CHAPTER 3 – RESEARCH DESIGN

3.1 Introduction

Chapter 3 reviews the research design and collection of data. The chapter also reviews the data processing procedures and the description of the research techniques applied in this study, giving more detail of the theoretical framework of the design, construction and use of the class tasks developed to teach the understanding and application of concepts of evidence. The data collection techniques, the sample, and the research design are described.

3.2 Research design

A qualitative research paradigm was applied to this study. In this study, two sets of tests were used to gather data so that the results from one form of data could help to inform and refine the other data. Furthermore, learners’ responses from both the written diagnostic test and Common Tasks for Assessment (CTA) investigation task were analyzed qualitatively. In this case, the nature and extent of learners’ understanding (and therefore application) of ideas about evidence were presented in a meaningful way. Leedy (1997) contends that when using qualitative method, data speak more clearly and forcefully, allowing for better understanding of their nature and interrelations. This view was echoed by Verma and Mallick (1999) who maintain that the process of using qualitative methods ensures that the conclusions drawn are meaningful, precise and representative. Hence the qualitative methods provided the richest data to answer the research questions and helped the research to understand the problem under investigation. Further, the use of free-response questions enabled learners to reply in detail. Leedy (1997) asserts that the free-response technique has an advantage over multiple-choice tests in the sense that it provides a full account of learners’ understanding of a concept. Therefore this technique provided an opportunity for learners in this study to present their ideas without the researcher providing alternative answers which could have influenced responses, which is the main limitation of interviews (Verma and Mallick, 1999). Free-
response questions also minimized guessing and encouraged honest responses from the respondents.

The study presented here examined changes in learners’ understanding of the concept of validity due to a specific teaching intervention. Although the study presented here may look like an experimental study it is not really experimental because the researcher did not do any statistics to see if there is a significant difference. Gott and Duggan (1996) suggested that learners’ understanding (and therefore application) of ideas about evidence depends, in part, on exposure or knowledge of ‘open-ended investigations and projects’. The purpose of the intervention in this study was to test the validity of this hypothesis.

According to Charles (1995), in this type of research design, subjects are measured in terms of a pre-test, exposed to a stimulus (intervention), and then re-measured in terms of a post-test. In this study the researcher began by pre-testing the extent of understanding of the concept of validity among the subjects. The subjects were then exposed to some class tasks designed to teach some ideas about evidence. After exposing the subjects to the class tasks, the same diagnostic test was administered to these subjects in order to evaluate changes in the understanding of the concept of validity due to the intervention. Responses given in the post-test permitted the researcher to measure the later extent of understanding of concept of validity in experimental design for the whole group.

3.3 Data collection techniques

This section examines the techniques used to collect data in this study. The study consisted of three distinct stages.

1. A diagnostic investigation of learners’ understanding of the concepts of evidence in experimental design, before and after exposure to some open-ended investigation tasks.

2. Development and implementation of class tasks designed to teach concepts of evidence and confront some of the difficulties identified in stage 1.

3. A diagnostic investigation of learners’ application of concepts of evidence or process skills taught in stage 2, to the CTA investigation task.
Testing understanding is no easy matter. Diagnostic tests were used as research instruments in this study because they were appropriate instruments to collect the information needed. Moreover, other researchers (Lubben and Millar, 1996; Rollnick et al., 2002; and Albers et al., 2003) have used written diagnostic probes to explore learners’ understandings of validity and reliability of measured data and they have been proved usable. This consideration made a diagnostic test attractive as a means of obtaining data regarding learners’ understanding and application of concepts of evidence in measured data. Although other methods such as interviews could have richly tested understanding, they are time consuming since they each last from fifteen to forty minutes each. Moreover, interviews also take a long time to analyze and they require expertise if they are to be conducted successfully and the data interpreted satisfactorily (Posner and Gertzog, 1982; Schumacher and McMillan 1993). Therefore the only practicable way of assessing understanding or thinking processes of forty four learners was to use a written diagnostic test that could be given to a whole group at a time.

