CHAPTER 3

METHODOLOGY

3.1 Chapter Overview

This section highlights the overall structure and approach to the research and the means used to collect data that is relevant to the study, in particularly the manner the data was gathered, analysed and subjected to interpretation in the two phases of the study: the modelling of the ER mental model and the trialling of this mental model. The research conceptual framework of the study is also depicted in this section as a 'scaffolding' to the study as a whole. Pilot study and how the findings are to be reported will also be discussed.

3.2 The Research Design

Qualitative framework was employed in this study. Denzin and Lincoln (1994) stated that qualitative research is "multi-method in its focus, involving an interpretative, naturalistic approach to its subject matter" (p. 2). This study was a case study that explores how the expert raters differed from that of novice raters in doing the task of assessing and giving feedback on students' writing. According to Creswell (1994), a case study focuses on the exploration of a single entity or phenomenon by a researcher. The researcher had chosen this research design as it studies a phenomenon in its natural setting and its findings are interpreted in terms of the meanings people bring to them. From this perspective, the researcher would be able to

build “a complex, holistic picture, formed with words” (Creswell, 1994, p. 1). A case study with the qualitative framework was also the researcher’s choice because it was best suited to answer the research questions which aimed to describe what was happening in a particular situation (Bouma, 1996). According to Ritchie (2003), a case study is contextual and it:

> …is concerned with identifying what exist in the social world and the way it manifest itself. A major feature of qualitative methods is their facility to describe and display phenomena as experienced by the study population, in fine-tuned detail and in the study participants’ own terms. It therefore offers the opportunity to ‘unpack’ issues, to see what they are about or what lies inside, and to explore how they are understood by those connected with them. (p. 27)

The case study would enable investigation to be carried where “the subject areas in which the phenomena that need to be studied are innately intricate or conceptually difficult to relate. The complexity may lie in the nature of the subject itself…the intricacy relates to the level of unpacking that is needed to formulate a position, view or belief” such as in the case where cognitive processes are focus of the study (Ritchie, 2003, p. 32).

In the context of this study, the participants were recognised as expert raters who had certain knowledge and skills which were highly proceduralised through many years of experience to doing similar task. Hence, by using the case study, the complex mental process of the participants could be understood in the context of the case. Moreover, the case study employed a multiple data collection method to improve the reliability of the information in the case (Anderson, 1990; Yin, 1994; Stake; 1995). The qualitative paradigm for this study enabled the researcher to describe the mental processes of the participants from the data collected through verbal protocol analysis, interviews, and observations/field notes. These data collection methods would provide

the multiple sources of evidence to triangulate the reliability of the information (Yin, 1989).

3.3 Research Conceptual Framework and Phases of the Study

To achieve the research objectives, this study was carried out in two phases. The **first phase** of the study is the **modelling of the ER mental model**. It focused on mapping out the lines of reasoning to form a comprehensive mental model of how the expert and novice raters would assess and evaluate writing (including giving feedbacks) using the verbal protocol analysis technique, as discussed in sub-section 2.3.1. This process is related to interpreting and codifying human thinking.

Mulhern (1989) highlighted the success behind the use of verbal protocol analysis to elicit appropriate forms of intellectual activities through problem solving, as it involved taking as complete a record as possible of what an individual did in solving a problem. The information drawn from the data, pertaining to the knowledge states and conceptual operators used, and the lines of reasoning, was crucial to gain insights on the participants' personal justification on the decision and action taken while doing the process of "thinking aloud".

The findings in phase one was later used as a scaffold in **phase two**, which would be the **trialling out the ER mental model** in the form of training sessions with a group of beginning teachers. In this phase, using the mental model of the expert raters, the researcher explored how the novice raters could be trained effectively in acquiring the skills of assessing and giving feedback on students' written work. The main concern for the trialling process was to gain prior insight into the effects and implications of

using such a conceptual model to gap the differences between an expert rater and a novice rater in terms of expertise in assessing and evaluating writing.

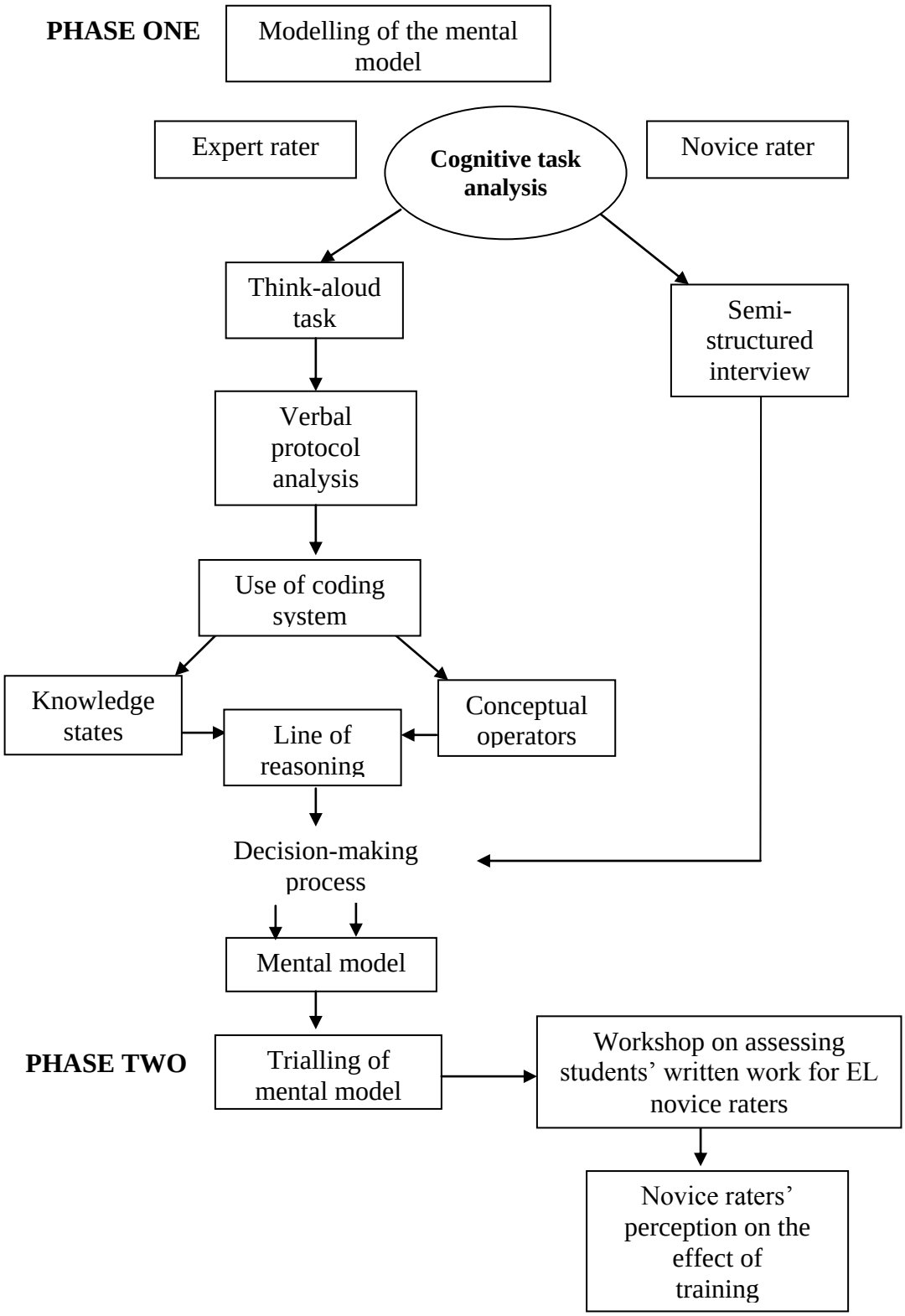

Figure 3.1 Research work plan

The quantitative data was derived from the qualitative data, especially from the interviews with the participants, whereby a simple percentage and frequency count were used to analyze the data in a more objective manner. This form of data was useful in the triangulation process of the findings to form a clearly picture of the phenomenon in study. The data was also in the form of qualitative data derived from interviews to gain insight on the participants' perception of the effectiveness and implications of the intervention using the formulated conceptual model in phase one of the study. The research conceptual framework that would serve as the scaffolding to the study is depicted in Figure 3.1 above.

