools participated in the study. The results showed that teachers had problems with interpreting and determining measures of dispersion (i.e., interpretation of variance and standard deviation; calculating variance and standard deviation with grouped data); representing and interpreting data on graphs or plots (i.e. cumulative frequency graph (ogive), box-and-whisker plot, scatter plot and line of best fit.); and, determining the five number summary (i.e., lower quartile, middle quartile and upper quartile with grouped data and when total number of the data values (n) is even).

Keywords: data-handling, teaching, problems, pre-service and in-service teacher training

Introduction

Before 1994 data handling was not part of school curriculum in the South African education system, instead it was only taught in universities and technikons (now known as universities of technology). Data handling was also not taught in the colleges of education where a considerable number of the current crop of teachers got their pre-service training from (see, for example, Mokgaphame, 2001; Owusu-Mensah, 2008). Most teachers therefore only saw data handling for the first time when it was introduced in school mathematics (Makwakwa & Mogari, 2012) where it is a component of statistics covered under the fourth learning outcome (LO4) of mathematics in the outgoing National Curriculum Statement (NCS) Grades 10-12 (Department of Education [DoE], 2003). In the incoming curriculum, Curriculum and Assessment Policy Statement (CAPS), data-handling is referred to as statistics forms part of the main 10 topics in Grades 10-12 (Department of Basic Education [DoBE], 2011:9).

In an attempt to strengthen the content knowledge of teachers in statistics there has been a series of in-service training programmes (DoE, 2008), which, unfortunately, seem to be have had little effect on improving the teaching of data handling (Mahlobo, 2009; Makwakwa & Mogari, 2012). Owusu-Mensah (2008) has noted that most in-service training programs do not address the teachers' actual needs; because they tend to speculate on teachers' difficulties and deficiencies. Some of the in-service workshops rely only on self evaluation

method of identifying the needs and weakness of a teacher (Owusu-Mensah, 2008). Perhaps, it is for this reason that there is no notable teacher change for teachers who attended such in-service training programmes. This study intends to address the question: What problems do grade 11 mathematics teachers experience when teaching of data handling? In this paper, word problem is defined as an obstacle that hinders teachers' success to generate effective learning environment when teaching data handling. These obstacles could amount to teacher's difficulty in explaining concepts in data handling, teachers' inability to reason about the data, teachers' difficulty to use statistical formulae in calculations, teachers' inability to choose a correct formula, teachers' difficulty to interpret statistical results, etc.

Literature Review

Several studies have been conducted to investigate problems experienced by teachers in the teaching of data handling (see, for example, Atagana et. al., 2009; Atagana, Mogari, Kriek, Ochonogor, Ogbonnaya, Dhlamini & Makwakwa, 2010; Countinho, 2008; Da Silva & De Queiroz Garegae, 2008; Jacobbe, 2008; Makwakwa & Mogari, 2012; Wessels & Nieuwoudt, 2011). For example, Atagana et al (2009) investigated perceived difficult topics by learners and teachers in the teaching and learning of mathematics in some of the provinces in South Africa. Atagana et al adopted a descriptive survey design involving a teacher questionnaire and found that teachers had problems with data handling. Atagana and his team undertook a similar study with improved objectives and a much more geographically diversified sample of teachers in 2010, and still came up with same results (Atagana et al, 2009).

Wessels and Nieuwoudt (2011) studied the profile of mathematics teachers' statistical knowledge, beliefs and confidence with a view to developing an appropriate in-service teacher education programmes to improve the knowledge of statistics for grades 8 and 9 teachers. They found that teachers show high levels of confidence in teaching most statistics topics and 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 also noted that it is in aspects that involve sampling and probability topics (probability language, basic probability calculations) where teachers showed lower levels of confidence.

Garegae (2008) studied challenges teachers experience in the teaching of statistics in Botswana. Her study used an open-ended questionnaire that was completed by 23 Heads of Department (HOD) and 30 teachers from senior schools. The study aimed to solicit HODs' and teachers' experiences of teaching statistics. The data collected were analyzed using Tesch (1990) and Bogdan and Biklen's (1992) techniques. Garegae's study showed that the teachers could not:

- i. explain statistics concepts to learners;
- ii. solve statistics problems taken from past examinations papers;
- iii. develop appropriate statistics activities for learners;
- iv. figure out syllabus objectives; and
- v. relate the teaching of statistics to learners' real-life experiences.

Garegae (2008) also found that teachers who were not outrightly trained to teach statistics encountered more problems.

