RESOURCING LEARNER ERRORS AND MISCONCEPTIONS IN THE TEACHING AND LEARNING OF STATISTICS AT GRADE 11 LEVEL.

Name Phaison Jonhera

Student Number 1532350

Protocol Number 2017ECE049M

Supervisor Dr Makonye

A research report submitted to the Faculty of Humanities, School of Education, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Master of Education.

28 September 2018

Declaration

I declare that this research report is my own, unaided work. It is being submitted for Master of Education degree at the University of the Witwatersrand, Johannesburg, South Africa. It has never been submitted before for any degree or examination in any other university.

SIGNATURE

28 September 2018

Dedication

Firstly, I dedicate this study to my wife, Chies and children, Deon, Divine and Delight. You have been my sole source of courage and strength throughout the course of this study. You stood by me in good times as well as when my moral was at its lowest level. You forego all the luxuries in life while the available, meagre financial resources were channelled towards funding for this study. Last, but not least, this study is also dedicated to my late mother, Mildred and my father, Boniface, for bearing me and making sure that I get basic education. Your words of encouragement in my quest to attain education will forever linger in my memory. I never imagined that your support and sacrifice would take me this far.

Acknowledgements

My thanks are due to my supervisor, Dr Makonye for his professional guidance, expertise, support and encouragement in putting up ideas together towards the successful completion of this study. I also wish to extend my appreciation to the National Research Fund, through Dr Pournara of the Wits Maths Connect Project of the Witwatersrand University, for providing funds towards my studies.

I am also indebted to the following institutions and individuals: Gauteng Department of Education, Mr Theta and Mrs Khumalo, the Provincial Education Authority, District Director and my school principal respectively, for according me the opportunity to conduct this study at their school. Last but not least, I shall forever be grateful to the Grade 11 mathematics class of 2017, which participated generously, resolutely and flawlessly, throughout the data collection process. Without such commitment from this wonderful group of young, future leaders of our great nation, this study would not have come to fruition.

This study is also dedicated to my fellow, part-time master's students at the Witwatersrand University

Last but not least, I wish to express my appreciation to Tapiwa Chingono and Michael Mangoyi for taking time off their busy schedules in order to edit drafts of this report.

Abstract

Access to mathematical concepts has for long been a major talking point, not only in South Africa, but throughout the world. Sadly, debate on ways of assisting learners to improve their understanding of mathematical concepts appear not to be yielding desirable results. One of the major barriers to effective and meaningful mathematics learning is errors which usually arise from misconceptions. To my knowledge and experience, learners hold a lot of misconceptions in the area of variability of data. In spite of this, little ground has been covered in misconception research around the study of statistics. In this regard, the study aimed at establishing the nature of errors which arise from misconceptions which learners hold particularly in representing and interpreting variability of data on ogives, frequency polygons and box plots. In addition, the study sought to establish the extent to which teaching intervention could assist towards error minimisation. Using ideas of constructivism and sociocultural theory, the study involved a group of eighteen Grade 11 learners and the researcher. It was conducted at a township high school in Gauteng, South Africa. The researcher was involved in the delivery of ten lessons on statistics. Upon completion of the topic, two tasks were assigned to learners prior to as well as after conducting a constructivist-based intervention lesson. The rationale for conducting the lesson was to determine the extent to which remediation could assist in minimising incidences of these errors. In order to capture as much data as possible, particularly on the thinking behind the errors, an interview was conducted on a sample of four learners.

The study found that most errors made by learners emanated from application of irrelevant prior knowledge (conceptual errors) which led to the use of unsuitable methods (procedural errors). Through interviews, it emerged that misconceptions in the study of statistics arise from poor language proficiency. Learners struggled to express themselves in giving verbal responses to questions posed to them by the interviewer.

The study recommends that teachers should always use the language of teaching and learning all the time so as to enable learners to improve their communication skills as well as enhance understanding of statistical concepts. Moreover, this study recommends that future studies be carried out on the statistical content knowledge of teachers as this could be yet another source of misconceptions held by learners in statistics.

Key words: Error, misconception, teaching intervention

TABLE OF CONTENTS

Declaration………………………………………………………………………………ii

Dedication iii

Acknowledgements iv

Abstract v

TABLE OF CONTENTS vi

LIST OF FIGURES ix

LIST OF TABLES x

CHAPTER 1: BACKGROUND AND CONTEXT OF THE STUDY 1

1.1 Introduction 1

1.2.1 Traditional and contemporary teaching and learning approaches in mathematics. 1

1.2.2 A brief history of the South African Education System 2

1.2.3 Mathematics Education in South Africa 4

1.3 Problem statement 6

1.4 Research aims 8

1.5 Research questions 8

1.6 Significance of the study 9

1.7 Conclusion 9

1.8 Definition of terms 10

CHAPTER 2: LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK 11

2.0 Introduction 11

2.1 Teaching and learning of mathematics 11

2.2 Errors and misconceptions 12

2.3 Errors and misconceptions in statistics 13

2.3.1 Box and whisker plots 14

2.3.2 Histograms and bar graphs 15

2.3.3 Measures of location 15

2.3.4 Measures of variability 16

2.4 Teacher mathematical knowledge 16

2.5 Learner mathematical proficiency 19

2.6 Conceptual framework 21

2.6.1 Intervention teaching 21

2.6.2 Multiple External Representations (MERs) 21

2.6.3 Benefits of MERs....22
2.6.4 Application of MERs in error remediation....23
2.7 Conclusion....26
CHAPTER 3: THEORETICAL FRAMEWORK....27
3.0 Introduction....27
3.1 Sociocultural perspective....27
3.1.1 Sign, tool and mediated activity....27
3.1.2 The controversy of learning and development....29
3.1.3 Spontaneous and scientific concepts....30
3.1.4 Zone of Proximal Development....31
3.2 Constructivism....32
3.2.1 Information processing theory....33
3.2.2 Weak constructivism....33
3.2.3 Radical constructivism....34
3.2.4 Social constructivism....34
3.3 APOS theory....35
3.3.1 Action....35
3.3.2 Process....36
3.3.3 Object....36
3.3.4 Schema....37
3.4 Conclusion....38
CHAPTER 4: RESEARCH METHODOLOGY....39
4.0 Introduction....39
4.1 Research design....39
4.2 Research setting....41
4.3 Sample discussion....41
4.4 Methods....41
4.4.1 Pre- and post-intervention tasks....42
4.4.2 Interview....42
4.4.3 Limitations in data collection....43
4.5 Summary of intervention lesson....43
4.6 Analytical protocol....45
4.7 Description of the analytical protocol....47
4.8 Reliability and Validity....48

4.9 Ethical considerations 49
4.10 Conclusion 49
CHAPTER 5: DATA ANALYSIS AND DISCUSSIONS 50
5.0 Introduction 50
5.1 Analysis style 50
5.2 Pre-intervention task diagnostic analysis 51
5.3 Pre- and post-intervention and interview data analysis 58
5.4 Discussions 63
5.5 Conclusion 65
CHAPTER 6: FINDINGS AND RECOMMENDATIONS 66
6.1 Introduction 66
6.2 Research findings 66
6.3 Limitations of the study 67
6.4 Recommendations 68
6.4.1 Future researchers 68
6.4.2 Practice of teaching 68
6.4.3 Teacher development 69
REFERENCES 70
APPENDICES 79
APPENDIX 1: ETHICS CLEARANCE LETTER 79
APPENDIX 2: GDE LETTER OF APPROVAL 80
APPENDIX 3: INFORMATION SHEET FOR PRINCIPAL AND SGB CHAIR 81
APPENDIX 4: INFORMATION SHEET FOR LEARNERS 82
APPENDIX 5: LEARNER CONSENT FORM 83
APPENDIX 6: INFORMATION SHEET FOR PARENTS 84
APPENDIX 7: PARENT’S CONSENT FORM 85
APPENDIX 8: TASK SHEET FOR LEARNERS 86
APPENDIX 9: MODEL SOLUTIONS 90
APPENDIX 10: INTERVIEW PROTOCOL 94
APPENDIX 11: EXEMPLAR INTERVENTION LESSON PLAN 95
APPENDIX 12: VIGNETTES FROM PRE-INTERVENTION TASKS 99

LIST OF FIGURES

Figure 1: Pedagogical and Content Knowledge....................18
Figure 2: Conceptual framework for using MERs in error remedial instruction....................24
Figure 3: Mediated activity....................28
Figure 4: APOS Theory....................38
Figure 5: Analytical protocol....................46
Figure 6: Vignette 1 for Que 1.1 and 1.2 (JPM3)....................51
Figure 7: Vignette 6 for Que 1.3 (JPM 7)....................52
Figure 8: Vignette 9 for Que 1.4 (JPM 7)....................53
Figure 9: Vignette 11 for Que 2.1 (JPM 13)....................54
Figure 10: Vignette 14 for Que 2.2 (JPM 13)....................55
Figure 11: Vignette 18 for Que 2.3 (JPM 7)....................55
Figure 12: Vignette 20 for Que 2.4 (JPM 13)....................56
Figure 13: Vignette 21 for Que 2.4 (JPM 7)....................56
Figure 14: Vignette 23 for Que 2.5 & 2.6 (JPM 13)....................56
Figure 15: Vignette 25 for Que 2.7 (JPM 3)....................57
Figure 16: Vignette 26 for Que 2.7 (JPM 3)....................57
Figure 17: Vignette 27 for Que 2.7 (JPM 14)....................57

LIST OF TABLES

Table 1: Summary of use of MERs with different functions in error remediation 23

Table 2: Marking codes 51

Table 3: Correct/incorrect responses for pre-intervention task.. 58

Table 4: correct/ incorrect responses for post-intervention task 61

CHAPTER 1: BACKGROUND AND CONTEXT OF THE STUDY

1.1 Introduction

This chapter is dedicated to introduction of the study. This is done by briefly revisiting the earlier methods of mathematical instruction as well as integrating the gradual shifts towards the present methods. The following also constitute the current chapter of the report: statement of the research problem, research aims and research questions as well as the significance of the study.

1.2.1 Traditional and contemporary teaching and learning approaches in mathematics.

There have been remarkable changes in the teaching and learning of mathematics. Traditionally, teaching and learning was inspired by the school of behaviourism. English and Halford (1995) suggest that Edward Thorndike was the main proponent of behaviourism, through his connectionism approach. He believed that mathematics should be taught as a habit that has connections with habits already acquired. One behaviourist teaching method which is still famous in mathematics, is drill and practice. In this instance, learning is done mainly through repeatedly following short and specific steps until these are mastered by the learner. This method of teaching and learning is favourable for reinforcement and examination preparation purposes. In other words, it works well once concepts are introduced to the learner. Although he emphasised the content of specific subject matter, Thorndike paid minimal attention to the nature or structure of mathematical thinking (English & Halford, 1995). In other terms, for behaviourists computational skills take the centre stage. This is referred to as relational understanding or rules without reason (Skemp, 1976). A child who promptly gives the answer "12" to the structure "7+5" would not be regarded as having demonstrated knowledge of combination unless he or she understood why 7 plus 5 equals 12.

With time, behaviourism became unpopular and now plays second fiddle to such contemporary learner-centred theoretical perspectives as constructivism. Piaget (1964), widely credited for developing the constructivist perspective asserts that conceptual knowledge cannot be transferred and carefully packaged from one person to another. Learners should take charge in active construction of their own knowledge, with minimum supervision from the teacher. Constructivism has brought a wave of sweeping changes in the way mathematics teaching is planned and executed. New visions for school mathematics as articulated in such influential reports as National Council of Teachers of Mathematics (NCTM) (1980) call for a shift in emphasis from routine skills and factual knowledge to posing and solving a variety of math-related problems. Borasi (1994) argues that reasoning, communicating mathematically and appreciating the value of mathematics are key to mathematics teaching and learning. As a result of this shift in focus, it is now expected that students are given opportunities for solving problems that resemble those encountered in real life settings (NCTM, 1980, 1989). In the Netherlands, the Freudenthal Institute developed the mathematics instruction known as the Realistic Mathematics Education (RME). In this approach, contexts and models are used as resources to assist learners come up with problem solving strategies (schemes) as well as generating big ideas (structuring) in the process of solving mathematical problems (Fosnot, 2005).

In the same vein, NCTM (1991) envisioned a rich mathematics curriculum. The aims of such a curriculum encompasses developing learners' ability to reason, solving problems and communicating mathematically. It apparently discourages the over-use of pencil-paper computations and algorithms. In other terms, mathematics should enable learners to approach every problem situation with an interrogative mind and to exhaust every strategy necessary to arrive at a solution. To achieve this goal, teachers should develop learning environments which stimulate mathematical thought processes as well as promote reasoning and disciplined understanding (National Research Council, 1991).

In the next section, focus is on how the South African Education system has evolved over the years.

1.2.2 A brief history of the South African Education System

During the apartheid era, the South African Education System was differentiated along racial lines. There was a separate education system for the elite, white minority while the black majority had their own, Bantu Education System. The former was superior to the latter, and learners who were privileged enough to acquire it managed to break into lucrative careers while black learners were restricted to low paying and menial jobs. However, the advent of democracy in 1994 meant that changes in the prevailing education systems were inevitable and top on the priority list of the Government of National Unity (GNU). One of most notable changes to the school curriculum brought about by the new, democratic government was the curriculum approach known as Outcome Based Education (OBE). This curriculum approach was implemented in 1998, with one of its major aims being to improve educational outcomes (Chisholm, Motala & Vally, 2003). This approach was the bedrock of the new curriculum, referred to as Curriculum 2005. This curriculum was focused on learner centredness, outcomes determining the process of learning and continuous rather than summative assessment. The objective was to move away from the traditional, rote learning and promote critical thinking (Chisholm et al., 2003). However, the new curriculum faced criticism from a variety of stakeholders. Its main weaknesses were lack of practicability (Chisholm et al., 2003) and vagueness in terms of what had to be taught in different grades (Govender, 2017).

Against a background of severe criticism of curriculum 2005, a curriculum review in 2000 was inevitable. The revised curriculum was subsequently implemented in 2002. In her preamble to the current Curriculum and Assessment Policy Statement (CAPS), Basic Education Minister, Angie Motshegka summed up all the aforementioned curricula revisions:

> In 1997, we introduced the outcomes-based education to overcome the curricular divisions of the past, but the experience of implementation prompted a review in 2000. This led to the first curriculum revision: Revised National Curriculum Statement (RNCS) Grades R-9 and the National Curriculum Statement Grades 10-12 (2000). Ongoing implementation challenges resulted in another review in 2009 and we revised the RNCS (2002) to produce the current CAPS curriculum. From 2012, the two 2002 curricula, for Grades R-9 and Grades 10-12, are combined in a single document and will simply be known as the National Curriculum Statement Grades R-12. The National Curriculum Statement Grades R-12 builds on previous curriculum but also updates it and aims to provide clearer specification of what is to be taught and learnt on term-by-term basis (Department of Basic Education, 2011).

Thus, the journey to development of a comprehensive and quality curriculum document for South Africa has been long and arduous. At this stage, it is worth noting how all these curriculum reviews affected the content as well as pedagogy of mathematics. The following section focuses on mathematics curriculum issues.

1.2.3 Mathematics Education in South Africa

The society regards mathematics as one of the most difficult subjects to learn and teach in South African schools. Much of these difficulties stem from the differing mathematics content taught in the various grades and numerous changes to the maths curriculum since 1994 (Govender, 2017). Besides high failure rates, societal views towards mathematics are cemented by a number of other factors. As alluded to before, educational systems of the apartheid era were aimed at promoting classes in society. In 1994, mathematics was compulsory up to grade 9. In grade 10, learners selected a minimum of 6 subjects, which could include mathematics, even though it was not compulsory. Until 2007, mathematics was done at either standard or higher grade level. The former was for the relatively weak students while the latter channelled learners towards professional courses which required mathematics. It also positioned students from privileged families for employment in such prestigious and challenging fields like engineering and medicine while those from racially prejudiced societies had to settle for low paying and menial jobs. Moreover, due to a critical shortage of mathematics teachers, a significant number of teachers who never did mathematics at school (because mathematics was not compulsory) was hired to give instruction in mathematics. Undoubtedly, these teachers had and still hold serious content gaps which are easily passed on to learners

To redress these challenges, the recently introduced CAPS mathematics curriculum aims at enabling learners to develop skills of collecting, organising and analysing qualitative data to evaluate and critique conclusions (DBE, 2011). The old mathematics syllabus had the following learning outcomes: number, operations and relationships; patterns, functions and relationships; space and shape; measurements as well as data handling. In the current CAPS syllabus, the above learning outcomes have been subdivided into ten content areas. These content areas are assessed in two papers (paper 1 and paper 2).

In its endeavour to promote value in the subject, DBE (2011) argues that mathematics should provide learners with skills to communicate appropriately by using descriptions in words, symbols, diagrams, graphs, and tables. In this regard, at the end of the Grade 11 statistics course, learners are expected to demonstrate the following skills: represent measures of central tendency and dispersion in univariate numerical data; calculate variance and standard deviation of sets of data manually (for small sets of data) and using calculators (for large data sets), represent results graphically (skewness on box and whisker diagrams and frequency polygons) as well as identify outliers.

For South Africa, the vision of having citizens who are mathematically literate and competent has faced a myriad of challenges. At global level, South African learners continue to struggle in mathematics as compared to those from other countries. Results of the Trends in International Mathematics and Science Study (TIMSS) show that in 2015, approximately three in five or 61% of South African grade 5 numeracy learners did not exhibit minimum competency in basic mathematical knowledge (Reddy, Visser, Winnaar, Arends, Juan, Prinsloo, & Isdale, 2016). As a result, South Africa was ranked second from last out of 48 countries drawn from across the globe, who participated in the study. Further, at national level, learners continue to struggle in mathematics with matric pass rates in 2015 and 2016 National Senior Certificate (NSC) examinations standing at 49, 1% and 51, 1% respectively. On average, national pass rates in mathematics have been tumbling since 2013 (DBE, 2016).

This study is mainly focused on one content area which is statistics. The selection is influenced by the unimpressive matric results over the years. Historically, statistics in the South African Mathematics Curriculum for primary school level was reduced to tasks in which learners were given small organised sets of data to represent on a specific graph (Wessels, 2008). Alternatively, mean, median and mode of small data sets had to be determined according to prescribed formulas, without real understanding from the teachers' or learners' side of the appropriateness of these formulas. This resulted in the production of ill-prepared students for statistics at tertiary level and adults who were not statistically literate (North & Zewotir, 2006).

Moreover, as alluded to above, a significant number of mathematics educators currently in the system did not do statistics in their formal training. In her report on 21 universities in South Africa, Wessels (2008) says that most mathematics teachers, especially in the Intermediate and Senior Phase, do not have sufficient statistical content knowledge or pedagogical content knowledge to teach statistics with confidence. The report further says that no information about the level of professional development of in-service teachers in statistics and statistics education in South Africa has yet been released by the Department of Education. In concluding her report, she bemoaned the lack of formal training in statistics by the majority of teachers in an increasingly statistics-driven world. This poses a serious challenge of content gaps on the part of teachers and inevitably, on learners.

Against the above, grim background, successive curricula of RNCS, NCS as well as CAPS, each brought about a broadened data handling content. It was postulated that all these changes would encourage statistical reasoning (Wessels, 2008). Optional Assessment Standards in the old NCS, particularly, in Learning Outcome 4 (Data handling and probability) became compulsory after 2010 (Department of Education, 2007). Thus, it is not surprising that learners continue to perform dismally in statistics.

In spite of the afore-mentioned policy changes and other historical factors, the researcher proposes that the major contributory factors of low attainment levels in mathematics in general and statistics, in particular, are that learners hold numerous misconceptions and make errors at some stage in the process of solving mathematical problems. The DBE (2016)'s annual diagnostic report of National Senior Certificate (NSC) examinations pointed out that errors and misconceptions were part of the factors affecting learner performance in mathematics in South Africa. Recently, the DBE (2016) reported that during the 2016 NSC mathematics examination, most learners demonstrated lack of knowledge of the concept of skewness in the context of a histogram. Furthermore, many candidates failed to draw an ogive because they could not distinguish between frequency and cumulative frequency. Arguably, errors and misconceptions in statistics are due to a number of factors chief of which are: lack of adequate teacher content and pedagogical knowledge in statistics.

Moreover, according to the mathematics annual plan, statistics is the last topic to be taught during term 4. During this time, learner attendance is at its lowest. It is also time for revision in order to prepare for final examinations. Therefore, teachers usually do not have time to adequately prepare and deliver concepts on statistics. This gives rise to errors and misconceptions because during this time, learning is usually by rote.

1.3 Problem statement

Research in statistics has received less attention than other mathematical topics. Most early research-work was carried out by psychologists rather than by statistics educators (Batanero, Godino, Vallecillos, Green & Holmes, 1994). Research done thus far has reported findings on how learners struggle to present and interpret data via such diagrams like histograms, dot plots, box plots and other frequency graphs (Lem, Kempen, Ceulemans, Onghena, Verschaffel, &Van Dooren, 2014; Pierce & Chick, 2013). Studies on both primary and secondary school students have documented difficulties in reasoning about qualitative data when it is presented in the aggregate, as in histograms, line plots and other frequency graphs (Cooper and Shore, 2008). Students have an easy time interpreting simple histograms, but they face difficulties reading the data when the bars contain intervals of values (DelMas, Garfield & Ooms, 2005). From the researcher's experience, most learners struggle to interpret data presented in compound and stacked bar graphs in which more than one set of data is presented. For instance, when data on income distribution across various age groups is presented on a compound or stacked bar graph, most learners struggle to make meaningful comparisons.

Moreover, Lem et al. (2014) reported on the commonly held area misinterpretation of box plots in which most learners think that the larger area in a box plot represents a larger proportion of data. Learners, generally face problems with learning data analysis. At the middle grades, misconceptions tend to focus more broadly on interpretation and selecting appropriate representations of data. At secondary school level, misconceptions have been more prevalent in measures of central tendency and its interpretation (Capraro, Kulm, & Capraro, 2005).

Locally, learners struggle in representing and interpreting data. In bivariate data, candidates cannot distinguish between line of best fit and least squares regression line (DBE, 2016). This misconception arises in the classroom, that is in the way this concept is taught. From the researcher's own experience, most teachers do not impress on the learners that the least squares regression line must pass through the point of reference,$(\bar{x}; \bar{y})$, just like any other straight line in geometry. Rather, they tend to draw line of best fit, which is just an intuitive line. DBE (2016) also reported widespread misconceptions in the concept of skewness in the context of a histogram. Also, most candidates confused the concept of range in statistics with that of range in functions. Most candidates reportedly failed to accurately draw the ogive; instead of plotting against upper class values, they plotted against either middle values (like in frequency polygon) or lower values of intervals. Some candidates could not link the 'number of learners' with the frequencies in the table, while some were under the impression that frequency and cumulative frequency were the same.

From the researcher’s experience in marking National Senior Certificate examinations in South Africa, there are so many challenges faced by candidates in answering questions on spread of data. Annual national diagnostic analyses have over the years reported on commonly recurring errors and misconceptions in graphical representation as well as in interpreting statistical data. Most learners still face difficulties in identifying skewness when data is presented on a histogram. Also, quite a number of candidates are unable to interpret the box and whisker diagram (Department of Basic Education, 2012 & 2016).

Arguably, there is not much literature on errors arising from misconceptions which learners hold in statistics, particularly, in the study of data variability. In this regard, the present study seeks to investigate the nature of errors emerging from misconceptions held by learners in representing and interpreting statistical data variability. Focus will be on three visual displays, namely cumulative frequency curve, frequency polygon as well as box and whisker plot. The rationale for selecting these three diagrams is that they complement each other in representing variability in a given set of data. Whilst it is impossible to read the five number summary on a frequency polygon, one can easily do so on a box plot and a cumulative frequency curve. Moreover, an ogive on its own does not depict skewness unless it is used in conjunction with either a box and whisker plot or frequency polygon.

1.4 Research aims

The aims of this study are:

1. to identify the errors learners make in statistics and the misconceptions behind these errors.
2. to determine the extent to which intervention assists in diminishing the identified errors.

1.5 Research questions

With respect to Grade 11 Statistics, this research study seeks to answer the following questions:

1. What are the common errors emerging from misconceptions held by learners in representation and interpretation of variability using frequency polygon, cumulative frequency curve and box plot?
2. To what extent does intervention assist in diminishing the identified errors?

1.6 Significance of the study

In South Africa, and indeed, throughout the world, use of statistics continue to make huge strides in such fields as health sciences, natural sciences, humanities, legal, agriculture and forestry, technology, marketing and management (Glencross & Binyavanga, 1996). Against this background, nothing much is being done to prepare young people for a statistics-driven world. For instance, analysis of learner performance in assessments shows that statistics remains a challenging topic for most learners both at school as well as tertiary levels. Thus, it is incumbent for teachers, researchers and other stakeholders to work together towards enhancing learner understanding of statistical concepts, and eventually boost achievement in this key mathematical strand. In this regard, it is my submission that by highlighting common errors and misconceptions learners make in statistics, findings of this study will be of use for teachers in lesson planning. Along the same lines, findings of this study could be used by teacher development units and service providers in equipping teachers so that they can be able to handle errors and misconceptions.