3.4 Sampling Method

As a qualitative study, this research used non-probability samples for selecting the population for study. Richie, Lewis and Elam (2003), in a non-probability sample, stated that "units are deliberately selected to reflect particular features of or groups within the sampled population" (p. 78). The chances of selection for each unit are unknown as the sample is not intended to be statistically representative. Instead, the characteristics of the population are used as the basis of selection. The selection of participants was also criterion referenced or purposive (Patton, 2002). The participants were chosen because they have particular features or characteristics which would enable detailed exploration and understanding of the central themes and puzzles which the researcher wished to study.

In this study, the researcher studied the verbal protocol analysis of the "purposefully select participants" (Creswell, 1994, p. 148) made up of expert raters and novice raters who would be differentiated from each other by working experience and the in-

service training/professional development programs that they have undergone. It was this features that would make them well-suited to small scale, and in-depth studies.

In deciding the participants for each group, the researcher had taken *homogeneity* of the participants into consideration to enable the samples chosen to give a detailed picture of particular phenomena. The two groups in this study, though homogeneous in its own category, were also heterogeneous samples in the sense that they varied widely from each other. The purpose of having these two groups of participants was to identify central themes which would cut across the variety of cognitive processes being under studied. Although purposive selection had involved quite deliberate choices in the selection of participants, this should not suggest any bias in the nature of the choices made. The process of purposive sampling required clear objectivity so that the sample stood up to independent scrutiny.

3.4.1 The Sample Size

There were eight participants in phase one of this study, four participants from each group: the expert raters and novice raters. In phase two, five participants were involved in the interview after they attended the intervention workshop. In this study the sample size was small which is usual for any qualitative study.

As mentioned by Richie, Lewis and Elam (2003), there are three main reasons for this. Firstly, if the data are properly analyzed, there will come a point where very little new evidence will be obtained from each additional fieldwork unit. This will happen because phenomena need to appear once to be part of the analytical map. Thus, there will be a point of diminishing return where increasing number of sample size no longer contributes new evidence.

Secondly, statements about incidence or prevalence are not the concern of qualitative research. Thus, there will be no requirement to ensure that the sample is of sufficient scale to provide estimates, or to determine statistically significant discriminatory variables. Finally, the type of information that qualitative studies yield is rich in detail. There will be probably vast information that will be contained in each unit of data collection. Thus, it will be fair to the analysis of the data collected that the sample size be kept to a reasonably small scale.

3.4.2 The Research Participants

This section presents a brief profile of the participants in phase one and phase two of the study. Eight participants were involved in phase one of the study, which were consisted of four expert raters and four novice raters. Phase two of the study consisted of five beginning teachers.

(a) Participants in Phase One of the Study

Name of school	Code name of participant	Gender	Experience in teaching English	Experience as rater/ level	Rater status
SMK Melur	NR1	Male	1 year	Nil	Novice
	NR2	Female	1 year	Nil	Novice
SMK Seroja	NR3	Male	3 years	Nil	Novice
SMK Mawar	NR4	Female	3 years	Nil	Novice
	ER1	Female	10 years	PMR & SPM	Expert
SMK Teratai	ER2	Male	18 years	SPM	Expert
	ER3	Male	13 years	SPM	Expert
SMK Melati	ER4	Male	13 years	SPM	Expert

Table 3.1 The pseudonym names of the schools and the data of the informants in phase 1 of the study

Through purposive sampling method, eight ELT teachers were selected based on teaching experience and in-service courses attended. The names of the schools, the

expert raters and the novice raters were all kept as pseudonyms. The data of the participants were shown in Table 3.1.

(i) The Expert Raters

According to Erisson and Charness (1994), an expert refers to a person with special skills and knowledge acquired through experience, rather than inherent talent. In this study, the expert raters were graduate ESL teachers who had at least ten years of teaching experience in the English subject and were qualified raters for the SPM English – Paper Two examination. Some of these raters were specially trained by the Malaysian Examination Syndicate (or *Lembaga Peperiksaan Malaysia)* to mark the English Paper Two in the SPM (*Sijil Peperiksaan Malaysia* or Malaysian Certificate of Examination) Examination. Thus, the expert raters could be identified from the novice raters in terms of the teaching experience in writing and the in-service courses they had attended. They were expected to be able to give better feedback on students' writing as compared to the novice raters. In fact, the study of expert/novice differences in other domains has continued to reveal the impact of knowledge on pattern recognition, and the precision of expert performance results from specialized schemata that drive performance (Egan & Schwartz, 1979).

(ii) The Novice Raters

The **novice raters** in this study were four beginning ESL teachers in their first three years of teaching appointment after completing teacher preparation program. In Malaysia, the teachers are trained in either teacher education colleges or universities. The two types of beginning ESL teachers who graduated from teacher training institutions were teachers who had completed the Degree in Education Program

(Program Ijazah Sarjana Muda Pendidikan or PISMP) and Postgraduate Education Course *(Kursus Perguruan Lepasan Ijazah or KPLI).*

PISMP and KPLI are recognized nationally as basic requirements for teaching primary and secondary schools respectively. Since this study was focused on beginning ESL teachers who were teaching in the secondary schools, the respondents would comprise graduates from the Postgraduate Education course. Another type of beginning teachers in this study would be those who had graduated from the education program which was carried out by faculties or centres of Education in local public universities (MOEM, 2001). In the local universities, the duration of the course for undergraduate degree programs in education is four years. The graduates from this course would also be teaching in secondary schools.

(b) Participants in Phase Two of the Study

For the purpose of getting the participants' feedback regarding their perception of the workshop that they have attended, the researcher managed to interview five participants out of twenty five participants who came for the workshop. Initially, this workshop was designed to cater for the needs of the so-called novice raters but apparently the requirement set (not more than three years experience of teaching English) for this workshop was unheeded by the principals of some of the participating schools though none of the participants was involved in marking SPM English Examination paper and had more than 5 years of teaching experience in their respective school.

This situation however has proved to be of an advantage to the workshop as the more experienced ones were able to contribute more to the discussion during the workshop

sessions. The novice raters who participated during phase 1 of the study could not attend the workshop due to some technical reasons which were beyond the researcher's control though special invitations to attend the workshop had been extended to them. The profiles of the participants/informants, based on the personal data form completed by the novice raters when they registered for the workshop, are summarised in Table 3.2 below.

Name of school	Name of participants	Gender	Experience in teaching English	Experience as rater/ level
SMK Orkid	Ms. Kim	Female	1 yr	Nil
SMK Cempaka	Mdm. Mary	Female	2 yrs	Nil
SMK Matahari	Mr. John	Male	2 yrs	Nil
SMK Bunga Raya	Mr. Chris	Male	3 yrs	Nil
SMK Kemboja	Mr. Gerald	Male	2 yrs	NIl

Table 3.2 The pseudonym names of the schools and the data of the informants in phase 2 of the study

3.5 Instrumentation

This section starts off with a discussion on a researcher as the main instrument in collecting qualitative. The rest of the research instruments used in the two phases of the study, like the verbal protocol analysis instruction and task, and semi-structured interview guidelines are also presented. This section ends with discussing the validation of the instrument, with respect to the procedure of instrument validation.

3.5.1 Researcher as a Research Instrument

In collecting the qualitative data, the main instrument was the researcher. As the main instrument, the researcher needed to be alert and sensitive to what happens in the field that involved social relationship as well as personal feelings (Neuman, 1997). The researcher found it necessary to include personal and subjective experiences as part of the field data which were valuable for interpreting events in the field. Thus, the

researcher needed to build rapport with the participants to be able to get all the necessary information from them.

3.5.2 Verbal Protocol Analysis Instruction and Task

In phase one of the study, collection of qualitative data is also facilitated by a simple instruction for verbal protocol analysis (Appendix 6a). This instruction was read to the participant prior to the verbal protocol analysis session before they were given the task stimulus in form of student(s) written work. The students' sample writings were taken from nine form four students (with different levels of English Language Proficiency based on their semester's examination marks). Prior to that, the researcher has sought permission from one of the teachers, who was also involved in the study, to use the students work as stimulus for the study.

The sample writings were produced by the students in a normal writing lesson. The topics for the writing task were those related to the themes suggested in the Form Four English Language Syllabus. The students' written work was collected by the researcher to be marked by three independent raters using the marking scheme for continuous writing taken from a school (see appendix 9).

The moderated marks of the nine sample writings by the three independent raters are shown in Table 3.3 below. The raw score marks were converted to percentage to determine the students' proficiency level (see Table 3.4).