In this study, a survey using written diagnostic questions was undertaken to elicit learners’ understanding (and therefore application) of some key ideas about empirical evidence. In this study, two sets of written diagnostic tests were used to explore learners’ understanding and application of aspects of scientific procedure or empirical evidence. One test was used to elicit learners’ understanding of key ideas about empirical evidence and the other to investigate learners’ application of aspects of scientific procedure to an investigation task. The former was based on a strategy followed by Albers et al. (2003) who investigated the understanding of concepts of validity by Foundation University learners. The subjects were tested before and after exposure to some class tasks designed to teach some ideas about evidence. After exposing the subjects to the class tasks, the same diagnostic test was administered again. Responses given in the post-test permitted the researcher to measure the later extent of understanding of ideas about evidence in experimental design for the whole group. The latter was Common Tasks for Assessment (CTA) investigation task (which was used as a diagnostic test in this study) given to 2004 Grade 9 Natural Sciences learners. It is important to note that the CTA investigation task was not specifically designed to test learners’ understanding, but the researcher changed
its purpose in order to check whether learners would apply their knowledge of concepts of evidence tested and taught in stage 1 and stage 2.

Although written tests have some disadvantages, they were used in this study to collect data from learners as they have several advantages as long as they are carefully designed. Written tests are relatively economical (per capita) compared to other instruments like interviews and observations (Cohen et al. 2000). Furthermore, they (written tests) can be administered to all subjects at the same time, so data can be collected quickly from large samples, unlike the situation with interviews or observations. Another advantage of using written diagnostic tests was that they are easier to validate since all learners answer exactly the same questions. According to Lubben and Millar (1996), all aspects of scientific procedure or empirical evidence can be explored adequately in practical contexts. What diagnostic probes offer, in addition, is the possibility of getting relatively large samples of learners to respond to identical stimuli in a way which is impossible to engineer in a practical investigation setting (Lubben and Millar, 1996) or with an interview. Moreover, learners are quite familiar with this method of assessment. This means that the use of a diagnostic test tends to be a more natural approach because learners write tests all the time at school, unlike interviews, and other methods, which if used have the disadvantage of being encountered for the first time by the learners, possibly leading to their discomfort. Consequently, learners would not be intimidated, as they would be with interviews, and this would enable them to explain their cognitive thinking freely. In addition, learners’ written-report data (Schuster, 1983) pose less threat to being interpreted according to the researcher’s prejudices, and provides a reflection of learners’ understanding of concepts and not the researchers’ expectations. Schuster (1983) also warns that the researcher’s interpretations can be influenced by what they expect to hear. Therefore, learners’ self-written reports might be expected to improve the reliability and validity of results.

### 3.3.1 Diagnostic investigation of learners’ understanding of concepts of validity

The learners’ test (which appears in Appendix A) consisted of three pencil – and – paper probes based on those developed for a similar study with Foundation University students.
(Albers et al., 2003). Although Albers et al. (2003) used five probes in their research, only three of the five probes were required to elicit the information sought in this study. The three probes explored learners’ understanding of concepts related to experimental design, in particular, they were the independent variable (IV); the most important control variables (CVs); and determining a suitable interval and range for the IV in an experiment. These probes were purposefully chosen because they deal with ideas mentioned in the CTA (i.e. identifying control variables in an experiment; choosing the independent variable and suitable range and interval of the independent variable). The probes were designed around a single experimental situation in which learners were asked to suggest an action. Figure 2 (adapted from Albers et al., 2003) below shows the first page of the diagnostic test, which explains the context to the learners.

Figure 3. Page 1 of the diagnostic test showing the context (Albers et al., 2003)

A solar cell is a small electrical power supply. It produces a current from a radiation source, for example from a lamp or from the sun.

Some students were provided with a solar cell and an ammeter connected across the solar cell with two leads. They were also given a clock, a thermometer and a light intensity meter. They were told they could do their experiment inside or outside.

