CHAPTER 3

3.0 RESEARCH METHODOLOGY AND PROCEDURE

3.1 Introduction

This chapter discusses the research methodology and procedures adopted for collecting data. It starts with a description of the research design, followed by the research method, and ends with an outline of the statistical techniques used to address issues of validity and reliability of the instruments used for the collection of data.

3.2 Assumption of PCK development during classroom practice

It was assumed that competent mathematics teachers would have developed their PCK, which enables them, through classroom teaching, to improve learners' performances at the Senior Certificate Examination over time. Observing the participating teachers prepare and teach a lesson in an assigned topic would enable the researcher to determine how they developed their topic-specific PCK in statistics teaching.

3.3 Research design and method used in this study

3.3.1 Research design

The study adopted a descriptive research design using the case study research method. Descriptive research investigates and describes a case about the current situation of an event or how it has happen in the past (Mayer & Fantz, 2004). It is used to tease out possible antecedents of an event that happened in the past. It is assumed that the competent mathematics teachers have developed adequate PCK, which enables them to improve their learners' performance in the Senior Certificate Examinations over time. A descriptive research design was considered appropriate for the nature of the topic under investigation because this study intends to investigate how the teachers developed their PCK over time.

3.3.2 Research method

This study used a qualitative research approach utilising a case study method. Creswell (2008) defines the case study method as 'an empirical inquiry that investigates a contemporary phenomenon within its real-life context when the boundaries between phenomenon and context are not clearly evident and when multiple sources of evidence are used'. This study sought to investigate how competent mathematics teachers developed their PCK in teaching statistics in their statistics lesson.

There is some criticism of the use of case study research methods. 'Critics believe that a small number of cases cannot offer adequate grounds for establishing reliability or generality of findings' (Yin, 1984). Others feel that intense exposure to the study of a case biases the findings (Yin, 1993 & 1994; Feagin, Orum & Sjoberg, 1991). Some argue that case study research is useful only as an exploratory tool (De Vos, 2000). However, researchers continue to use the method successfully in carefully planned practical studies of real-life situations, issues and problems (Soy, 2006). Soy (2006) argued that successful use of case studies in conducting investigations in scientific studies, despite the criticisms, has many benefits, such as providing a rich and detailed account of the case in a real-life context. The case study was chosen for this research in order to provide a rich and detailed account in a real-life context of how the mathematics teachers develop their PCK in statistics teaching. It is considered adequate and conventional in the field of the author's research interest, as it is used to collect information in order to gain greater insight into and understanding of the way in which PCK may have been developed by competent teachers.

This study is a qualitative one that uses both quantitative and qualitative data. The quantitative data was gathered through the conceptual knowledge exercise for teachers and concept mapping. The participants' performance in these exercises involved their marks (expressed in percentages). Interview schedules, observations of lessons, teacher questionnaires, teachers' written reports, video recordings, and document analysis were used to collect qualitative data. The individual teacher's PCK and its development in data handling teaching/statistics constituted the unit of analysis in this study.

3.4 Population and sample description

3.4.1 Study population

The population of the study comprised Grade 11 mathematics teachers in Tshwane North District, Gauteng, South Africa. There are twelve high schools in Tshwane North District. With a criterion of 70% for learners' performance in the Senior Certificate Examination in Mathematics for a period of two years, seven schools were identified from which the participating teachers were selected. The identification of the schools was followed by interviews with the principals, peers and subject specialists at the Department of Basic Education (DoBE) to identify the willing participating teachers.

3.4.2 Study sample

The teachers in the main study were selected, through a process of elimination, according to certain criteria: learners' performance in mathematics in the Senior Certificate Examination; recommendations by school principals, subject specialists at the Department of Education and peers; and competence in statistics through performance in a statistics test. Tshwane North Education District Cluster 3, Gauteng Province, comprises twelve schools. Of these schools, only seven had scored a minimum of 70% mathematics pass rate for two consecutive years in the Senior Certificate Examination. Mathematics teachers from these schools were invited to volunteer for the project. Six teachers from six separate schools indicated their willingness to participate. The researcher requested recommendations from principals, peers and subject specialists from the Department of Basic Education (DoBE) for these teachers. Based on their recommendations, six teachers were selected. Finally, the six teachers wrote the conceptual knowledge exercise in statistics. The top four scorers were selected for the main study. Table 3.1 summarises their performances, and their demographic profiles are described in section 4.3.

Table 3.1: Schools and teachers that participated in the main study

S/NO	SCHOOL	NSC RESULTS	TEACHER
1	School A	81%	Teacher A
2	School B	94%	Teacher B
3	School D	93%	Teacher D
4	School E	98%	Teacher E

3.5 Research instrument used for collecting data

3.5.1 Development of research instruments

3.5.1.1 Teacher conceptual knowledge exercise in statistics

The conceptual knowledge exercise was adopted to collect data in this study.

The National Curriculum Statement for Mathematics for the Senior Phase of the Further Education Training (FET) bands for Grades 10–12 and the prescribed textbooks were reviewed and analysed. The aim was to ascertain the targeted knowledge, competence and skills for developing the test items based on the mathematics assessment taxonomy. A large number of multiple-choice test items were initially formulated by the researcher from sources such as public examinations, locally prepared past examinations and tests, selection tests, achievement tests and textbooks in mathematics. The items were designed in line with Bloom's Taxonomy and the South African Mathematics Assessment Taxonomy, as indicated in the examination guidelines of the NCS (DoE, 2008) and Table 3.2. The competencies tested according to Bloom's Taxonomy included knowledge, comprehension, analysis, synthesis, application and evaluation (DoE, 2010). The levels of the mathematics assessment taxonomy are knowledge (level 1); applying routine procedures in familiar contexts (level 2); applying multi-step procedures in a variety of contexts (level 3); and reasoning and reflecting (level 4) (DoE, 2010). Comprehension and application of Bloom's Taxonomy were used to design the conceptual knowledge exercise, in line with the mathematics assessment taxonomy. The mark allocation was the total mark allocated to all items that were developed according to levels. For instance, all marks allocated to level 1 questions that test knowledge in any mathematics test or examination must not exceed 20 out of the total mark of 100 for the examination or test.