Rater	Sample writing								
	1	2	3	4	5	6	7	8	9
Rater 1	42% (21)	38% (19)	50% (25)	56% (28)	46% (23)	42% (21)	38% (19)	36% (18)	36%(18)
Rater 2	40% (20)	36% (18)	52% (26)	48% (24)	44% (22)	44% (22)	36% (18)	40% (20)	36%(18)
Rater 3	42% (21)	40% (20)	54% (27)	44% (22)	46% (20)	48% (24)	32% (16)	36% (18)	38%(19)
Average Marks	41.3% (20.7)	38% (19)	52% (26)	55.4% (24.7)	45.3% (21.7)	46.7% (22.3)	35.3% (17.7)	37.3% (18.7)	36.7% (18.3)

Table 3.3 Moderated marks of the sample writings by independent raters

Next, the researcher grouped the students' writing according to the different levels before selecting two pieces of writing to represent the lower two different level of English proficiency (the elementary and intermediate levels) as shown in the Table 3.4:

Range of Marks	**English Proficiency Level**
40 and below	Elementary
41 – 59	Intermediate
60 and above	Advance

Table 3.4 Proficiency level taken from a school report card

The reason why only two pieces of writing were chosen was because none of the sample writings was rated under advance level of English proficiency. Based on the moderated marks, the final grouping of the sample writings is shown in Table 3.5:

English Proficiency Level	Sample Writings
Elementary (40 and below)	2, 7, 8, 9
Intermediate (41- 59)	1, 3, 4, 5, 6,
Advance (60 and above)	Nil

Table 3.5 Students' proficiency level

After considering the sample writings with the lowest and the highest moderated marks, the researcher decided to choose sample writing 2 and 4 to be used as stimulus in the verbal protocol analysis.

3.5.3 Semi-Structured Interview Guidelines

For both phases of the study, the researcher had also made use of the semi-structured interview guidelines which consisted of open-ended questions that would allow the researcher to explore the subject matter or related themes deeper, and gathered observational notes by conducting observations as an observer. According to Yin (1994), interviews are a useful source of collecting evidence for the case study as they will give the participants' constructions of the reality around them and may help provide important insights on how participants interpret some piece of the world in their own words. Interviewing the participants is necessary when required information about feelings, belief, perceptions and opinion cannot be observed (Merriam, 2001).

In this study, the researcher referred to Patton's (2002) five types of questions - experience/behaviour questions, opinion/values questions, feelings question (affective questions), knowledge questions, and sensory questions as in Appendix 7. In phase

two of the study, some of the questions asked on how the novice raters perceived the effectiveness of the workshop (Appendix 8) are as follows:

- What was your initial expectation of the workshop?
- To what extent do you think you have learnt or acquired some knowledge and skills on assessing and giving feedbacks to students' writing through this workshop?
- To what extent do you think that this workshop is useful in your own classroom practice especially in assessing and giving feedbacks to students' writing?
- How confident are you in assessing and giving feedbacks to your students' writing after attending this workshop?
- If given the chance to improve the content of the course, what would be your recommendation(s)?

3.5.4 Validation of Instrument

Validation of instrument in this study focused on (a) whether the verbal protocol analysis instruction was clear to the participants as they think-aloud while assessing and giving feedback on students' writing so that the cognitive processes could be inferred from the participants' verbal protocols, and (b) assessing the suitability of the students' sample writing that were used in the problem solving task.

To accomplish validity on the first issue, the feedback from the participants during the pilot study ascertained the clarity of the VPA instruction and was no change made to the original instruction given to them. On the suitability of the sample writings to be used in the VPA, the researcher had addressed this through the selection process discussed in earlier part of section 3.5 and the final two pieces of writing were chosen

because none of the sample writings was rated under advance level of English Proficiency.

3.6 Pilot Study

To ensure reliability and validity of the data, there are two aspects of this study that were piloted. Firstly, the verbal protocol analysis instrument had to be piloted for validation purposes, in terms of clarity of instruction. Secondly was the methodology of the study which was heavily dependent on VPA or think-aloud verbalisation. Therefore, a pilot study was carried out to trial out the research data collecting procedure and to investigate the feasibility of implementing the data analysis procedure. It had given the researcher the opportunity to see whether the study was feasible to be carried out in the manner it was planned or any amendment that needed to be made either in the data collecting procedure or the data analysis procedure before the actual study could be carried out.

Based on the pilot study, some modifications were made to the procedures in collecting the data. One of the modifications is on the instructions in which it was made clearer to the participants as during the pilot study the respondents showed confusions on some of the steps. The interview questions were also modified so that they were aligned to the research objectives.

3.6.1 The Instrument

The Verbal Protocol Analysis Instruction (Pilot Study) (Appendix 6a) was piloted on two participants: an expert rater and a novice rater. Prior to that, sample writings were taken from nine selected students who had different levels of English Language proficiency. The topics for the writing task were those related to the themes suggested

in the Form Four English Language Syllabus. Later, these samples were given to the expert and novice raters who marked and assessed them (also giving comments/feedbacks), as they usually would have done with their students' writing, through verbal protocol analysis. The researcher decided to add extra notes to remind him of the procedure prior to reading the VPA instruction to the participants (see Appendix 6b). There was no change to the VPA instruction as the two participants were in agreement that the instruction was clear to them.

3.6.2 Think-aloud Verbalization

The pilot testing of this data collection method served two purposes. The first purpose was to give the researcher a clearer view of the difficulties that might be faced during the actual data collection process. The second purpose was to provide an opportunity for the researcher to practice facilitating the think-aloud verbalization session and later practice the technique of analysing verbal protocols.

Prior to the verbal protocol analysis by the participants, the researcher had explained the purpose of the think-aloud session and what would be done with the results. The researcher had also explained the basis on which the participant was picked to participate. He then sought permission from the participants to have the session video-recorded. After that, the verbal instruction was read out to the participant. When asked if the instruction was clear to them, both participants agreed that they understood what was required of them. However, they said they were not familiar with VPA though they had the confidence to carry it out.

At this juncture, the researcher found it necessary to give the participant a trial so that he/she would get use to this rather awkward process of thinking aloud. When the

participant was ready, he was then provided with a student's written work. A pencil was provided, and his/her verbalization was video-recorded. During the session(s), the researcher took notes of any salient behaviour or action that might useful cues in the analyzing of verbal protocol analysis. Towards the end of the VPA, the researcher found it necessary to remind the participant to write comments on the student's written work as a feedback and give his/her rating of the writing as a whole. This was also noted down in the modified VPA instruction (Appendix 6b) later.

Later, the video-taped sessions were transcribed before it could be subjected to further analysis, using a coding scheme, to identify the knowledge states and conceptual operators used by the expert and novice raters in drawing the lines of reasoning embedded in their responses to students' writing. The VPA data analysis guidelines as discussed in sub-section 3.8.1(e) did not require any change as the researcher did not encounter any problem in identifying the knowledge states and conceptual operators related to the raters' VPA.

This pilot study had provided some feedback to the researcher on what issues would arise during the actual data collection. Thus, the researcher would be better equipped with reasonable solution to address the issue(s) to ensure the validity and reliability of the data collected is maintained.

3.7 Procedures of the Study

Procedures involved in this study included gaining access to the study sites, data collection procedures and ethical procedure. These three aspects of procedures are discussed in the following sub-sections.

3.7.1 Gaining Access to Study Sites

For phase one of the study which is the trialling of the ER mental model, the researcher gained access to the sites of study by obtaining written approval from the Education Planning and Research Development Unit (EPRD), Federal Education Ministry (Appendix 1) and Sarawak State Education Department (Appendix 2). Then, the researcher obtained the approval from the principals of the schools concerned to carry out the study with the teachers and students (Appendix 3, 4 and 5).

3.7.2 Data Collection Procedures

This study employed a multiple data collection method to enhance the accuracy of the information in the case study and thus improve its reliability (Yin, 1994 and Stake, 1995). Both qualitative and quantitative data were collected in the two phases of the study: modelling of the ER mental model and trialling of this mental model. The data were collected through verbal protocol analysis, which is the think-aloud verbalisation of the task of assessing and giving feedbacks to students' writing, interviews, and field notes and observations.

(a) Verbal Protocol Analysis

Phase one of this study is the **modelling of the ER mental model**. The main data was collected through the verbal protocol analysis method. The researcher studied the transcript of the following "purposefully select participant" (Creswell, 1994, p. 148) as they 'think-aloud' in assessing and giving feedback on students' writing. Collecting the protocol data involved presenting the problem tasks to participants along with verbalization instructions (Appendix 6b), and video-taping of the verbal protocol analysis sessions.