The students were asked to find out how the current $I$ changes with the amount of radiation.
Two stages were involved in the administration of the diagnostic test, i.e. a pilot study and the main study. The pilot study was carried out, during July of 2004, to establish whether the students understood what was required of them, to find out how long it would take the learners to complete the diagnostic test, and to improve the wording of the questions prior to the main study. Moreover, the pilot study was conducted because it (the diagnostic test) had previously only been used in a university context. For this reason, the written test was administered to approximately 40 learners of whom several were interviewed informally. The primary aim of the interview was to establish the reflection of the learners on the written test, in particular the language used in the test, and the context involved. Two University lecturers and two science educators were consulted to establish appropriate wording of the questions. Treagust (1998) points out that one of the important skills of planning and preparing a diagnostic test is to construct and apply it in such a way that the task falls within the range and ability of the target group. We also need to bear in mind the warning by Mossom cited by Rambuda (2002) who pointed out that the learners in the lower grades might be too immature to cope with tasks that require the application of higher-level skills. When modifying the instrument, it was ascertained that the wording of the investigation task was within the range and ability of the targeted subjects. This was achieved by the substitution of ‘difficult’ words by less difficult ones, or the removal of words, or by sentence reorganization. Consequently it was found from the pilot study that learners had little difficulty answering the diagnostic test. It was also discovered that learners could take about 45 minutes to complete the test. Once the pilot study had been completed, the main study was carried out with another class early in August 2004, following the same procedure as in the previous study (Albers et al, 2003). The same test was administered to the same class in October 2004, after the learners were taught some aspects of empirical evidence through open-ended investigation tasks. Learners worked individually under test conditions, supervised by the researcher.

3.3.2 Class tasks (Intervention)

It is being proposed that the understanding (and therefore the application) of ideas about evidence on measured data can be best learned through open-ended
investigations/projects because they allow learners to carry out a whole task with the autonomy to put into practice their understanding of, and apply, ideas about evidence (Gott and Duggan, 1996). For this reason, one of the main objectives of this study was to engage learners in open-ended investigation tasks in order to help them understand and apply ideas about evidence. During August and September of 2004 learners were taught the aspects of empirical evidence through open-ended investigation tasks. In essence, the approach adopted was:

- To help learners construct the knowledge about the procedures of science via a sequence of open-ended projects, in order to make learners familiar with different procedures ‘scientists’ employ during projects.
- To ask the learners to plan their own investigations in order to assess if the taught procedures were put into operation.

Selection of tasks:

Millar et al (1994) advise that when one designs or select tasks, he/she should bear in mind the following important research –related criteria:

- Tasks should allow a variety of approaches, which would illustrate different decisions about strategy, data collection and data handling/interpretation.
- Tasks should, ideally, be such that differences in children’s declarative knowledge of the relevant science concepts and in their understanding of aspects of scientific evidence might be expected to lead to differences in their actions, as well as to differences of interpretation of results.
- The set of investigations used should cover a range of contexts.

Following these criteria, three class tasks (Appendix D) were selected to help learners understand (and therefore apply) ideas about evidence. The tasks certainly have face validity, in that they were selected and developed directly from the Millar et al.’s (1994) protocols. The tasks used in this study were:
Task 1: *The water roller* (Solomon and Lee, 1992): This activity includes measuring by hand spans, and recording the numbers. That is a useful mathematical or estimation exercise. The central scientific concept in the activity is ENERGY. It would however be possible to go through the whole activity without using the word “energy” once. But this might be considered a wasted opportunity.

Task 2: *Heat and electricity* (Lubben et al., 1995): This task investigated the heating effect of an electric current. In this activity learners classify appliances and then investigate the heating effect of an electric current.

Task 3: *The pendulum*: This is a class activity, used to investigate a child’s ability to control variables and to deduce the effects of weight, length, and amplitude on the period of a swing, from a series of matched demonstrations.

In order to impose some limits on the extent of likely variation in performance, all the learners in the class were given the same tasks. In other words, the researcher presented the learners with a task, rather than allow them to choose their own investigation task. That is, the learners were given a problem to be investigated and were then required to propose an appropriate experimental procedure. This included the identification of variables that had to be controlled, a plan about the quantities to be measured and the recording results. The design of the investigation tasks was centered on the issue of variables, which is broadly defined as something which changes or can change in a scientific investigation (Keys, 1998).

*Administration of tasks*

During the teaching of the three investigation projects, the researcher took the primary role for all instructions. In starting the intervention, the researcher introduced and explained the purpose of the projects. He then started by reminding learners that science is not just about learning and understanding facts but is also about a way of working (investigating). Thereafter, he explained what processes of scientific enquiry were by means of examples. Points made included:

- Identifying a problem and devising a plan to address this problem
- The need to control variables and to change just one at a time
• Ensuring that comparisons are fair
• Repeating experiments
• Selecting a sensible interval between readings
• Deciding when readings are appropriate
• Interpreting data

The researcher then introduced the activities where the above points would be dealt with in more detail. Learners were reminded that the purpose of these activities was not to find the correct answer but to provide them with the experience of carrying out their own investigation and to discuss how they do this so that their abilities and understanding in this area can be improved.