Cardoso (2007) applied an activity organized into three stages to 29 high school teachers, viz., a data set; two distributions represented by tables; and two represented graphically. Teachers were asked to analyze the data through association between the mean and standard deviation and between median and quartiles; to calculate the mean and standard deviation and the median and quartiles; and to explain the meaning of the results obtained after calculating the summaries (the mean, standard deviation, median and quartiles). A discussion ensued with the teachers on the values of summaries and their meaning, which showed that teachers had difficulty in answering and giving a clear critical analysis and meaning. The interviewed teachers could not even make an oral or written analysis to justify the results. Cardoso also found that teachers wrongly analyzed concepts like mean and median; and some teachers attributed symmetry to all data distributions; leading to the confusion between median and mean.

Da Silva and De Queiroz e Silva Coutinho (2008) explored how Brazilian secondary school teachers rationalised variation of a univariate distribution. The results showed that none of the teachers showed a complete reasoning, which would relate to the understanding of mean, deviations from the mean, and the interval of k standard deviations from the mean and estimation of frequency in this interval. The teachers' predominant reasoning about variation was verbal and understood standard deviation to be a measure of sample homogeneity. Da Silva and De Queiroz e Silva Coutinho reckoned verbal reasoning about variation did not enable teachers to teach their learners the meaning of standard deviation, and this restricted them to teaching algorithms. According to one of the teachers they were only taught how to calculate the mean, median and the standard deviation, as a result the teacher was not aware of that these three concepts were somehow related (Da Silva & De Queiroz e Silva Coutinho, 2008).

Jacobbe (2008) examined three elementary school teachers' understanding of mean and median in USA during an 18month course. Jacobbe found that two of the three teachers had procedural knowledge of the measures of centre and only one had conceptual understanding of the measure of centre. Clearly from the studies discussed there are

problems embedded in the teaching of data handling in schools, and it is against this backdrop that the current study is pursued.

Methodology

Study design and Sample

The study followed a descriptive survey research design on a convenient sample of nine grade 11 statistics teachers of public schools in the Gauteng province, South Africa. 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 with funds to attend the educational development programmes.

Instruments and Data collection

Data were collected using classroom observation instrument and semi-structured interviews. Teachers were first observed teaching and thereafter interviewed. Each teacher was observed at least once.

Validity and reliability of instrument

Classroom observation instrument was validated by three established researchers, where they scrutinised its aspects and decided on its suitability for the purpose of the study. Convergent validity was then used to test the validity of teacher interviews and classroom observation instrument by comparing the data obtained through these two procedures (Cohen, Manion, & Morrison, 2007). The respective reliabilities of the semi-structured interviews and classrooms observation instrument were determined by using test-retest derived from their repeated use (Gay & Airasian, 2003).

Data Analysis

1. Classroom observation

According to the Gauteng province teacher work schedule for mathematics grade 11, data handling is supposed to be taught over two weeks. However, it appeared that some schools could not comply with the schedule. Only two schools managed to teach data handling over the required period. The codes T1, T2, T3, and so on, were used to identify the teachers. The teacher who was observed first was referred to as T1, the teacher who was observed second was coded T2, and so on. Therefore, teachers' codes ran from T1 to T7, where T7 was observed last. The duration of mathematics lesson in schools of T1 & T2; T3 & T4 was 40 minutes; lessons taught by T5 & T6 were 30 minutes each; and T7's lesson was 45 minutes long.

In T1's class, it was noted that the teacher had already started teaching data handling. In fact, T1 started teaching data handling ahead of its scheduled time set in the work schedule. However, other six teachers taught data handling within the stipulated time frame. This arrangement posed a challenge during classroom observations because some lessons occurred simultaneously across the schools and rearranging the teaching periods was not possible. It was not even possible to arrange for after school lessons since most of learners stayed far away. Only lessons on Ogive, Measure of central tendency, Five number summary, Stem and leaf plot, Box and whisker plot, Measure of dispersion, and Scatter plot & line of best fit were observed. The results of classroom observation are presented in a Table 1 according to the data handling aspects taught.

Table1: Implementation of an observation schedule during teachers' lessons

Topics observed & Total time spent in class (in minutes)	Classroom observations feedback						
	T1	T2	T3	T4	T5	T6	T7
Topic: Ogive (Cumulative frequency graph)	*	*	x	x	*	*	*
Total time for topic observed	160	40	x	x	70	35	45
Topic: Measure of central tendency (mean, mode, median)	x	*	*	*	*	*	*
Total time for topic observed	x	80	40	80	70	70	45
Topic: Five number summary (minimum number, Q1, Q2, Q3, Maximum number)	*	*	*	*	*	*	*
Total time for topic observed	160	40	40	40	70	35	45
Topic: Stem and leaf plot	x	√	√	√	√	√	√
Total time for topic observed	x	80	40	40	35	70	45
Topic: Box and whisker plot	x	√	√	√	√	*	√
Total time for topic observed	x	40	40	40	35	35	45
Topic: Measure of dispersion (variance and Standard deviations)	x	*	*	*	*	*	*
Total time for topic observed	x	40	40	40	35	70	45
Topic: Scatter plot and line of best fit	x	x	*	*	x	x	√
Total time for topic observed	x	x	80	40	x	x	90