Prior to the verbal protocol analysis by the participants, the researcher explained the purpose of the think-aloud session and what would be done with the results. The researcher also explained the basis on which the performer was picked to participate. He then sought permission from the participants to have the session video-recorded. The researcher had decided on video recording over audio recording as the task would involve reference to visual stimuli or reference materials. After that, a verbal instruction was read out to him/her as in Appendix 6b.

Protocol analysis instructions generally involved telling the participants to think-aloud while performing the task and instructions given emphasized talking continuously, being complete, and not worrying about saying irrelevant things. They were not allowed to refer to any sorts of marking scheme and error-correction guidelines but merely analyzing and assessing the sample writings, using their existing knowledge and experience. By doing this, the researcher would be able to gain insights into how the expert rater could be differentiated from the novice rater in terms of their mental cognitive processes.

In the initial stage before the actual VPA task is carried out, the researcher gave the performer a trial run so that he/she would get use to this rather awkward process of thinking aloud. The idea of trial session like this was also emphasized by Lee (2001) and Pugalee (2004) in their studies. When the participant was ready, he was then provided with three target writings with no writers' name visible. A pencil was provided, and their verbalizations were video-taped. During data collection, the researcher took an unobtrusive role though he may give "keep talking" reminders after long pauses.

The context provided by the writer would not be in focus, thus raters' assessment feedback would not be not influenced by the writer's personal variables like catering to his/her proficiency level, individual needs and perhaps his/her personal development of writing skills. This research was confined to just analyzing and assessing the text itself, while exploring the mental cognitive processes of the participants in the assigned specific task.

During the actual data collection, participants were individually video-taped in an enclosed room arranged by the school authority. Later, interview questions were posed if deemed relevant to explore the participants' mental processes as they embarked on the task and these. These interview sessions were also video-taped. The video recordings of the verbal protocols and interview session were converted into a typewritten verbatim transcript. The data collection activity in phase one of the study ended with a session of verification of their video transcripts by all the participants (also known as member checks as discussed in sub-section 3.7.4), to ensure credibility of the data collected.

(b) Interviews

The data also came in the form of qualitative data derived from interviews, especially in **phase two** of the study – **the trialling of the ER mental model**. In phase one of the study, the interview data gathered was discussed briefly under sub-section 3.5.3. In this study, the researcher referred to Patton's (2002) five types of questions - experience/behaviour questions, opinion/values questions, feelings question (affective questions), knowledge questions, and sensory questions as in Appendix 7.

Five selected participants were interviewed two week after they had attended the workshop to give them time to use the strategy they have learnt in their actual practice. The implementation of the training workshop was discussed thoroughly in sub-section 7.2.3. Before the interview session, the researcher briefly informed the participant of his purpose, and make assurances (if necessary) of what that would be said in the interview would be treated confidentially. After the participant gave his or her consent, the information was tape-recorded and notes were taken during the session, especially on the non-verbal expressions.

During the interview sessions, which lasted about fifteen to twenty minutes, participants were asked on how they perceived the effectiveness of the workshop through the open-ended questions as in Appendix 8. The questions above are just a guide to enable the researcher to stay focus on information that helps to answer the research question. The researcher also posed related questions to validate the participants' response or to probe further into related thematic concern.

Employing a semi-structured interview method was an advantage as the researcher did not want to control the content of the interview too rigidly, apart from managing time more efficiently. Opened-ended interview allowed the participant to shape the content of the interview. The researcher would encourage the participant to talk in the area of interest and then the researcher would probe more deeply, picking up on the topics and issues that were initiated by the participant.

Immediately after the interview, the researcher went through the interview notes with the participant to validate the data taken during the session. The recorded interview session was later transcribed and the results were also shown to interviewees for

verification (member checks) and to find out if they had any additions or revisions to the content.

The qualitative data collected was used to construct the picture of how useful the ER mental model would be in helping the novice raters acquire the knowledge and skills of the expert raters in responding to the related task of assessing and evaluating the students' writing.

(c) Field Notes and Observations

During the duration of this study, the researcher had also collected qualitative data in the form of field notes and observations. The researcher took down field notes on anything that he observed during the verbal protocol analysis sessions or any complimentary remarks from critical friends (in this case they were my fellow colleagues) that would give rich data for later reconstruction of the phenomena being under studied.

In this study, there are two types of observations that were employed – unobtrusive and obtrusive. In unobtrusive observations, the researcher would not interact with the person observed (especially when he/she is doing the think-aloud process). After the participant was given the instruction on the specific task, he or she would be left alone to perform the necessary task so as to not interfere in any way with the actions of the person being observed.

During the verbal protocol analysis, the researcher merely observed and recorded the actions and verbalizations of the task performer or rater. There was a possibility that the participant being observed would not perform in typical manners. Thus, to remove

the Hawthorne effect, the researcher found it necessary to assure the participant that there would be no right or wrong method in assessing and giving feedbacks to students' writing. They would only be required to perform the task assigned to them in the manner they would usually do using the method(s) they were familiar with (usual practice).

On the other hand, there were certain tasks that would not be directly observable because they primarily involved mental operations. The result of mental operations may be obvious at the end of the problem solving task but the mental operations would be missed if we rely on unobtrusive observation. Thus, when this happened, the researcher would employ obtrusive observation where he asked the participant about what he/she was thinking at certain stage of the verbal protocol analysis or why a certain action was made. However, in this study, the researcher would only resort to obtrusive observation after the verbal protocol analysis was completed.

Regardless of whether the unobtrusive or obtrusive technique is used, the researcher would comply with the characteristics of a good observer as postulated by Norris (1984) that include:

- not allowing emotions to interfere with good judgement
- being alert to every situation and carefully considering the observation
- having no conflict of interest
- being skilled in observation techniques and skilled in observing task operation
- understand the operation being observed
- not having preconceived notions about the outcomes
- using as precise techniques as possible, and

- allowing adequate time and opportunity for observation

By adhering to the above guidelines, the researcher hopes that the quality of observation in this study is upheld thus would ensure rich data to be collected for the final analysis.

3.7.3 Ethical Procedures

The most important thing to be considered before the qualitative data could be collected in the two phases of the study was to have the informants or agree to participate in the verbal protocol analysis task and to be interviewed. Each participant was given a consent form (Appendix 4) that explains their rights as research participants. In the consent form, the researcher made it known to each the participant all the possible threats, consequences of being his participants and his professional obligation as a government servant. The researcher gave them the assurance that he would preserve the confidentiality and anonymity of participants.

In due course where the researcher would have to reveal the data to the authority concerned, he would make sure that the participants would be informed to get their consent. And when necessary, the researcher would destroy any part of the data that the participants want to withhold, in particularly the information that would jeopardize his or her basic rights. The researcher only accepted those who agreed to these terms and conditions as participants.

3.7.4 Credibility of the Study

According to Mertens (1998), credibility refers to the internal validity of the study. There are various ways to achieve credibility of a qualitative study. In this study, member checks and triangulation were used.

Member checks involved verification of data collected and analysed with the participants of a study (Merten, 1998). In this study, it was usually done at the end of the data collection session. After the researcher had transcribed the verbal protocols, he asked the participants to validate the content as they meant them to be and to ensure the true value of the data. For this study, there were few instances where the participants disagreed with the researcher's description, especially with the choice of certain words or phrases of the interview transcripts. This was expected as the interviews were not video-taped or audio-taped upon special request from workshop participants (CPs).

Another way to achieve credibility of data was triangulation (Patton, 2002) which involved checking data that had been collected from different sources or methods for consistency of evidence across sources and methods of data collection. According to Maxwell (2005), achieving consistency of evidence is a way to rule out validity threat. In the first phase of this study, the interview protocols, as a follow-up after the VPA session, were triangulated against the transcribed think-aloud protocols. This was indeed very important to get the clear picture of the assessing strategies and processes of each participant deployed during the problem solving task. In the second phase of the study, the interview protocols were triangulated against the observational data during the implementation of the training sessions, so as to ascertain the novice raters' perception of the effects of the workshop on their actual practice in their respective school.

3.7.5 The Trialling of the Expert Rater Mental Model

In the second phase of the study, trialling out the ER mental model in the form of training sessions (workshops) with a group of novice raters (beginning teachers) is described in detail. In phase one, the mental model has captured the essence of assessment using VPA and is based on assessment feedback from the expert raters. In phase two, it should be made clear that the trailing of the ER mental model is by all means not for the purpose of making differentiation between the expert and the novice raters or to claim that the model is applicable to all contexts. The ultimate purpose of the workshops is the check if the teachers are following the identified cognitive processes based on the mental model. It focused on how the novice raters can be trained to be more sensitive to the operations of the ERs that are captured in the model. It is, therefore, not trying to correlate the mental model and the outcome of the workshops.