Table of specification 3.2: Mathematics assessment taxonomy and marks allocation

LEVELS OF ASSESSMENT	ASSESSMENT TAXONOMY	MARKS ALLOCATION
1	Knowledge	20
2	Applying routine procedures in familiar contexts	25
3	Applying multi-step procedures in a variety of contexts	30
4	Reasoning and reflecting	25

(DoE, 2010)

Table of specification 3.3: Showing competency and skills and marks allocated

COMPETENCE	ABILITIES	SKILLS DEMONSTRATED	QUESTION	MARKS ALLOCATED	TOTAL
Comprehension (understanding)	Applying routine procedures in familiar contexts	Grasping (understanding) the meaning of informational concept/materials	1, 2, 3, 6, 11, 13, 15, 20	5 for each item	40
Applications	Applying what was learnt in the classroom in solving problems in familiar or other situations by using routine, multi-step procedures	Solving problems using required skills or knowledge	4, 5, 7, 8. 9, 10, 12, 14 16, 17, 18, 19	5 for each item	60
				TOTAL	**100**

(DoE, 2010)

The conceptual knowledge exercise included 40% of the questions designed to test comprehension and consisted of level 2 and 3 questions in statistics (ref Table of specification 3.2). Examples of items measuring comprehension knowledge are 1, 2, 3, 6, 11, 13, 15 and 20 (ref Table of specification 3.3). Below is an example of the levels 2 and 3 questions.

Use the frequency distribution table below to answer question 2

Interval	*0-4*	*5-9*	*10-14*	*15-19*	*20-24*
Frequency	*3*	*5*	*7*	*4*	*1*

2 *Estimate the mode of the distribution*

The remaining 60% tested application knowledge at levels 3 and 4, where participants had to apply higher-order thinking to solve problems in statistics (ref Table of specification 3.2). Examples of items measuring application of knowledge are 4, 5, 7, 8, 9, 10, 12, 14, 16, 17, 18 and 19. The question below is an example of levels 3 and 4 questions.

4 *The mean height of three groups of students consisting of 20, 16 and 14 students is 1.67m, 1.50m and 1.40m respectively. Find the mean height of all the students.*

The conceptual knowledge test was designed to determine how well the teachers could demonstrate that they had adequate content knowledge of the topic by applying routine and multi-step procedures, as well as reasoning and reflection. Initially 30 multiple-choice test items were developed in statistics from the sources indicated, each with five possible responses. Only one of the five options was correct. These items were scrutinised by mathematics experts at the DoBE, and national examiners in NCS mathematics (ref Appendix XXII). The responses from the reviewers were used to modify the test items that formed the first draft of the instrument. For example, item 4 asked, 'The mean heights of three groups of students consisting of 20, 16 and 14 students are 1.67 m, 1.50 m and 1.40 m respectively. What is the mean height of all the students?' The item was modified to 'Find' instead of 'What', as previously used in the question.

- **Scoring the test items**

One mark was allocated to each item. The total mark for the 20 items was therefore 20 marks. While the comprehension part of the question was 8 marks, the application part was 12 marks. For the correct answer to each question, one mark was awarded in both the comprehension and application parts of the question. The marks were later converted to 100 marks. Selection of participants for the concept map and qualitative aspect of the research was based on performance in the conceptual knowledge exercise. A teacher had to score a minimum of 70% to be adjudged to have adequate subject matter content knowledge of statistics in school mathematics.

3.5.1.2 Concept mapping for teachers

The NCS was used to compile the list of contents of statistics in school mathematics. The topics according to the NCS for Grades 10 to 12 are stem-and-leaf; mode, median and mean of ungrouped data; frequency table of grouped data; range, percentiles, quartiles; inter-quartiles and semi-quartile range; bar and compound bar graphs; histograms; frequency polygons; pie charts; line and broken line graphs; box-and-whisker plots; variance, mean deviation; standard deviation; ogives; five number summaries; scatter plots; lines of best fit (DoE, 2010) (ref Appendix XXIV).

The participating teachers were required to use the topics listed above to construct a concept map. The question states:

(a) *Arrange the topics in each grade on how you think they should be taught in grades 10, 11 and 12.*

(b) *With an arrow, show how you can teach these topics sequentially in each grade. For example, you observe morning before afternoon and before evening. Therefore;*

Morning → *afternoon* → *evening*

For example, in measures of central tendency, the mode is taught first, followed by the median and the mean. Therefore, the memorandum for question (a) should be:

Table 3.4: Table showing the list of statistics taught in grades 10, 11 and 12 (if any)

GRADE 10	***GRADE 11***	***GRADE 12***
Mode, median, mean, ranges, (ungrouped data), frequency table, bar and compound bar graphs, histogram, frequency polygons, pie charts, line and broken line graphs. mode, median and mean (grouped data), quartiles, inter-quartiles and semi-inter-quartile range	*Five number summary, box and whisker diagrams, ogives, variance and standard deviation, scatter diagrams, lines of best fit*	*N/A*

An example of how question (b) should be answered for grade 10 is:

Mode → *Median* → *Mean Ranges* → *(Ungrouped data)*
Frequency table Bar → *and Compound bar graphs* → *Histogram* → *Frequency* →
Polygon → *Pie Charts Line and broken line graphs.*
Mode → *Median* → *Mean (Grouped data)*
Quartiles → *Inter-quartile and semi-inter-quartile ranges*

- **Scoring of concept mapping**

A rubric was designed by the researcher to indicate how to evaluate the concept map drawn by the participants. It allocated marks to the number of topics that were correctly arranged, and deducted marks for incorrect arrangement of topics (ref Appendix XXV).