Furthermore, remedial teachers stand to benefit from the findings of this study in that errors and misconceptions identified in learners' tasks as well as interviews, will be used as resources for remedial teaching. Besides, findings of this study will be of great use to school administrators, who handle a lot of data almost on a daily basis. Pierce and Chick (2013) argue that school principals and teachers are expected to use statistical information to inform school policy as well as to improve teaching practices and programmes.

Finally, this study seeks to set the stage for future, large scale studies on how errors and misconceptions in statistics develop over time. It is my submission that, unless sources of errors and misconceptions are identified, all efforts to minimise and possibly, eradicate them will be futile.

1.7 Conclusion

In this section, the discussion was centred on background and context of the study. Findings from a number of previous, international research studies on errors and misconceptions were discussed. Moreover, findings from local researchers on how errors and misconceptions hinder learners from recording gains in the study of mathematics were also highlighted in this section of the study. This culminated in the statement of the problem. Research questions as well as aims also form part of this chapter. Finally, rationale for carrying out the study as well as its limitations, concludes the first chapter of the report.

1.8 Definition of terms

Error: a wrong answer due to planning that are systematic in that they are applied regularly in some circumstances (Olivier, 1989);

Misconception: any sort of fallacies, misunderstanding, misuses, or misinterpretations of concepts, provided that they result in a documented systematic pattern of error (Cohen, Smith, Chechile, Burns & Tsai, 1996); a faulty line of thinking that causes a series of errors due to incorrect underlying premise (Nesher, 1987).

Teaching intervention: a set of actions that, when taken, have demonstrated ability to change a fixed educational trajectory (Methe & Riley-Tillman, 2008).

CHAPTER 2: LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK

2.0 Introduction

This chapter begins with a brief history of how mathematics teaching and learning has evolved over time. Afterwards, the notions of error and misconception are defined and exemplified, particularly in the context of the study of statistics. Nature of teacher knowledge and as well as learner mathematical proficiency will also be part of the discussion in the present chapter. In closing the chapter, I propose a conceptual framework for carrying out intervention teaching to rectify identified errors.

2.1 Teaching and learning of mathematics

In the past, mathematics teaching and learning was nothing more than shoving of facts on learners, memorisation of formulae, theorems, axioms and regurgitation of these in assessments. Autonomy in the classroom lay solely with the teacher, while the learner remained in the periphery of the instructional set-up. Many conceptions of mathematical reasoning were confined to formal proofs and other forms of deductive reasoning (Kilpatrick, Swafford, & Findell, 2001). During the twentieth century, drill and practice was the primary focus of mathematics education with Edward Thorndike being the main proponent, with his connectionism or associationism theory (English & Halford, 1995). Through connectionism, instruction that focused on formation of necessary bonds and habits was considered particularly important. (English & Halford, 1995). Thrust was put on competence rather than nature or structure of thinking and learning. As such, competence was used as a yardstick for measuring the extent to which instructional objectives were met.

Since the turn of the twentieth century, there has been a remarkable change to views and approaches to the teaching and learning of mathematics, from the traditional, behaviouristic, teacher-centred methods towards, learner-centred methods. These basically emphasise on structure and meaning. The advent of Gestaltists and Constructivists brought about monumental changes in mathematics education practice and research. One famous gestaltist, Wertheimer, argued that, if learners are adequately guided, they can derive the formulae for finding areas of various plane shapes, instead of blindly following readily made rules

(English & Halford, 1995). Nowadays, the dominant school of thought influencing most curriculum policies and decisions, both locally (CAPS) as well as abroad is constructivism. Confrey (1990) is of the view that a constructivist classroom is a community in which there is negotiated discourse through collaborative group work; students are expected to discuss, collaborate, challenge, reflect, negotiate, and renegotiate meanings. On the other hand, the teacher's role is to firstly acquire an adequate model of students' understanding and to assist them structuring more appropriate and internally consistent understanding. NCTM (1989) highlights the importance of children being actively involved in their learning, that is, they should 'construct, modify and negotiate ideas by interacting with the physical world, materials and other children'. In South Africa, the National Curriculum Statement (NCS) (Grade R-12) is also guided by constructivists' ideas in that one of its aims is to produce learners who are able to identify and solve problems and make decisions using critical and creative thinking (DBE, 2011).

2.2 Errors and misconceptions

Formation of misconceptions and commission of errors are inevitable in the teaching and learning of mathematics. This is probably due to the nature of the subject of mathematics as well as the widely held view that mathematics is the most difficult subject in the school curriculum. It is imperative, at this stage, to clarify the relationship between errors and misconceptions so that readers and future researchers are not confused by the two terms. Essentially, Olivier (1989) clarifies that errors occur due to students' existing conceptual structures; it is these underlying beliefs and principles which are referred to as misconceptions. In other words, errors emerge as a result of misconceptions. Errors, especially those which are systematic are usually a consequence of student misconceptions (Sarwadi & Shahrill, 2014).

Having explained the intricate relationship which exist between errors and misconceptions, the following are definitions of an error and a misconception as well as categorisations of these two notions. Olivier (1989) defines an *error* as a wrong answer due to planning. Cox (1975) defines errors by distinguishing between systematic and random errors. An error is systematic when there is a repeatedly occurring incorrect response that is evident in a specific algorithmic computation. On the other hand, random errors give no evidence of a recurring incorrect process of thinking or recording. Donaldson (1963) categorised errors into structural, arbitrary and executive. *Structural* errors are those errors which arise from some failure to appreciate the relationships involved in the problem or to grasp some principle essential to solution. A learner might use frequency, instead of cumulative frequency, in drawing an ogive. *Arbitrary* errors are those errors in which the subject behave arbitrarily by failing to take into account the constraints laid down in what is given. In this case, learners literary change the question to suit their prior knowledge. *Executive* errors are those errors which involve failure to carry out manipulations or procedures, despite understanding the principles involved (Orton, 1983). Despite learners' understanding of the concept of an outlier of a data set and its identification, they could still fail to locate outliers. This is a clear demonstration of executive error.

The notion of misconception denotes "a line of thinking that causes a series of errors all resulting from an incorrect underlying premise, rather than sporadic, unconnected and non-systematic errors" (Nesher, 1987, p.35). Thus, errors, particularly those which demonstrate some form of a pattern in the way they occur are a manifestation of misconceptions. Misconceptions arise because learners do not come to the class as blank slates (Resnick, 1983) but with theories constructed from their everyday life experiences, no matter how inaccurate such theories might be. It is these incomplete and inaccurate "theories" which may be viewed as misconceptions. Olivier (1989) identified three types of misconceptions. The first one, misconception of overgeneralisation is whereby a learner applies a familiar rule to inappropriate problem situations. Secondly, misconception of interference is whereby previously learnt rules interfere with new skills, for instance, addition schema is constructed and developed prior to the multiplication schema. As a result, learners with misconception of interference add even where they are supposed to multiply. Lastly, misconception of meaning is whereby a learner conceives open expressions like $3x - 1$, as incomplete and requiring further simplification to $2x$ (in most instances). Students become emotionally and intellectually attached to their misconceptions and often find it difficult to accept new concepts which are unfamiliar and dissimilar to their misconceptions (Mohyuddin & Khalil, 2016).

2.3 Errors and misconceptions in statistics

Literature reveals that the majority of learners and teachers alike, have misconceptions in visual representation as well as making sense of statistical data. Statistics education is characterised by deep rooted misconceptions that students often hold (Chance, Ben-Zvi, Garfield, & Medina, 2007). In the following section, I look at some of the commonly appearing errors in data representation and interpretation of data variability, in particular, on box and whisker plots, histograms and bar graphs.

2.3.1 Box and whisker plots

Of all visual displays of data in statistics, the box and whisker plot is one of the most confusing and misunderstood diagrams for both teachers and learners. In their study, Lem et al. (2014) cite a number of challenges learners face with regards to interpretation of box plots. Students tend to interpret area of the box as representing the number or proportion of observations in a certain interval. In other words, most students perceive a longer or larger box as a representative of a higher concentration of data, when in actual fact, the opposite is true. Students fail to interpret parts of the box plot as measures of centre or spread. Instead, they regard the median as a cut point rather than a summary where the data was centred. Bakker, Biehler & Konold (2004) concur that in most statistical diagrams, one can easily read data frequencies, unlike on a box plot, wherein each of the four major components contains roughly 25% of the data. Thus, the absence of frequency on a box and whisker plot takes away frequency from the discourse. Subsequently, this can be a potential source of new misconceptions as well as strengthening of existing ones.

Along the same lines, students interpret the interquartile range as the spread of the middle half of the data rather than as a measure of spread that is a property of the whole data set (shape summary interpretation). DelMas, Garfield and Ooms (2005) argue that box plots remove much of the detail from a set of data to make certain features, like central tendency, variability and skewness, stand out. Understanding how the abstract representation of a "box" can stand for abstract aspect of a data set (a specific, localised portion of its variability) is no small task. A box plot contains much information in a single display; it allows for the comparison of distributions by median, quartiles, minimum, maximum, inter-quartile range and range, all at once. (Bakker et al., 2004). The selection and synthesis of all these various aspects is not an easy task for most learners. Moreover, box plots generally do not allow perceiving individual cases; they operate differently from every other data displays which learners encounter on a daily basis; the median is not as intuitive to students as we once suspected; and quartiles divide data into groups in a way that few students (or even teachers) really understand (Bakker et al., 2004).

2.3.2 Histograms and bar graphs

Apart from box and whisker plots which do not have anything to do with frequency, learners equally struggle with most frequency-related graphs. In most instances, the main challenge is reading the scale as well as distinguishing between data values and the corresponding frequencies. Examples of data displays with high records of misconceptions on frequency are histograms and bar graphs. Delmas et al. (2005) reported on students who could not identify what the vertical scale was measuring. They reported that most students indicated that the vertical scale was measuring score values. In other terms, there appears to be a confusion among students with regards to frequency and data value. Similarly, most students struggle more with grouped data as compared to ungrouped data, especially if the widths of the groups are not uniform. In the same study, Delmas et al. (2005) concluded that most students found it difficult to read frequencies for specified values on a histogram when values are grouped into intervals.

Another common misconception in bar graphs and histograms is to think that variability in the data is represented in a histogram by variability in the heights of the bars (Whitaker & Jacobbe, 2017). In other words, histograms that appear flat are perceived as representing little or no variability at all. In addition, students cannot properly use histograms to find such measures of spread as the median. Cooper & Shore (2008) reported on students who failed to calculate median from histograms, with the majority finding the centre of the x-axis and ignoring the vertical axis altogether.

2.3.3 Measures of location

Of the three measures of location, the mean is probably the most applicable measure in real life situations yet it is the most misunderstood. Errors and misconceptions are even rampant if the mean is weighted. Pollatsek, Konold, Well & Lima (1991) reported on errors found in combining two weighted means as if they were simple means. A typical scenario is when there are ten people in an elevator, four women and six men. The average weight of the women is 90 kilograms and that of men is 80 kilograms. What is the average weight of the ten people in the elevator? In this case, the widely held misconception is to add the two weights and divide the sum by two. Generally, most learners do not recognise situations when weighted mean must be computed.

The concept of mean is not properly developed in most learners. From researcher's own experience in the practice of teaching, learners know how to find mean but cannot explain what it is. If students have only the computational knowledge of the mean, they are likely to make predictable kinds of errors in all but the most transparent problems (Batanero et al., 1994). Experience also reveals that most learners do not have much knowledge on median apart from it being the central value of an arranged data set. In going from the definition of the median as 'middle value of the distribution' to its calculation, there are many steps which are not always sufficiently stated or sufficiently understood (Batanero et al., 1994).

2.3.4 Measures of variability

The concept of variability is prone to different interpretations and promotes development of errors and misconceptions, if not properly taught. For most learners, variability is how much the values differ from each other, rather than from some fixed value such as the mean. Moreover, the words used: variation, dispersion, diversity, spread, fluctuation among others, are open to misinterpretations (Loosen, Lioen, & Lacante, 1985). Misconceptions are stubborn and so they may even extend beyond high school. Mevarech (1983) notes that, even at university level, students have a tendency to assume that the data set and the operation of variance together constitute a group structure.

2.4 Teacher mathematical knowledge

Teacher's mathematical knowledge of content as well as how to teach the content is paramount in minimising and remedying errors and misconceptions. Generally, there is a high prevalence of errors and misconceptions in mathematics. It is, thus, important to equip teachers with ways of identifying and analysing errors so as to take appropriate remedial action. Shepard (2009) is of the view that error analysis requires professional judgement to recognise exactly what learners do not understand, their reasoning behind the error, how that may affect their learning and which instructional practices could provide affordances (or constrain them) to address learner difficulties. With reference to statistics education, Batanero et al. (1994) argue that teachers much increase their knowledge of both subject matter of statistics as well as appropriate ways to teach the subject. Furthermore, this preparation must include knowledge of the difficulties and errors that students experience during the learning of statistics.

Shulman (1986) identified two broad categories of teacher knowledge. These are Pedagogical Content Knowledge (PCK) as well as Subject Matter Knowledge (SMK). He defines PCK as follows:

> It represents the blending of content and pedagogy into an understanding of how particular topics, problems, or issues are organised, represented, and adapted to the diverse interests and abilities of learners, and presented for instruction (p.8).

In other terms, PCK entails knowledge of how to structure and present academic content for teaching, the kinds of conceptions and misconceptions students are likely to encounter in particular classroom settings as well as strategies of how to respond appropriately (Son & Sinclair, 2010). The intersection between pedagogy and content (PCK) (see Figure 1 below) contains within it "the most regularly taught topics, in one's subject area, the most useful forms of representation of those ideas, the most powerful analogies, illustrations, examples, explanations and demonstrations (Shulman, 1986). On the other hand, SMK is the amount and organisation of knowledge in the mind of the teacher. This type of knowledge would be equal to that of a subject specialist. In a nutshell, SMK is necessary in teaching but it is not a sufficient condition because it needs to be blended with pedagogy. In the context of the current study, teacher knowledge of statistical concepts alone is of no benefit to learners. For them to have access to subject matter, suitable teaching methods must be used at all times.

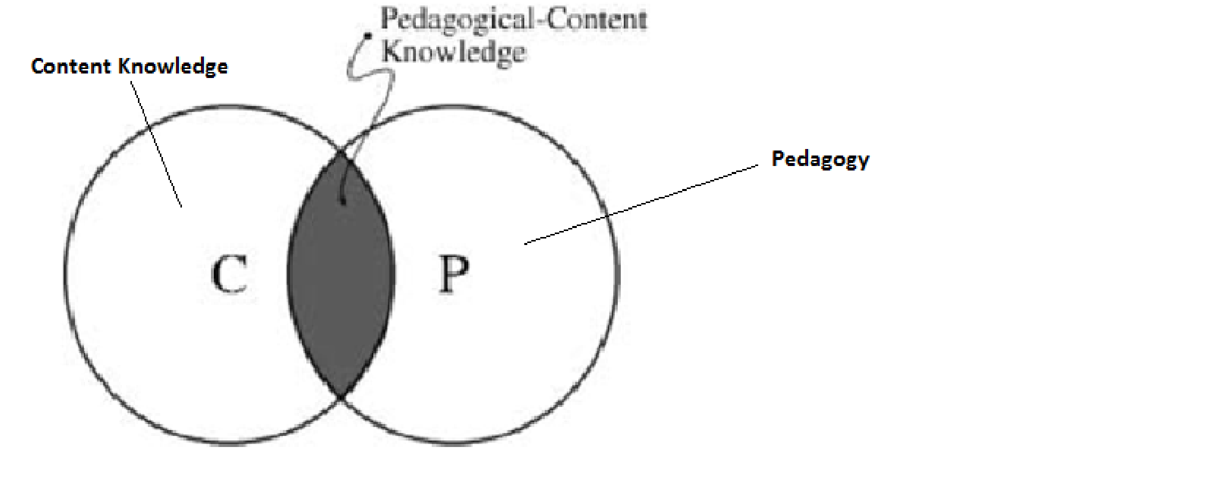


Figure 1: Pedagogical and Content Knowledge *(adapted from Mishra & Koehler, 2006).*

Refining Shulman (1986)'s teacher knowledge categorisation, Ball, Thames & Phelps (2008) came up with the following domains: Common Content Knowledge (CCK), which is the knowledge and skill used in settings other than teaching. This domain includes such attributes as knowing what to teach as well as recognising wrong answers from learners and in textbooks. The second knowledge domain is Specialised Content Knowledge (SCK), which is the mathematical knowledge and skill unique to teaching. It is present in most everyday tasks of teaching and is very useful in recognition of nature of errors (Ball et al., 2008). Whereas teachers' knowledge of what counts as the explanation of the correct answer enables them to spot the error, looking for patterns in learners' errors enables them to interpret learners' solutions and evaluate their plausibility (Shalem, Sapire & Sorto, 2014). CCK and SCK both elaborate on the specialisation of subject matter knowledge (Ball et al., 2008) and are important from the perspective of examining teacher explanation of learner errors (Shalem et al., 2014). Thirdly, Knowledge of Content and Students (KCS) is knowledge that combines knowing about students and knowing about mathematics. When selecting examples to use in class, teachers could use KCS to predict what leaners would find interesting and motivating. On the other hand, when selecting tasks in class, teachers use KCS to anticipate what students are likely to do with it and whether they will find the task easy or difficult. The last knowledge domain is Knowledge of Content and Teaching (KCT). This is a combination of teaching and knowing about mathematics. Sequencing of content in instruction, identifying methods and procedures of instruction and evaluation of instruction are all functions of KCT. KCS and KCT elaborate on specialisation involved in the teaching of mathematics from the perspective of learners, curriculum and pedagogy (Ball et al., 2008). These two also elaborate on Shulman (1986)'s PCK (Shalem et al., 2014).

Teacher knowledge, in all its forms, gathered through several years of hard work at tertiary institutions as well as in-service training (INSET) has a special role to play in the teaching and learning of mathematics. It is through the way teachers impart their knowledge to learners that the latter draw inspiration and develop the love for mathematics which is key to excellence in learning mathematical content. "The teacher of mathematics plays a critical role in encouraging students to maintain positive attitudes towards mathematics" (Kilpatrick, Swafford & Findell, 2001, p.132).

Furthermore, teachers' ability to identify and address learner errors and misconceptions, to a great extent, depends upon how they harness and harmonise various knowledge domains. The tasks that teachers engage with in errors analysis, such as sizing up the errors or interpreting the source of its production are possible because of the mathematical reasoning which teachers derive from the knowledge domains (Salem et al., 2014). While student errors used to be perceived as stumbling blocks to effective teaching and learning, nowadays, errors and misconceptions are seen as vital evidence of thinking on the part of learner. Makonye and Nzima (2016) argue that errors and misconceptions have to be welcome and be used as a starting point in teaching and learning to help learners construct and restructure their knowledge. In other terms teachers must embrace errors and use them as catalysts in their day to day practices. NCTM (2000) submits that recognising and responding to students' errors appropriately is one of the main tasks of teachers in the teaching of mathematics. In other terms, how teachers respond to learner errors depends upon the nature of errors at hand.

Moreover, when analysing learners' tasks, teachers should apply their knowledge and skills to move away from the traditional, superficial "right or wrong" form of analysis. Ball et al. (2008) reckon teachers should use errors and misconceptions as windows into student understanding, endeavouring to help student understanding of the conceptual basis of their errors. Thus, without the afore-mentioned domains of knowledge, teachers are incapacitated and might be unable to engage and effectively deal with some of the common sources of mathematical errors.

2.5 Learner mathematical proficiency

Essentially, every mathematical idea can be understood at many levels and facets. Failure to understand at any of these facets usually results in the formation of errors and misconceptions. Thus, understanding is key to effective and meaningful learning of mathematics. Learning with understanding is more powerful than simply memorising because the organisation improves retention, promotes fluency, and facilitates learning related material (Kilpatrick et al., 2001).

In their framework, Kilpatrick et al. (2001) are of the view that proficiency in mathematics could be used to capture what is believed to be necessary for anyone to learn mathematics successfully. They categorise mathematical proficiency into five, interwoven strands, namely: procedural fluency, conceptual understanding, adaptive reasoning, strategic competency, as well as productive disposition. Conceptual understanding refers to "comprehension of mathematical concepts, operations, and relations" while procedural fluency refers to "skill in carrying out procedures flexibly, accurately, and appropriately'' (Kilpatrick et al., 2001, p.116). On the other hand, adaptive reasoning refers to the capacity to think logically about the relationships among concepts and situations. Task items which measure learners' adaptive reasoning usually require them to justify and explain their solutions. Thus, adaptive reasoning is closely related to conceptual understanding and procedural fluency, without which, learners would not be able to justify their solution strategies. In the context of the present study, in order to find the number of learners whose scores were within one-standard deviation from the mean, a solution strategy could require procedural fluency in carrying out computations of mean and standard deviation on a calculator. At the same time, adaptive reasoning enables the learner to justify and explain their solution.

Fourthly, strategic competency refers to the ability to formulate mathematical problems, represent them, and subsequently solve them. Lastly, productive disposition is the tendency to see sense in mathematics, to perceive it both as useful and worthwhile, to believe that steady effort in learning mathematics pays off, and to see oneself as an effective learner and doer of mathematics (Kilpatrick et al., 2001). As alluded to before, these strands are interwoven, affecting each other in particular, procedural and conceptual understanding are always working concurrently. As a child gains conceptual understanding, computational procedures are remembered better and used more flexibly to solve new problems. In turn, as procedures become more automatic, the child is enabled to think about other aspects of a problem and to tackle new kinds of problems, which leads to new understanding (Kilpatrick et al., 2001).

In the following section, I present the conceptual framework for this study.

2.6 Conceptual framework

Conceptual framework for the present study focuses solely on intervention teaching aimed at identified errors. Thus, the main concepts of the conceptual framework are intervention teaching and functions of Multiple External Representations (MERs) as well as their application to error remediation. These are explained in detail in the following section.

2.6.1 Intervention teaching

There are various definitions of intervention teaching. In this study, intervention teaching refers to "a set of actions that, when taken, have demonstrated ability to change a fixed eduactional trajectory" (Methe & Riley-Tillman, 2008, p. 37). Thus, intervention teaching can be regarded as a specific program or set of steps aimed at assisting a child to improve in an area of need. If the teacher realises that upon completion of instruction, learner errors still exist, strategies should be put in place so as to minimise these errors and misconceptions. Intervention is arguably, one of the strategies which the teacher must consider so as to help learners who are in need. Intervention should always form part of planning for every classroom practitioner because there are always those learners who need re-teaching for them to grasp mathematical concepts. Ashlock (2002) notes that when students make errors and misconceptions, teachers should recognise the errors, prescribe an appropriate instructional focus, and implement an effective and efficient reteaching plan. This implies that for it to be effective, an intervention program should be intentional, targeting a specific area of weakness. It must also be structured, running for a certain period of time and subjected to review as informed by progress made. During intervention, the teacher should try as much as possible to involve learners because it is their last opportunity to understand problematic concepts. Allowing the learner to inspect the resolution path at each step is one of the key features of remediation (Leite, Marczal, Pimentel & Direne, 2014).

2.6.2 Multiple External Representations (MERs)

A new element that has been introduced in the error remediation process is external representations. According to Leite et al. (2014), these can be understood as constant steps in a process of taking possession of stable states to obtain information that can then be used in a more flexible way for other purposes. Leite et al. (2014) are in support of a study that links external representations to the errors committed by learners. They further argue that such a study is possible after elaborating a broad error classification that allows for a more precise remediation, since this would enable error mapping. In this study, I propose a remedial program framed by Multiple External Representations (MERs) theory. According to Ainsworth (2006), MERs is a cognitivist theory that supports the use of techniques to present, organise and transmit knowledge. MERs are catergorised according to the function they perform as follows: complementary, interpretation-constraining and deeper-understanding construction. Complementary function of MERs provides support or complement a cognitive process. In simple terms, one external representation supplies information in a learnng environment that is lacking in other external representation. Construction of deeper understanding means that by using MERs, different elements of learning material are more easily connected, providing abstraction of learning material. Lastly, interpretation-constraining function constrain possible interpretations that are not relevant to certain concepts. The use of the latter function means that the misinterpretation of one representation is mitigated by use of another. In order to allow for a more precise remediation and error mapping, I first elaborate on the error classification to be used through out this study.

2.6.3 Benefits of MERs

There are benefits which can be drived from using MERs in providing remediation to learners. Ainsworth (2008) argues that use of MERs in learning environments have shown that students can benefit from the properties of each representation and ultimately achieve a broader understanding of the subject being taught. They also allow the learner to review facts, rules, concepts and/ or fragments that have been forgotten. In this study, by over-laying box plots onto histograms (see lesson plan in APPENDIX 11), it was visible that the same number of data points is represented in each of the four parts of a box plot and that a wider part implies that the data are less densely populated in that interval.