The data collected through interviews was analyzed to gain insights on the participants' perception on the effectiveness and implications of the training using the formulated conceptual model in phase one of the study. It is hoped that through this workshop, the novice raters of students' writing would have a renewed self-awareness of their crucial role as assessors, evaluators and facilitators to guide their students in developing their writing skills. This awareness can help them to see the importance of understanding their students' writing skill better and support the latter for learning and improvement.

(a) Training Workshops

Based on the ER mental model (refer to Figure 4.1) that was constructed in Chapter 4, the expert raters seem to focus a lot on knowledge states such as content and

development of paragraph, apart from grammar. As presented in Sub-section 4.2.4, it is well-acknowledge that in assessing a piece of writing, a rater needs to be able to identify whether or not the students are in control of content, syntax, grammar, mechanics, organisation, word choice and choice of expression.

Due to time constraint, this 2-day workshop gave emphasis on choice of expression, especially on wordiness and variety of sentences. It also exposed participants to marking symbols and allow them to explore related assessing strategy to develop their assessing skill. In terms of the conceptual operators to be given emphasis during the workshop, the ER mental model suggested that the meta-reasoning aspect needed to be given top priority by the raters. Thus, emphasising on content aspect of writing would employ a high level cognitive process such as meta-reasoning to help raters in the problem solving of assessing and giving feedback on students' writing.

More emphasis was also needed to be given to conceptual operators like data exploration and data explanation, as they complimented the extensive use of meta-reasoning during the problem-solving task. Based on the finding of the present study discussed in sub-section 4.2.3(c), this pattern of cognitive processes had enable the expert raters to articulate more concrete and in depth comments on how they assessed writings and be more objective in their decision making.

(b) Contextual Background of Target Group

The training workshop in this chapter was meant primarily as a short training programme for novice raters of students' written work. These novice raters were beginning teachers of English, who were preparing their students for the *Sijil Pelajaran Malaysia* (SPM) Examination. Any future training for this type of

programme would include the refinement of partnership in practice with school relating to training arrangement and communication network which support this arrangement at various levels in school and teacher training institutes.

The novices/participants brought with them a wide and varied range of knowledge, expertise and attitudes. Acquiring the necessary skills in assessing and evaluating writing remained a main concern of theses participants who were directly involved in developing the students' writing skill in their respective schools. They were the ones who would continue to oversee the following aspects of the training entrusted to them: ability to assess students' pieces of prose in accurate Standard English, and the relevance and creativity of students' response to the writing task. With preparing students for the SPM examination in focus, particular emphasis was placed on the participants' assessing skill for content and language use in their students' writing.

The focus of language use would deal with the participants' ability in identifying errors, the knowledge that they should have from their own professional and academic training. However, it was one of the objectives of this training to ensure that the participants were able to assess writing effectively. Through the researchers' observation of twenty four years teaching English at various levels in schools and a teacher training institution, it can be concluded that there are practically three groups of teachers, especially beginning English teachers: the ones who know and can spot errors, the ones who know but fail to spot the errors, and the ones who do not know and cannot spot the errors. Some teachers focus too much on errors identification that they ignore the content aspects of writing: relevant and precision of ideas, and appropriate choice of expressions that make the piece of writing interesting.

The teachers' obsession of errors identification is some kind of 'red pen syndrome' which would not be necessarily useful to help students to be real good writer in true sense. This is also a need for teachers to help their students to develop inquisitive and creative mind which is very crucial in producing quality writings. Thus, with the SPM Examination being in focus, this training would help participants in coming better assessors and evaluators of their students' ability to:

- write sentences of various lengths and types,
- use a variety of sentence structures an aid to meaning,
- use a wide range of vocabulary with precision,
- punctuate accurate and effectively,
- write in paragraphs which are appropriately linked and show precision,
- spell accurately over the full range of vocabulary,
- respond with relevance and precision to the chosen topic, and
- engage and sustain the interest of the reader.

(c) Aims of the Training Workshop

The training workshop sought to reinforce competences such as knowledge, skills and awareness which participants had in widely varying degrees. It aimed to introduce participants to the appropriate assessment techniques and strategies to enable them to assess and give feedback on students' writing effectively and to become more competent and informed assessors of learning.

The training package consisted of a combination of activities in varying form and length of treatment (refer to training materials in Appendix 14). They were merely *sample* activities which could be carried out in order to enable the development of

more effective assessing and evaluating skills among the novices. The training workshop hoped to achieve the following aims:

- to enable participants to review and evaluate their current practices in terms of the assessment of students' writing skills and abilities
- to help participants to understand the necessary components to good writing, especially on choice of expression, especially on wordiness and variety of sentences
- to familiarize participants with strategies for assessing writing
- to develop awareness among the participants of their own strength and weakness in assessing and giving feedback on students' writing
- to provide hands-on experience in the assessing and giving feedback on students' sample writings through collaborative tasks and feedback sessions

Knowledge, skill and personal awareness are the main elements of effective process of assessing and evaluating writing. These competences will prove useful beyond the immediate demands of effective English teachers. They are elements which constitute professional learning, growth and development.

(d) Implementation Strategies

This training workshop encouraged critical self-assessment of the participants' own work and classroom practices. It involved objective analysis of teacher-designed assessment instruments, analysis of requirement of high-stakes public exams and their impact on novice raters' classroom practices. It also involved the engagement of novice raters in collaborative tasks, workshops, presentations, feedback sessions and discussions.

Reflective grids were recommended for use in all materials. Course participates could either 'grid' after every activity or at the end of the whole session. They could use the 'griddled' notes as a basis for starting or ending a session to show its use and to enable novice raters to develop an understanding of the rationales behind the learning experiences the induction and training sets up for them (Borg, 1995, p. 17).

At the beginning of a session, the trainer (Tr) had to establish the group rules, distribute the reflective grids as group or pair work, and then outline the session. During a session, the trainer gave clear instruction in an active and interactive style of training, and drew attention to ground rules when necessary. Trainer summarised or rounded up at the end of the whole session or activity to link it to the next session or activity. This helped trainer to crystallize points raised and tied up loose ends together. It also helped to make sense of what the novice raters had gone through and made important links to what they had done or were going to do. Trainer reinforced this by producing summary notes and thinking questions.

(e) Feedbacks on Workshop

To gain further insights on the workshop, five selected participants were interviewed one week after they had attended the workshop. During the interview sessions, which lasted about fifteen to twenty minutes, participants were asked on how they perceived the effectiveness of the workshop through open-ended questions which were already discussed in sub-section 3.7.2(b). The researcher also posed related questions to validate the participants' response or to probe further into related thematic concern.

3.8 Data Analysis Procedure and Data Interpretation

This section discusses the procedure involved in the analysis of qualitative and quantitative data in the study.

3.8.1 Qualitative Data

The case study employed a multiple data collection method to ensure the reliability of the information in the case study (Anderson, 1990; Yin, 1994; Stake, 1995). Basically in this study, the researcher has two sets of qualitative data: the data from the verbal protocol analysis data through the participants' verbalization of the task and also the interviews data that was taken from the participants after their training workshop that used the expert mental model derived from the earlier phase of the study. In the first phase of the study, before any data analysis could be carried out on the video-taped verbal protocol analysis and interviews of the participant, they had to be transcribed first.

After the researcher had transcribed the video-taped verbal protocol and interviews, he showed the transcripts to the participants to validate the content as how they meant them to be. Actually, these transcripts are the main data of this study that would help the researcher to increase his own understanding of the 'mental processes' of the participants in their task of assessing and evaluating writings, and would enable the researcher to present what he have discovered to others.

a) Verbal Protocol Analysis Data

The video-taped verbal protocol analysis of the expert and novice raters was transcribed and coded using a set of standard conventional transcription symbols. The transcription layout was principally based on conventions used by Jefferson (1978)

with modification where appropriate. The participants were allowed to read through the piece of writing before doing the thinking aloud. However, participants would occasionally interject a thought before reading the complete data group.

Thinking-aloud statements would be included in [brackets] within the paragraph when this happen. Syntactic breaks in thinking aloud would be designated by commas, periods, or semicolons as appropriate. Pauses of short duration would be designated by an ellipsis (…). Typing the word "pause" within parentheses would designate pauses longer than 5 seconds. Abrupt changes in diction or corrections in speech would be indicated by a double dash (--).

The transcripts were analyzed using a certain coding technique. The coding technique depended on the nature of the task being analyzed, the degree to which a model of task performance have already been formulated, and the issues to be addressed in the analysis. In this instance, since the categories are particular to the task domain, the task is assessing and evaluating sample writing where no specific rules could be given.