As indicated in Appendix XXV, marks were allocated for the number of topics that were correctly arranged, and deducted for incorrect arrangement of topics in each grade. The mark allocation for the concept mapping exercise was 25 marks in each grade for question a). The combined mark for Grades 10 and 11 for question a) was 50. Question a) requested the participating teachers to 'Arrange the topics in each grade on how best they can be taught in Grades 10, 11 and 12'. No mark was allocated for Grade 12, as the topic is not taught in that grade. The same scoring system was applied to question b), in which the participants were requested to 'With an arrow, show how you can teach these topics sequentially in each grade. For example, you observe morning before afternoon and before evening. Therefore, in a sequential order, it is

Morning → Afternoon → Evening. A teacher who scored less than 60 marks could be regarded as not having the knowledge of the curriculum that would inform his or her insight into the topic. The reason for allocating the same mark is that each question required approximately the same time to solve.

3.5.1.3 Interview schedule for teachers

The purpose of the semi-structured interview was to gain some insight into mathematics teachers' content knowledge and educational background that may have enabled them to develop their topic-specific PCK in statistics. The semi-structured interview schedule was based on several literature sources on PCK (e.g. Jong, 2003; Jong et al., 2005; Van Driel et al., 1998; Rollnick et al., 2008)). To this end, questions were developed to address the teachers' teaching experience, qualifications, educational background and professional development, knowledge of instructional strategies, and preconceptions in teaching and learning statistics. The questions were grouped according to the components of PCK being assessed in this study. This approach has been used by several researchers (Jong, 2003; Jong et al., 2005; Van Driel et al., 1998; Rollnick et al., 2008) in the fields of mathematics and science education. The distribution of the questions is shown in Table 3.5. The questions are indicated in Appendix XXVI.

Table of specification 3.5: Item specification table for the interview

PCK components	Subject matter content knowledge	Instructional skills and strategies	Learning difficulties	Workshop
Number of items	1–9	10–13	14	15–20

Questions 1 to 9 were used to assess the teachers' subject matter content knowledge and demographic profile in statistics teaching. For example, question 1 asked:

> '*What university/college did you attend?*'

They were then asked to indicate the course they had studied in their disciplinary education programme and their understanding of the nature of statistics in school mathematics.

Questions 10 to 13 probed the instructional strategies that they used for teaching statistics and why they employed these strategies. For example, in question 12, participants were asked:

> *'If the learners have any problem in understanding the topic based on the instructional approach, what do you do to help them to understand?'*

Question 14 was used to determine the learning difficulties that teachers themselves think learners have about the topic. For example, the teachers were asked:

> *'What learning difficulty do you remember experiencing as a pupil and as a university student or from teaching experience in statistics?'*

Questions 15 to 20 focused on workshops that the teachers had attended. For instance, the teachers were asked:

> *'Have you ever been to a mathematics workshop or teacher development programme?'*

The data related to workshops were used to triangulate data on teachers' content knowledge.

Prior to the validation of the teacher structured interview schedule, it was given to three secondary school Grade 11–12 mathematics teachers for comments about the categories and educational background for developing PCK. Their comments were used to review the questions before the pilot study.

3.5.1.4 Lesson observation schedules

The lesson observation schedules (ref Appendix XXIX) were standard ones recommended by the Provincial Department of Education for normal classroom practice (DoE, 2010). The schedule was therefore adopted for gathering data for assessing instructional knowledge used in teaching statistics, which is the major focus of this study. The purpose of using the standard lesson observation schedule was to collect data from real-life situations and to assess how well the teachers prepared for lessons, as well as to check for consistency in their implementation of plans (Vistro-Yu, 2003 & DoE, 2010) (ref Appendix XXIX).

3.5.1.5 Teacher questionnaire

The teacher questionnaire was designed to assess teachers' PCK in terms of their knowledge of instructional skills and strategies, learners' conceptions in teaching and learning statistics, and learning difficulties. The teacher questionnaire (ref Appendix XXVIII) consisted of 16 questions designed to triangulate data collected during lesson observation. Questions 1 to 9, 12, 13, 15, and 16 were used to assess the instructional strategies that the teachers used in classroom practices in statistics teaching. An example of the questions focusing on instructional skills and strategies is:

> *How did you identify the prior knowledge (preconceptions) which the learners bring to the class about statistical graphs?*

Questions 10, 11 and 14 were used to determine the learning difficulties that learners have with the topics in statistics teaching (ref Table 3.6) (ref Appendix XXVIII). An example of the questions is:

> *What is it about statistics that makes the learning easy or difficult?*

Table of specification 3.6: Item specification table for the questionnaire

PCK components	Instructional skills and strategies	Learning difficulties
Number of items	1–9, 12, 13, 15, 16	10, 11, 14

The questionnaire focused on what the teachers actually does while teaching, namely their strategies or approach and methods (items 7–11, 15–16) and contents of the lessons (item 2). Other information related to how the teacher identified learners' preconceptions and learning difficulties (items 4–6, 10, 17), how these difficulties were resolved (items11, 12, 14), and how the lessons were evaluated (items 13, 15 and 16) (ref Appendix XXVIII). As regards teachers' instructional strategies and skills, participants were requested to indicate the duration of the lesson, topic, and essential prior knowledge (ref Appendix XXVIII). In addition, participants were requested to indicate how learners responded to the class activities, homework and assignments (ref Appendix XXVIII). For instance, the teachers were asked, 'How did learners respond to class activities, homework and assignments?' Knowledge of learners' conceptions and learning difficulties was assessed by asking the teachers to indicate how they identified learners' preconceptions and misconceptions, if any, as well as learning difficulties in the context of teaching (ref Appendix XXVIII). For example, the participating teachers were asked, 'How did you identify the prior knowledge (preconceptions) that the learners bring to the class about statistical graphs?' Table 3.6 displays how the questions were distributed according to the various components of PCK, namely instructional strategies and learning difficulties, and how the components were assessed. The questionnaire was administered to the participants immediately after the last lesson had been observed.