In cultural development, Vygotsky (1978) argues that a child undergoes several stages in the use of signs. At first, the child relies upon external signs and as he or she grows, the whole operation of mediated activity (for example memorising) begins to take place as a purely internal process. In the late stages, the child naturally begins to memorise more and better than ever. At its highest level, the child stops relying upon external signs in memorising. This process of internal reconstruction of external operations is referred to as internalisation. Thus, in the context of the present study, it was envisaged that use of a variety of diagrams in a single representation (MERs) would go a long way in assisting learners to visualise and interpret the seemingly abstract concept of data variability. With time and through practice, learners would not need multiple representations to make meaningful interpretations in the future because they shall have internalised the concept of data variability.

2.6.4 Application of MERs in error remediation

Table 1 below summarises how the above stated three functions of MERs can be matched with and applied to error remediation. Information presented in this table was used to draw a conceptual framework for carrying out error analysis for purposes of remediation.

Error type	**MER function**	**Remediation**
Procedural	Interpretation constraining	Showing that although a solution strategy may be right, the deficiency lies in the use of information.
	Deeper-understanding construction	Explaining to the learner that his or her mistake was choosing the wrong operator, like adding instead of subtracting
Conceptual	Complementary	Proposing other ways of presenting a problem, so that learners get a mathematical formula instead.
	Deeper-understanding construction	Introducing a rule or theorem, in order to allow the learner to reorganise or generalise the concept.
	Interpretation constraining	Demonstrating to the learner that the adopted logic does not lead to the solution of the problem.

Table 1: Summary of use of MERs with different functions in error remediation *(adapted from Liete et al., 2014)*

Based on the table above, a conceptual framework is proposed on how errors identified could be addressed through the use of a teaching intervention informed by MERs functions identified above. Leite et al. (2014) are of the view that use of error remediation based on

MERs requires a conceptual framework that helps identify the errors committed by learners, classify these into appropriate categories associated with a MER function, and finally offer appropriate and adequate support to the learner. The conceptual framework and its description are presented below.

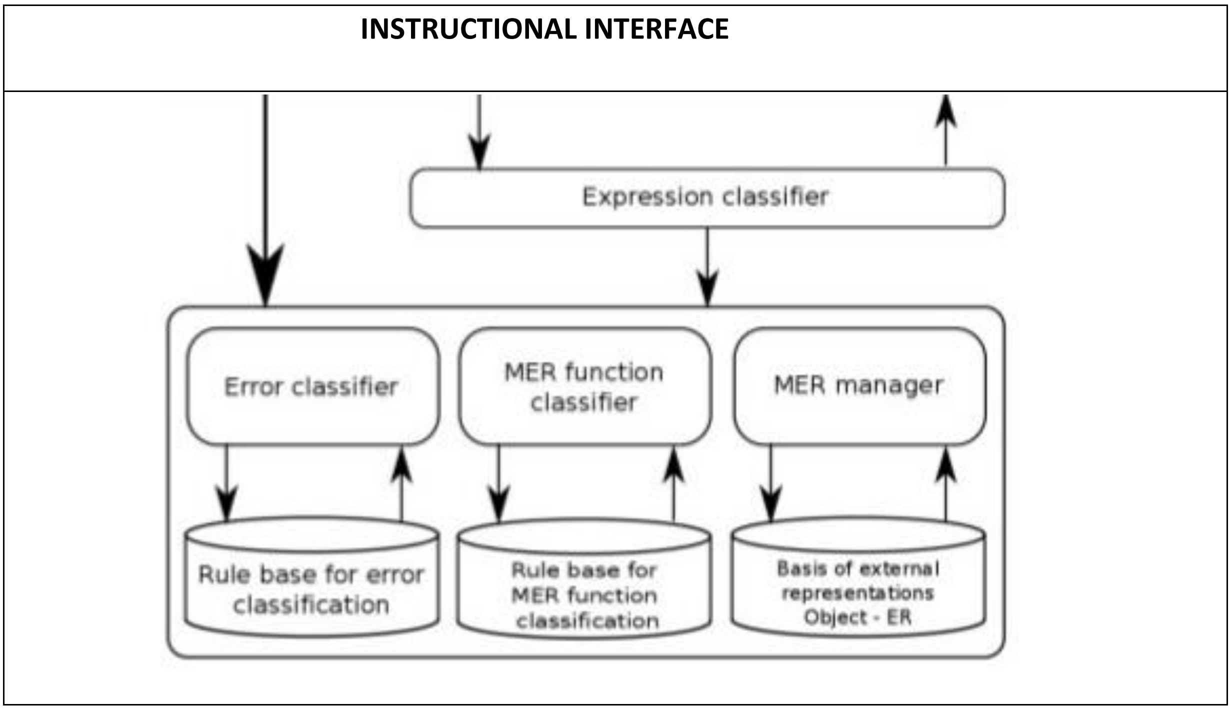


Figure 2: Conceptual framework for using MERs in error remedial instruction *(adapted from Leite et al., 2014).*

Figure 2 above represents the proposed conceptual framework for remediation based on MERs as well as its modules namely, error classifier, MERs classifier and MERs manager. Other components of the framework are the rule for error classification, rule base for MER function as well as the basis for external representations.

Basically, instructional interface represents the learner's current step along the resolution path. It enables the error to be remediated before the final answer is reached. This also determines the type of MER to be employed so as to rectify the identified error. According to Leite et al. (2014), the aim of the expression classifier is to establish a link between the object of learning (that is, representing and interpreting variability) and the system. This module is responsible for the initial communication and receives data to determine the degree of correctness of the learner's answer. If the answer is correct, the other modules will not be launched, otherwise the error classifier module will be activated.

Error classifier module categorises errors made by learners, for which purpose the incorrect data or expression is identified by the expression identifier module. This module receives an error from the expression classifier module and classifies it using the rules contained in the rule base for error classification, which includes the error classification presented in this study (conceptual and procedural errors).

The purpose of the rule base for error classification is to lay the basis for classifying the error detected (Leite et al., 2014). The action and number of attempts are stored as individualised remediation elements for the learner so as to track him/ her during the error resolution process. As alluded to above, the action to be taken depends on the learner's current resolution step. In addition, the number of attempts assist in determining the adequacy of the MER used for remediation.

Leite et al. (2014) posit that MER manager is probably the most relevant of all the modules in that it determines the type of remediation needed for the learner to improve his or her problem solving strategy. It receives the following inputs: error, action, number of attempts and MER function. In this module, it is not necessary to continue storing the error type as it has already been used for MER classification. The role of the MER manager is to decide which MER to use to remediate error once it has been classified.

Lastly, the MER manager has a sub-module Object-ER, which is responsible for searching for external representations according to basic criteria, such as error persistence, success of certain external representations in previous situations, and the degree of complexity of the situation faced by the learner. In essence, this is a crucial point in that the teacher has to demonstrate his or her PCK as well as SMK based on previous experience in teaching and learning of variability through the use of the three representations of ogive, box and whisker diagram as well as frequency polygons. If inappropriate MERs are selected, then this might actually cement any existing misconceptions.

Details of how the conceptual framework was implemented in the intervention lesson are summarised towards the end of Chapter 4.

2.7 Conclusion

In the opening stages of this chapter, a short history of teaching and learning mathematics was presented. The notions of error and misconception were discussed, particularly in the context of teaching and learning of statistical concepts of measures of location and variability. Teacher knowledge domains as well as learner mathematical proficiency were also highlighted. In the closing section, the conceptual framework for the study was presented and described. In the next chapter, methods of gathering data for this study form the object of discussion.

CHAPTER 3: THEORETICAL FRAMEWORK

3.0 Introduction

In this chapter, broad ideas of Piaget's constructivism as well as Vygotsky's sociocultural theoretical perspectives are discussed in detail. Apart from these, stages of concept formation will be discussed according to a special constructivist theory of Action, Process, Object and Schema (APOS). This theory will be used to draw an analytical tool to analyse data for this study.

3.1 Sociocultural perspective

In recognition of the growing influence of cultural and social factors towards development of consciousness in human beings, sociocultural theory has lately received a lot of attention from educational practitioners. Russian psychologist, Levy Vygotsky is widely recognised as the founder of this theoretical perspective. In the following sections, the writer discusses some of key ideas on how concepts are formed in children.

3.1.1 Sign, tool and mediated activity.

Vygotsky (1987) argues that tools and signs are indispensable in human cultural development. Signs are used to remember, compare things, selecting and reporting among others. They act in a way analogous to the role of tools in labour. All higher order mental processes, such as reasoning and problem solving, are mediated (accomplished through and with the help of) by psychological tools such as language, signs, and symbols (Woolfolk, 2010). From Vygotsky (1978)'s psychological point of view, tools and signs complement each other in constituting the mediated (indirect) activity as shown in the Figure 3 below.

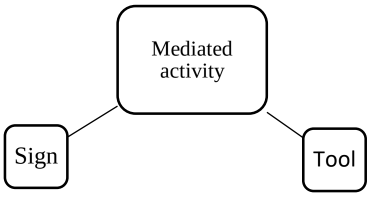


Figure 3: Mediated activity *(adapted from Vygotsky, 1978)*

The role of signs and symbols, such as human speech, written language and algebraic and mathematical symbols is to serve as carriers of both meaning and sociocultural patterns. Vygotsky (1978) argues that the tool is externally oriented and serves as the conductor of human influence on the object of activity. It is a cultural artefact which is useful in controlling human behaviour from outside. In other words, it is the means by which man masters and triumphs over nature. Psychological tools (symbol systems) are devices for mastering mental processes and these include: language, various systems for counting, mnemonic techniques, algebraic symbol systems, works of art, writing, schemes, diagrams, maps, graphs, and mechanical drawings. They allow people in society to communicate, think, solve problems and create knowledge (Woolfolk, 2010). In statistics, usage and knowledge of such psychological tools like number systems have a far reaching impact on child development as well as in the society in general. For example, the extension of the number system to irrational and complex numbers led to rapid advancements in scientific discoveries, engineering and astronomy among other disciplines. "Language is the most important symbol system in the cultural tool kit as it is used to fill up the tool kit with other tools" (Woolfolk, 2010, p.44).

Vygotsky (1978) argues that signs are aimed at mastering oneself, internally oriented and change nothing in the object of activity. It is not the tools or signs, in and of themselves, which are important for thought development but the meaning encoded in them. For instance, language as one of the most vital tools in thought development, is indispensable in that it assists in conveying messages and interpretation thereof, among individuals.

In its cultural development, a child undergoes several stages in the use of signs. Initially, the child relies upon external signs and as he or she grows, the whole operation of mediated activity (for example memorising) begins to take place as a purely internal process. In the late stages, the child naturally begins to memorise more and better than ever. At its highest level, the child stops relying upon external signs in memorising. This process of internal reconstruction of external operations is referred to as internalisation (Vygotsky, 1978).

It is crucial to note that concepts and language develop in children in the same, afore-mentioned process of changes in sign operations. Moreover, Vygotsky (1978) opines that every function in the child's cultural development appears twice. Firstly, on the social level, and later, on individual level. This pattern of development justifies the use of teaching aids drawn from learners' immediate environment, which would be discontinued once the learner has reached the formal operation stage. Alternatively, cultural development first occurs between two or more people (inter-psychological) and then inside the child (intra-psychological). This is the case in the development of logical memory as well as formation of concepts. He referred to this primacy of social world over the individual as the general genetic law of cultural development. Society is the bearer of cultural heritage, without which the development of the individual mind is highly impossible.

Cultural-historical theorists, Cole and Wertsch (1996) define artefacts as objectifications of human needs and intentions already invested with cognitive and affective content. They distinguished between three hierarchical levels of the notion of artefacts. Primary artefacts are those such as needles, clubs, bowls, which are used to make things (reality). Secondary artefacts are representations of primary artefacts and of modes of action using primary artefacts (beliefs or traditions). Lastly, the highest level in the artefacts classifications is that of tertiary artefacts (imagined worlds). It is these artefacts which are indispensable in the mediation of both social and individual processes in the child's cultural development. Furthermore, these cultural artefacts are products of the mediated activity.

3.1.2 The controversy of learning and development

There is a lot of debate with regards to which precedes the other between learning and development. Instruction in the form of collaboration and co-operation, is the major educational driving force of development (Vygotsky, 1987). Learning takes place long before the child attends school; any learning a child encounters in school always has a previous history. For example, children begin the study of arithmetic in school, yet by then, they already know how to add, subtract and share objects amongst themselves. Also, before formal schooling starts, children already possess a lot of skills through imitating adults as well as through instruction on how to perform certain tasks (Vygotsky, 1978). In a family setup, boys usually model their fathers while girls derive satisfaction from complementing their mothers by performing those tasks which their mothers perform. Human activities take place in cultural settings and cannot be understood apart from these settings. Our mental structures and processes can be traced to our interactions with others, which, in turn, creates our cognitive structures and thinking processes (Woolfolk, 2010).

Piaget (1964) opines that learning is subordinated to development and not vice-versa while Vygotsky (1978) argues that it is an active process that does not have to wait for readiness. In other words, properly organised learning results in mental development and sets in motion a variety of developmental processes that would be impossible apart from learning. Learning pulls development to higher levels and other people, for instance teachers, play an important role towards achievement of cognitive development (Woolfolk, 2010). The only good kind of instruction is that which marches ahead of development and leads it (Karpov, 2003). In other words, for learning to be effective in generating the needed development, it should be carefully planned and executed.

3.1.3 Spontaneous and scientific concepts

Vygotsky (1978), classifies concepts as either scientific or spontaneous (everyday). The former is introduced to a child by the teacher (theoretical learning) while the latter are those acquired outside contexts in which explicit instruction is in place. In other words, spontaneous concepts arise through the child's social interaction with elder members of the society, including immediate family and friends (empirical learning). Scientific concepts are highly coherent and hierarchical, coupled with a high degree of generality. Essentially, presence of spontaneous concepts in the child is necessary in the subsequent understanding of higher order, scientific concepts. Karpov (2003) argues, "Despite their 'unscientific' nature, spontaneous concepts play an important role in children's learning as a foundation for the acquisition of scientific concepts." Hence, there is a need to incorporate spontaneous concepts in day to day teaching so that scientific concepts, especially those which are mainly abstract, are concretised as far as possible to enhance understanding of content by learners. Along the same lines, Vygotsky (1987) cautions that teachers should guard against direct instruction because in so doing, the child only absorbs words (verbalism) through memory rather than thought. Instruction should be designed to foster conscious awareness of conceptual form and structure and thus, allow for individual access and control over acquired scientific concepts. Effective teaching is characterised by co-operation and collaboration.

3.1.4 Zone of Proximal Development

Vygotsky (1978) says that the transition from the process of learning to that of cultural development results in the formation of Zone of Proximal Development (ZPD). This is the area between the child's current development level as determined by independent problem solving and the level of development that the child could achieve through adult guidance or in collaboration with more capable peers (Woolfolk, 2010). ZPD can also be viewed as the 'magic middle,' somewhere between what the student already knows and what the student is not ready to learn (Berger, 2006). In simple terms, this is a point where a child is about to find a solution to a problem but not without provision of some clue, reminder, structure, help and encouragement. In the ZPD region, effective learning takes place mainly through the twin-instructional processes of collaboration and co-operation (Vygotsky, 1978). The students learn by imitating a more knowledgeable other (MKO), who must be more competent, otherwise he or she might fail to discern the learners' ZPD and development of misconceptions may ensue. The teacher or MKO mediates to the learner society's treasured knowledge through social interaction or formal learning with the aid of the language tool (Makonye & Luneta, 2011)

Vygotsky (1978) further argues that the interaction with the members of a particular culture gives rise to higher forms of mental development. Through interaction with teachers, senior and former students of mathematics (MKOs), learners stand to derive a higher understanding of mathematics. Essentially, through both PCK and SMK, teachers must locate learners' ZPDs so as to avoid assigning too difficult tasks to learners because this may negatively affect the learning environment. Assigning a learner in Grade 8 to find mean age of a group of, say, 30 learners in a particular grade might be a difficult task because at this stage most learners still use manual method of finding mean. On the other hand, this is an easy task for a Grade 11 learner who can comfortably make use of the statistical mode of a calculator to compute the mean. Be very cautious and avoid giving learners either too simple or too difficult tasks. Failure to correctly identify learners' ZPDs by assigning inappropriate tasks leads to formation of errors and misconceptions because learning will be mainly through rote.

During the process of solving problems, learners need to be scaffolded which is a process of assisting a learner to master a concept or skill within his or her reach (Vygotsky, 1978). This enables the learner to complete assigned tasks within a reasonable time period and avoid frustration. In this way, they will be motivated to perform assigned tasks in the future. Learners' ZPDs must be accurately and appropriately determined to identify the prior knowledge that the learner could use.

3.2 Constructivism

Since the turn of the last century, mathematics education has undergone a major shift away from the traditional, transmission model of instruction, towards constructivism, the notion that students must construct their own concepts (Schmittau, 2003). Autonomy in the classroom moved from the teacher to the learner while the role of the teacher was relegated to that of a facilitator. Constructivism evolved as researchers' interest in a child's reasoning went beyond a simple diagnostic view of errors to understanding the richness of student strategy and approach (Confrey & Kazak, 2006). In other terms, constructivists view learners not as empty vessels, but rather as endowed with concepts on which teachers can plan new instruction.

One of the major proponents of constructivism, Piaget, posits that all learning involves the interpretation of phenomena, situations, and events, including classroom instruction, through the perspective of the learners' existing knowledge (Smith, Disessa, & Roschelle, 1993). In other words, through own experiences, learners actively build new knowledge in the learning process. White and Gunstone (1992) indicate that in order to understand a concept, some information about that concept must be in existence in some way or another in the experience of the participant. In primary school, learners are taught how to group items according to a given criterion like use of a pictogram to show scholar transport preferences. Building on this knowledge, and at a later stage, learners are taught how to display data such as the one mode of transport (above), on bar graphs or pie charts. According to Piaget (1964), learners are actively engaged in a continuous process of adaptation of information in their mental structures (schema). He further argues that the information is continually reorganised in learners' mental structures through assimilation, accommodation and equilibration. During assimilation, new ideas are fitted to what a child already knows (existing schema) while accommodation involves the restructuring of existing schemata to fit new information. However, if a learner is confronted with information that does not deal with the current mental structure, it then gives rise to cognitive conflict (Piaget, 1964). Cognitive conflict compels the learner to modify the entire, existing mental structure through assimilation and accommodation until a balance is reached, that is, equilibration (Piaget, 1964). Thus, the process of equilibration is a major driving force behind knowledge construction as well as opening up avenues for effective learning of mathematics concepts.

Basically, there are four, constructivist-related positions with regards to mathematics education namely: information processing theory, weak constructivism, radical constructivism, and social constructivism. These are briefly discussed in the following section.

3.2.1 Information processing theory

It is largely based on the metaphor and sometimes the conscious model of the mind as a computer. This actively processes information and data, calling up various routines and procedures, organising memorization and retrieval of data (Ernest, 1996). In other words, the human mind is taken as a computer which is given a command or instruction, processes it and then performs the task. This theory represents a shift from the traditional empiricist metaphor of mind as passive to a complex mechanical (or rather electronic) metaphor of mind-as-computer. Proponents of this version of constructivism posit that true knowledge of the state of affairs in the world may be possible, as it is a certainty in mathematical knowledge (Ernest, 1996). Ashlock (1982) argues that an important outcome of this perspective in terms of learning theory (and pedagogy) is that it accounts for student 'error patterns' in mathematics, which is the crux of the present study. In the next section, I turn my focus to another constructivist position of mathematics education, namely, weak constructivism.

3.2.2 Weak constructivism

This form of constructivism, sometimes referred to as trivial constructivism, is premised on the idea that knowledge is not passively received but actively built up by the cognizing subject (Ernest, 1996). Although weak by nature, this form of constructivism represents a major shift away from the traditional, classical behaviourist, transmission model of teaching. Moreover, knowing is active, that it is individual and personal, and that it is based on previously constructed knowledge. All human knowledge is constructed by each individual and also knowing involves active mental processing and learning is more interactive, involving the selection, processing and assimilation of information accordingly. Knowledge is constructed to match the world, or the eternal verities of mathematics, and not as a recursive construction based on previous constructions, satisfying inner constraints (Ernest, 1996). On one side, weak constructivism emphasises that all individual knowledge is constructed while on the other, there is a realm of objective knowledge, that is, mathematical truths and facts about the world (Ernest, 1996). It is probably due to these two, somewhat opposing epistemological aspects in one paradigm that this form of constructivism is considered weak. Without doubt, these two views cannot co-exist. Metaphorically, weak constructivism views the mind as a "soft computer"

3.2.3 Radical constructivism

From Piaget's original ideas on constructivism, von Glasersfeld, derived radical constructivism (Ernest, 1996). It is premised on two principles namely; knowledge is not passively received but actively built up by the cognising subject. Also, the function of cognition is adaptive and serves the organisation of the experiential world, not the ontological reality. In other words, learners are not empty vessels, but they always have some concepts which they bring into the learning environment. It is the role of the teacher to discover this prior knowledge and plan instruction accordingly. Von Glasersfeld (1989) argues that learning occurs as a result of the interaction between the child's existing ideas and new ideas. In this view, ideas are interpreted, organised and understood in the light of children own current knowledge. Followers of radical constructivism argue that its epistemology is purely fallibilist, sceptical and anti-objectivist (Ernest, 1996), such that there is no ultimate true knowledge about the state of affairs in the world or in the realm of mathematics.

3.2.4 Social constructivism

According to Ernest (1996), social constructivism regards individual subjects and the realm of the social as indissolubly interconnected. Human subjects are formed through their interactions with each other (as well as by their individual processes. The social constructivist model of the world is that of a socially constructed world which creates and is constrained by the shared experience of underlying physical reality. Social constructivism awareness of the social construction of knowledge suggests a greater pedagogical emphasis on discussion, collaboration, negotiation and shared meanings.

One common metaphor for the four varieties of constructivism discussed above is that of carpentry, which is about the building up of structures from pre-existing pieces, possibly specially shaped for the task. Thus, teachers, can draw from this common metaphor in adopting constructivism and infusing it into their instructional planning and implementation. Constructivism is indispensable in that it assists educational practitioners to uncover structure and meaning in students' responses, which is key in dealing with learner errors and misconceptions (Smith et al., 1993). Subsequently, through knowledge of pre-existing concepts, teachers can plan their instructional programmes accordingly, navigating with their learners from the known to the unknown.

Besides the theoretical views discussed thus far, this study draws from constructivist views of Action, Process, Object and Schema (APOS) (Dubinsky, 1991; Arnon, Cottrill, Dubinsky, Oktac, Fuentes, and Trigueros, 2014).

3.3 APOS theory

Basically, this theory explains how knowledge and mathematical structures are formed in learners' minds. It comprises four stages in the development of mathematical concepts. The major tenet of this theory is that individuals tend to deal with perceived mathematical problems by constructing mental actions, processes, objects and organising these into schemas to make sense of the situations and solve problems (Dubinsky, 1991). It is in the process of knowledge construction that faulty lines of thinking emerge, which may eventually escalate into misconceptions. Each of the four stages of the APOS theory is explained in the following sections.

3.3.1 Action

An action is a physical or mental transformation of a mathematical entity in response to outside stimuli. Actions may require initiation or mediation by a teacher or peer to direct steps that are explicitly taken towards a goal. It is the beginning stage of learning and making sense of a mathematical situation. Thus, an action conception leads to an operation external to a learner's mind. In the context of statistics, action can be promoted, by representing learners in a class according to the mode of transport they use to come to school. In this instance, learners can take turns in saying out which mode of transport they use, while one learner is recording through use of tallies. Data obtained can then be used to draw up frequency tables and other statistical diagrams, at later stages of concept development. In another scenario of demonstrating the idea of median age in a class of say, 25 learners, the teacher can instruct the class to form a queue based on the chronological order of their ages. In this case, the median age is that of the learner who occupies the central position of the queue. Thus, external stimuli are used to represent the mathematical entity of median.

3.3.2 Process

When an individual reflectively repeats an action they may interiorise it into a process. Interiorisation occurs when someone can carry out an action mentally, without the need for an external stimuli. When someone reflects upon an action without actually engaging with it, they are said to have interiorised that action into a mental process (Aineamani, 2015 ; Cottril et al., 1996). A concept is viewed as a process or procedure semi-external to the learner and it invokes the metaphor of a verb. A process leads to an operation internal to a learner's mind. A learner with a process conception of the median can be given a cumulative frequency curve and asked to find the median. The learner does not see the central data point like the previously mentioned learner because he or she has internalised the concept of median. In this case, the learner makes use of the median position, $\frac{n+1}{2}$, and to interpolate it from the curve.

3.3.3 Object

When a mathematical entity is 'seen as an object', it is seen as if it were "a real thing that exists in space and time" (Sfard, 1991, p.4). A mathematical object is structural and invokes the metaphor of a noun. In real life, percentage and average pass rates of learners in a school leaving examination, median weight of the growth of a baby and extreme ages of learners (very young or very old) in a mathematics class are mathematical objects in the sense of APOS theory. Essentially, an object is obtained from understanding the process in totality and realising that transformations can act on it. Cottrill, Dubinsky, Nichlos, Schwingenndorf, Thomas and Vidakovic (1996, p.4) posit, ‘An object is constructed through the encapsulation of a process…[It] is achieved when the individual becomes aware of the totality of the process, realises that transformations can act on it’. Thus, when a process becomes an object, it becomes a thing, it achieves permanence, it becomes an entity in its own right and it has become a noun (Davids, 1984). This transformation is referred to as the encapsulation of a process into an object (Dubinsky, 1991). This reification (or encapsulation) enables the learner to view a familiar mathematical expression in a totally new light (Ainemani, 2015). In the context of this study, a learner with object conception of measures of central tendency is able to select the most appropriate measure in any given data distribution. For example, at a company where the highest paid executive earns R750000 annually while the least paid shop floor worker earns R64 000 annually, a learner who has objectified the mean is able to tell that it is not the best measure of location in this case. This can potentially mislead data consumers. In this situation, the median could be the most suitable measure of location to use.