The protocol analysis used in this study follows Newell and Simon's (1972) guidelines. Based on these guidelines, the first step in analyzing a protocol is to break down the transcript into short segments or phrases. Each segment would reflect an "assertion" made by a subject at a given instance. The boundaries of segments could often be determined by noting syntactic cues such as clauses, sentences and pauses in the protocol. This step in analyzing the protocol would yield a topic representation in which each segment would address a particular instance of reasoning behaviour on the task.

The topic representations were coded depending upon the analyst's decision as to how a particular topic segment maps onto the constructs or categories of the **coding scheme**. The coding scheme (see Figure 3.2 below) the researcher used to code topic representations is based on Newell and Simon's (1972) work. The scheme suggests that protocols consist of two types of semantic elements: the knowledge states and conceptual operations. In this study, the researcher had also included the third element- the lines of reasoning that is introduced by Hassebrock and Prietula (1992).

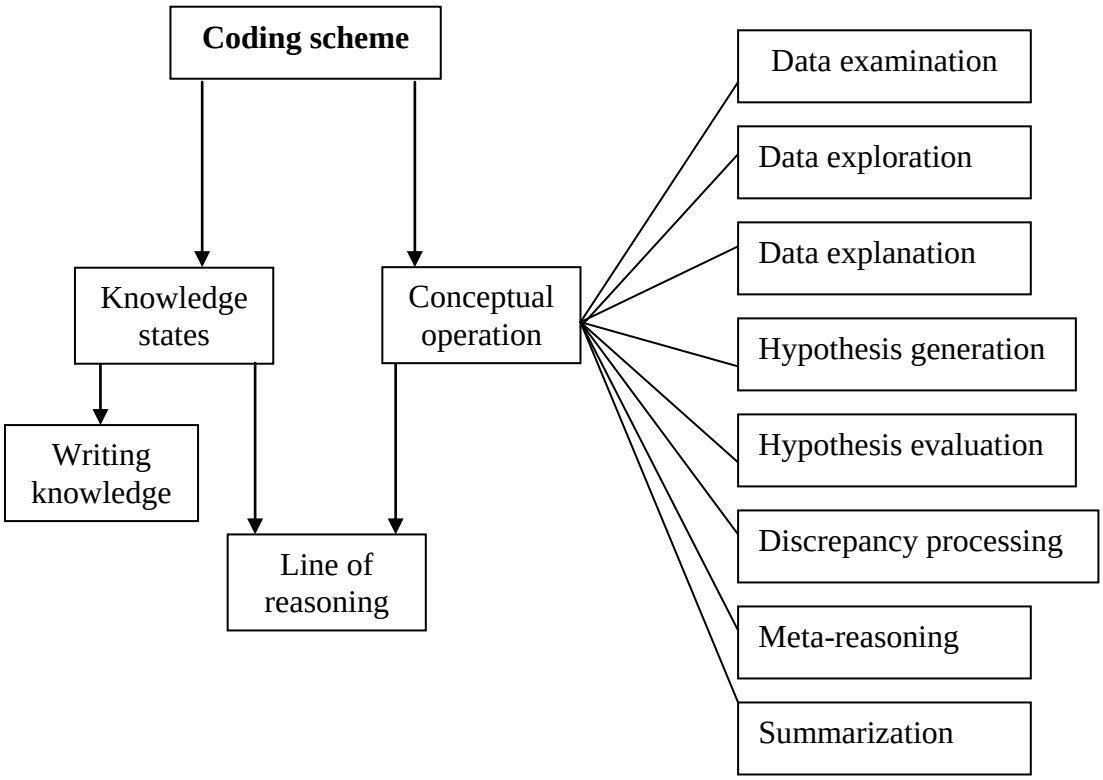

Figure 3.2 The coding scheme for assessing writing (adapted from Hasserbrock and Prietula, 1992, p. 662)

A given topic segment could consist of knowledge states which are units of information or knowledge retrieved from memory and used by an individual. Second, these knowledge states are associated with conceptual operations which are inferential processes that produce new or modify existing knowledge states. The general goal of

the analysis was to trace the sequence and patterning of knowledge states and conceptual operations used in response to a problem-solving task.

The third element is known as the line of reasoning which can be viewed as a problem-specific representation of problem-solving knowledge. Figure 3.2 above summarizes the components of the scheme to assess students' writing.

The verbal protocol analysis was accomplished through a series of procedures designed to identify knowledge states and conceptual operations produced by an expert or novice when solving a problem. The non-verbal behaviour of the participants was also available as the data for this study was video-taped. Non-verbal behaviours like head nods, body postures, gestures and facial expression of the participants during the process of verbal protocol analysis were noted where necessary in the analysis in order to emphasise and/or disambiguate unclear, incoherent meanings of the discourse (Jariah, 1999).

(i) Knowledge States

The knowledge states in this study refer to a type of protocol representation (Newell & Simon, 1972) which identifies units of writing knowledge that were used by the expert and novice ESL teachers or raters in this study. This included their knowledge of clear writing, and evaluating and assessing skills needed in the task of rating and assessing writing.

As stated by Victori (1999), the writing knowledge is important for student writers and language learners of English to write in such a way that it makes comprehending the written text easy for the audience or the reader. Apart from that, it serves as a repertoire of tools for students to construct an effective piece of writing. As such,

writing in a specific language calls for observance to text conventions of the particular language. In this study, the writing knowledge that a teacher/rater should have/use in assessing writing are categorised into five divisions as in Table 3.6.

Grammar	• grammaticality
Mechanics	• spelling • punctuation
Content	• clarity • completeness • exemplification • non-English terms' equivalents • avoidance of translation • reasonable length • wordiness
Organisation	• central idea of text • development of paragraphs • use of discourse markers • cohesion • coherence
Vocabulary	• choice of words • introduction of key terms, concepts, and individuals

Table 3.6 Writing knowledge used in assessing writing

(1) Grammar

One of the most salient areas of difficulty for L2 writers is grammar (Leki, 1992). Grammatical errors, especially those which effects communication are likely to cause comprehension problems (Hewings, 2005). Hughes (2005) asserted that grammar is crucial in the writing process as it serves as the linkage between the intended meaning (by the writer) and the perceived meaning (by the reader or audiences). These global errors affect overall sentence organization and would probably hinder communication. Readers should not be burdened to try hard to decipher what has been written. In fact too many ungrammatical sentences would give a bad impression to the whole text even if the message may be understood by the reader (Xie & Jiang, 2008). Some of

the common grammatical errors (Fournier, 2004; Hughes & Heah, 2006) are errors related to:

- Adjective
- Adverb
- Article
- Capitalization
- Clause
- Colloquial Phase(s)
- Comma Splice
- Comparative/superlative
- Conjunction
- Double Negative
- Ellipsis
- Preposition
- Incomplete Sentence
- Incorrect Verb Form
- Subject/Verb Agreement
- Infinitive/Split Infinitive
- Pronoun Case/Number
- Punctuation
- Redundant Word(s)
- Relative Pronoun
- Repeated Words or Punctuation
- Run-on Sentence
- Sentence Variety
- Similar Words
- Spelling
- Split Words, and
- Tenses/Tense Shift

(2) Mechanics

Discussion on mechanics would include spelling and punctuation. Celce-Murcia (2001) stated that while language mechanics are often taken for granted by the students, they are in fact very crucial in conveying the message that they intend to send across. Inappropriate spelling and lack of appropriate punctuation will very likely make our writing fuzzy and unclear. This elements in writing assessment have given due emphasis in the holistic rubric for the continuous writing section provided by the Malaysian Board of Examination (2006) for the Malaysian Certificate of

Examination (MCE) or *Sijil Peperiksaan Malaysia* (SPM) as discussed in sub-section 2.2.1(a).

(3) Content

Discussion on content of a piece of writing would include elements of clarity, completeness, exemplification, non-English terms' equivalents, avoidance of translation, reasonable length, and wordiness (Oshima & Hogue, 1997; Savage & Mayer, 2006). Olson and Raffeld (1987) define content feedback as teachers' comments pertaining to concepts in and structure of students' writing.