3.5.1.6 Teacher written reports

The teachers' structured written reports (ref Appendix XXVII), in which they recorded what made the lessons easy or difficult, were used to assess instructional strategies and learners' learning difficulties after a four-week period of teaching statistics. The purpose of the teachers' written reports was to determine what (for the teacher) made the lessons easy or difficult, and to triangulate other data related to how the teachers developed their PCK over time. The written reports were compiled from teachers' and learners' portfolios, as well as learners' workbooks. For instance, the participating teachers were asked, 'How did learners respond to classroom activities as well as homework or assignments?'

The teachers' portfolios contained information such as a formal programme of assessment in mathematics for Grade 11, mathematics assessment tasks (standardised tests, assignments, investigations or projects and examination papers), tools for assessments (memoranda, checklists, rubrics, etc), and model answers for all assessment tasks. The learners' portfolios contained continuous moderation reports, a summary of marks, tests, examinations, and assessments (DoE, 2010).

Table of specification 3.7: Item specification table for the written reports

PCK components	Instructional skills and strategies	Learning difficulties
Number of items	5 and 6	1–4 and 7-9

Nine questions were formulated as guidelines for the teachers in compiling the report. Questions 1 to 4 and 7 to 9 were used to examine learning difficulties, and questions 5 to 6 were used to determine instructional skills and strategies (ref Appendix XXVII). An example of an item focusing on learning difficulties is:

What learning difficulties do you identify in learners when teaching statistical graphs?

An example of questions focusing on instructional skills and strategies is:

How did the learners respond to classroom activities as well as homework or assignments?

The reports were given to experienced mathematics teachers in Grades 11 and 12, who were asked to comment on the questions guiding the report for normal classroom practice (ref Appendix XXVII). Their comments were used to review the report guidelines before use in the pilot study. For example, comment on every task in statistics was checked, marked, had comments and suggestions for motivation and improvement to any learning difficulty that learners might have encountered.

3.5.1.7 Document analysis

In this study the documents analysed in terms of teachers' compliance with curricular recommendations for teaching and learning school statistics were the learners' class workbooks, learners' and teachers' portfolios, and the NCS for mathematics. The purpose of the analysis was to triangulate the data, using the teacher interviews, questionnaires, lesson observation and written reports on how teachers developed their PCK in statistics teaching. At the end of the four weeks' teaching, these documents were made available to the researcher.

The learners' workbooks contained completed, written classwork, homework, and remedial work. Teachers' portfolios for example included work samples and reflective commentary by the teachers as to what had made the lesson easy or difficult, and intervention strategies adopted to address learners' learning difficulties, if any (ref Appendix XXI).

The NCS policy documents gave an indication of whether the teachers were adhering to policy recommendations for teaching and learning, such as the work schedule to be used for teaching statistics according to grade, resources, and assessment plans. It is assumed that a teacher with adequate knowledge of the curriculum would be able to design good teaching strategies in line with the curricular goals. In practice, this requirement meant checking for consistency in the implementation of lesson plans according to the NCS.

3.5.1.8 Video recording

The purpose of the video recording was to record the teachers' teaching (lessons), which would enable the researcher to triangulate the data collected from the lesson observations. The duration of the lessons observed ranged from 40 to 45 minutes for each of the eight lessons. The transcribed protocols (ref Appendix V-XII) were used to gain insight into teachers' content knowledge and how it was used, including the instructional strategies demonstrated in the lessons on statistical graphs.

3.6 Validation of the research instruments

Validity tells us whether an instrument measures or describes what it is supposed to measure or describe. It means that whatever scores were obtained from the instrument should make sense, be meaningful, and enable the researcher to draw conclusions from the sample of the population under investigation (Creswell, 2008). The test validity of an instrument could involve construct validity, content validity, and criterion validity (Creswell, 2008). In this study, content validity was chosen to validate the test instrument (conceptual knowledge exercise). The purpose was to determine whether the test covered the content of the domain that it was supposed to measure. The instrument was meant to assess the subject matter content knowledge in statistics (the domain) that the selected mathematics teachers possessed, which, it was assumed, enabled them to develop PCK. The other instruments such as the concept map exercise and semi-structured interview schedule were validated as follow.

3.6.1 Validity and reliability of the concept map

The purpose of the concept mapping exercise (ref Appendix XXIV) was to assess the participating teachers' knowledge of the school statistics curriculum. In this study, although the major instruments used for assessing teachers' school statistics content knowledge were the statistics conceptual knowledge exercise and teacher lesson observation, the concept map exercise was further used as an addendum to that assessment. A concept map is a viable means of gathering information on a person's conceptual knowledge of a topic (Novak & Canas, 2006). The concept mapping exercise required the participating teachers first to list the given school statistics topics according to the grades for which those topics are taught, namely Grades 10, 11 or 12; and second to arrange them in the order in which they should be taught in a conceptually logical and sequential fashion. The assumption was that ability to arrange the topics for teaching in a hierarchical manner for each grade level provided an indirect indication that the teachers had adequate knowledge of the statistics topics in the mathematics curriculum and the conceptual relationships among them.