3.3.4 Schema

When actions, processes and objects are revisited and a learner has a bird’s eye view of them, they form a schema of mathematical entity. This schema is organically linked to other schemas. Analysis and reflection on a schema can generate yet another cycle of actions and processes so that new, more advanced mathematical objects and schemas can be formed. A typical example of action conception of function is when an individual is limited to thinking about formulas involving letters which can be manipulated or replaced by numbers and with which, calculation can be done. We think of this notion as preceding a process conception, in which a function is thought of as an input-output machine. In the beginning, an individual is restricted to certain kinds of formulae, then reflects on calculations (thinking about a process). Afterwards, he or she may go back to an action interpretation, perhaps with more sophisticated formulae, thereby developing further process conceptions. Dubinsky (1991) explains that, in reality, when someone is developing a mathematical concept, it does not necessarily happen in a linear sequence. Rather it is dialectical as shown in Figure 4 below.

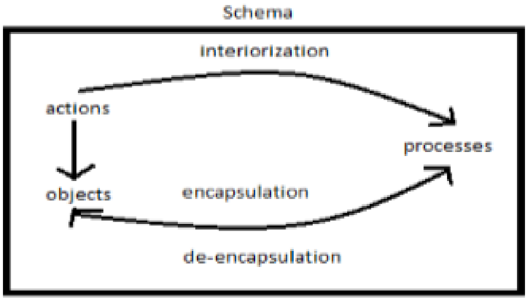


Figure 4*:* APOS Theory *(adapted from Asiala, Brown, DeVries, Dubinsky, Mathews, & Thomas, 1997).*

An example of application of APOS theory is the progression of a learner in understanding regression in statistics. An action conception of regression would be illustrated when the learner requires some external rule , say $\hat{y} = 0.736 + 0.912x$, and is able to find the y values by substituting x values into the rule. As they interiorise these actions, learners can think of the given rule, along lines “we take any value from the one variable multiply by 0,912 and add 0,736 to get the value of the other variable”. Learners do not need to perform these calculations for a particular x value but can imagine this process within themselves. When an individual thinks of using this regression equation to extrapolate values of y based on x-values, or to comment on the strength of association between variables x and y based on the regression co-efficient, r, then the learner has “encapsulated the process of the regression equation into an object.

3.4 Conclusion

In this section, major themes emerging from Vygotsky’s sociocultural perspective were highlighted. Some of these include the role of tools and signs in mediating learning, scientific and everyday concepts, learning and development as well as the ZPD. Moreover, themes from Piaget’s theory as well as such contemporary constructivist tenets like that of Dubinsky’s APOS theory were also discussed. In the next chapter, the researcher discusses the research methodology used in this study.

CHAPTER 4: RESEARCH METHODOLOGY

4.0 Introduction

In this chapter, discussion is centred on research design, research setting, sample and method of sampling, methods of collecting data, intervention, analytical protocol, reliability and validity as well as ethical issues.

4.1 Research design

For a long time, research has been polarised between two, dominant research paradigms, namely qualitative and quantitative research. The antagonism between qualitative and quantitative research proponents has become known, loosely, as the "paradigms war." (Muijs, 2011). Quantitative research involves explaining phenomena by collecting numerical data that is analysed using mathematically based methods (Cotterill, 2012). This research paradigm is founded on positivist/ realist world view that the truth is out there, waiting to be discovered and that the world works according to some fixed laws of cause and effect (Mujis, 2011). Fundamentally, it is the task of the researcher to use objective research methods to uncover the "truth" (Muijs, 2011). In quantitative research, the relationship between phenomena under study is investigated in terms of generalizable causal effects relationships which will, in turn, allow for predictions to be made (Gelo, Braakmann & Benetka, 2008). Theories about these laws can be tested through scientific thinking (Muis, 2011).

In my own view, there is no objective truth lying somewhere, in a static world, waiting to be discovered. Knowledge is being generated from time to time, with old theories being discarded and replaced by new ones. Moreover, it does not make sense to explain all phenomena in terms of causal –effect relationships. In most instances, such a relationship cannot apply. This takes us to an alternative paradigm, namely qualitative research.

Creswell, cited in Leedy (1997) describes a qualitative study as, 'An inquiry process of understanding a social or human problem, based on building a complex, holistic picture, formed with words, reporting detailed views of informants and conducted in a natural setting'. The world or reality is not fixed, single, agreed upon, or measurable phenomenon that it is assumed to be positivist. Ideally, this research paradigm is subjective and interpretive. Thus, it can be done on people's lives, behaviour, lived experiences, emotions, feelings, social movements, cultural phenomena and interactions between nations. Ernest (1997) opines that research in qualitative paradigm adopts a bottom-up perspective, using particular and concrete instances to suggest, evoke, illustrate, and describe general cases. Thus, an advantage of adopting a qualitative approach is that specific situations are studied with the goal of building knowledge and contribute towards theory.

On the nature of reality, proponents of the qualitative research approach view it as a multiple, socially and psychologically constructed phenomenon, where the knower and the known are inextricably connected to each other (Gelo et al.,2008). Along the same lines, we need to acknowledge the uniqueness of human beings, that is, other people have realities independent of ours. Essentially, other people's realities along with ours are what we strive to understand through qualitative research (Ernest, 1997). Hence, the researcher-participant relationship is key in qualitative research. Merriam (1998) concurs that all qualitative research is characterised by the search for meaning and understanding.

The 'paradigm war' between qualitative and quantitative research followers brought about another, modern category of researchers. This category of researchers is known as pragmatists. They subscribe to whatever works in either qualitative or quantitative research. Their research paradigm is known as mixed methods. Muijs (2011) submits that mixed methods research is a flexible approach, where the research design is determined by what we want to find out rather than by any, pre-determined epistemological position. However, mixed methods research has its own weaknesses. One of the weaknesses is that, carrying out a sophisticated quantitative study as well as an in-depth qualitative investigation simultaneously is not an easy task (Scott & Morrison, 2005).

This study adopts a qualitative approach in the form of a case study. According to Opie (2004), a case study is an in-depth study of a single instance, in an enclosed system where certain features of social behaviour or activities in particular settings together with other factors influence the situation. The key understanding is that meaning is socially constructed and negotiated by the individual and the social world around them (interpretivist view). The rationale for opting for this approach is that the study was not intended to look for generalisations of results but to highlight, through in-depth exploration, the nature of frequently appearing errors probably emerging from misconceptions held by learners in statistics.

4.2 Research setting

This study was done at a township school in the Vaal Triangle area of the Gauteng Province. The school has an enrolment of 500 learners ranging from Grade 8 to Grade 12. Most of the learners stay within a walking distance from school, while a few use taxis and trains as mode of transport. The learners are mainly from low to middle income families and the school is a non-fee paying institution. There is a very rich cultural diversity in the school, with learners coming from a wide range of South African as well as Lesotho cultural backgrounds. The school is also part of the government nutrition programme.

4.3 Sample discussion

The sample for this study was made up of 18 Grade 11 learners, consisting of 11 girls and 7 boys. Basically, these learners constituted the mathematics class which was taught by the researcher at the time of the study. The method of sampling that was used to determine subjects for this study is purposive sampling. In this method, there is a deliberate selection of participants that meet criteria relevant to the research study (Cohen & Manion, 1994). Furthermore, in a qualitative study, the sample is usually "a small, non-probability sample that is non-representative" (Makonye & Luneta 2013, p. 119). Representativeness of sample was not a factor in the determination of subjects of the study because the findings of this study are not for extension of theory. Moreover, the rationale for selecting the above-mentioned sample is that the subjects were readily accessible and most importantly, willing to participate in the study outside normal school time.

4.4 Methods

In this study, the leading objective was to uncover misconceptions learners had which potentially resulted in their making of errors regarding variability of data, in particular, representations of ogive, frequency polygon as well as box and whisker plots. In order to get first-hand data on the challenges which learners were encountering, the researcher was personally involved in the classroom delivery of content. This was in compliance with one of the key tenets of interpretivist qualitative research that the researcher must be immersed in the research setting and have first- hand experience in the data gathering process. Besides delivering content, the researcher was also actively involved in assessment of learners, both formative and summative. It must be noted that assessment was not for grading purposes but to identify and categorise errors.

The researcher planned for and conducted ten lessons, with each lesson lasting for 45 minutes. After delivering the lessons, two sets of summative assessment tasks were administered in order to identify errors. In between these two tasks, an intervention lesson was conducted, with the main objective being to minimise errors identified in the first summative assessment task. To better understand the thinking behind errors which were committed in the first summative assessment, four learners were interviewed. Cooper and Shore (2008) argue that students need to talk out their understanding of graphs in order to glean more information and interpret appropriately. After intervening, the second summative assessment was administered to ascertain the extent to which intervention impacted on learner errors as identified in the first summative assessment task.

As part of arrangements made between the researcher, school governing body, principal, parents and learners, prior to the commencement of the study, all the lesson were conducted after normal contact time so as not to interfere with the school curriculum delivery programmes. The methods used in this study were pre-and post-intervention tasks as well as an interview.

4.4.1 Pre- and post-intervention tasks.

As part of the data collection process, two tasks were given to learners. These tasks were prepared by the researcher based on the concepts covered during the lessons. As alluded to above, the first task was given after teaching while the second task was given after an intervention was done. The second task was a derivation of the first one, with minor alterations to items being done while retaining the concepts. The rationale for administering the second task was to establish whether or not intervention had assisted in minimising the identified errors.

4.4.2 Interview

Besides tasks, a semi- structured interview was administered in collecting data. This was done as a follow-up to the task in order to get the thinking behind learner responses. Semistructured interview involves asking questions and then probing more deeply with open-form questions to obtain additional information (Gall, Gall & Borg, 2010). There are a number of advantages of using interview as a data collection method. Firstly, an interview was preferred because responses can be followed-up, probed, clarified, and elaborated to achieve more specific and accurate responses. Secondly, an interview allows the researcher to note both verbal and non-verbal behaviour and the interviewer has an opportunity to motivate the respondent (McMillan & Schumacher, 2010). Lastly, an interview helps to gain better insight into categories of errors and misconceptions that respondents exhibit and find out causes thereof (Makonye & Fakude, 2016).

4.4.3 Limitations in data collection

Due to time constraint, interview data was not transcribed. Instead, the researcher reported on findings from interview data in its audio form. As reported earlier own, lessons for this study were conducted in the afternoon. Usually, most learners were tired by then and thus, were prone to make errors due to a limited concentration span. As a result, this was a potential limiting factor in obtaining valid data on the nature of errors made by learners. High rates of absenteeism among certain individual learners negatively impacted progress of the data gathering process. If a learner was absent during any given lesson, the researcher would have to cater for him/ her to ensure fair summative assessment for all of them. In the following section, a summary of the intervention lesson is presented.

4.5 Summary of intervention lesson

A single intervention lesson (see detailed Lesson Plan in APPENDIX 11) was conducted after administration of the first task. The objectives of this lesson were to enable learners to draw a box and whisker plot, ogive and histogram as well as interpret data based on these diagrams in a given context. These objectives were motivated by a high error rate in learner pre-intervention task scripts. Learners generally struggled to answer questions on box and whisker plots as well as making interpretations. Analysis of data from pre-interventon tasks as established by the analytical tool developed for this study, along with the error classifier module of the conceptual framework were used as starting points both in lesson planning as well as delivery. In introducing the lesson, a short video on how to draw a box and whisker plot was played so as to remind learners as well as to capture their attention. Error remediation was provided as informed by the Multiple External Representation (MER) classifier module of the framework. In particular, the complementary and interpretation constraining functions of MER were adopted through the use of more than one representation in a single solution. In the first activity (see lesson plan in APPENDIX 11 for details), learners were divided into six groups. Based on a given set of data, the groups were assigned a task of drawing a cumulative frequency curve and a box and whisker plot on a single graph paper. They were instructed to extend lines from the five summary on the box and whisker plot to the ogive (as shown on the lesson plan). This was done so that the learners could realise that the two diagrams were different in nature but could be used to depict variability in different ways to enhance learner understanding. Bakker et al. (2004) is of the view that when aggregate plots like box and whisker diagrams are introduced, it is recommended that they initially be accompanied with representations that still allow students to see an individual case (Bakker et al., 2004). After moving around to check the group solutions, the researcher demonstrated the solution on the smartboard while the groups which had failed to obtain the correct solutions wrote remedials in their scripts.

In the second activity (see detailed lesson plan in APPENDIX 11), learners were asked to draw a box plot and a histogram on a single graph paper. As in the previous activity, this was done so as to show the learners that these two graphs played the complementary function of MER as detailed in Chapter 3. The use of the two representations enabled the majority of learners to visualise the abstract and often elusive concept of whisker as represented by shorter, extreme bars on the histogram. From literature, learners generally focus much on the box (Bieler et al., 2004; Lem et al., 2014) and fail to realise that 50% of data values is concentrated in the whiskers. In this particular instance, it was interesting to note the increased level of learner participation throughout the course of the lessons. Groups were competing against each other, making very reasonable arguments and trying to sell their ideas to others. The researcher had solutions done and hidden on the smartboard so as to allow learners time to attempt in pairs before a whole-class discussion was done.The researcher made use of the Geogebra function of the smartboard to draw graphs (see lesson plan in APPENDIX 11).

In the next section, the researcher presents the analytical protocol for the study as well as how it will be used in analysing learner pre-intervention tasks.

4.6 Analytical protocol

Knowledge construction	Common errors
Action - e.g. from a given set of raw data learners were asked to organise the data so as to calculate the quartiles of data	**Procedural** - e.g. in drawing frequency polygon, learners plotted frequencies against upper class values, instead of mid-points of data intervals
Process . :- requires manipulations and / or computation to complete them - e.g. $SD = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$ to find standard deviation.	**Conceptual** - e.g. learners gave cumulative frequencies instead of frequencies
Object : - concepts exist as noun, as a 'thing' : - e.g. learners begin to see standard deviation as "the average spread around the mean" (Newbold et al. (2010, p.77).	**Poor adaptive reasoning** - :- e.g. based on two box and whisker plots displaying pass-rates for two schools, learners were to make comparisons of the two graphs and tell which school performed better than the other.
Schema : - actions, processes and objects all synthesised into a new entity : - shortcomings in this category lead to all the three error categories	

Figure 5: Analytical protocol (*adapted from Kilpatrick et al., 2001, & Dubinsky, 1991*).

4.7 Description of the analytical protocol

To begin the data analysis process, the study revisits the four-stage process of concept development (APOS), as described in chapter 2 above. As detailed in Chapter 2, stages of concept formation are matched with errors that are likely to happen at each stage, as informed by Kilpatrick et al. (2001)'s mathematical proficiency framework (see Chapter 3). At the first stage, learners who require external stimuli to build concepts (action) require a supporting environment, resourced with learning aids, drawn from real life situations (Dubinsky, 1991). It is at this stage that learning starts and failure to provide requisite aids leaves learners badly exposed to errors and misconceptions. Moreover, this is the stage where the language tool is vital in that there is a lot of verbal interaction amongst the learners as well as between the teacher and the learners in the process of mediation of learning.

In the next stage, that is, process stage, focus is mainly on operations and procedures (Sfard, 1991) and much of the reasoning involved is formulaic. Learners at this stage are said to have "interiorised the action into a process". In this study, tasks which involved manipulation of formula like $\sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}}$, to find standard deviation were classified as process tasks. At any given stage, learners (including the capable ones), are bound to make mistakes, say in substituting data into formula or feeding data into a calculator in order to compute mean or standard deviation. Errors committed at this stage were predominantly procedural in nature. This could be an indicator of lack of conceptual understanding.

Thirdly, when a mathematical entity becomes an object, it is seen as a 'real thing that exist in space and time' (Sfard, 1991). At this stage, all the processes are reified into an object, a transformation referred to as 'encapsulation of process into object' (Dubinsky, 1991). A learner who is at this stage of concept development views a mathematical problem as a noun. It is possible that a learner moves from this stage all the way back to the action stage, like when presented with a task to find outliers on a scatter plot, some learners might resort to use the formula $(Q_1 - 1.5 \times IQR;\ Q_3 + IQR \times 1.5)$. Whilst there is nothing wrong in using the formula, in essence, this child shall have moved back to action by using a formula to respond to an external stimulus. In this particular instance, a learner who has reified the concept of outlier would proceed to inserting a line of best fit onto the scatter plot and read off any outlier, if it exists. At this stage of concept development, errors made are mainly conceptual as well as those emanating from poor adaptive reasoning.

Lastly, the highest stage in the process of concept formation is schema. This is reached when the learner has mastered actions, processes, objects and has a bird's eye view of them (Dubinsky, 1991). Each schema is organically linked to other schemas. Analysis and reflection on a schema can generate yet another cycle of actions and processes so that new, more advanced mathematical objects and schemas can be formed. If learners are able to compare data on maternal death rates in Africa and Asia over the past five years, they are said to have mastered the schema on graphical objects. Individuals who do not possess this schema cannot participate in the discourse of making sense of data and will lag behind in understanding phenomena around them.

Experience reveals that learners who have attained this stage of concept formation can tackle high order questions but are equally prone to making procedural oversights, probably due to complacency. As alluded to above, analysis and reflection usher in new cycles of actions, processes and objects. Essentially, when someone is developing the understanding of mathematical idea, it does not necessarily happen in a linear process; rather it is dialectical (Dubinsky, 1991). Hence, in the figure above, three arrows link schema to all the three error categories

4.8 Reliability and Validity

Reliability and validity of research instruments is important for credible results. Plumb and Spyridakis (1992) say that reliability is the extent to which instruments elicit the same answers from the same people at different times. On the other hand, validity refers to the degree to which an instrument actually measures what it is designed to measure. To enhance validity of task sheets, questions were sampled from previous examination papers from one of provincial as well as the national education department. Moreover, a colleague checked for spelling and grammatical errors for both interview protocols as well as task sheet items. This pre-moderation was done to make sure that cognitive levels of items were appropriate for subjects. After administration and marking of task sheets, a few scripts were sampled for post-moderation to ensure that marking was fairly done. It was conceived that this would further enhance validity of the instrument. In order to ensure validity, due consideration must be given to the source of test items as well as moderation of the tool (Makonye & Nzima, 2016).

After conducting the two intervention lessons, a task sheet was again administered in order to establish whether or not intervention had made any impact in the minimisation of errors and misconceptions. This post-intervention task sheet had a similar structure to the one administered prior to intervention. However, there were minor alterations of the figures which were made in generating the post-intervention tool.

4.9 Ethical considerations

McMillan and Schumacher (2010) maintain that it is imperative for researchers to obtain permission to enter any particular field and also ensure the confidentiality and anonymity of the participants, thus encouraging the latter's free choice of participation. In this study, a number of ethical issues were considered and addressed. Firstly, consent was obtained from parents and respondents (learners) to collect data through task sheet as well as an interview. Throughout the study, pseudonyms were used instead of real names so as to ensure confidentiality of learners. In the same vein, the exact location of place where the study was conducted was not revealed. The Gauteng Department of Education (GDE) was also approached for permission to carry out the study at one of its schools. The requisite documentation together with the research proposal were submitted at the GDE head office. Permission was also granted with the proviso that the permission letter (see APPENDIX 2) be submitted to the District Office where the school is located and that both the school principal and school governing body be approached for permission. After the process was completed, ethical clearance (protocol number, 2017ECE049M, attached in APPENDIX 1) was granted by the University Ethics Committee.

4.10 Conclusion

In the opening stage of this chapter, the research design was identified. Also, research setting, sample and sampling method, methods of collecting data, intervention strategy and analytical protocol were discussed. In closing, reliability, validity and ethical considerations were highlighted.

CHAPTER 5: DATA ANALYSIS AND DISCUSSIONS

5.0 Introduction

In this chapter, research data is analysed so as to come up with themes and patterns which would provide possible answers to questions which were put forward in the opening chapter of the study. In this analysis, task sheet data is put into categories based on the analytical framework which was discussed in chapter 4. In the next section, I present the style adopted in data analysing.

5.1 Analysis style

Qualitative data analysis is “a relatively systematic process of coding, categorising, and interpreting data to provide explanation of a single phenomenon of interest” (McMillan & Schumacher, 2010, p.367). In this study, errors made by learners were predetermined either as procedural or conceptual. In this regard, inductive data analysis is the style which was adopted in this study in analysing data. Inductive analysis is “the process through which qualitative researchers synthesise and make meaning from the data, starting with specific data and ending with categories and patterns” (McMillan & Schumacher, 2010, p.367). In other terms, inductive style of analysing data aims at achieving some degree of universality in explanation of phenomena in diverse situations. In this way, more general themes and conclusions emerge from the (particular) data rather than being imposed prior to data collection.

In this study, a sample of 3 pre-intervention tasks scripts were analysed, per question, in order to answer the first research question: what are the common errors and misconceptions learners have in visual display and interpretation of variability through frequency polygon, cumulative frequency curve, and box and whisker plot? The rationale for only analysing pre-intervention was to capture the nature of errors emanating from misconceptions in learners before any intervention was done. In doing so, a two-way analysis of both concept development as well as the category of errors demonstrated is suggested. It was felt that looking at errors alone without considering the concept being developed would not assist much in tracing their sources. Moreover, it is proposed that by looking at stage of concept development, suggestions on possible course of action to take in their minimisation could be made. Borasi (1994) is of the view that error analysis in mathematics education can be viewed as a motivation tool and a perfect starting point for creative mathematical exploration. It is valuable in problem solving as well as problem posing activities.

5.2 Pre-intervention task diagnostic analysis

The analysis was done on task sheets of three learners, who were given pseudonyms of JPM 3, JPM7 and JPM 13 in order to protect their identity. Codes were used in marking task sheets for convenience of collating data as well as analysis purposes. The codes were as used as indicated on the table below:

Code	**0**	**1**	**2**
Decision	No attempt	Correct	Incorrect

Table 2: Marking codes

Question 1

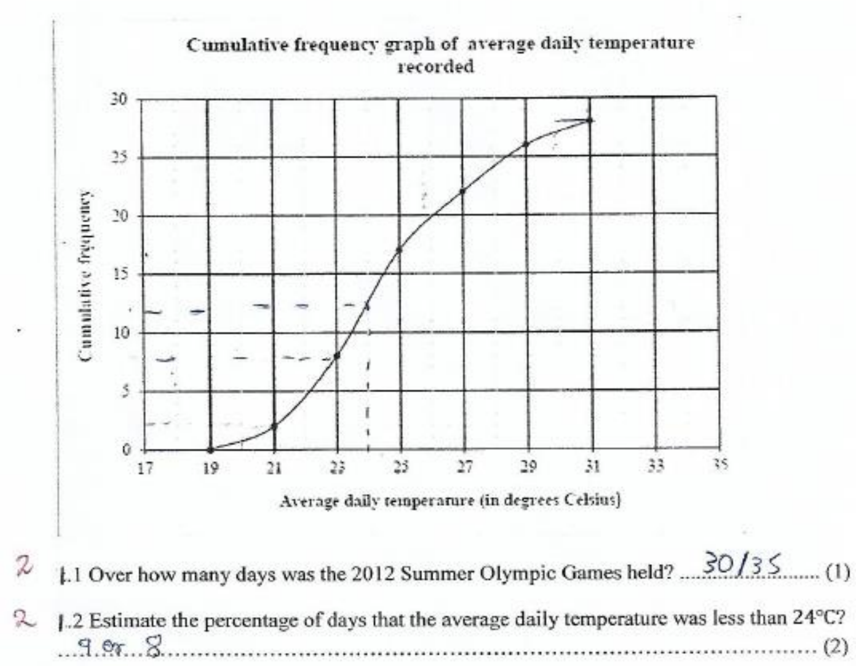
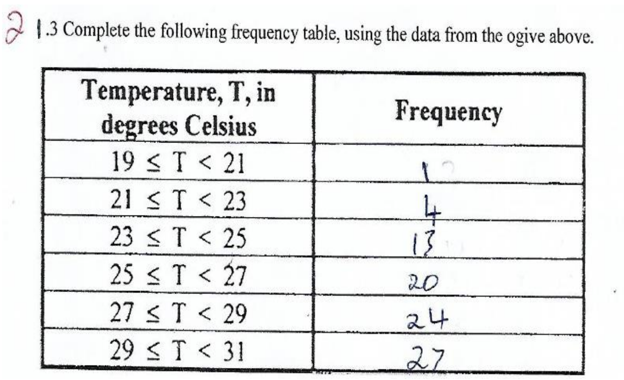


Figure 6: Vignette 1 for Que 1.1 and 1.2 (JPM3)

1.1 **Process task**.Generally, these three learners made conceptual errors, though in different ways. JPM3 could not link the total number of days over which the 2012 London Games were played to the highest value of cumulative frequency (vertical axis). The other two learners, JPM13 and JPM 7 (see vignettes in APPENDIX 12) appear to have guessed their answer because there is neither evidence of numerical calculations in the build-up to the solution nor graphical manipulation.