The first element in the discussion of content in writing is **clarity.** Clarity in writing is at macro level when meaning or idea in a sentence is involved, for instance, a confusing sentence or the wrong choice of expression is an indicator that clarity is affected in that sentence. According to Caulk (1994), clarity in a sentence can be achieved by making the right choice of words (and/or expression), using the correct punctuation mark where it is needed, being grammatically accurate, using correct spelling, being unambiguous and showing connection between what has been said before and what is to follow. Specifically, grammatical aspects, mechanics of writing, choice of words etc may contribute to clarity but they are at micro level (vocabulary). The next element, which is **completeness,** refers to the flow or thoughts of a text which should be complete. Lane and Lange (1999) emphasised that completeness is one of the key indicators between good writer and bad writers. Good writers tend to make full use of this element to produce a solid piece of work while bad writers often left sentences dangling as fragment. A sentence should have a subject and verb, at its most basic level.

Another element of good writing is **exemplification,** which helps reader to understanding the content better. Writing empty phrases or digressing from the topic at hand would obscure clear writing. Statements that we make can be clarified using examples to validate our point of view, convince our readers and provide them with useful insight of the topic (Paquot, 2008). Therefore, appropriate exemplification helps us to have clarity in our writing.

As for the next element, which is **non-English terms' equivalent** in a writing, give one can avoid misinterpretation of what one is trying to convey. When we introduce non-English words in our writing, we need to give their equivalent expressions in English within brackets, or as a brief translation following the first mention of these non-English terms.

Another feature of a good writing related to content is **avoidance of translation.** Translation from one language to another is rather common. It could be direct and literal or indirect but nevertheless accurate. It is the direct translation from the source language to the target language that causes problems in the sense that the literal translation sometimes appears awkward and biased. Although translation cannot be fully eliminated by bilinguals while performing in the teaching and learning process, it should be the last resort. According to Smith (1994), good writers should be able to think in the target language rather than depending on direct translations. This would allow better fluidity in the writing process. Erroneous co-occurrence of words may be the result of word-for-word translation. For example, Tamil-speaking students in India who write 'drive the cow' are obviously translating from their native language-Tamil. The equivalent of 'drive' in Tamil can be applied to cow, horse, bus and car.

Finally, a good writing would always have sentences of **reasonable length**. The sentences that make up a written text ought to be of **reasonable length**. If the sentences are somewhat long, readers can lose track of the writer's train of thought. It is not always easy for all readers to process long sentence. However, the sentences need not always be short and simple. A good piece of writing would have a combination of simple, compound and complex sentences.

(4) Organisation

Discussion on the organisation of a piece of writing would include the elements of central idea of text, development of paragraphs, use of discourse markers, cohesion, and coherence. The element of central idea in writing makes the whole text clear. This is also known as the *thesis statement*. The thesis statement is the main idea, opinion or theory in a whole piece of writing (The Cambridge International Dictionary of English, 1995). This thesis statement or central idea has to be developed over the various paragraphs that make up the entire text, with each paragraph developing a particular point. Thus, the thesis statement is different from a topic sentence. All the paragraphs are subordinate to one main idea that is expressed by the thesis statement.

Another element of good writing related to organisation is a good **development of paragraphs**. According to the Cambridge International Dictionary of English (1995), a paragraph "usually deals with a single event, description, ideas, etc." A series of related paragraphs make up a text. A good piece of writing would usually have paragraphs developed around a topic sentence. The *topic sentence* in a paragraph is a single sentence that provides the main idea for the whole paragraph. The remaining sentences serve to support and revolve around the topic sentence. Even though it can appear anywhere in the text but very often it is the first sentence in a paragraph. The

topic sentence is the main sentence which sums up the point developed in the paragraph.

In addition to that, a well-organised writing contains the **use of discourse markers** or conjunctions appropriately as they signal different meanings (Fraser, 1999). Discourse markers help to provide a clear flow of ideas from the beginning until the end. In English, a conjunction does not appear at the beginning of a paragraph as it usually refers to something immediately before it. Some writers have the inclination to begin a paragraph with a conjunction when writing in English as a second language. The discourse markers should be varied and one needs to avoid using certain conjunctions more than is necessary.

Coherence is another element related to the organisation of writing. It is operationally defined by Halliday and Hassan (1976) to refer to both the unity of a text in terms of a logical continuation of sentences within paragraphs and the unity of a text as a whole, viz. in terms of how it maintains a clear focus on the topic at hand and signals transitions between topics. Coherence in writing can be enhanced through:

- Meaningful language
- Connectedness of ideas/flow of discourse
- Relevance of ideas
- Use of cohesive ties
- Appropriate genres
- Familiarity to the writing convention of the target language, etc.

Finally, the element of **cohesion** in writing is seen as an important factor contributing to the coherence of a text (Halliday & Hassan, 1979). Cohesive texts are also likely to

contain both grammatical and lexical cohesion. Halliday and Hassan (1979) have identified five types of ties to achieve cohesion. They are reference, substitution, ellipsis and conjunction which contribute to grammatical cohesion and repetition of lexical items, expression of sense relation and collaboration which help to realize lexical cohesion. The effective use of cohesive links contributes to variety and clear writing.

(5) Vocabulary

Discussion on vocabulary would include choice of words, and introduction of key terms, concepts, and individuals. In a sentence, clarity at micro level (vocabulary) may be affected by a wrong **choice of word**, or one that is imprecise (Fournier, 2004; Hughes & Heah, 2006). Writers should also avoid archaic or obsolete words which have become rare in common language use. The 'vanquished' may be more appropriate in literary usage. 'Defeated' is the word in current use. Nowadays we describe our heroic soldiers as brave, not valiant. Finally, writers needs to make clear key terms, concepts (Fournier, 2004) and the individuals they are referring to by adding helpful identifying phrases, especially when these terms and people are initially introduced in a text .

(ii) Conceptual Operation

In this study, the researcher used the eight types of conceptual operations and their specific operators used by Hassebrock and Prietula (1992), which were discussed in detail in sub-section 2.3.1(b) in the previous chapter, for analyzing verbal protocols of the expert and novice raters. They are data examination, data exploration, data explanation, hypothesis generation, hypothesis evaluation, discrepancy processing, meta-reasoning, and summarization. They are used to characterize distinct segments

of a raters' problem solving behaviour. To give a more detailed representation of knowledge and reasoning behaviour required by the task, each basic conceptual operation is further analyzed and interpreted into raters' lines of reasoning.

(iii) Lines of Reasoning

Based on the analysis of the knowledge states and the conceptual operations used by the raters in assessing and giving feedback on students' writing, the next procedure would be to draw out the raters' lines of reasoning (LoR) (refer to sub-section 2.3.4) which is analogous to the development of a problem representation or a mental model (Hassebrock & Prietula, 1992). The knowledge content of a LoR would serve to discriminate expertise as well as individual differences between the expert and novice raters.

(vi) Steps in Processing and Analysing the Verbal Protocol Analysis

A sequence of steps would be undertaken to prepare the verbal protocols for analysis and interpretation. The video recording of a case would be converted into a typewritten transcript.

Step One

The analysis would proceed in a sequential manner beginning with the first group of data. Knowledge states that would be identified in this first step of analysis would be represented directly on the transcript. For each data group, the researcher would underline all the knowledge states which would be contained in the transcript.

Step Two

Each knowledge state of writing identified in the first step would be associated with one of the possible basic conceptual operations. Each identification of a basic

operation would represent a discrete problem solving segment of assessing and giving feedback on students' writing.

Step Three

After identifying each basic operation associated with a protocol segment, the researcher would then identify the specific operators used to produce the knowledge state or states within the segment. A sample transcript fragment and coding of a novice rater from the pilot study is shown in Table 3.7:

A wrong use of word (phrase) and preposition is found in the third paragraph. Despite having a flow in the story, there are also confusion of words, wrong tenses and a spelling error.
[(1) Wrong tense used, ('change' instead of 'changed') is used in the third paragraph… (pause).] [(2) Wrong (use of) preposition…[(3)The story has a flow (coherence); from the home to the Damai Resort…Again preposition…sit 'in' supposed to be sit 'under']…. [(4) Confusion of words ('to reading newspaper')… (pause)] Err….. [(5)Again tenses….tenses again… the student tend to use…tends to mix up between past tense; … sleep supposed to become 'slept'… becomes 'sleeped'...] [(6) Spelling error ('muzium' instead of 'Museum')… (pause). [(7) Tense… ('to called' instead of 'to call')]
2.1 Data-examination: identify → (wrong tense) 2.2 Data-examination: identify → (wrong (use of) preposition) 2.3 Data-meta-reasoning: self-evaluation → [The story has a flow... (coherence)] 2.4 Data-examination: identify → (preposition error) 2.5 Data-explanation: infer → [confusion of word...(choice of word)] 2.6 Data-examination: identify → (tense-Past Tense) 2.7 Data-examination: identify → (spelling error) 2.8 Data-examination: identify → (wrong tense)

Table 3.7 Example transcript fragment and coding of a novice rater from the pilot study

Step Four

Lines of reasoning that help in deciding the types of feedback a rater would give on students' writing would be identified after the analysis of knowledge states and conceptual operations. Coding for a LoR would be carried out in two steps.