A given set of criteria was used by a mathematics specialist from the Department of Education and two university lecturers in the Mathematics Education Department (ref Appendices XXIV and XXV) to content validate the concept map exercise and memorandum. First, the experts had to ascertain whether the concept map exercise would allow the mathematics teachers to list the topics according to Grades 10, 11 and 12 and arrange them in logical order, such that one topic formed the basal knowledge of the next for each of those grades. Second, they were required to ascertain whether the memorandum (expected answers to the concept mapping exercise) was appropriate for answering the concept mapping exercise. The experts' responses (mathematics specialist from the Department of Education and two university lecturers in the Mathematics Education

Department) showed unanimous agreement that the concept map exercise contained adequate information for assessing teachers' content knowledge of the statistics topics in the various grades and the ways in which they should be taught in a logical and sequential order. In addition, all the raters agreed that the memorandum was adequate and appropriate for assessing the concept mapping exercise.

The reliability of the concept map was determined as follows. The concept map exercise and memorandum were given to four school mathematics teachers that did not participate in the study and who were physically located outside the study site to avoid contamination. There were consistencies in the responses of the mathematics teachers with the anticipated answers (memorandum) of the concept mapping exercise. In other words, the responses of the respondents (mathematics teachers) were consistent with the idea of listing the statistics topics according to grade and the way in which they should be taught in a logical hierarchical and sequential order. The consistency in the responses of the teachers indicated that the concept mapping exercise is reliable enough for assessing the teachers' knowledge of statistics in the school mathematics curriculum (Bush, 2002; Barriball & White, 2006). Where necessary, their responses were used to review the concept mapping exercise and memorandum before they were used for the main study.

3.6.2 Validity and reliability of the interview schedule

The purpose of the semi-structured interview (ref Appendix XXVI) was to assess the educational backgrounds that may have enabled the mathematics teachers to develop their assumed topic-specific PCK in statistics (Jong, 2003; Jong *et al.,* 2005; Van Driel *et al.,* 1998; and Rollnick *et al.,* 2008).

The schedule was validated by a mathematics expert in the Department of Education and two mathematics education specialists from a university, using a specific set of criteria. The raters were requested to establish whether the interview schedule contained appropriate information to determine teachers' mathematics educational background for developing PCK as defined in statistics teaching (ref Appendix IV). Their responses showed unanimous agreement that the schedule contained the necessary information for assessing how the participating teachers developed their topic-specific PCK (Bush, 2002; Barriball & White, 2006).

To ascertain the reliability of the interview schedule, it was used with some school mathematics teachers who were not participating or involved in the study. The interest was in determining the extent to which the schedule was likely to yield consistent responses from them (Bush, 2002) in terms of assessing the mathematics teachers' educational background that may have enabled them to develop their topic-specific PCK in statistics teaching (ref Appendix XXVI). The responses of the pilot teachers were identical and consistent in terms of the items selected for the interview schedule. The reliability of the instrument was thus generally assured and, where necessary, the respondents' comments were used to review the schedule.

3.7 Pilot study

Purpose of the pilot study

The purposes of the study were:

- To test the validity and reliability of the test instruments
- To test the logistics feasibility for administration of the instruments
- To improve the design of the research instruments and methodology for the administration of the main study
- To check that the instructions given to investigators were comprehensible
- To check the timing for the administration of the instruments

3.7.1 Subjects used in the pilot study

The subjects used in the pilot study were two willing mathematics teachers at high school level who did not participate in the main study. They were selected from their schools' performance in mathematics in the Senior Certificate Examination for at least two years. The participating teachers had taught higher grade or optional mathematics for a minimum period of three years. One of the participants had a BSc (Hons) in mathematics and the others had BEd degrees in mathematics education. All the participants had taught mathematics at high school for a minimum of five years. The schools from which the participants were selected had shown consistent pass rates of 70% and above in mathematics for at least two years.

3.7.2 Administration of the pilot study

The researcher applied for permission to administer the test to the teachers from the Provincial Department of Education (ref Appendix III). Permission was granted and the teachers participated voluntarily in the exercise (ref Appendix I).

The conceptual knowledge exercise, concept mapping, lesson observation schedule, lesson plan schedule, questionnaire, interview schedule, teachers' written report guide, video recording and document analysis schedule were administered to the participants during the pilot study. The conceptual knowledge exercise was administered to the participants in a classroom at the centre where the cluster meeting took place. Before the conceptual knowledge exercise was administered, participants were informed of their right to participate voluntarily or withdraw from the research process if they wished to, and were informed of their role, the aims and objectives of the research, and how their privacy would be maintained. The time for the completion of the conceptual knowledge exercise ranged from 45 to 55 minutes.

3.7.3 Result of the pilot study

3.7.3.1 Conceptual knowledge exercise

As indicated in Section 3.5.1.1, three mathematics lecturers (raters) from the university assessed the first draft of the conceptual knowledge exercise for content validity. Content validity was obtained by determining the extent to which the raters agreed with the researcher (test developer) and whether the test covered the entire content of statistics in school mathematics adequately according to the NCS (ref Appendix XXIX). The raters were asked to rate each question in terms of *sureness* **(**with rating levels of; 1 = not very sure; 2 = fairly sure; and 3 = very sure) and *relevance* **(**with rating levels of; 1 = low/not relevant; 2 = fairly relevant; 3 = highly relevant), with a maximum of three marks for each question. By indicating 'sureness', one had no doubt that the instrument measured the content knowledge of the chosen topic. By indicating 'relevance', one had no doubt that the item was a measure or determinant of content knowledge of the chosen topic (ref Appendix XXIX). The raters' responses demonstrated an overall average of 97% agreement (for the first draft) on the extent to which the test items covered the curriculum. Furthermore, based on their comments, the final items agreed upon totalled 20.