1.2 **Process task**. JPM3 demonstrated that he knew how to manipulate the graph to get the estimate of the percentage of days (procedural fluency). However, he went on to write wrong values. This is a demonstration of procedural error. The other two learners managed to manipulate the graph and obtain the number of days but could not convert to percentage. This is a conceptual error



1.3 Complete the following frequency table, using the data from the ogive above.

Temperature, T, in degrees Celsius	Frequency
$19 \leq T < 21$	1
$21 \leq T < 23$	4
$23 \leq T < 25$	13
$25 \leq T < 27$	20
$27 \leq T < 29$	24
$29 \leq T < 31$	27

Figure 7: Vignette 6 for Que 1.3 (JPM 7)

1.3 **Process task.** JPM 3 and JPM 13 made procedural and conceptual errors. Firstly, they could not realise that one can get frequencies from an ogive by successive subtraction of cumulative frequencies at each stage. This is basically the inverse operation of obtaining cumulative frequency from frequencies. As a result, this is lack of both conceptual as well as procedural knowledge. Secondly, JPM7 showed lack of consistency in his values. In other words, it could not be established how he obtained the values in the frequency column.

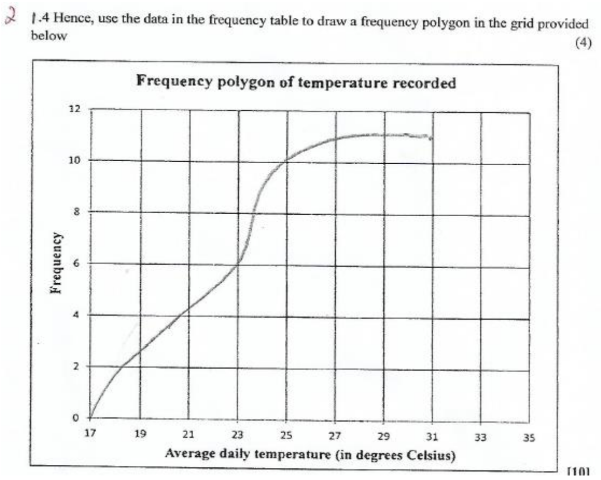
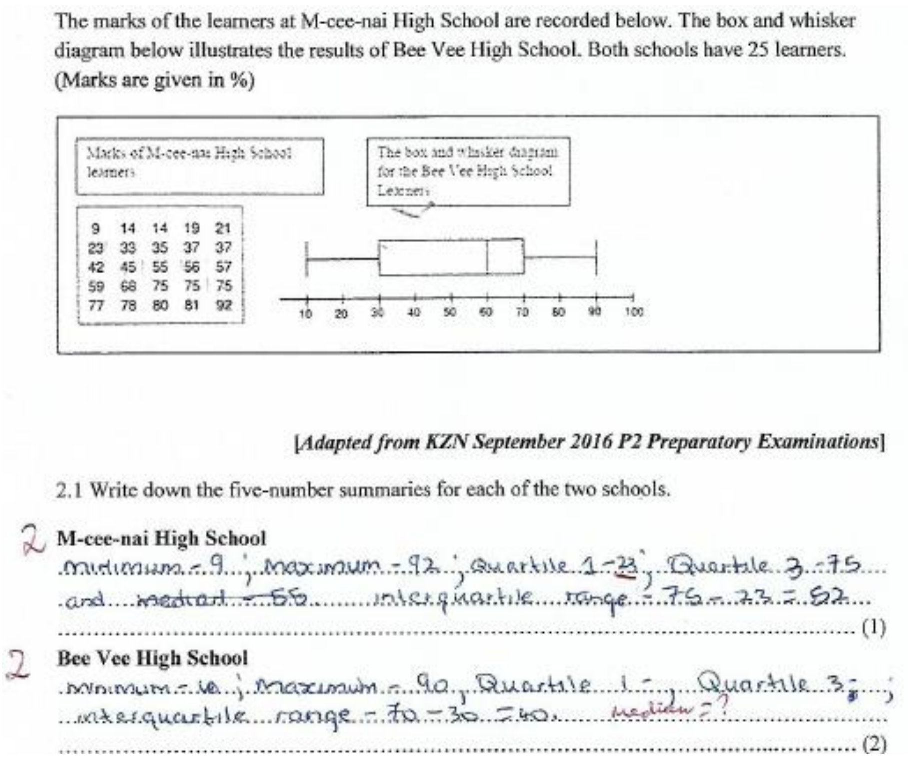


Figure 8: Vignette 9 for Que 1.4 (JPM 7)

1.4 **Process task**. In this case, JPM 3 demonstrated some prior knowledge of frequency polygon. However, his plotting of values on the horizontal axis was not done in the middle of class interval, as required. He plotted at the end of the class interval, as if he was drawing an ogive. This was more of a procedural error than conceptual. From a constructivist perspective, his concept of an ogive interfered with that of a frequency polygon and this resulted in him plotting at the end points of horizontal axis. Learners JPM13 and JPM 7 were both not sure of what they were doing. JPM 13 abandoned the plotting process while JPM7 just drew a curve with no evidence of plotting. This is a conceptual error.

Question 2


The marks of the learners at M-cee-nai High School are recorded below. The box and whisker diagram below illustrates the results of Bee Vee High School. Both schools have 25 learners. (Marks are given in %)

Marks of M-cee-nai High School learners

9	14	14	19	21
23	33	35	37	37
42	45	55	56	57
59	68	75	75	75
77	78	80	81	92

[*Adapted from KZN September 2016 P2 Preparatory Examinations*]

2.1 Write down the five-number summaries for each of the two schools.

2 **M-cee-nai High School**
minimum = 9; maximum = 92; Quartile 1 = 23; Quartile 3 = 75 and ~~median = 55~~ interquartile range = 75 – 23 = 52 (1)

2 **Bee Vee High School**
minimum = 10; maximum = 90, Quartile 1 = ; Quartile 3 = ; interquartile range = 70 – 30 = 40. median = ? (2)

Figure 9: Vignette 11 for Que 2.1 (JPM 13)

2.1 Process task. JPM 3 could neither come up with five number summary from the given raw data nor from the box and whisker plot. This is both conceptual as well as procedural errors error. JPM 13 showed that she knew the five number summary but lacked precision and confidence as demonstrated by the cancelling of the correct value of median. In this instance, a decision could not be made regarding the nature of error. JPM 7 left out some values of the five-number summary. Some of the values he gave were wrong. He also included the mode, which is not part of the summary. This is a demonstration of lack of conceptual as well as procedural understanding of the five-number summary. Thus, he made conceptual and procedural error in that instance.

2 2.2 Draw a well-labelled box and whisker diagram that represent M-cee-nai High School marks on the scaled line below. (3)

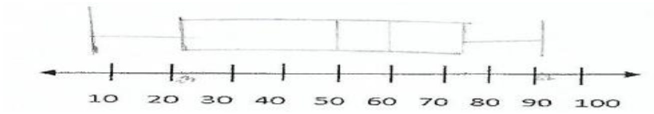
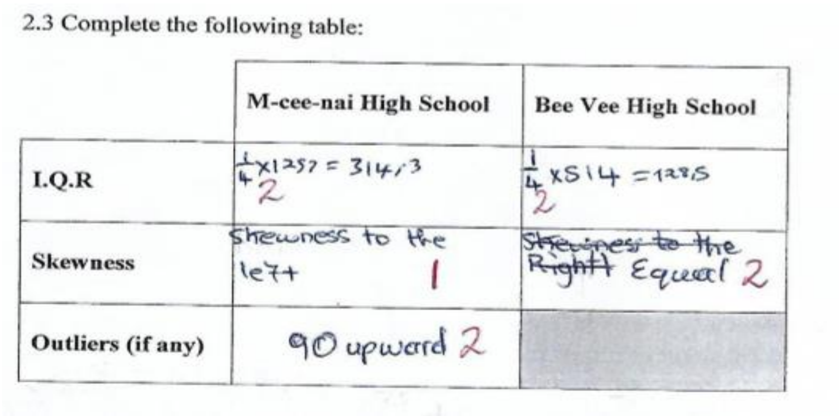


Figure 10: Vignette 14 for Que 2.2 (JPM 13)

2.2 **Process task.** JPM3 did not even attempt to find solution to this question, which indicates a gap in conceptual understanding. JPM 13 showed that she knew what a box and whisker is as well as how to draw it. The only drawback for her was that she had obtained a wrong value of median in item 2.1 above, so she ended up not having a perfectly drawn graph. JPM 7 made both procedural as well as conceptual errors in that all the values at key points of the box plot, apart from the central value, were wrong.



2.3 Complete the following table:

	M-cee-nai High School	Bee Vee High School
I.Q.R	1/4 x 1257 = 314,3 2	1/4 x 514 = 128,5 2
Skewness	Skewness to the left 1	~~Skewness to the Right~~ Equal 2
Outliers (if any)	90 upward 2	

Figure 11: Vignette 18 for Que 2.3 (JPM 7)

2.3 In this item, finding IQR and applying standard deviation and mean in finding outliers were coded as **Process tasks** while commenting on skewness was coded as an **Object task.** JPM3 showed no knowledge of IQR (conceptual error). He could use the provided box and whisker plot to interpret data but still failed to do so using the box plot he had drawn for M-cee-nai School (Procedural). He had no understanding of the concept of outliers (conceptual error). JPM 13 demonstrated knowledge of IQR, though in the second instance, she got it wrong since wrong values had been obtained in 2.2. She could not use the provided box plot to comment on skewness of data (conceptual error and poor adaptive reasoning). Moreover, this learner had no idea on outliers (conceptual error). JPM7 had no understanding of the concept of IQR (conceptual error) since he got the first answer correct but was not decided in the second part of the question as shown by constant cancelling (conceptual, procedural and poor adaptive reasoning). Lastly, the learner had no knowledge of outliers (conceptual error).

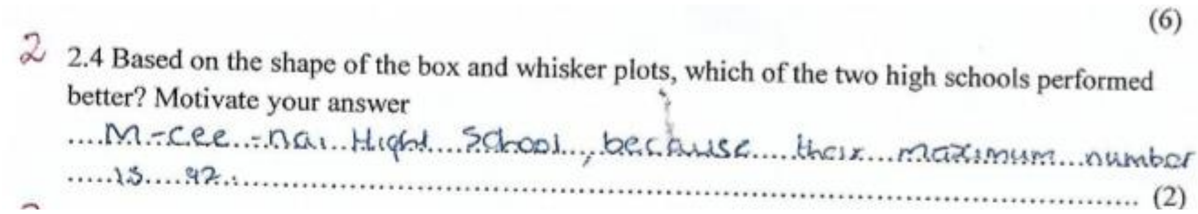

(6)
2 2.4 Based on the shape of the box and whisker plots, which of the two high schools performed better? Motivate your answer
....M.-cee..-nai..High...School..., because.....their...maximum...number
.....is....92.. (2)

Figure 12: Vignette 20 for Que 2.4 (JPM 13)

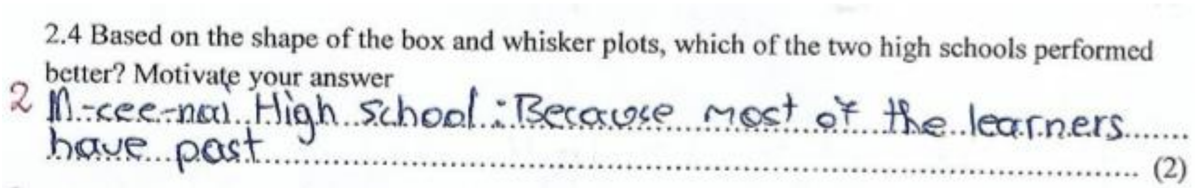

2.4 Based on the shape of the box and whisker plots, which of the two high schools performed better? Motivate your answer
2 M.-cee-nai..High..school..: Because..most..of..the..learners......
have..past.. (2)

Figure 13: Vignette 21 for Que 2.4 (JPM 7)

2.4 **Object-based task.** In this item, only two responses were analysed because the other learner managed to get the correct response. JPM 13 could not use box plots to make comparisons of the matric pass rates and decide on the school which performed better than the other (conceptual error and poor adaptive reasoning). It appears that this learner had some challenges with language as well. JPM 7, likewise could not make the comparison and reach a conclusion (conceptual error and poor adaptive reasoning).

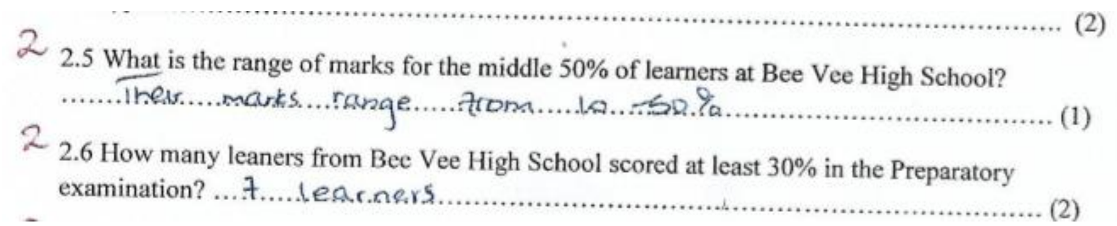
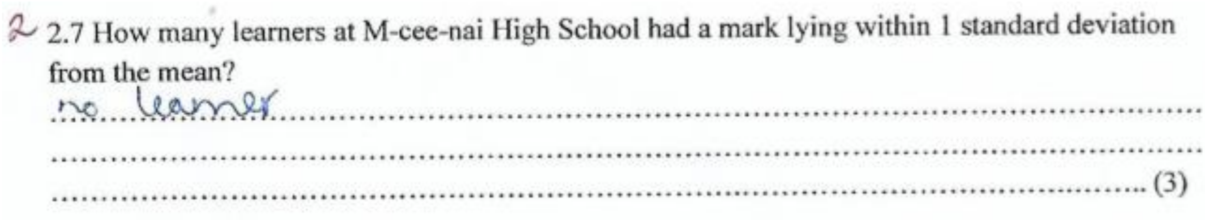

.... (2)
2 2.5 What is the range of marks for the middle 50% of learners at Bee Vee High School?
.......their.....marks....range.....from.....10..-50%.................. (1)
2 2.6 How many leaners from Bee Vee High School scored at least 30% in the Preparatory examination? ...7.....learners.................................. (2)

Figure 14: Vignette 23 for Que 2.5 & 2.6 (JPM 13)

2.5 **Process task**. All the three learners did not understand the concept of ‘range’ (conceptual error)

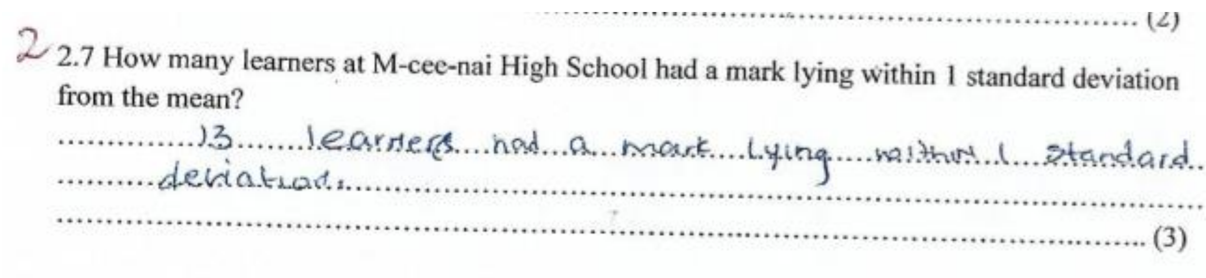
2.6 **Process.** All the learners failed to realise that 30% was the first quartile of the data and so 75% of the data lies above this value. This is lack of both conceptual and procedural understanding. (conceptual and procedural error).


2 2.7 How many learners at M-cee-nai High School had a mark lying within 1 standard deviation from the mean?

no learner

(3)

Figure 15: Vignette 25 for Que 2.7 (JPM 3)

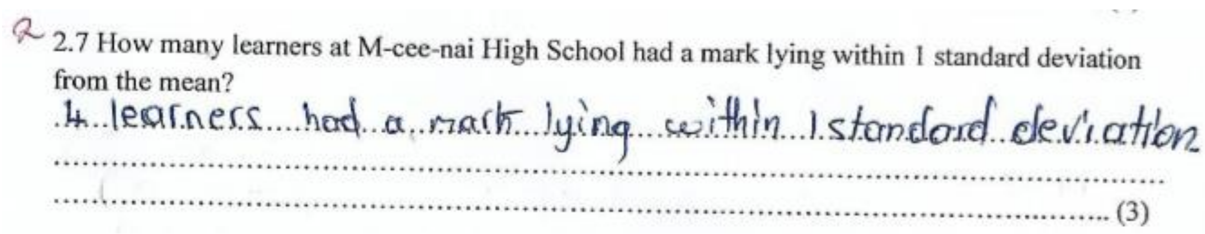

(2)

2 2.7 How many learners at M-cee-nai High School had a mark lying within 1 standard deviation from the mean?

13 learners had a mark lying within 1 standard deviation.

(3)

Figure 16: Vignette 26 for Que 2.7 (JPM 3)


2 2.7 How many learners at M-cee-nai High School had a mark lying within 1 standard deviation from the mean?

4 learners had a mark lying within 1 standard deviation

(3)

Figure 17: Vignette 27 for Que 2.7 (JPM 14)

2.7 Object. All the three learners demonstrated lack of knowledge of the concepts of mean and standard deviation (conceptual and procedural error), which were necessary to identify the number of learners lying within 1 standard deviation of the mean (poor adaptive reasoning).

In the following section, a summary of the analysis is made through comparing error responses made in pre- and post-intervention tasks, along with learner responses in intervention.

5.3 Pre- and post-intervention and interview data analysis

In this section, a comparison is drawn between correct and incorrect responses made in both pre- and post-intervention tasks in order to establish the effectiveness of intervention in the minimisation of errors. Moreover, transcripts of interviews were used to analyse the thinking behind the identified errors and misconceptions.

Item	Focus	No. correct	%	No. incorrect/blank	%
1.1	Reading cumulative frequency	5	28	13	72
1.2	Converting cumulative frequency to percentage	3	17	15	83
1.3	Read frequency from an ogive	2	11	16	89
1.4	Drawing frequency polygon	2	11	16	89
2.1.1	Find five-number summary	6	33.3	12	66.7
2.1.2	Read 5-no. summary on boxplot	7	39	11	61
2.2	Drawing a boxplot	3	16.7	15	83.3
2.3.1	Finding interquartile range	4	22.2	14	77.8
2.3.2	Finding interquartile range	5	27.8	13	72.2
2.3.3	Commenting on skewness	7	46.7	11	53.3
2.3.4	Commenting on skewness	6	33.3	12	66.7
2.3.5	Identifying outliers	2	11	16	88.9
2.4	Box plots: applications	2	11	16	88.9
2.5	Spread of data: range	1	6	17	94
2.6	Spread of data: lower quartile	0	0	18	100
2.7	Application of standard deviation	0	0	18	100

Table 3: Correct/incorrect responses for pre-intervention task

Overall, Table 3 reflects a mean percentage incorrect response of 80.4% across items in pre-intervention tasks. Only 5 learners could correctly read the total number of days taken at the 2012 London Olympic Games, which translates to 72% rate of incorrect responses. In item 1.2, only 3 of the 18 learners could correctly read and calculate the percentage number of days when average daily temperature was less than 24℃. For illustrative purposes, the following is a transcription of the interviewer (researcher) (R) and four learners, JPM 4, JPM 5, JPM 6 and JPM 12, with regards to their thinking behind their answers to item 1.2.

R: What is your understanding of "percentage number of days when temperature was less than 24 ℃?"

JPM5: Sir...I failed that one.

R: If you are given another chance, is there a way which you think could be used to answer this question?

JPM5: It is there sir…

R: Could you please explain.

JPM 5: Sir, I realised it late that this question was easy…I told myself that they needed percentage and I did not write number of days.

R: (*interjects*) How many days did you get from your graph?

JPM5: They are 13?

R: How did you get that number?

JPM5: ok, you go to 24 degrees and you go up the y-axis and when you get to the curve, you turn left and straight ahead on the y-axis you will get 13. That's how I got it.

R: So how do you then get percentage from this figure?

JPM 5: I multiply by 100

R: Before you multiply by 100, there is something you must do to 13...

JPM 5: I was supposed to first divide 13 by number of days and then multiply by 100 to get percentage

R: I think you have answered my question

R: Let us have others' views

JPM 6**:** I started at 24 degrees and moved up and turned left when I reached the graph. This took me to 13 on the vertical axis.

R: How did you get percentage from this value?

JPM 6: I only stopped at number of days and I did not know what to do with it to find percentage.

JPM 12: I took 13, which is my answer in item 1.1 and I divided it by 35 which is the total number of days and multiplied by 100.

JPM 4: I did not understand what this question wanted me to do.

From the above transcript, it is clear that learners were not sure of the demands of the question. A common stumbling block to most learners was that their thinking was processual. If they had objectified the concept of fraction as a portion of total days and an ogive as a "less than" graph, then there could have been very limited room for them to misinterpret the graph and make errors. However, for some, mathematical language appeared to have been the only barrier to successfully answering this item. Through rephrasing, probing and giving the learners time to reflect and talk about how they reasoned to get answers, some learners, like JPM5, realised that they could have easily got the correct answer. Cooper and Shore (2008) argue that students need to talk out their understanding of a graph in order to glean more information from it and interpret it more appropriately.

Item 1.3 recorded even more percentage incorrect responses (89%) than the previous two items. Most learners could not manipulate cumulative frequency to obtain frequency values. As a result, 89% of learners could not accurately draw the frequency polygon. However, it is worth noting that a significant number of respondents were aware that a frequency polygon is obtained by plotting mid-class values against frequencies. In items 2.1.1 and 2.1.2, 66, 7% and 61% respectively, got incorrect answers. Moreover, while a fair number of learners (6) were able to find the 5-number summary of raw data as well as reading from the given box and whisker plot, 83.3% of learners could not accurately draw a box and whisker plot.

Generally, learners struggled with making use of the two box and whisker plots in commenting on learner performance at the two schools (item 2.4). Of the 18 learners, 16 failed in the interpretation of learner performance based on box plots. However, as shown in the following interview transcript, 2 learners observed and made quite insightful comments on learner performance at the two high schools as displayed by the box and whisker plots. JPM 4 confirmed what Lem et al. (2014) alluded to above that, naturally, most learners pay more attention to the median on a box and whisker plot. In other words, to have a better mental picture of a box plot, learners should look at it holistically rather than focusing on individual parts.

R: Based on the shapes of your box and whisker diagrams, comment on the overall performance of learners at each of the two high schools.

JPM 5: (takes a deep sigh…) median of Bee Vee is 60% while that of M-cee nai High School is 55% which means 50% of learners at Bee Vee High School got more than 60% while that at M-cee nai High School scored more than 55%. Therefore, Bee Vee performed better than M-cee nai High School

R: JPM 4, what do you have to say?

JPM 4: I noticed that 50% of M-cee nai is lower than that of Bee Vee High School. Also, the other box and whisker has an outlier.

After writing post-intervention tasks, learners' responses were tabulated as shown in **Table 4** below. It was envisaged that this would assist in assessing the impact of intervention in minimisation of identified errors and misconceptions.

Item	Focus	No. correct	%	No. incorrect	%
1.1	Reading cumulative frequency	6	33.3	12	66.7
1.2	Converting cumulative frequency to percentage	1	6	17	94
1.3	Reading frequency from an ogive	4	22	14	78
1.4	Drawing frequency polygon	6	33.3	12	66.7
2.1.1	Calculating five-number summary	8	44.45	10	55.55
2.1.2	Reading 5-no. summary on boxplot	9	50	9	50
2.2	Drawing a boxplot	7	39	11	61
2.3.1	Finding interquartile range	3	17	15	83
2.3.2	Finding interquartile range	3	17	15	83
2.3.3	Commenting on skewness	9	50	9	50
2.3.4	Commenting on skewness	8	44.4	10	45.5
2.3.5	Identifying outliers	2	11	16	88.9
2.4	Box plots: applications	3	17	15	83
2.5	Spread of data: range	0	0	18	100
2.6	Spread of data: lower quartile	0	0	18	100
2.7	Application of standard deviation	1	6	17	94

Table 4: correct/ incorrect responses for post-intervention task

According to data in the Table 4, there was a mean percentage incorrect response of 74. 96% as compared to 80.4% recorded in pre-intervention tasks. This translates to a marginal reduction rate of 5.44% between errors made in pre- and post-intervention tasks. This is a confirmation of the claim that students can doggedly hold on to mistaken ideas even after receiving instruction designed to dislodge them (Smith et al., 1993).