The generative model of learning science investigation processes (Hackling and Fairbrother, 1996) provides maximum opportunities for children to generate and test their own science ideas and thus, was used as an instructional template during the teaching of three investigation tasks. The model involves three teaching phases (exploration phase; investigation phase; and reflection phase). During the exploration phase, learners work in groups to interact with materials having to do with a topic, raise questions and formulate a problem for investigation. During the investigation phase, learners plan and carry out investigations and interpret their findings. During the reflection phase, learners report their findings to others and reflect on their investigative processes.

Each project lasted about two days (two 1 ½ lessons each) and involved learners working in groups to carry out the three phases of generative learning: exploration, investigation, and reflection. The groups met for one and half hour sessions every other day, resulting in approximately 3 hours devoted to each project. All three activities were done in the third block of 2004 over a period of three weeks. Each project was initiated by a teacher – directed activity. The exploration activities are also described in some detail here, since they may influence the investigations learners designed.

The first investigation project was the water roller and was taught in the first week of August 2004. The scientific concept on which the project was based was energy. This activity included measuring by hand spans, and recording the numbers. It was therefore an estimation exercise. The project lasted about two days and involved students working
in groups to carry out the phases of generative learning described above; exploration, investigation, and reflection. In this project, the researcher led the class in a “hands-on” exploration with plastic drinks bottles, yoghurt pots and margarine tubs, tray and short planks of wood. The learners were asked to discuss what factors affect the actual distance moved by the yoghurt pot. They were then required to design their own experiment to investigate the effect of the number of planks of wood and the volume of water on the actual distance moved by the yoghurt pot. During the project, learners were required to formulate their hypothesis, to design their own investigation and to work according to their designs. The learners were also involved in the evaluation of their own work in order to help them reflect on what they did well, what they did poorly and what they needed to do differently to improve. The researcher and learners discussed the procedures used in the learner designed experiments.

The other two investigation projects were taught during the last two weeks of August 2004. Science topics on which the projects were based include heat and electricity, and changes in substances. In the project on heat and electricity, the researcher engaged learners in a discussion related to electrical appliances that are used to provide heat. Learners classified appliances and then investigated the heating effect of an electric current. Learners also worked in groups for this activity and collected the required equipment from the teacher. The teacher demonstrated how to make the immersion heater by winding a coil of wire round a pencil. There were two stages to the activity. In the first stage learners gained experience of setting up the apparatus and noted that a coil of wire passing an electric current will heat up some water. In the second stage they investigated the effect of coiling on the heat produced. When introducing this stage, the teacher reviewed factors influencing resistance. This activity required learners to design their own experiment and thus consider which factors to vary (and to what degree) and which to keep the same.

For the other project on the ‘pendulum’, the researcher demonstrated how to make a pendulum. He then engaged the class in a discussion to find the factors that would affect the period of the swing. Learners were then asked to design their own experiment to investigate factors which affect the period of a swing. Subsequently to the exploration activities, each student group worked together to generate and select a question for
investigation (as they did with the other projects). The groups spent several minutes (during each period) reading and summarising background information on their topic, designing investigation plans, implementing their plans, creating data tables, grouping their data and writing rough draft and final reports.

For each project, the researcher checked designs before learners started their experiments. This check was done to ensure safety. Students were then allowed to follow their own designs if considered safe to do so. The researcher noted different designs and watched how learners carried out their investigation, collecting points to include in the later discussions. The groups were brought together occasionally for whole class instruction or discussion on a particular ‘science investigation process’ such as planning investigations, conducting investigations and processing data. After each investigation, learners had the opportunity to reflect on and evaluate their own work through the use of the self – evaluation checklist and discussions with their group members and the researcher/teacher. During the reflection phase of each project, all groups were required to make a poster to summarise their investigations and make presentations to the class. Throughout the entire process, the researcher circulated through the groups and facilitated the learners’ thinking about their investigations. However, this facilitation took the form of asking the learners questions, rather than directing their actions. Although this phase was part of the research, the researcher did not collect data or enough evidence for analysis in this phase.