Additionally, the instrument was given to some Grade 11 and 12 mathematics teachers who would not participate in the conceptual knowledge exercise in order to identify difficult and confusing terms or phrases and these were modified or rephrased.

- **Scoring the conceptual knowledge exercise**

Marks were allocated for correct responses or correct choice of options, and no mark was allocated for a wrong or omitted choice, or a choice of more than one response per item. The total correct score was determined out of 20 and the percentage of the score was calculated. Both the raw score and percentage score were analysed to determine the reliability of the exercise. Other test characteristics, including the item response pattern, discrimination, and difficulty indices, were determined and are discussed below.

a) Item response pattern

The analysis of the conceptual knowledge test showed that in some of the items, only 1 or 2 or 3 or 4 (both) participants chose the items, as shown in Table 3.6. For instance, all the participants chose option E of item 1, which is the correct answer to the item. In item 3, three participants chose option B and only one participant chose option D. In items 12 and 13, only one participant chose option D in each case. One of the participants wrote 'no answer' and the other two left the question unanswered. This may be due to bad distracters. Such ambiguous items were discarded. All the participants answered items 1, 2, 5, 8, 9, 19, 11, 14, 15, 16, 17, 18, 18 and 20 correctly. These items tested participants' knowledge in statistics in school mathematics according to the NCS. While items 1, 2, 11, 15, and 20 tested comprehension, items 5, 8, 9, 10, 14, 16, 17 and 18 tested application**.** These items seemed to be easy for the teachers**.** The items were modified by replacing and rephrasing the questions and were therefore considered for inclusion in the main study. Few participants answered items 2, 3, 4, 6, 7, and 19 correctly. These items tested teachers' knowledge of measures of central tendency in statistical graphs of grouped data. While items 2, 3 and 6 tested comprehension, in which the teacher applied routine procedures to solve graphing problem in familiar context, items 7 and 19 tested application, in which the teacher applied his knowledge of statistics to solve familiar or other situations by using routine or multi-step procedures**.** These items were considered difficult. At the end of the review exercise, based on test characteristics (item response pattern), 20 items testing statistics in school mathematics (measures of central tendency and spread) were selected for the main study.

Table 3.8: Item response pattern of the conceptual knowledge exercise from the pilot study test items

OPTION	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	0	0	0	0	3	3	4	0	0	4	0	0	0	4	4	0	4	1	4
B	0	0	3	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	2	0
C	0	0	1	3	4	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0
D	0	4	0	0	0	0	1	0	0	0	0	1	1	4	0	0	0	0	1	0
E	4	0	0	1	0	1	0	0	4	0	0	0	0	0	0	0	0	0	0	0

b) Reliability of the conceptual knowledge exercise

Reliability is the extent to which a test produces similar results when administered under constant conditions on all occasions (Cohen, Manion & Morrison, 2007). It refers to the ability of a researcher to obtain the same response each time a test is administered. Principally, there are three types of reliability: stability, equivalence and internal consistency reliability (Creswell, 2007). While stability and equivalence can be examined by test-retest procedures (to give the same test to the same group on different occasions), internal consistency can be examined using the Kuder-Richardson split half procedure (KR-20, KR-21) or coefficient alpha (Creswell, 2007). Reliability in terms of stability and internal consistency of the conceptual knowledge exercise was established in this study using the Kuder-Richardson split half procedure (KR-20, KR-21).

In measuring the stability using the test-retest method, the scores of two tests from two similar groups were correlated. The correlation coefficient must be significant at 95% or a higher confidence interval (Cohen, Manion and Morrison, 2007). In this study, similar groups of teachers were used to pilot test the instrument in order to establish the reliability of the instruments.

The correlation (r) of the two equivalent groups was determined with Window SPSS Version 17.0 as shown in Table 3.11.

$$r = \frac{N\sum XY - \left(\sum X\right)\left(\sum Y\right)}{\sqrt{\left[N\sum X^2 - \left(\sum X\right)^2\right]\left[N\sum Y^2 - \left(\sum Y\right)^2\right]}}$$

Where:

r = the correlation between the two half (even numbered and odd-numbered) items

N = total number of scores

$\sum X$ = sum of scores from the first half test (even-numbered items)

$\sum Y$ = sum of scores from the second half test (odd numbered items)

$\sum X^2$ = sum of the squared scores from the first half test

$\sum Y^2$ = sum of the squared scores from the second scores from the second half test

$\sum XY$ = sum of the product of the scores from the first and second half test

Applying the Spearman Brown prophecy formula to adjust the correlation coefficient, (R) was obtained to reflect the full-length exercise (Creswell, 2007; Gay, 1987; Gall and Borg, 1996);

$$R = \frac{2r}{1+r}$$

Where:

R = estimated reliability coefficient of the full length exercise

r = the correlation between the two half length exercises

The actual correlation (r) between the two half-length exercises was found to be 0.70. Hence, the reliability coefficient (R) of the test is 0.81. The reliability coefficient is within the limit of the acceptable range of reliability 0.70–1.00 (Adkins 1974; Hinkle, 1998). The exercise that was developed can therefore be considered reliable for use in the main study.

c) Discrimination index

The discrimination index is a measure of the effectiveness of an item in discriminating between high and low scorers on the whole test (Tristan, 1998). Once a discrimination index of an item has been computed, the value can be interpreted as an indication of the extent to which overall knowledge of the content area is related to the responses on an item. Therefore it is considered that the ability of a test taker to answer an item correctly depends on the level of knowledge that the test taker has about a subject or topic.