While the overall error reduction margin appears to be insignificant, certain task items, like drawing box plots (from 83.3% down to 61%) and commenting on skewness of data (from

66.7% down to 45.5%) actually recorded a huge reduction in errors committed. This can be attributed to the use of MERs, specifically, on how to draw and interpret data on box plots.

Clearly, most learners struggled to read frequency from an ogive and this problem persisted even to the actual drawing of frequency polygon. Only 4 out 18 learners managed to correctly read frequencies from an ogive. Moreover, as revealed by the following short transcript of the interview, some learners still had no knowledge of certain statistical concepts. They demonstrated that they had not objectified such statistical concepts like the mean, despite possessing computational skills on how to find it either manually or through use of a calculator. Skemp (1976) referred to this as rules without reason.

In the following short transcript, the learner demonstrated that he had no knowledge of the concept of standard deviation despite being able to compute it from given data set.

R: There is a measure known as IQR. What is your understanding of this number?

JPM 4: To be honest sir, there is somewhere IQR confuses me.

R: How is IQR used in making sense of data?

JPM 12: It’s not important sir.

Notably, none of the learners were able to notice that a score of 30% coincided with the lower quartile mark for Bee Vee High School. As a result, they failed to find the number of students who performed above the 25^{th} percentile of the data set (item 2.6).

Also, the following interview transcript reveals that this particular learner had no conceptual understanding of what he was writing in the tasks. This was conspicuous in the case where there was general lack of coherence in his responses. In some instances, his responses were self-contradicting. In explaining how he drew a box plot, he referred to “five number summary” simply as “numbers”. Because he did not have conceptual understanding of the five number summary, he indicated that he used all the 25 numbers to draw a box plot.

R: How did you draw the box plot?

JPM 4: I used the numbers.

R: What is the name of these numbers?

JMP 4: To be honest sir, i have forgotten.

R: How many are these numbers? JPM4: They are 25 sir.

As alluded to above, certain post-intervention items recorded significant gains in terms of error reduction. Before intervention, there was an error rate of 72% in reading cumulative frequencies. This fell to 66,7% after intervention. The following items also recorded significant error reduction (shown in brackets): reading frequencies from ogives (89% down to 78%), drawing box plots (83.3% down to 61%), interpretation of box and whisker plots, in particular skewness (53,3% down to 50% and 66,7% down to 45.5%). Overall, these gains could be attributed to intervention lesson which was administered, in particularly, the use of MERs to assist in the interpretation of spread of data.

In the next section, a discussion is made of what emerged from research data in the light of literature as well as theoretical framework.

5.4 Discussions

From the above analysis, it is evident that errors constitute a major part of the day to day learners' written activities as well as verbal interaction in the classroom. An error is a significant part of the learning process if it is dealt with diagnostically (Karpov, 2007; Sarwadi & Shahrill, 2014). Most of the errors and misconceptions which emerged in learner tasks as well as those displayed in students' responses in interview were mainly of conceptual nature. In other terms, most learners tried to apply their existing knowledge in a situation where such knowledge was not required. This confirms Piaget's view that all learning involves the interpretation of phenomena, situations, and events, including classroom instruction, through the perspective of the learners' existing knowledge (Smith, et al., 1993). Learners' knowledge of frequency was inappropriately applied to display variability of data on box plots, which do not have anything to do with frequency. Moreover, from a constructivist point of view, their experience on frequency presented a cognitive conflict because new knowledge could not fit in their existing schema (Piaget, 1964). Ideally, this should present a perfect opportunity for teachers to reflect on their practice and plan instruction carefully in making use of errors as springboards to effective teaching and learning.

Whilst both procedural and conceptual errors were committed in answering tasks, it was sometimes difficult to distinguish between these two types of errors. Perhaps, this is because knowledge of basic computation skills cannot be separated from conceptual understanding and forms the foundation for mathematical thinking (Wu, 1999). Procedural and conceptual understanding are always working concurrently. For example, as a child gains conceptual understanding, computational procedures are remembered better and used more flexibly to solve new problems. In turn, as procedures become more automatic, the child becomes aware of other aspects of a problem as well as how to tackle them, which leads to new understanding (Kilpatrick et al., 2001).

Interview data reveals that generally learners lack one of the most vital tools in thinking and communication, namely that of language. From a socio-cultural perspective, all higher mental activities like reasoning and problem solving are accomplished through and with the help of psychological tools like language (Vygotsky, 1987). Language is the bedrock for all verbal communication which, in turn fosters thinking through sharing of points of view. Virtually, all the four learners who were interviewed struggled to express themselves due to linguistic challenges. Perhaps, this was a major contributory factor to the prevalence of most errors and misconceptions which were in both pre- and post-intervention task sheets. Along the same lines, by using Vygotsky's notion of language, both as a tool and a sign, one can guide own work through verbal thought in solving problems which are seemingly insurmountable for many. Based on the interview data, it can be concluded that nothing should be left to chance when assessing learners. Every learner's effort and work should be scrutinised so as to get the underlying thinking behind the learner's answers. By taking one learner, JPM5, through her solution to one of the tasks during the interview, the learner was able to realise where she made a mistake in answering the question. Cooper and Shore (2008) argue that students need to talk out their understanding of graphs in order to glean more information from them and interpret it more appropriately.

Analysis of both pre- and post-intervention learner responses to task sheets as well as interview data show that the majority of learners had not made adequate advances on the APOS stages (Dubinsky, 1991; Arnon et al., 2014) in constructing appropriate statistical conceptions. This manifested either in the application of very long and tedious methods to solve seemingly simple problems or use of inappropriate methods or formulae on a given problem situation. Usually, it is through use of long methods of solving problems in which learners end up developing errors and misconceptions. For instance, despite being able to find quartiles of a given set of raw data, no learner managed to make use of the obtained quartiles in commenting on skewness of data (encapsulation of process into object). In some instances, learners inappropriately used formula for finding outliers to number of data points lying within a given number of standard deviation from the mean. Furthermore, they were unable to make use of their box plots to compare matric pass rates at the two, given schools. This was also reflected in learner responses during interviews.

Comparison of correct and incorrect response rate for both pre- and post-intervention tasks shows that intervention yielded very insignificant positive change of 5,44% mean error reduction. Overall, this slight change appeared to reinforce radical constructivist view that there is no ultimate, true knowledge about the state of affairs in the world or in the realm of mathematics (Ernest, 1996). However, diagnostic analysis of errors which learners made after intervention actually reveals that there was a huge decrease in errors. In particular, errors on drawing box plots decreased from 83.3% to 61% after intervention while those on how to comment on skewness of data plummeted from 66.7% to 45.5%. The objectives of the intervention lesson (see intervention lesson plan attached in APPENDIX 11) were to assist learners to draw and interpret box plots through use of MERs

5.5 Conclusion

In open stage of this chapter, the analytical tool proposed and detailed in Chapter 4 was used to analyse scripts of three learners in order to categorise errors as well as identify stages of concept formation. Secondly, the rate of commission of errors for both pre- and post-intervention activities were tabulated for the purpose of comparing the effectiveness of the intervention strategy towards minimisation of errors. This was complemented by data collected through interview. Throughout the current chapter, literature and theoretical frameworks were used to illuminate the data analysis process. In the next chapter are the major findings of the study as well as recommendations.

CHAPTER 6: FINDINGS AND RECOMMENDATIONS

6.1 Introduction

This study problematized that learners have a lot of challenges on representing and interpretation of data variability in statistics. Review of related literature has revealed that no research has been done specifically on nature of errors and misconceptions on representation and interpretation of variability through use of box and whisker plots, frequency polygons as well as ogive. To assist in solving the aforesaid problem, the following research questions were put forward:

1. What are the common errors emerging from misconceptions held by learners in representation and interpretation of variability using frequency polygon, cumulative frequency curve and box plot?
2. To what extent does intervention assist in diminishing the identified errors?

Sociocultural as well as constructivist perspectives informed the study while analytical resources were drawn from Dubinsky (1991) and Arnon et al. (2014)'s APOS theory as well as Kilpatrick et al. (2001)'s Mathematical Proficiency Framework. A constructivist framework of MERs was adopted to remediate the identified errors emanating from misconceptions. In the following section, the main findings of the study are presented.

6.2 Research findings

With respect to the first research question, the study revealed that:

Most errors made by learners emerged as a result of application of irrelevant prior knowledge (conceptual errors) which led to use of unsuitable methods (procedural errors). In some of the learner responses, verbal or non-verbal, the errors which they made were so intricate that it was not easy to classify into either of the afore-stated categories. It also emerged, through this study that, instead of labelling learners as dull, errors and misconceptions are vital cues which teachers should take note of and use as starting points in planning for intervention.

It also emerged that language is a huge barrier to learner understanding of statistical concepts. Whilst it was not conspicuous in written tasks, language proved to be a major stumbling block in expression of mathematical ideas during interviews. Learners struggled in answering questions posed to them by the interviewer even if they had an idea on what was required of them. Without adequate understanding of language, learners can easily misinterpret questions, lose confidence in working out solutions and in verbal engagements with others during group discussions. Undoubtedly, this shuts out any real opportunity for learners to correct existing misconceptions and minimise errors.

With respect to the second research question, the study found that:

After an intervention lesson, directed specifically on how to use multiple methods to represent data and make interpretations, there was a remarkable decrease in errors in post-intervention tasks. Thus, if teachers are adequately empowered on how to effectively plan and deliver intervention lessons directed at specific problems, most of the frequently appearing mathematical errors could be minimised. By using a number of diagrams to represent the same data set, learners were able to overcome the pervasive misconception that a longer box implies more data in the central half of the data set (area misinterpretation). In the same respect, through intervention, learners were able to pair box plots to histograms. This is a good sign that they realised that shorter box on a box plot implies bigger bars on a histogram which, overall, means that data is highly concentrated in that section of data set.

6.3 Limitations of the study

Firstly, the use of 18 learners from the same class, most likely, posed a serious threat to external reliability of the research findings. If resources permitted, this could have been averted by making the sample more diverse, through extending it to learners from other schools in the district. Secondly, conducting the study on one’s own learners, could be another threat to external reliability in that the researcher’s background knowledge of individual subjects might have interfered with objectivity in analysis of errors. Moreover, error mapping done in this study was limited to only two forms namely conceptual and procedural. By analysing learners’ scripts as well as interviewing them, it became evident, however, that learner errors are varied and therefore cannot be reduced to the two aforesaid categories of errors. Mapping of more categories of errors could have yielded more credible outcomes had it been done. Kilpatrick et al. (2001) state that all the five strands of mathematical proficiency are interconnected and must work together if learners are to learn successfully. Each strand of mathematical proficiency should be developed in synchrony with the other. Furthermore, since misconceptions originate from early stages of learning (Smith, diSessa & Roschelle, 1993) and continue to manifest along an instructional program, a longitudinal study on errors and misconceptions in statistical graphs could have been more appropriate. However, since the study was qualitative, this was a single, major mitigating factor against the above stated reliability and validity threats.

6.4 Recommendations

6.4.1 Future researchers

After a single intervention lesson was conducted by the researcher, this study found that there was a remarkable reduction in errors amongst learners, particularly, in items on box and whisker plots. Thus, this study recommends that, in future, large scale studies be conducted on errors and misconceptions, preferably, under experimental conditions to confirm or refute this key finding of the study.

6.4.2 Practice of teaching

According to the findings of this study, generally, errors and misconceptions are unavoidable in every learning endeavour. In this regard, teachers should note that errors and misconceptions present a perfect opportunity for them to reflect on their practice and make good use of pedagogical and subject matter knowledge to maximise learning gains. In other words, a class in which there is no tolerance of learner errors and misconceptions is not conducive for effective learning to take place. Nobody is perfect and learning is a life-long undertaking. From a constructivist perspective, virtually all new knowledge can be traced back to some pre-existing ideas, despite how inappropriate they might be.

6.4.3 Teacher development

Literature in this study revealed that most of the teachers of statistics did not receive formal training in this subject. Successive curricula reviews since the dawn of democracy in South Africa did not prepare teachers to implement these curricula changes. In addition, learners are exposed to a lot of information due to rapid advancement in information and communication technology. All these have a bearing on prior knowledge which, if not properly managed, could lead to development of misconceptions. This implies that teachers must be prepared to assist learners by channelling them towards information which is relevant to their studies. In this regard, this study recommends that in-service training be put in place to provide teachers with tools to tackle some of the aforesaid challenges. Provision of necessary training empowers teachers, in terms of both content and pedagogy, to plan for lessons and deliver confidently. An adequately trained educator does not view errors and misconceptions as obstacles but springboards for learner progression towards higher order concepts like objects and schema. Moreover, educators need to be equipped with current assessment techniques so that learners are assessed at all cognitive levels.

REFERENCES

Aineamani, B. (2015). Teaching algebra with understanding. In S. Maoto, B. Chigonga, & K. Masha (Eds.*), Proceedings of the 21st Annual National Congress of the Association for Mathematics Educators of South Arica* (AMESA), 1, 57-68. Polokwane, South Africa: AMESA.

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*, 131-152.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16* (3), 183- 198.

Ainsworth, S. (2008). The educational value of multiple representations when learning complex scientific concepts. *Theory and Practice in Science Education*, 191-208.

Anselm, S., & Corbin, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, California: Saga Publication.

Anwar, L., Budayasa, I. K., Amin, S. M., & Haan, D. D. (2014). Eliciting Mathematical Thinking of Students through Realistic Mathematics Education. *Journal on Mathematics Education*, *3*(1), 55-70.

Arnon, I., Cottrill, J., Dubinsky, E., Oktac, A., Fuentes, S.R., Trigueros, M. et al. (2014). *APOS theory: A framework for research and curriculum development in mathematics education*. New York, NY: Springer.

Ashlock, R. B. (2002). *Error patterns in computation* (8th ed.). New York: Merill.

Asiala, M., Brown, A., DeVries, D., Dubinsky, E., Mathews, D., & Thomas, K. (1997). *A Framework for Research and Curriculum Development in Undergraduate Mathematics Education*. MAA Notes.

Bakker, A., Biehler, R., & Konold, C. (2004). Should young students learn about box plots. *Curricular development in statistics education: International Association for Statistical Education*, 163-173.

Ball, D., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of teacher education*, *59*(5), 389-407.

Batanero, C., Godino, J., Vallecillos, A., Green, D., & Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematical Education in Science and Technology*, 25(4), 527-547

Bell, J. (2005) *Doing Your Research Project*. Maidenhead: McGraw-Hill

Berger, K.S. (2006). *The developing person through childhood and adolescence* (7th ed.). New York: Worth.

Borasi, R. (1994). Capitalizing on errors as "springboards for inquiry": A teaching experiment. *Journal for Research in Mathematics Education*, *25* (2), 166-208.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, *2*, 155-192.

Capraro, M. M., Kulm, G., & Capraro, R. M. (2005). Middle grades: Misconceptions in statistical thinking. *School Science and Mathematics*, *105*(4), 165-174.

Champagne, A. B., Gunstone, R.F., & Klopfer, L.E. (1985). Consequences of knowledge about physical phenomena. In L. H. T. West & A.L. Pines (Eds.). Cognitive structure and conceptual change (pp. 29-50). New York: Academic.

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2007). The Role of Technology in Improving Student Learning of Statistics. *Technology Innovations in Statistics Education 1*(1): 1-24.

Chisholm, L., Motala, S., & Vally, S. (Eds.) (2003). *South African Education Policy Review*. Johannesburg: Heinemann.

Cohen, L., & Manion, L. (1994). *Research Methods in Education*. New York: Routledge.

Cohen, S., Smith, G., Chechile, R. A., Burns, G., & Tsai, F. (1996). Identifying impediments to learning probability and statistics from an assessment of instructional software. *Journal of Educational and Behavioural Statistics, 21*(1), 35-54.

Cole, M. and Wertsch, J. (1996). Beyond the Individual-Social Antinomy in Discussions of Piaget and Vygotsky. *Human Development, 39*(5), 250-256.

Confrey, J. (1990). *What constructivism implies for teaching. Journal for Research in Mathematics Education. Monograph, 4*, 107 – 210.

Confrey, J., & Kazak, S. (2006). A thirty-year reflection on constructivism in mathematics education in PME. In P. Boero, & A. Gutierrez (Eds.), *Handbook of research on the psychology of mathematics education: Past, present and future* (pp. 305–345). Rotterdam: Sense Publishers.

Cooper, L. L., & Shore, F. (2008). Students' misconceptions in interpreting centre and variability of data represented via histograms and stem-and-leaf plots. *Journal of Statistics Education*, *16*(2), 1-13.

Corbin, J., & Strauss, A. (2008). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, CA: Sage.

Cotterill, N. (2012). Doing Quantitative Research in Education with SPSS. Nurse Researcher, 19(4), 1-11.

Cottrill, J., Dubinsky, E., Nichlos, D., Schwingenndorf, K., Thomas, K., & Vidakovic, D. (1996). Understanding the limit concept: Beginning with a coordinated process scheme. *Journal of Mathematical Behaviour*, *15*, 167-192.

Cox, L. S. (1975). Systematic errors in the four vertical algorithms in normal and handicapped populations. *Journal for Research in Mathematics Education*, 202-220.

Creswell, J. W. (2007). Five qualitative approaches to inquiry. *Qualitative inquiry and research design: Choosing among five approaches*, *2*, 53-80.

Davis, R.B. (1984). *Learning mathematics: The cognitive science approach to mathematics education.* Norwood, N.J: Ablex Publishing Corporation.

De Lange, J. (1996). *Assessment: No Change Without Problems*. The Netherlands: Freudenthal Institute.

DelMas, R., Garfield, J., & Ooms, A. (2005). Using assessment items to study students' difficulty reading and interpreting graphical representations of distributions. In *Fourth Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4)*

Department of Basic Education (2011). *Curriculum and Assessment Policy Statement: Further Education and Training Phase, Grades 10-12.* Pretoria: DBE.

Department of Basic Education (2016). National Senior Certificate Diagnostic Report. Pretoria: DBE

Department of Education (DoE). (2002). *National curriculum statement grades 10-12 (General), Mathematics.* Pretoria: Department of Education.

Department of Education (DoE). (2007). *Revised national curriculum statement grades 10-12 (schools), mathematics: Subject assessment guidelines*. Pretoria: Department of Education.

Donaldson, M. (1963). *A study of children's thinking*. London: Tavistock Publications.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D.O. Tall (Ed.), *Advanced mathematical thinking* (pp. 95-126). Dordrecht: Kluwer.

Edwards, T. G., Özgün-Koca, A., & Barr, J. (2017). Interpretations of Boxplots: Helping Middle School Students to Think Outside the Box. *Journal of Statistics Education, 25*(1), 21-28.

English, L. D., & Halford, G. S. (1995). *Mathematics Education*. Mahwah, NJ: LEA. Chapter 1 (pp. 1-20)

Ernest, P. (1996). Varieties of Constructivism; A framework for comparison. In L.P. Steffe, P. Nesher, P. Cobb, B. Sriraman, & B. Greer (Eds.). Theories of mathematical learning (pp. 335-350). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Ernest, P. (1997). The Epistemological Basis of Qualitative Research in Mathematics Education: A Postmodern Perspective. *Journal for Research in Mathematics Education. Monograph*, 22-177.

Fosnot, C. T. (2005). Constructivism revisited: Implications and reflections. *The Constructivist*, *16*(1), 1-17.

Gall, M.P., Gall, J.P. & Borg, W.R. (2010) Educational Research: An Introduction. 8th Ed. Boston MA: Pearson

Gelo, O., Braakmann, D., & Benetka, G. (2008). Quantitative and qualitative research: Beyond the debate. *Integrative psychological and behavioural science*, *42*(3), 266-290.

Glencross, M. J., & Binyavanga, K. W. (1996). The role of technology in statistics education: A view from a developing country. *Role of Technology, 1,* 301-307.

Govender, V. (2017, February 5). The startling changes to the South African mathematics curriculum since 1994. *BusinessTech online.* Retrieved from https://businesstech.co.za/news/lifestyle/153517/the-startling-changes-to-the-south-african-mathematics-curriculum-since-1994/

Karpov, Y. V. (2003). Vygotsky's doctrine of scientific concepts. In A. Kozulin, B. Gindis, V. Ageyev and S. Miller (Eds.) *Vygotsky's Educational Theory in Cultural Context*. New York: Routledge

Kilpatrick, J. (1987). What constructivism might be in mathematics education. In J.C. Bergeron, N. Herscovics, & C. Kieran (Eds.), *Proceedings of the 11th PME International Conference, 1,* pp.1-27.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). Adding it up: Helping children learn mathematics. Washington DC: National Academy Press. Chapter 4 (pp. 115- 155).

Kozulin, B. Gindis, V. Ageyev and S. Miller (Eds.) *Vygotsky's Educational Theory in Cultural Context.* (pp. 225-245). New York: Routledge.

Leedy, P. (1997). *Practical research, planning and design*. New York: McGraw-Hill.

Leite, M.D., Marczal, D., Pimentel, A. R. & Direne, A.I. (2014). A conceptual framework for error remediation with multiple external representations applied to learning objects. *Journal of Technology and Science Education,4*(3),155-166. Retrievedfromhttp://dx.doi.org/10.3926/jotse.111

Lem, S., Kempen, G., Ceulemans, E., Onghena, P., Verschaffel, L., & Van Dooren, W. (2014). Combining multiple external representations and refutational text: an intervention on learning to interpret box plots. *International Journal of Science and Mathematics Education*, *13*(4), 909-926.

Lem, S., Onghena, P., Verschaffel, L. & Van Dooren, W. (2013). The heauristic interpretation of box plots. *Learning and Instruction, 26*(4), 22-35.

Loosen, F., Lioen, M., & Lacante, M. (1985). The standard deviation: some drawbacks of an intuitive approach. *Teaching Statistics, 7*(1), 2-5.

Mahn, H. (1999). Vygotsky's methodological contribution to sociocultural theory. *Journal for Special Educators*, *20*(6), 341-350.

Makonye, J. P., & Luneta, K. (2013). Learner mathematical errors in introductory differential calculus tasks: A study of misconceptions in the senior school certificate examinations. *Education as Change*, *18*(1), 119-136.

Makonye, J., & Fakude, J. (2016). *A Study of Errors and Misconceptions in the Learning of Addition and Subtraction of Directed Numbers in Grade 8*. SAGE Open, 6(4),

Makonye, J.P., & Nzima, S. (2016). Eliciting Learner Errors and Misconceptions in Simplifying Rational Algebraic Expressions to Improve Teaching and Learning. *International Journal of Educational Sciences*, *12* (1), 16-28.

Mason, J. (1990). *Qualitative Researching*. London, Sage.

McMillan, J.H., & Schumacher, S. (2010). Research in Education: Evidence-based inquiry. (7th ed.). Upper Saddle River, New Jersey: Pearson.

Merriam, B.S. (1998). Qualitative research and case study applications in education. San Francisco, CA: Jossey-Bass Inc.

Methe, S.A., & Riley-Tillman, T.C. (2008). An informed approach to selecting and designing early mathematics interventions. *School Psychology Forum: Research into Practice*, 2, 29-41.

Mevarech, Z.R. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics, 14*(4), 415-429.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*: *A sourcebook.* Beverly Hills: Sage Publications.

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

Mohyuddin, R.G. & Khalil, U. (2016). Misconceptions of Students in Learning Mathematics at Primary Level. *Bulletin of Education and Research, 38* (1), 133 – 162.

Muijs, D. (2011). *Doing Qualitative Research in Education with SPSS*. SAGE: LA.

National Council of Teachers of Mathematics. (1980). An agenda for action: Recommendations for school mathematics of the 1980s. Reston, VA: NCTM.

National Council of Teachers of Mathematics. (1989). Evaluation standards: Curriculum and evaluation for school mathematics. Reston, VA: NCTM.

National Council of Teachers of Mathematics. (1991). *Principles and standards for school mathematics.* Reston, VA: NCTM

National Research Council. (1991) Counting on you: Actions supporting mathematics teaching standards. Washington, DC: National Academy Press.

Nesher, P. (1987). Towards an instructional theory: the role of student's misconceptions. *For the learning of mathematics*, *7*(3), 33-40.

Newbold, P., Calson, W.L., & Thorne, B. (2010). Statistics for Business and Economics (7th Global ed.). New York, Pearson.

Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

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 Statistics Institute and International Association for Statistical Education.

Olivier, A. (1989). Handling pupils' misconceptions. *Pythagoras, 21,* 10-19.

Opie, C. (2004) (Ed.). *Doing Education Research- A Guide to First Time Researchers*. London: Sage Publications.

Orton, A. (1983). Students' understanding of integration. *Educational Studies in Mathematics*, *14*(1), 1-18.

Peng, A., & Luo, Z. (2009). A framework for examining mathematics teacher knowledge as used in error analysis. *For the learning of mathematics*, *29* (3), 22-25.

Piaget, J. (1964). Development and Learning. In Ripple, R.E., & Rockcastle, V.N. (Eds.) *Piaget rediscovered*. Ithaca: Cornell University Press.

Pierce, R., & Chick, H. (2013). The Statistical Literacy Needed to Interpret School Assessment Data. *Mathematics Teacher Education and Development*, *15*(2), n2.

Plumb, C., & Spyridakis, J. H. (1992). Survey research in technical communication: Designing and administering questionnaires. *Technical communication*, 625-638.