3.3.3 Diagnostic investigation of learners’ application of concepts of evidence to the CTA investigation task.

A CTA investigation task (which appears in appendix F) was administered in November 2004, four days after it was administered to the Grade 9 Natural Sciences learners as part of a public exam. It should be noted that the subjects in this study were not aware that the CTA investigation task written by the Grade 9 learners would be used to assess their application of concepts of evidence in measured data. The CTA investigation task was selected because it required learners to apply their understanding of scientific evidence. In this investigative task, learners were asked to investigate how the length of a pencil
lead affected the resistance of different pencil leads. Learners were required to respond to a series of questions about the problem investigated in the activity, method or procedures used to investigate the problem, identify control variables, independent variables and dependent variable, in the investigation task, as well as interpret data obtained in the investigation task. Three experts were consulted to establish face validity of this test. The instrument was then piloted with a group of six learners who were not involved in the main study in order to establish whether the learners understood what was required of them. This was also done to find how long it would take for the learners to complete the test, and to establish whether there would be any problems in administering the instrument. In the pilot study the researcher specifically asked the respondents to report on any problems or queries, which struck them when they were answering the questions. It was established that there would be no problems administering the test and that the test could be completed successfully within a period of one hour.

3.4 Population and research sample

The originally targeted learner population was all Grade 9 Natural Sciences learners in one school. However, due to time constraints and inaccessibility of these subjects, the study was conducted with Grade 10 Physical science learners of one school in the Gauteng Department of Education (GDE). Grade 10 was chosen because it represents a transitional phase between learners in the Senior Phase and the learners in the FET phase. In 2004, the year in which data for this study was gathered there were four Grade 10 Physical Science classes in the school, two of which were taught by the researcher. Each class had 40 to 45 pupils (boys and girls). It was not possible to involve all the Physical Science learners in this project due to time constraints and large population. This would have been an expensive exercise in terms of time to administer the instrument and analyze data. From the four Grade 10 physical science classes in the school, one class was selected. According to Borg (1981), the size of the samples and the procedures used in selecting the samples determine the degree of confidence with which the researcher can apply the research findings to the population. This view was echoed by Charles (1995) who maintains that samples are a necessity in research where findings are
intended to be generalized to the population. Convenience sampling was therefore used in the selection of subjects for this study. Convenience sampling involves the nearest or most accessible individuals to serve as respondents (Cohen et al., 2000). However, the researcher was quite aware of the biases and limitations in terms of generalizing findings of studies involving such samples. Notwithstanding this, the present study was not interested in drawing generalizations rather it seeks to generate deep understanding of factors that affect pupils’ understanding and application of concepts of evidence in experimental design. Moreover, it was assumed that these learners could be knowledgeable and informative about the application of science process skills and concepts of evidence. Most of the learners in this class had been introduced to open-ended projects in their Grade 9 Natural Sciences course. At Grade 9 level, there is mandated syllabus on this aspect in the Natural Sciences learning area. 42 out of 45 learners in this were engaged in these investigation projects at Grade 9 level by the researcher, unlike in other three classes where there were more than ten repeaters. The learners were consulted and invited to participate. From this contact it was determined that the sample comprised 45 learners, representing about 28% of the Grade 10 physical science learners at the school. Of the 45 learners who participated in the study, 24 were female and 21 were male.

3.5 Reliability and Validity of Analysis

In this study, considerable attention has been paid to ensuring the reliability and validity of the results. The propositions have been developed based on existing theories. Related earlier research has been used when developing the constructs and the derived variables and measurement items. Data collection and analyses methods have been selected carefully on the basis of literature review. However, the small sample of 43 learners required some compromises in statistical methods and analyses. Finally, the results and conclusions have been carefully analyzed to ensure their feasibility.
3.5.1 Reliability

Reliability refers to the degree to which results are consistent across repeated measurements (Bollen, 1989). The analysis in the present data is based on the primary data collected from the two main research instruments (written probes and a CTA investigation task). Several methods were used to ensure the reliability of self-reported data. First, the main methods used to minimize errors and biases, and thus increase reliability, were to use experts to verify the coding of learners’ responses to both the written probes and the CTA investigation task. Second, reliability of data was also increased by documenting everything precisely, especially the data but also phases and processes of the study. For example, code allocation for the probe responses has been done carefully as evident from the quoted learners’ responses in chapter 4. Finally, the review and piloting of the data collection instruments also increased the reliability.