The following statistical formula was used to determine the (DI) of the conceptual knowledge exercise.

$$D = \frac{R_H}{n_H} - \frac{R_L}{n_L} \text{ OR } D = \frac{R_H - R_L}{N} \text{ (if } n_H = n_L\text{)}$$

where:

D = item discrimination index

R_H = number of teachers from the high scoring group who answered the item correctly

R_L = number of teachers from the low scoring group who answered the item correctly

n_H = total number of high scorers

n_L = total number of low scorers

The discrimination index of each item was obtained by subtracting the proportion of low scorers who answered the question correctly, from the proportion of high scorers who answered the question correctly (Trochium, 2001). The discrimination index is a measure of the quality of the items in the exercise and identifies the teachers who possess the desired competency as well as those who do not. The discrimination index ranges from -1.0 to +1.0. If the discrimination index is positive, it means that more test takers in the higher group answered the item correctly than the test takers in the lower group.

Table 3.9: Summary of discrimination indices of the test items

Item no	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Discr index	1.0	0.5	0.5	1.0	1.0	1.0	0.5	0.0	1.0	1.0	1.0	0.0o	0.0	1.0	1.0	1.0	1.0	1.0	0.0	1.0

In this study, a discrimination index range of 0.5 to 1.0 was considered appropriate for the inclusion of items in the test instrument (Haladyna, Downing & Rodriguez, 2002). It was therefore necessary to choose more difficult items since the researcher was interested in assessing the content and competency of the teachers in the topic. All the items (e.g. questions 8, 12, 13, and 19) outside the range 0.5 to 1.0 were modified, replaced, or discarded. The overall discrimination index was 0.7, which was within the acceptable range of values for the test characteristics.

(d) Index of difficulty

Another statistical technique that was applied to determine the quality of the test was the index of difficulty. The index of difficulty is given by:

$P = \frac{R*100}{n}$ where; P = index of difficulty

n = total number of teachers in the high and low scoring groups

R = number of high and low scoring teachers who answer the item correctly

The index of difficulty was determined by calculating the proportion of the participants taking the test who answered the item correctly (Nitko, 1996). The larger the proportion, the more students who have learned the content measured by the item (Haladyna, et al., 2002). A test with an overall index of difficulty of more than 0.8 is considered too easy (Nitko 1996). In this study, a difficulty index range of 0.4 to 1.0 was considered appropriate for the inclusion of an item in the test instrument, since participants were assumed to be competent in this topic and were currently teaching it. It was therefore necessary to modify, replace, simplify, or discard items that were outside this difficulty index range. Table 3.10 below summarises the difficulty indices of the tests items.

Table 3.10: Summary of difficulty indices of the test items

Item no	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Diff. index	1.0	0.3	0.5	0.8	1.0	0.8	0.5	0.0	1.0	1.0	1.0	0.0	0.0	1.0	1.0	1.0	1.0	1.0	0.0	1.0

The above analysis shows that about 80% of the items had difficulty indices of between 0.5 and 1.0. This is within the acceptable range. The items cover mode, median, mean, pie charts, histograms, double bar graphs, ogives, variance, standard deviation and scatter diagrams. Therefore, most of the items were retained for use in the main study. Items 2, 12, and 13, with difficult indices of less than 0.4, cover some aspects of grouped data and double bar graphs. The overall difficulty index was 0.7, which was within the acceptable range of values for the test characteristics. Table 3.11 shows the summary of the test characteristics for the conceptual knowledge exercise.

Table 3.11: Summary of test characteristics

Test characteristics	Range of values for test characteristics	Results from pilot study
Reliability	0.70 to 1.00	0.81
Discrimination index	0.3 to 1.0	0,7
Index of difficulties	0,4 to 1,0	0,7
Content validity	0,97	0,7

3.7.3.2 Concept mapping

With reference to Section 3.5.1.2, the concept map and method of assessing (memorandum) the responses of the participating teachers were validated by mathematics specialists from the Department of Basic Education and two lecturers from the university (ref Appendix XXIV) on the content of statistics in school mathematics according to the NCS.

As explained in Section 3.5.1.2, it is assumed that if participants are able to group the statistics topics according to how they should be taught in each grade and in a sequential fashion, then they possess sufficient knowledge of the curriculum and how it should be organised for effective teaching. The raters' responses showed that, first, the list of contents of NCS statistics adequately covered the contents of statistics in school mathematics in accordance with the National Curriculum Statement for Mathematics (DoE, 2010). Second, the memorandum developed for scoring was said to be adequate to assess the performance of the participating teachers in the concept map exercise. Therefore, the concept map instrument was accepted for the pilot study. Using the rubric designed to evaluate the concept map drawn by the participants, the first participant scored 62% and the second participant 64% in the pilot study (see method of scoring in Section 3.5.1.2).

3.7.3.3 Lesson observation schedule

As discussed in Section 3.5.1.4, two mathematics lecturers at the university validated the lesson plan and lesson observation schedules adopted from the Provincial Department of Education. Criteria (ref Appendix XVI) were developed by the researcher by taking into consideration how a standard lesson in normal classroom practice is supposed to proceed (Ofsted, 2010). The two lecturers were questioned, via these criteria, to determine whether the lesson plan and observation schedule contained adequate information to assess normal classroom practice in compliance with the NCS. The validation confirmed that the schedules were the current ones used by mathematics teachers according to the NCS and contained the necessary information to assess normal classroom practice. The classroom observation schedule was used by the researcher during the observation of lessons (ref Appendix XVI).

The lesson observation schedule contained information such as planning, which involves the lesson topic, learning outcomes, assessment standard and resources. The second part described pedagogical issues, such as the introduction of the lesson, general class handling involving class organisation, discipline, interaction, movement, learning climate and involvement of learners in the lesson. Other pedagogical issues contained in the lesson observation schedule were the lesson development, consolidation of the lesson and description of actual teaching and learning. In the actual teaching and learning, the language used for teaching, questioning techniques, assessments, use of resources, knowledge of the teacher, and errors and misconceptions identified were included in the schedule. The teachers and learners' activities, as well as how the lesson was evaluated before the conclusion, were also contained in the lesson observation schedule (ref Appendix XXXII).