Pollatsek, A., Konold, C., Well, A.D. & Lima, S. (1991). Beliefs underlying random sampling. *Memory and Cognition*, *12*, 395-401.

Reddy, V., Visser, M., Winnaar, L., Arends, F., Juan, A. L., Prinsloo, C., & Isdale, K. (2016). *TIMSS 2015*: *highlights of mathematics and science achievement of grade 9 South African learners.*

Resnick, L.B. (1983). Mathematics and science learning: A new conception. Science, 220,477 -478.

Riccomini, P. J. (2005). Identification and remediation of systematic error patterns in subtraction. *Learning Disability Quarterly*, *28*(3), 233-242.

Sarwadi, H. R. H., & Shahrill, M. (2014). Understanding students' mathematical errors and misconceptions: The case of year 11 repeating students. *Mathematics Education Trends and Research*, 1-10.

Schmittau, J. (2003). Cultural-historical theory and mathematics education. In A. Kozulin, B. Gindis, V. Ageyev and S. Miller (Eds.) *Vygotsky's Educational Theory in Cultural Context* (pp. 65-82). New York: Routledge

Scott, D. & Morrison, M. (2005). *Key Ideas in Educational Research.* London, Continuum.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*(1), 1-36.

Shalem, Y., Sapire, I., & Sorto, M. A. (2014). Teachers' explanations of learners' errors in standardised mathematics assessments: original research. *Pythagoras*, *35*(1), 1-11.

Shepard, L.A. (2009). Commentary: Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practices*, *28*(3), 32-37.

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher, 15* (2), 4-14.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics teaching*, *77*, 20-26.

Smith, J. P., Disessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The journal of the learning sciences*, *3*(2), 115-163.

Son, J., & Sinclair, N. (2010). How Preservice Teachers Interpret and Respond to Student Geometric Errors. *School Science and Mathematics, 110* (1), 31-46

Stein, M.K., Smith, M.S., Henningsen, M.A. & Silver, E.A. (1993). *Implementing standard-based mathematics instruction: A casebook for professional development*. Reston: NCTM

Treffers, A. (1991). Didactical background of a mathematics program for Primary Education. In Len Screenland (Ed.), *Realistic Mathematics Education in Primary Schools*. Utrecht: Freudenthal Institute, Utrecht University.

Von Glasersfeld, E. (1989). Cognition, construction of knowledge, and teaching. *Synthese*, 80, 121-140.

Vygotsky, L. (1978). Mind in Society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Vygotsky, L. (1987). R.W. Rieber (Eds.), *The collected work of L.S Vygotsky*. Vol 1. *Problems of general psychology*. New York: Plenum.

Wessels, H. (2008). Statistics in the South African school curriculum: Content, assessment and teacher training. In *Proceedings of the 18th ICMI Study Conference and 2008 IASE Round Table Conference* (Vol. 30).

Whitaker, D., & Jacobbe, T. (2017). Students’ Understanding of Bar Graphs and Histograms: Results from the LOCUS Assessments. *Journal of Statistics Education, 25* (2), 90-102.

White, R. & Gunstone, R. (1992). *Probing Understanding*. London: The Falmer Press.

Williamson, D. (1989). The Box Plot: A Simple Visual Method to Interpret Data. *Annals of Internal Medicine, 110* (11), 916.

Woolfolk, A. (2010). *Educational Psychology*. Upper saddle River, Pearson Educational

Wu, H. (1999). Basic skills versus conceptual understanding. *American Educator, 23*(3) 14-19, 50-52.

APPENDICES

APPENDIX 1: ETHICS CLEARANCE LETTER

Wits School of Education

WITS UNIVERSITY

27 St Andrews Road, Parktown, Johannesburg, 2193 Private Bag 3, Wits 2050, South Africa. Tel: +27 11 717-3064 Fax: +27 11 717-3100 E-mail: enquiries@educ.wits.ac.za Website: www.wits.ac.za

22 September 2017

Student Number: 1532350

Protocol Number: 2017ECE049M

Dear Phaison Jonhera

Application for Ethics Clearance: Master of Education

Thank you very much for your ethics application. The Ethics Committee in Education of the Faculty of Humanities, acting on behalf of the Senate, has considered your application for ethics clearance for your proposal entitled:

Resourcing learner errors and misconceptions in the teaching and learning of statistics at Grade 11 level

The committee recently met and I am pleased to inform you that clearance was granted.

Please use the above protocol number in all correspondence to the relevant research parties (schools, parents, learners etc.) and include it in your research report or project on the title page.

The Protocol Number above should be submitted to the Graduate Studies in Education Committee upon submission of your final research report.

All the best with your research project.

Yours sincerely,

Wits School of Education

011 717-3416

cc Supervisor - Dr Judah Makonye

APPENDIX 2: GDE LETTER OF APPROVAL

GAUTENG PROVINCE
Department: Education
REPUBLIC OF SOUTH AFRICA

8/4/4/1/2

GDE RESEARCH APPROVAL LETTER

Date:	10 July 2017
Validity of Research Approval:	06 February 2017 – 29 September 2017 2017/165
Name of Researcher:	Johnhera P.
Address of Researcher:	3 Moseley Street
	CW 3
	Vanderbijlpark, 1911
Telephone Number:	078 782 0634
Email address:	1532350@students.wits.ac.za
Research Topic:	Resourcing learner errors and misconceptions in the teaching and learning of statistics at Grade 11 level
Number and type of schools:	One Secondary School
District/s/HO	Sedibeng West

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:

12/07/2017

Making education a societal priority

Office of the Director: Education Research and Knowledge Management
7th Floor, 17 Simmonds Street, Johannesburg, 2001
Tel: (011) 355 0488
Email: Faith.Tshabalala@gauteng.gov.za
Website: www.education.gpg.gov.za

APPENDIX 3: INFORMATION SHEET FOR PRINCIPAL AND SGB CHAIR

Dear: **School Principal**
SGB Chairperson

My name is **Phaison Jonhera.** I am a Masters of Education student in the School of Education at the University of the Witwatersrand, Johannesburg.

I am doing research on: **Resourcing learner errors and misconceptions in the teaching and learning of statistics at Grade 11 level.**

In doing the research study, I will initially teach the topic of statistics to Grade 11 learners during afternoon lessons running from Monday to Thursday over a period of two weeks. Each lesson will be 1-hour long. At the end of the series of lessons, participants will complete a questionnaire which will be followed by a short interview.

Since the study will be conducted in the afternoon, I have chosen your school because it is easily accessible for me.

I am kindly inviting your school to participate in this research, as detailed in this letter, and I am very confident that your institution presents opportunities for me to successfully complete this study.

The research participants will not be advantaged or disadvantaged in any way. There are no foreseeable risks in participating in this study. The participants will not be paid for this study.

Pseudonyms will be used to protect the identity of participants as well as that of the school. Your individual privacy will be maintained in all published and written data resulting from the study.

Data gathered in this study is solely for research report and not any other purposes and it will be destroyed within 3 to 5 years after completion of the project.

Yours sincerely,

Phaison Jonhera (Student)
Cell: 0787820634

Dr.Judah Makonye (Supervisor)
Tel: 0117173086
Email: Judah. Makonye@wits.ac.za

APPENDIX 4: INFORMATION SHEET FOR LEARNERS

Dear Learner

My name is **PHAISON JONHERA** and I am a Master of Education student in the School of Education at University of the Witwatersrand, Johannesburg.

I am doing research on: **Resourcing learner errors and misconceptions in the teaching and learning of statistics at Grade 11 level.**

The purpose of my study is to find out errors which learners at Grade 11 level make when representing and interpreting statistical data. I will teach statistics, particularly, use of diagrams to represent data as well as their interpretation during afternoon lessons from Monday to Thursday over a period of two weeks. Each lesson will be 1-hour long. At the end of series of lessons, participants will complete a questionnaire which will be followed by a short interview. Since this process will take place after school, I have chosen your school because it is easily accessible for me. The data gathered in this study will be for research only.

Would you mind giving me help by attending the lessons, after I have completed teaching the series of lessons?

Remember, this is not for awarding marks or financial rewards. Moreover, it is a voluntary exercise, meaning that you are free to discontinue with the study, should you decide to do so along the way.
I will not be using your actual name but will make one up so that no one can identify you. All information about you will be kept confidential in all my writing about the study. Also, all collected information will be stored safely and destroyed between 3-5 years after I have completed my project.

Your parents have also been given an information sheet and consent form, but at the end of the day it is your decision to join me in the study.

I look forward to working with you.

Thank you
Phaison Jonhera (Student)
Cell: 078 782 0634

Dr.Judah Makonye (Supervisor)
Tel: 011 717 3086
Email: Judah. Makonye@wits.ac.za

APPENDIX 5: LEARNER CONSENT FORM

I………………………………………………………. hereby **AGREE**/ **DO NOT AGREE** to participate in the study as detailed below:

Permission to be interviewed
I would like to be interviewed for this study. **YES/NO**

I know that I can stop the interview at any time and do not have to answer all the questions asked. **YES/NO**

Permission to be audio-taped
I would like to be audio-taped for this study **YES/NO**

Permission for questionnaire
I agree to fill in a questionnaire and answer sheet for this study. **YES/NO**

Furthermore, I understand that:

- My name and information will be kept confidential and safe and that my name and the name of my school will not be revealed;
- No harm to me and to my school will arise from my participation in this study;
- I do not have to answer every question and can withdraw from the study at any time without any sanctions being imposed on me;
- I can ask not to be audiotaped at any given time;
- No monetary compensation is offered for my participation in this study.
- All the data collected during this study will be destroyed within 3-5 years after completion of the project.

Sign______________________________ Date__________________________

APPENDIX 6: INFORMATION SHEET FOR PARENTS

Dear Parent

My name is **PHAISON JONHERA** and I am a Master of Education student in the School of Education at University of the Witwatersrand, Johannesburg.

I am doing research on: **Resourcing learner errors and misconceptions in the teaching and learning of statistics at Grade 11 level.**

The purpose of my study is to find out errors which learners at Grade 11 level make, in particular, when representing and interpreting statistical data. Initially, I will teach statistics, particularly, use of diagrams to represent data as well as their interpretation during afternoon lessons from Monday to Thursday over a period of two weeks. Each lesson will be 1-hour long. At the end of series of lessons, participants will complete a questionnaire which will be followed by a short interview. Since this process will take place after school, I have chosen your child's school because it is easily accessible for me. The data gathered in this study will be for no other purposes besides my research report.

Would you mind if your child can help me by attending the afternoon lessons for this research study?

Your child will not be advantaged or disadvantaged in any way. Therefore, she or he can withdraw from the study at any time without being punished in any way. There are no foreseeable risks in participating and your child will not be paid for this study.

Your child's name and identity will be kept confidential at all times and in all academic writing about the study. All research data will be destroyed between 3-5 years after completion of the project.

Yours sincerely,

Phaison Jonhera (Student)
Cell: 078 782 0634

Dr. Judah Makonye (Supervisor)
Tel: 011 717 3086
Email: Judah. Makonye@wits.ac.za

APPENDIX 7: PARENT'S CONSENT FORM

I, ____________________________________ the parent of ___________________________, hereby **GIVE/DO NOT** give permission for my child to participate in the study after school hours, as detailed below:

Permission to be interviewed
I agree that my child may be interviewed for this study. **YES/NO**

I know that he/she can stop the interview at any time and does not have to answer all the questions asked. **YES/NO**

Permission to be audio-taped
I agree that my child may be audio-taped **YES/NO**

I know that he/she can discontinue the audio-tapping anytime without being asked any question **YES/NO**

Permission for questionnaire
I agree that my child may fill in a questionnaire and answer sheet for this study. **YES/NO**

Furthermore, I am aware that:

- My child's name and information will be kept confidential and safe and that my name and the name of my school will not be revealed;
- He/she does not have to answer every question and can withdraw from the study at any time without any sanctions being imposed;
- He/she can ask not to be audiotaped, photographed and/or videotape;
- All the data collected during this study will be destroyed within 3-5 years after completion of this project;
- No monetary compensation is offered for the participation of my child in this study.

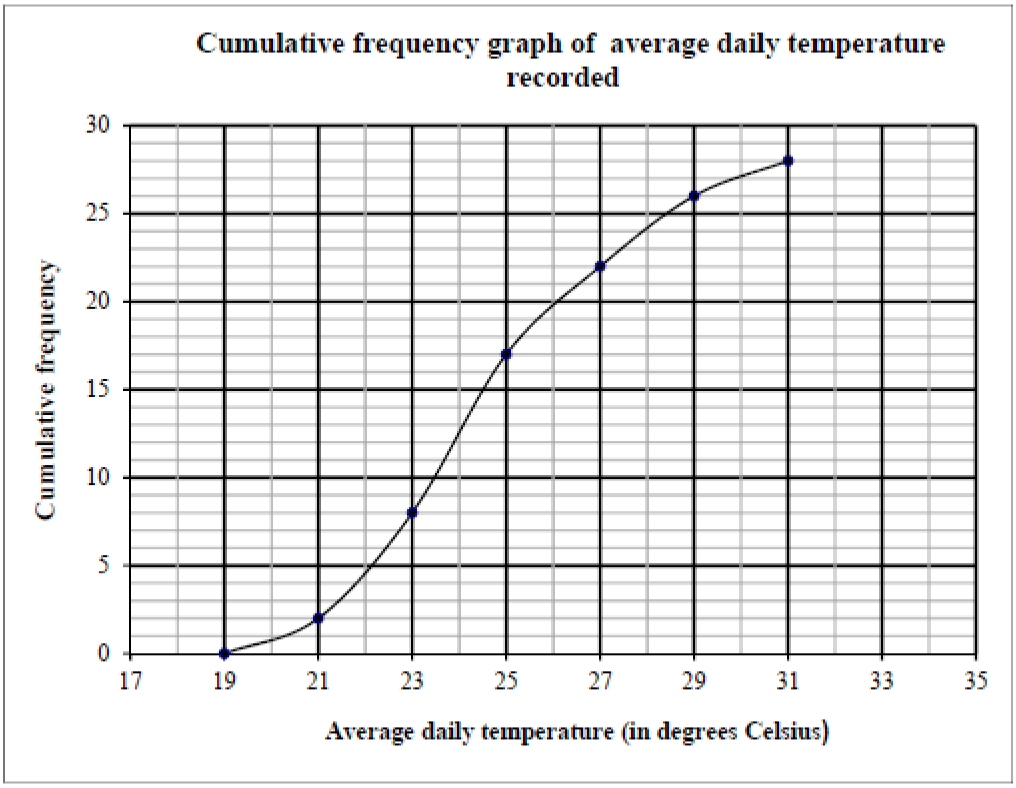
Sign_______________________________ **Date**_____________________________

APPENDIX 8: TASK SHEET FOR LEARNERS

INSTRUCTION: Please **DO NOT** write your name on any part of this task sheet.

QUESTION 1: [CUMULATIVE FREQUENCY, OGIVE & FREQUENCY POLYGON]

The 2012 Summer Olympic Games was held in London. The average daily temperature, in degrees Celsius, was recorded for the duration of the games. A cumulative frequency graph (ogive) of this data is shown below.



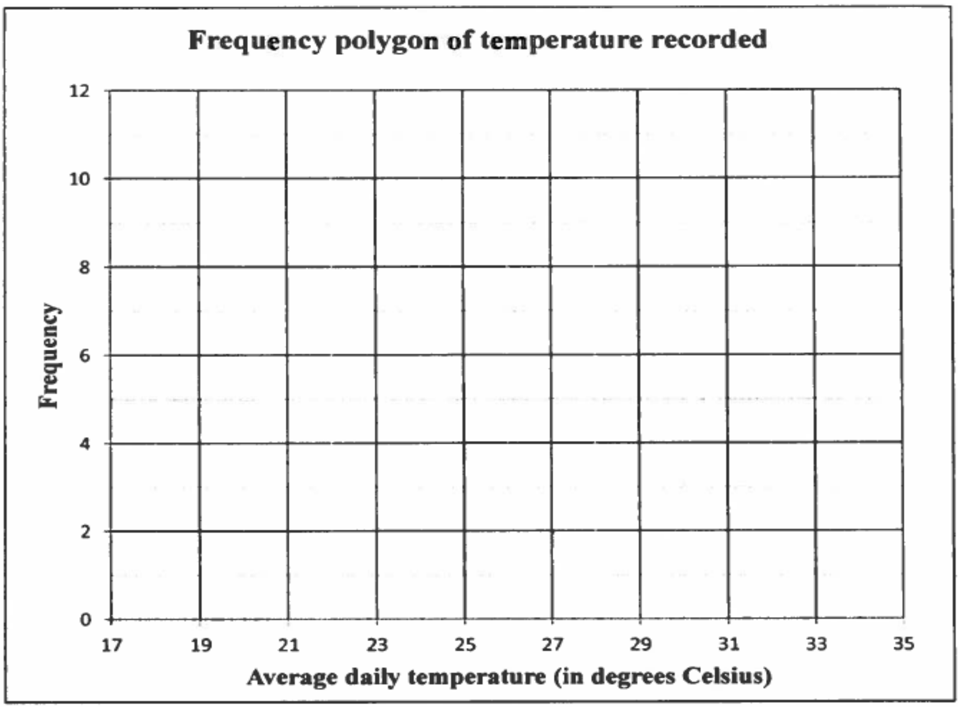
1.1 Over how many days where the 2012 London Summer Olympic Games held?
....

1.2 Estimate the percentage of days that the average daily temperature was less than 24°C?
………………………………………………………………………………………………

1.3 Complete the following frequency table, using the data from the ogive above.

Temperature, T, in degrees Celsius	Frequency
$19 \leq T < 21$	
$21 \leq T < 23$	
$23 \leq T < 25$	
$25 \leq T < 27$	
$27 \leq T < 29$	
$29 \leq T < 31$	

1.4 Hence, use the data in the frequency table to draw a frequency polygon in the grid provided below



<u>QUESTION 2</u>: [BOX & WHISKER DIAGRAM AND APPLICATIONS OF S.D]

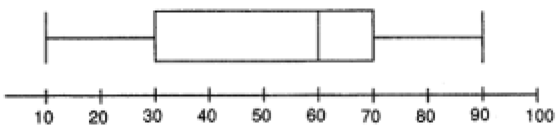
Two schools, M-cee-nai High and Bee Vee High are in a competition to see which school performed better in mathematics in the June Examinations.

The marks of the learners at M-cee-nai High School are recorded below. The box and whisker diagram below illustrates the results of Bee Vee High School. Both schools have 25 learners. (Marks are given in %)

Marks of M-cee-nai High School learners

9	14	14	19	21
23	33	35	37	37
42	45	55	56	57
59	68	75	75	75
77	78	80	81	92

The box and whisker diagram for the Bee Vee High School Learners



2.1 Write down the five-number summaries for each of the two schools.

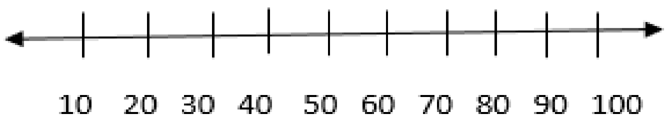
M-cee-nai High School

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………
……………………………………………………………………………………………..

Bee Vee High School

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………..
……………………………………………………………………………………………

2.2 Draw a well-labelled box and whisker diagram that represent M-cee-nai High School marks on the scaled line below.



2.3 Complete the following table:

	M-cee-nai High School	**Bee Vee High School**
I.Q.R		
Skewness		
Outliers (if any)		

2.4 Based on the shape of the box and whisker plots, which of the two high schools performed better? Motivate your answer
………………………………………………………………………………………………
………………………………………………………………………………………………
…….

2.5 What is the range of marks for the middle 50% of learners at Bee Vee High School?
………………………………………………………………………………………………

2.6 How many leaners from Bee Vee High School scored at least 30% in the Preparatory examination? ………………………………………………………………………………

2.7 How many learners at M-cee-nai High School had a mark lying within 1 standard deviation from the mean?
………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………

**** THANK YOU FOR YOUR PARTICIPATION IN THE STUDY****

APPENDIX 9: MODEL SOLUTIONS

MEMO

GAUTENG PROVINCE
REPUBLIC OF SOUTH AFRICA

Tel: 016 5948338
Tel/Fax: 016 594 1287
Moghakasecondary@
gmail.com

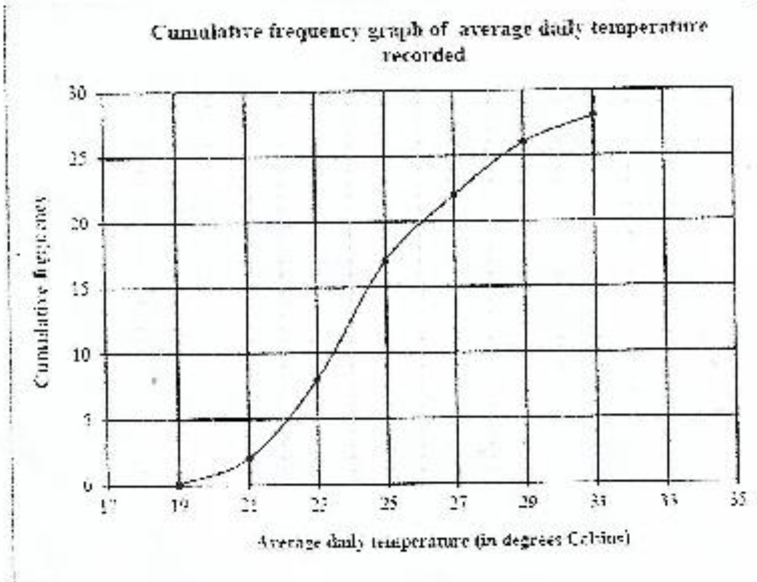
13776 Zone 11
Sebokeng
1983
P.O. Box 96
Sebokeng
1983

QUESTIONNAIRE FOR LEARNER

INSTRUCTION: Please **DO NOT** write your name on any part of this questionnaire

QUESTION 1: [CUMULATIVE FREQUENCY, OGIVE & FREQUENCY POLYGON]

The 2012 Summer Olympic Games was held in London. The average daily temperature, in degrees Celsius, was recorded for the duration of the games. A cumulative frequency graph (ogive) of this data is shown below.



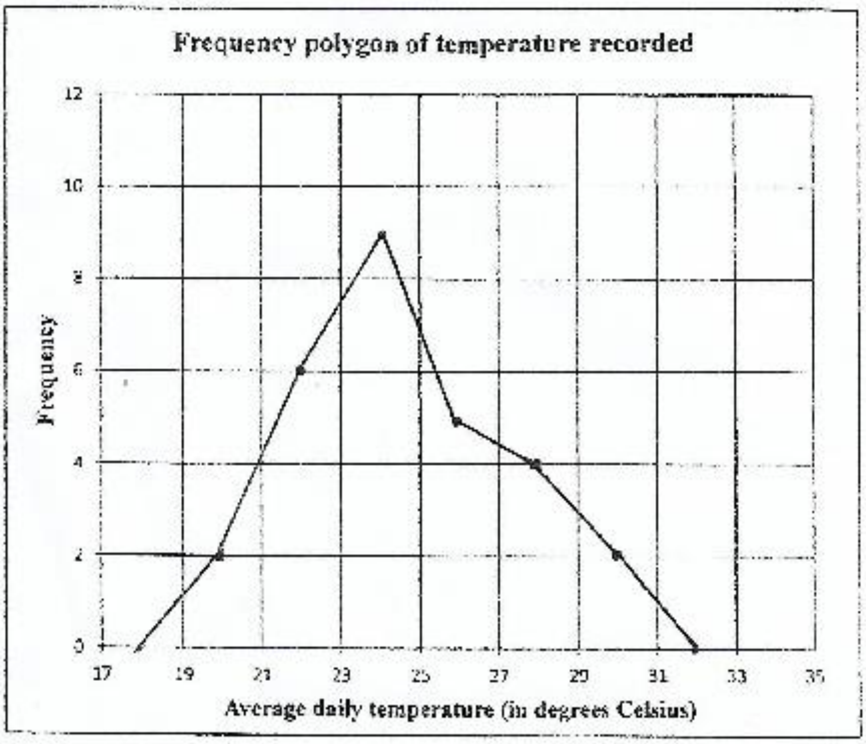
1.1 Over how many days was the 2012 Summer Olympic Games held? ....28........ (1)

1.2 Estimate the percentage of days that the average daily temperature was less than 24°C?
..13/28 x 100 = 46.4%.. (2)

1.3 Complete the following frequency table, using the data from the ogive above. (3)

Temperature, T, in degrees Celsius	Frequency
19 ≤ T < 21	2
21 ≤ T < 23	6
23 ≤ T < 25	9
25 ≤ T < 27	5
27 ≤ T < 29	4
29 ≤ T < 31	2

1.4 Hence, use the data in the frequency table to draw a frequency polygon in the grid provided below (4)



[10]

GAUTENG PROVINCE

Tel: 016 5948338
Tel/Fax: 016 594 1287
Moqhakasecondary@
gmail.com

13776 Zone 11
Sebokeng
1983
P.O. Box 95
Sebokeng
1983

QUESTION 2: [BOX & WHISKER DIAGRAM AND APPLICATIONS OF S.D]

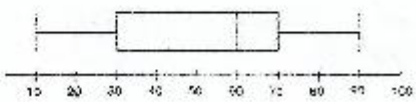
Two schools, M-cee-nai High and Bee Vee High are in a competition to see which school performed better in mathematics in the June Examinations.