√ observed with no problems

* Observed and problems manifested

x Not observed

1.1 Ogive (Cumulative frequency graph)

Table 1 shows that there were problems experienced in five classrooms taught by T1, T2, T5, T6 and T7. The problems emerged during lessons on cumulative frequency graph. Lessons on

this concept were not observed in lessons taught by T3 and T4 because of time table clashes. The following problems were noted in the observed lessons:

- (i) T1, T2, T5, T6 and T7 did not teach the learners that the cumulative frequency graph can also be used to estimate the cumulative percentages (of a less than or more than nature). T1, T2, T5, T6 and T7 taught learners to estimate the cumulative percentages (of a less than or more than nature) only with cumulative frequency table and T1 & T2 used only one example with ungrouped numeric data to teach ogive.
- (ii) T1, T2, T5, T6 and T7 did not teach learners how to construct cumulative frequency graph when the numeric data is grouped into intervals. T1, T2, T5, T6 and T7 did not teach learners how to use cumulative frequency curve to find an estimate of the median, lower quartile and upper quartile.

1.2 Measure of central tendency

According to Table 1 there were significantly common problems in the lessons taught by T2, T3, T4, T5, T6 and T7. The lessons were on measure of central tendency and the teachers did not define nor explain the mean concept (i.e. explaining its use, disadvantages and when it can be used). T2, T3, T4, T5 and T7 could only explain how to compute the mean and define the formula for computing it for ungrouped data. The Table also shows that lessons on the measure of central tendency taught by T1 were not observed because the teacher taught the measure of central tendency before the time set in the work schedule. The following problems were identified in the lessons:

- (i) T2, T3, T4, T5, T6 and T7 did not define nor explain the mean concept. T2, T3, T4, T5 and T7 only explained how to compute the mean and defined the formula for computing the mean for ungrouped data.
- (ii) T2, T3, T4, T5 and T7 did not teach how to compute the mean, modal and median with a set of numeric data when it was grouped into intervals.
- (iii) T6 did not teach how to compute the modal and median with a set of numeric data when it was grouped into intervals.
- (iv) T2, T3, T4 and T5 did not teach how to compute the mean using a calculator.

1.3 Five number summary

Table 1 indicates that T1, T2, T3, T4, T5, T6 and T7 encountered problems when teaching the five number summary concepts. In particular,

- (i) T1, T2, T3, T5 and T6 used only counting method to teach how to determine lower quartile (q1), median (q2) and upper quartile (q3).
- (ii) T1, T2, T3, T5 and T6 did not teach learners how to determine the quartiles (q1, q2, q3) with formulas of quartile positions. T4 and T5 had a problem of 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.

(iii) Even though T1, T2, T3, T4, T5 and T6 understood the meaning of quartiles (q_1 , q_2 , q_3) they could not explain the meaning of quartiles and interpret the quartiles to learners.

(v) The teachers did not teach how to compute the quartiles with numeric data when data has been grouped into intervals.

1.4 Stem and leaf plot

It is evident in Table 1 that T2, T3, T4, T5, T6 and T7 did not have problems with teaching stem and leaf plot concepts. T1's lessons on stem and leaf plot clashed with those taught by T3 and T4.

1.5 Box and whisker plot

Table 1 illustrates that only T6 had problems teaching the topic of box and whisker plot where he could not explain the measure of skewness observed in box-and-whisker plot and incorrectly explained that a box and whisker plot is negatively skewed when the median is close to lower quartile and is positively skewed when the median is close to upper quartile. T1 was not observed teaching box and whisker plot because the lessons clashed with those taught by T3 and T4.

1.6 Measure of dispersion (variance and standard deviation)

In table 5 we observe that there were problems identified in lessons on variance and standard deviation which were taught by T2, T3, T4, T5, T6 and T7. The table also shows that T1 was not observed teaching variance and standard deviation because he had already taught the concept as he did not follow the official work schedule. The following problems were noted in various lessons:

- (i) T2, T3, T4, T5, T6 and T7 did not teach how to interpret variance and standard deviation. They instead only taught learners to compute variance and standard deviation with a set of ungrouped numeric data.
- (ii) T2, T3, T4, T5 and T7 did not teach how to compute variance and standard deviation with a set of grouped data where data is grouped into intervals.
- (iii) T2, T3, T4, T5 and T6 did not teach how to compute variance and 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.

1.7 Scatter plot and line of best fit

What is also evident in Table 1 is that there were problems in the lessons on the scatter plot taught by only T3 and T4, where they did not explain, for example, what scatter plot is all about, why and when scatter plot is used; did not teach correct methods (i.e. median-median line and least squares regression line) for constructing a line of best fit; and misled learners when they told them that to construct a line of best fit learners should construct a positive slope between the points. T7's lessons were devoid of problems. T1, T2, T5 and T6 did not teach the scatter plot and line of best fit at all hence there is no record of their lessons.