3.5.2 Validity

Validity refers to the extent of use of a measurement instrument and how well it measures what it purports to measure (Nunnally and Bernstein, 1984:83). In order to improve validity of the present study, previously validated measures have been used when possible.

*Face validity*

Face validity (the extent a construct confirms a common understanding of the related concept) was ensured by extensive and development or selection of the constructs and measurement items used in previous literature - the measures were discussed and pre-tested in practice with science education specialists, as well as with academics with experience in the relevant fields. The measures are in line with common understanding.
Content validity

Content validity (the extent to which an empirical measurement reflects a specific domain of content) was demonstrated by having four experts (three university lectures and two science educators) rate the diagnostic test and assessment task as: (a) the representativeness of the items from the total pool of possible items for the content covered; and (b) the readability and clarity of the items. In fact, to ensure the validity of the data collection instruments (they measure what they are supposed to measure according to the aim of the study), they were given to experts to establish appropriate wording for the questions. The experts were consulted to establish whether the questions in the instruments were written in such a way that would make it possible for second language speakers or grade learners to understand. The five experts were also asked to check or verify the relevance of the content in the instruments to the grade 10 Physical science curriculum. With children of secondary school age, the contexts used in research instruments were explained orally by the researcher. In fact, the researcher and learners investigated discussed the experimental set-up in a whole group discussion. Although demonstrations would have increased content validity further, care was taken to ensure that all participants understood the experimental set-up in both the written probes and the CTA investigation task. This approach was adopted in the light of pilot experience which indicated that the experimental set-up was understood through a researcher led discussion.

External validity

External validity or generalizability refers the extent to which the results of the study can be generalized). The present study has certain limitations as regard to external validity. In this study, it may be argued that there is a limitation regarding the external validity as a result of sample selection. The sample of 43 learners was limited to a specific group of learners in a secondary school representing the characteristics of a fairly young, dynamic advanced – science environment. However, it is worth mentioning that the sample actually represents the population of grade 10 learners in that specific school at the time.
of the research. Generalization on statistical grounds is possible, but strictly valid to the study group. On the basis of theoretical or analytical generalization, the results may be considered generalizable to similar kinds of environments. However, to get fully established and obtain rigorous position the theory definitely calls for extensive evidence and requires future research. The external validity of the study is discussed in detail in chapter 6 ‘Limitations’.

3.6 Permission to conduct research in the school

Permission to conduct research was sought from the Gauteng Department of Education. The purpose of the research was explained and the information sought was made clear. The Head of the Gauteng Education Department was told that the study attempts to describe and explain the understanding and application of concepts of validity in experimental data, and was likely to provide useful information which could be of supportive nature to science educators in general and outcomes-based education in particular. Since the researcher intended to conduct the research in the school where he teaches, the Head of the Education Department referred the researcher to the principal of the school for permission to conduct the research. The researcher then arranged the meeting with the principal of the school. The purpose of the project was explained to the principal and the information sought was made clear. The principal of the school was also made aware that the permission was first sought from the head of the department of education. The principal also telephoned the Head of the Education Department for confirmation. The principal then asked the researcher to identify the sample so that they could be requested to participate in the study and a consent letter be sent to their parents. All the targeted subjects agreed to participate in the study. A consent form was sent to their parents which also requested that they hand them back to the school. Parents and participating learners were assured the learners participate voluntarily in the project and they were not going to be penalized for not participating since the project was not part of the assessment requirements. The following points were also made clear to participating learners and their parents.
• The names of the school and participating learners remain confidential in all respects;
• Collection of data by the researcher would take place outside normal tuition time of the school;
• A copy of the thesis would be available to the parents of participating learners on request.

Therefore participants and their parents gave informed consent for their participation.

3.7 Conclusion

This chapter reviewed the research design and collection of data. It reviewed the data processing procedures and the description of the research techniques applied in this study. The chapter also gave more detail of the theoretical framework, of the design, construction and use of the class tasks developed to teach the understanding and application of concepts of evidence. The data collection techniques, the sample, and the research design were also described. The following chapter discusses the analysis of data and results of the probes diagnostic test.