The marks of the learners at M-cee-nai High School are recorded below. The box and whisker diagram below illustrates the results of Bee Vee High School. Both schools have 25 learners. (Marks are given in %)

Marks of M-cee-nai High School learners

9	14	14	19	21
23	33	36	37	37
42	45	55	56	57
59	68	75	75	75
77	78	80	81	92

The box and whisker diagram for the Bee Vee High School Learners



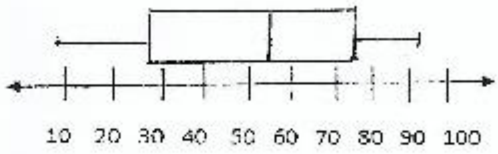
[Adapted from KZN September 2016 P2 Preparatory Examinations]

2.1 Write down the five-number summaries for each of the two schools.

M-cee-nai High School { 9; 28 ; 55 ; 75 ; 92 }
....
.... (1)

Bee Vee High School {10 ; 30 ; 60 ; 70 ; 90}
....
.... (2)

2.2 Draw a well-labelled box and whisker diagram that represent M-cee-nai High School marks on the scaled line below. (3)



2.3 Complete the following table:

	M-cee-nai High School	Bee Vee High School
I.Q.R	2.3.1 75 - 28 = 47	2.3.2 70 - 30 = 40
Skewness	2.3.3 Skewed to the left	2.3.4 Skewed to the left.
Outliers (if any)	2.3.5 28 - 1.5 × 47 = -42.5 75 + 1.5 × 47 = 145.5 No outliers	

(6)

2.4 Based on the shape of the box and whisker plots, which of the two high schools performed better? Motivate your answer

Bee Vee High School performed better. 50% of learners scored more than 60% compared to 55% for M-cee-nai (2)

2.5 What is the range of marks for the middle 50% of learners at Bee Vee High School?

30 to 70 OR 70 - 30 = 40% OR 30 - 70 (1)

2.6 How many leaners from Bee Vee High School scored at least 30% in the Preparatory examination? 75% × 25 = 18.75 ≈ 19 learners (2)

2.7 How many learners at M-cee-nai High School had a mark lying within 1 standard deviation from the mean?

50.28 ± 24.72

(25.56 ; 75). Marks lying within 1 S.d from $\bar{x}$; 33; 35; 37; 37; 42; 45; 55; 56; 57; 59; 68. Therefore there (3) were 11 learners whose marks lied within 1 S.d from mean.

[20]

*** THANK YOU FOR YOUR PARTICIPATION IN THE STUDY****

APPENDIX 10: INTERVIEW PROTOCOL

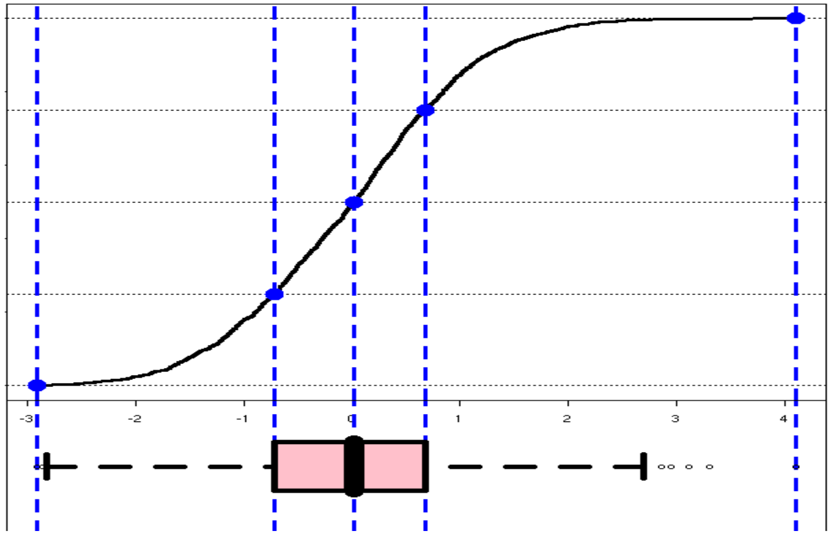
1. Briefly explain you used the ogive to obtain the total number of days over which the 2012 London Olympics were played?
2. How did you read off frequencies from the cumulative frequency curve (ogive)?
3. By considering the shapes of cumulative frequency curve (ogive) and a frequency polygon, do you think there is any link between the two? (Yes/No)
4. If yes, explain briefly.
5. Explain briefly how did you draw the box and whisker plot to represent the data for M-cee-nai High School?
6. What is your own understanding of the concept of inter-quartile range (I.Q.R)?
7. What do you think is its use in statistics?
8. Is it possible to read off number of observations (in this instance, learners) represented on a box plot?
9. If not, provide a possible explanation on why it is not possible to do so.
10. In your own view, how did learner performance at each of the two high schools affect the shape of the box and whisker diagrams?

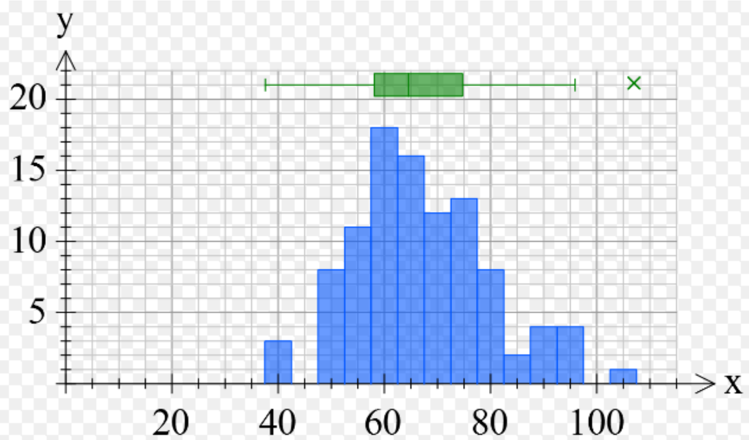
APPENDIX 11: EXEMPLAR INTERVENTION LESSON PLAN

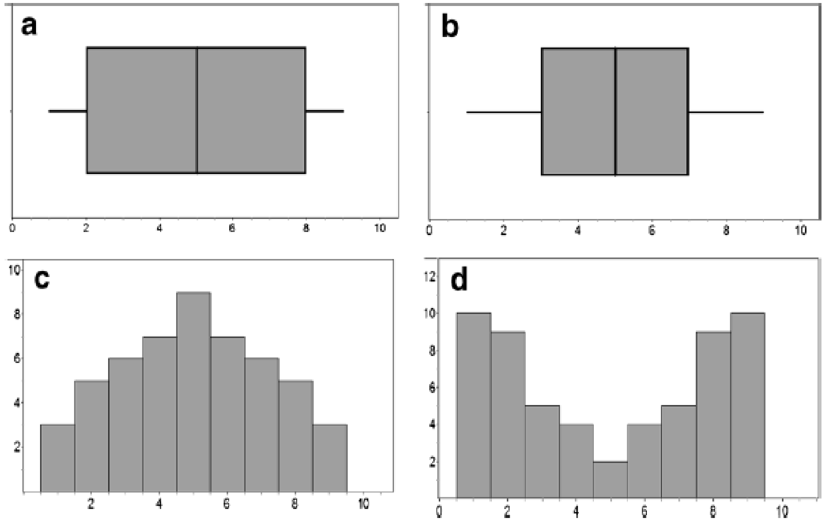
Grade	11	Subject	Mathematics	Topic	Variability of data	Duration	45 mins

Lesson objectives	At the end of the lesson, learners should be able to: • Draw box and whisker plot, ogive and histogram, • Use representations of box plot, ogive and histogram to interpret variability of data in a given context.

Teacher activities	Learner activities	Duration (minutes)	Resources needed
1. Teaching methods: Explanation, Demonstration and Question & Answer **2. Lesson development** **2.1 Introduction** Drawing of box and whisker plot and its interpretation: Remind learners how to quickly draw a box plot in quick, easy steps by playing the video on the link:			• Smartboard, internet (YouTube video)

https://www.youtube.com/watch?v=4M4BV9SLHvM **2.2 Main body (Lesson presentation)** First ask the learners to draw: 2.2.1 an ogive, 2.2.2 box plot, on a single graph paper (as shown in the model solution below). *After verification, demonstrate the solution on the smartboard.* 	**Activity 1** 1.1 Draw a well-labelled cumulative frequency curve from the data given in the table (supplied on the smartboard) 1.2 On this ogive, superimpose a box-plot and indicate all the key points	5 15	• GeoGebra on smartboard or smartphone • graph paper • Pencil • Pen • Grade 11 Oxford Successful Mathematics • Grade 11 Siyavula

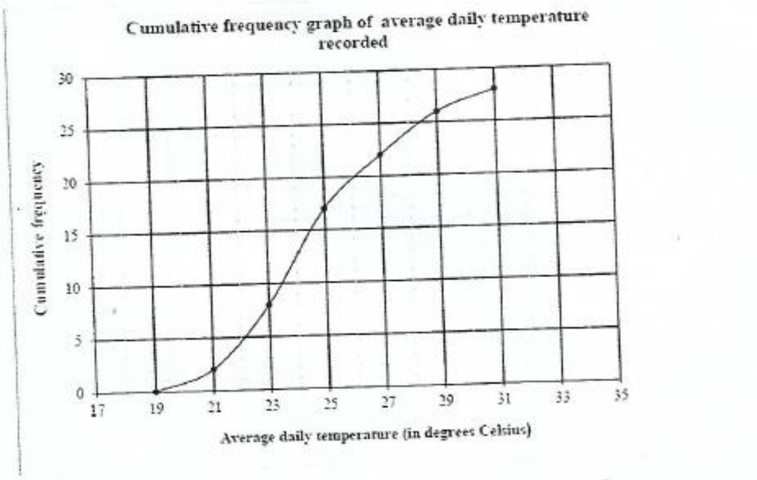
2.3 *The distribution of data is symmetric. All the three measures of location are centrally placed on the box plot.* 3. After demonstration, let learners attempt activity 2 on their own. 3.1  *3.2 The data is skewed to the right because the right whisker of the box plot is longer than the left whisker OR there are shorter bars on the right side of the tallest bar as compared to those to the left side.*	1.3 By comparing the two displays, comment on the distribution of data. **Activity 2** 1.1 Using data supplied on the smartboard, draw a histogram and a box plot on a single diagram. 1.2 Using your histogram as well as box and	15	• Graph paper • Pen • Pencil • ruler

<table>
<tr>
<td>3.3


Generally, it would appear as if (a) matches with (c) since most learners misinterpreted that a larger box represents more data around the median. In fact, (a) matches with (d) because a wider box symbolises low density of data in this interval and the opposite is true</td>
<td>whisker plot, comment on skewness of data, with justification(s).
1.3 Which of the following box plots and histograms represent same data distribution?</td>
<td></td>
<td></td>
</tr>
<tr>
<td colspan="4">Conclusion (2,5 minutes): ……

Reflections/ Evaluation (2,5 minutes): ……………………………………………………………………………………………………..</td>
</tr>
</table>

APPENDIX 12: VIGNETTES FROM PRE-INTERVENTION TASKS

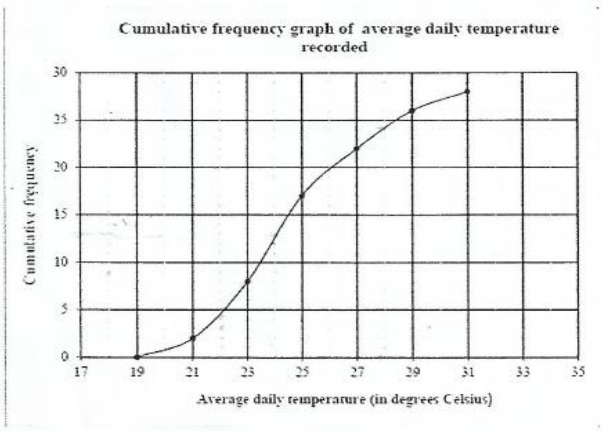


2 1.1 Over how many days was the 2012 Summer Olympic Games held? ..30................ (1)

1 1.2 Estimate the percentage of days that the average daily temperature was less than 24°C?
..................13...days.. (2)

Vignette 2: JPM 13

The 2012 Summer Olympic Games was held in London. The average daily temperature, in degrees Celsius, was recorded for the duration of the games. A cumulative frequency graph (ogive) of this data is shown below.



2 1.1 Over how many days was the 2012 Summer Olympic Games held? .70..days...... (1)

1 1.2 Estimate the percentage of days that the average daily temperature was less than 24°C?
....13.days.. (2)

Vignette 3: JPM 7

2 1.3 Complete the following frequency table, using the data from the ogive above.

Temperature, T, in degrees Celsius	Frequency
$19 \leq T < 21$	2
$21 \leq T < 23$	8
$23 \leq T < 25$	17
$25 \leq T < 27$	22
$27 \leq T < 29$	26
$29 \leq T < 31$	28

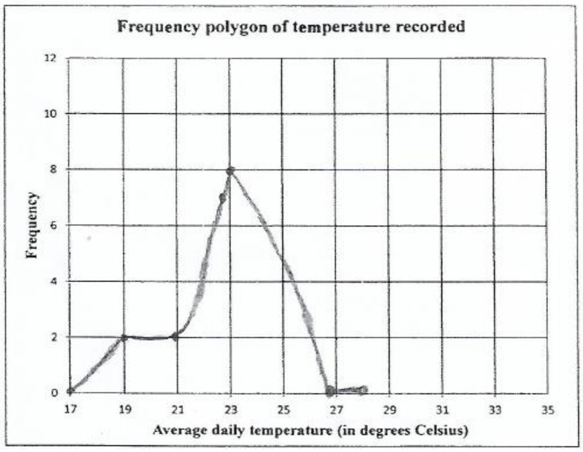
Vignette 4: JPM 3

2 1.3 Complete the following frequency table, using the data from the ogive above.

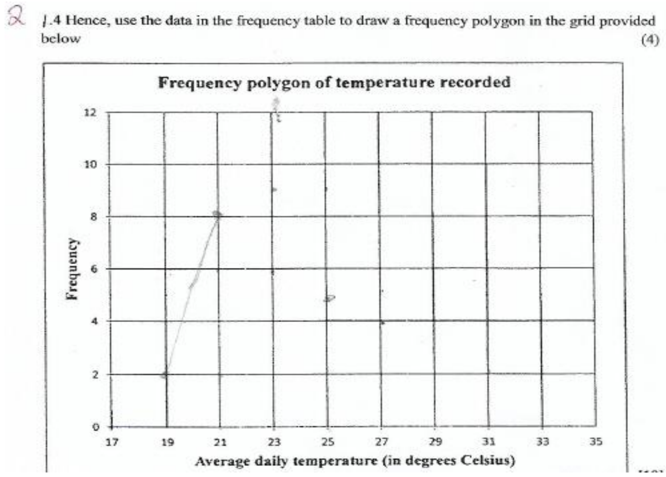
Temperature, T, in degrees Celsius	Frequency
$19 \leq T < 21$	2
$21 \leq T < 23$	8
$23 \leq T < 25$	17
$25 \leq T < 27$	22
$27 \leq T < 29$	26
$29 \leq T < 31$	28

Vignette 5: JPM 13

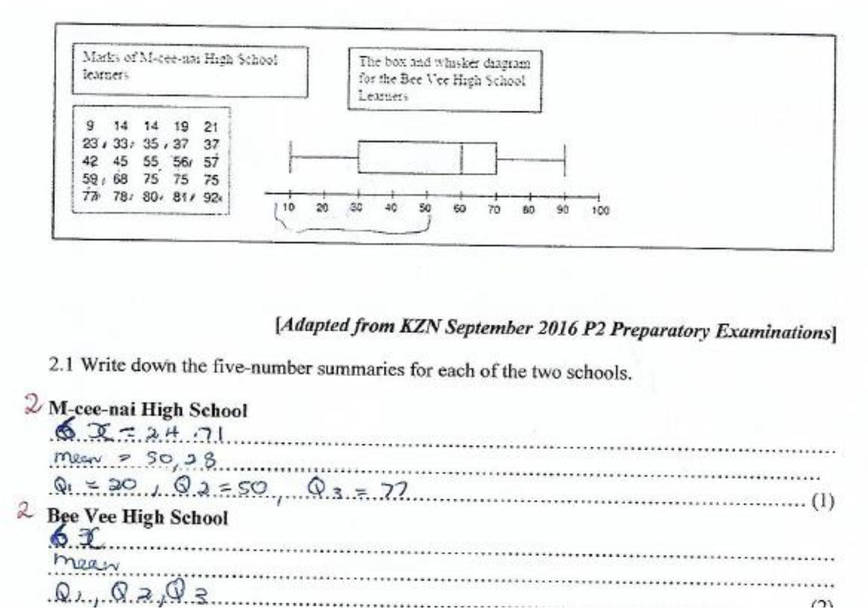
2 1.4 Hence, use the data in the frequency table to draw a frequency polygon in the grid provided below (4)



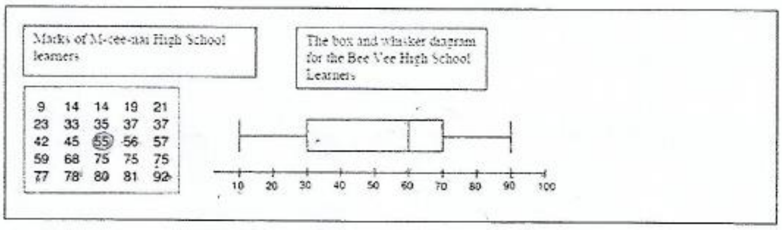
Vignette 7: JPM3



Vignette 8: JPM 13



Vignette 10: JPM 3



[*Adapted from KZN September 2016 P2 Preparatory Examinations*]

2.1 Write down the five-number summaries for each of the two schools.

2 **M-cee-nai High School**

Rage = 92 − 9 = 83 ~~med~~ ....

Mode = 75 ; 14 ; 37 ....

Mean = 55 .... (1)

2 **Bee Vee High School**

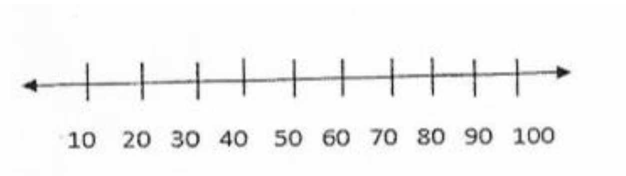
Rage = 90 − 10 = 80 .... Meadin = 60 ....

Mode = non ....

Mean = $\frac{50+60}{2}$ = 55 .... (2)

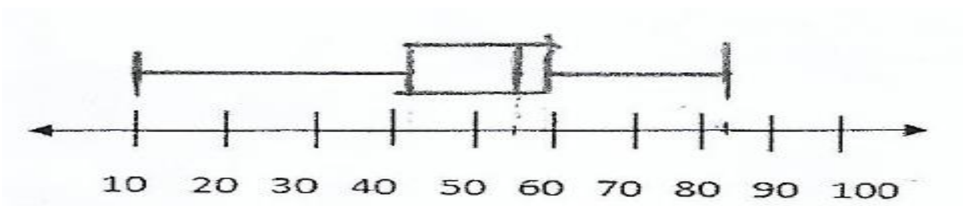
Vignette 12: JPM7

2.2 Draw a well-labelled box and whisker diagram that represent M-cee-nai High School marks on the scaled line below. (3

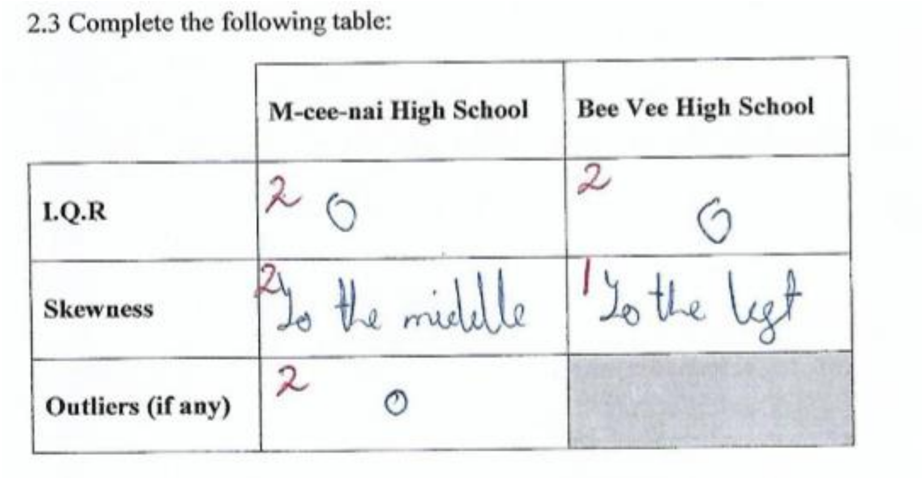


Vignette 13: JPM 3

2 2.2 Draw a well-labelled box and whisker diagram that represent M-cee-nai High School marks on the scaled line below. (3



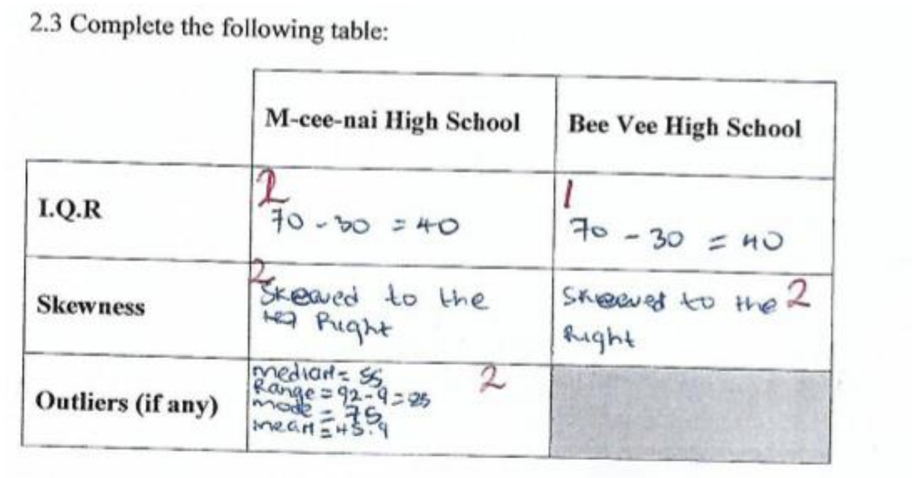
Vignette 15: JPM 7



2.3 Complete the following table:

	M-cee-nai High School	Bee Vee High School
I.Q.R	2 0	2 0
Skewness	2 To the middle	1 To the left
Outliers (if any)	2 0	

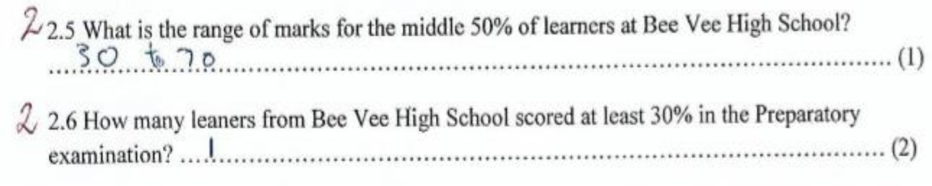
Vignette 16: JPM 3



2.3 Complete the following table:

	M-cee-nai High School	Bee Vee High School
I.Q.R	2 70 − 30 = 40	1 70 − 30 = 40
Skewness	2 Skewed to the ~~left~~ Right	Skewed to the Right 2
Outliers (if any)	median = 55 Range = 92 − 9 = 25 mode = 75 mean = 45.9 2	

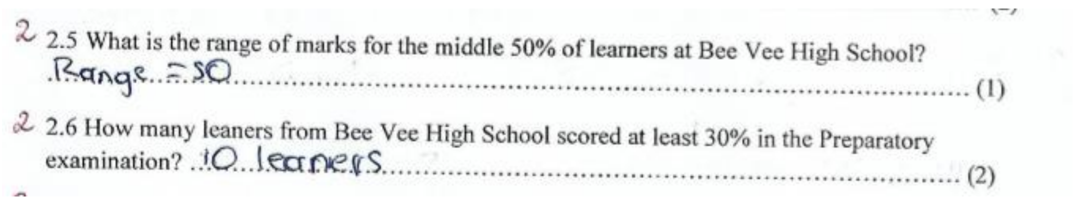
Vignette 17: JPM 13



2 2.5 What is the range of marks for the middle 50% of learners at Bee Vee High School? 30 to 70 .... (1)

2 2.6 How many leaners from Bee Vee High School scored at least 30% in the Preparatory examination? 1 .... (2)

Vignette 22: JPM 3



2 2.5 What is the range of marks for the middle 50% of learners at Bee Vee High School?
Range = 50 .... (1)

2 2.6 How many leaners from Bee Vee High School scored at least 30% in the Preparatory examination? 10 learners .... (2)

Vignette 24: JPM7