2. Interviews

Only T2, T3, T4, T6, T8 and T9 were interviewed. T8 and T9 were not observed teaching. Teachers T1, T5 and T7 were not interviewed because they did not want to be tape recorded. T8, T4, and T9 acknowledged that they encountered problems with histogram with grouped data; ogive; box-and-whisker plot; calculation of quartiles (Q1 & Q3); variance and standard deviation with grouped data; median when the total number of observation is even; and scatter plot. T2, T6 and T3, indicated that they did not experience problems with the teaching of data handling and this was contrary to what was observed in their lessons. It is not surprising as Robson (2002: p.310) noted that observation "provides a reality check", because "what people do may differ from what they say". This is one of the reasons Owusu-Mensah (2008) suggested that in addition to self-evaluation by teachers, there is a need for classroom observations to determine further in-service needs as it is not easy for some teachers to acknowledge their teaching problems.

Discussion

The data showed that T2, T3, T4, T5, T6, T7, T8 and T9 had problems with interpreting and calculating the measure of dispersion such as variance and standard deviation more especially with a grouped data. The results are consistent with the findings by Cardoso(2007), and Da Silva and De Queiroz e Silva Coutinho (2006). Also, the data showed that T1, T2, T3, T4; T5, T6 and T7 have problems with representing and interpreting the data of plot such as cumulative frequency graph, box and whisker plot and Scatter plot and line of fit. This finding can be compared with that by Bruno and Espinel (2009) who found large percentage of primary teachers having difficulties constructing histograms and the evaluation of graphs. Even though the study by Bruno and Espinel was not on ogive box-and-box plot, scatter plot and line of best fit, per se, its relevance is derived from the fact that the teachers battled with some form of graphs. Furthermore, T1, T2, T3, T4, T5, T6, T7, T8 and T9 had difficulties with five number summary. Unfortunately, there seems to be a dearth of literature 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). Invariably, it is not strange that teachers have problems with such aspects of statistics given that there is evidence of teachers with a deficient statistical

knowledge (see, for example, Atagana et al., 2009; Atagana et al., 2010; Wessels & Nieuwoudt, 2011). The data show that T2, T3, T4, T5, T6, and T7 have problems with interpretation and calculation of measure of central tendency such as mean and median more especially with a grouped data. The results are consistent with the findings by Jacobbe (2008), Da Silva and De Queiroz e Silva Coutinho (2008), and Cardoso (2007). Teachers do not teach the meaning and interpretation of the values of the measures of central tendency and measures of dispersion, they instead focus on algorithms and mechanical approaches (Da Silva & De Queiroz e Silva Coutinho, 2008; Cardoso, 2007).

In conclusion, the study has shown that there are problems embedded in the teaching of data handling in grade 11. Teachers are still not well grounded with some aspects of data handling. This implies that there is a need for more in-service training for teachers particularly on the aspects discussed in the current study. There is also a need to focus a great deal on improving the teachers' statistical content knowledge.

References

- 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. (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.
- Cardoso, R. (2007). *The mathematics' teachers and the exploratory data analysis in high school*. published Masters thesis. PUC-SP. Online: www.pucsp.br/pos/edmat.
- Cassim, I. (2012, February 9). *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 e-library.
- Da Silva, C. B., & De Queiroz e Coutinho, 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.
- 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. (2003). *National Curriculum Statement grades 10-12 (General). Mathematics*. Pretoria, South Africa.
- Department of Education. (2008). *National Curriculum Statement grades 10-12 (General) Subject Assessment Guidelines: Mathematics*. Pretoria: Department of Education.
- Education Foundation.(2010). *Education in South Africa*.Big Media Publishers, South Africa.
- Gay, L. R., & Airasian, P. (2003). *Educational research: Competencies for analysis and applications. Seventh Edition*. Upper Saddle River, New Jersey: Pearson Education, Inc.
- 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.
- 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.
- Mahlobo, R. K. (2009). *A model for an open-ended task-based approach in grade 11 mathematics classes*. University of North West. Potchefstroom : Unpublished PhD Thesis.
- Makwakwa, E. G., & Mogari, D. (June, 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 Tasks*. 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).
- Mokgaphame, P. M. (2001). *The management of OBE teacher training in the North Province*. Pretoria: University of South Africa: Unpublished Masters dissertation.
- 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*. Pretoria: University of South Africa: Unpublished Masters dissertation.
- Wessels, H., & Nieuwoudt, H. (2011). Teachers' professional development needs in datahandling and probability. *Pythagoras*, 32(1), 1-9.