3.7. 3.4 Interview schedule

The mathematics expert and the two lecturers were requested to assess the interview schedule to determine if it contained adequate information that will enable the researcher to gather data to gain an insight into the mathematics teachers' content knowledge and educational background for developing PCK in statistics teaching. Their responses to the items in the interview questions showed that the schedule contained adequate information needed to assess teachers' PCK. The items that were not well phrased were modified before they were used in the pilot study (ref Appendix XXVI).

3.7.3.5 Questionnaire for teachers

As explained in Section 3.1.5.2, the questionnaire for the teachers focused on what the teachers did while teaching, namely the strategies used or approach/methods applied, the content of the lessons, the nature of the topic, how the teachers identified the learners' preconceptions and learning difficulties, how the difficulties were resolved if they were, and how the lessons were evaluated. The two lecturers and the mathematics expert validated the designed questionnaire by the researcher with the aid of several sources on classroom practice (Leinhardt et al, 1990; Muijs & Reynolds, 2000; Cangelosi, 1996; Erickson, 1999; DoE, 2010). The two lecturers and the specialists were requested to assess if the questionnaire adequately covered what the teachers are supposed to do while carrying out effective teaching in classrooms with a set of criteria (ref Appendix XIII). Their reports showed that the questionnaire contained questions that are able to elicit from the teachers' information regarding the instructional strategies that mathematics teachers use during classroom practice. The comments from the mathematics specialists and English specialists were used to review the questionnaire before it was used in the pilot study (ref Appendix XXVIII).

3.7.3.6 Written report guide

The written report guide (ref Section 3.5.1.6) was validated by a mathematics specialist in the Department of Basic Education and two lecturers from the university. The three of them were requested to determine whether the written report guides could be used to collect data about what has made the lesson easy or difficult with a set of criteria (ref Appendix XIV). Their responses confirmed that the guide contained adequate questions to guide a mathematics teacher to write such a report. Their comments were used to revise the written report guide before it was considered for the pilot study (ref Appendix XXVII).

3.8 Main study

3.8.1 Subjects used in the main study

The selected four mathematics teachers at high school level in Tshwane North District were involved in the main study.

3.8.2 Administration of the main study

The procedure used in administering the pilot study was also used for the main study. The validated test instruments consisting of i) conceptual knowledge exercises; ii) concept mapping; iii) interview schedule; iv) lesson plan and observation schedule; v) questionnaires; vi) teacher written report guides, vii) and document analysis schedule were administered to the participants. The teachers taught for four weeks and eight periods of lessons were observed on scheduled dates by the researcher.

3.9 Data analysis and results of the main study

3.9.1 Quantitative data analysis

The scores obtained by the four teachers who wrote the conceptual knowledge and concept mapping exercises in the main test were scored as described in sections 4.2 and 4.4 of this study.

3.9.2 Qualitative data analysis

The qualitative data gathered from teachers using the teacher interview, questionnaire, written report and document analysis were analysed by coding and categorising their responses according to the theme in order to determine how the participating teachers developed their PCK in statistics teaching. The analyses were described in Section 4.7.

For the lesson observation, the duration of the observed lessons ranged from 40 to 45 minutes with each of the four participants, and the observation was conducted over four weeks of teaching statistical graphs. The purpose of lesson observation was to determine the subject matter content knowledge, knowledge of instructional skills and strategies as well as insight into learners' conceptions and learning difficulties that the teachers demonstrated in classroom practice over the period. The lesson observations were analysed using the format and content of the lesson observation schedule designed by the Department of Basic Education for normal classroom practice, as was done in the pilot study. The lesson observation reports were coded and categorised in order to determine the similarities and differences between the teachers' teaching methods in the assigned topic (statistical graphs).

The reports of the lesson observations for each of the participants allowed for individual lesson observation analysis and comparison of the instructional skills and strategies used for teaching school statistics. The similarities and differences in content knowledge, knowledge of learners' conceptions in the learning of statistics and learning difficulties that the participating teachers demonstrated and enabled them to develop topic-specific PCK in statistics teaching.

3.10 Ethical issues

Before the commencement of data collection for this study, the researcher applied for ethical clearance. The application was approved and the researcher was issued with a clearance letter (ref Appendices 1, 2, 3A & 3B).

The participants in this study were duly informed of the objectives of the study in writing and oral explanation before the tests were administered to them (ref Appendix I). All the procedures that involved the participants were explained to them. They were informed of their right to decline participation in the study if they so wished. The schools and participants were given codes to ensure that they remained anonymous to the public. The test scripts, interview schedule, responses to questionnaires, the CD for the video recording, and the written reports were kept in a safe place after the information was used for this study. The performance of the participants in the conceptual knowledge exercise was highly confidential. Participants and participating schools were promised access to the result on request. The study report will be submitted to the supervisor of the study, the Gauteng Department of Education, and the University of Pretoria.

3.11 Summary of the chapter

The piloting process of this study was conducted in two phases. The first phase consisted of development, administration and writing of the conceptual knowledge exercise with willing participants. The second phase was to administer and validate the research instruments. The feedback from the pilot study showed that the conceptual knowledge exercise and concept mapping needed to be modified before the main study was undertaken. Other instruments such as the interview schedule, the lesson observation schedule, the lesson plan schedule, the questionnaire and written report were found to contain adequate information that could be used to assess subject matter content knowledge, educational background, instructional skills and strategies as well as knowledge of learners' learning difficulties that teachers use in teaching statistics in school mathematics. The administration of the main study followed the procedure used in the pilot study.