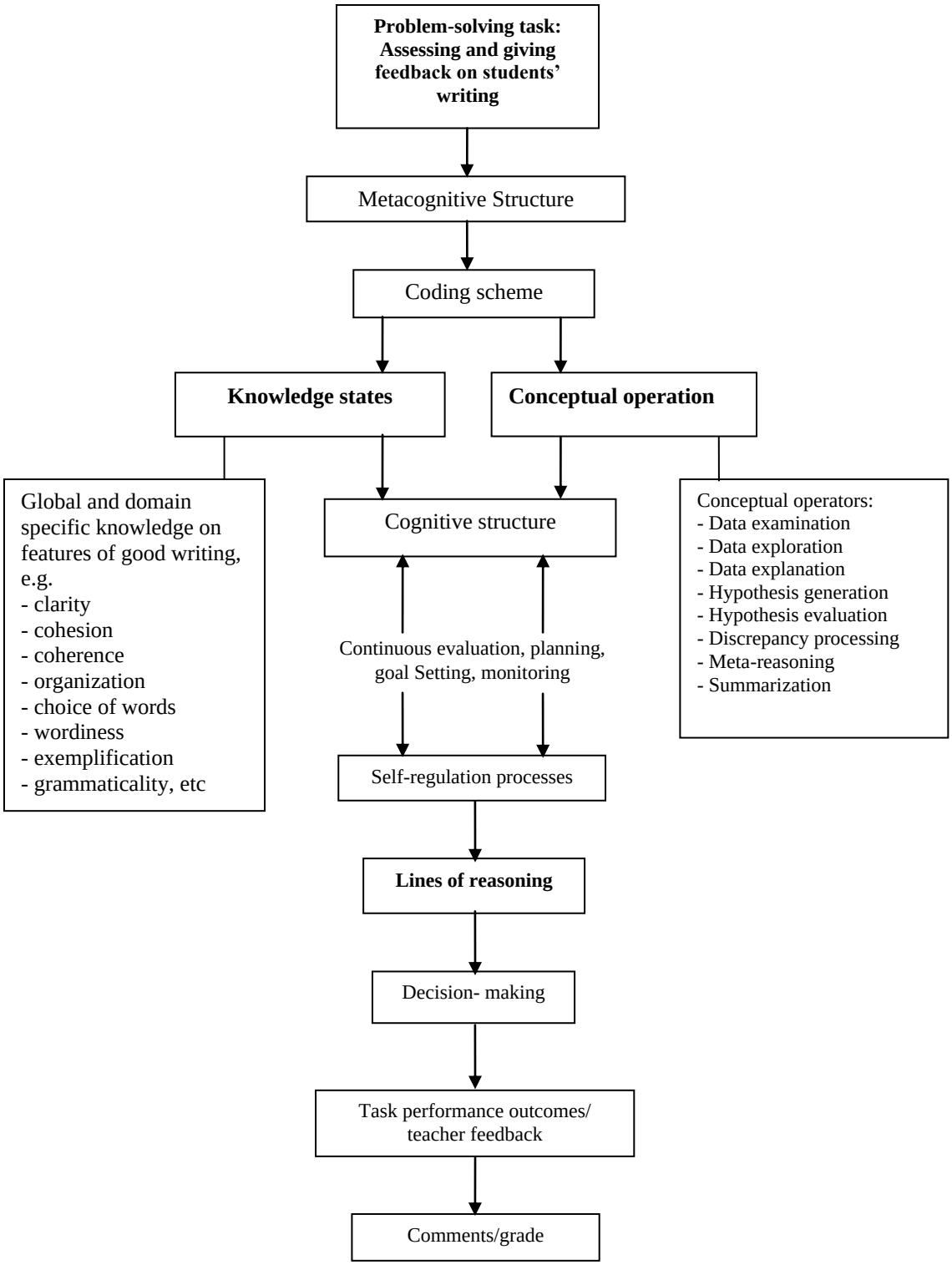

Figure 3.3 Canonical mental model for assessing and giving feedback on students' writing

First, a general canonical model of assessing and giving feedback on students' writing is constructed based on domain-theoretic principles (given in Figure 3.3 above). This canonical model reflects the process of assessing and giving feedback on students' writing. It provides a hypothesized qualitative description of the knowledge states and

the conceptual operation that a rater may use in the analysis of a problem (analyzing error(s) in a student's writing).

It is important to note that this canonical model does not refer specifically to how the knowledge state(s) may relate to the conceptual operation or operators in drawing the kind of decision involved in identifying the errors and deciding the appropriate comments or feedbacks; rather, it serves as a domain-theoretic guideline or template, to help determine specific reasoning contexts evidenced by the protocol (which does reflect knowledge states and operators). Specifically, this canonical model can used to facilitate the construction of other more specified canonical models, for instance, an expert rater may construct a certain model for describing how he identifies a specific error or vagueness in the writing which is related to 'clarity'.

Through verbalization of the task, the expert rater may try to reason out whether the sentence structure is being grammatically accurate, showing connection between what has been said before and what is to follow, and being unambiguous or not. In the process of doing do, he may used certain conceptual operators to accompany his lines of reasoning (which may differ from that of a novice rater). On the other hand, the novice rater may construct a different model where he may continually "be drawn away" by specific data to reason down a wrong path or unjustified ones.

The second step of LoR coding involves reviewing the protocol in the context of the relevant canonical model(s) to determine what elements of the canonical model, and what sequence, would be instantiated by the rater's reasoning process as evidenced by the knowledge states and operators. Although the canonical model would instantiate the necessary knowledge states involved in raters' assessing and giving feedbacks to

students' writing, the differences in the expert and novice raters' knowledge and experience can vary the sequence.

Apart from that, the relationship between the conceptual objects (the knowledge states, and conceptual operations) may be instantiated by the rater not as a simple conceptual operator, but as quite complex and detailed reasoning processes involving several intermediate knowledge states and operators functioning at multiple levels of abstraction. If the complex reasoning processes are successful, they would converge to form correct solution and eventually be chunked, readily available to be invoked as a higher level operator when similar situation occurs.

Thus, the LoR coding would reflect a summary trace of reasoning steps over a period of time in the context of a canonical model. After that, the researcher would construct a mental model showing the data cues at which a rater evoked knowledge states found in LoR. The application of this LoR coding in studying rater's reasoning in assessing and giving feedback on students' writing would reveal expert-novice differences in the use of knowledge states, conceptual operations and lines of reasoning.

(b) Interview Data

The second method of data collection was interviews. Interviews are useful source of case study evidence because they give the participants' construction of the reality around them and may help to provide important insights into a situation (Yin, 1994). Interview data in the form of qualitative data was collected from the participants in both phases of the study, the first one was collected immediately after each VPA session and the second one was conducted after the novice raters attended the training

workshops. They were interviewed two weeks after that to give them the opportunity to use the assessing strategies they learnt during the workshop.

To process the interview data that the researcher had accumulated, the data was read through a few times for certain words, phrases, patterns of behaviour, participants' ways of thinking, and events which were repeated or stand out (Bogdan and Biklen, 1992). After that, the researcher organised his data based on the coding system that he had created. Later, the researcher coded the data and identified emerging categories of issues or themes. Next, the researcher sorted out the data mechanically case by case according to the categories that he had identified – a crucial step in data analysis. Finally, further questions and possible routes of inquiry were devised to continue the effort in seeking for appropriate answers. The issue of verification has already been discussed in sub-section 3.7.4, where member checks and triangulation were used to achieve credibility of the data, and the study as a whole.

(c) Field Notes and Observations

Apart from verbal protocol data and the quantitative data, the researcher kept a research journal throughout the case study to note pertinent ideas, issues and decisions (Bouma, 1996). He used field notes and observations during the triangulation of data to reconstruct the phenomena being under studied. The field notes were evaluated to examine the participants' reactions to the verbal protocol activities, training activities and the materials presented in the training workshop.

During the protocol analysis session, the researcher would take note of the salient behaviour and interesting happenings. Notes were therefore made on how the participants understand the instruction given to them during the verbal protocol

analysis sessions, how they reacted to the certain aspects of the task in the verbal protocol analysis, how well they participated in the activity, points made during/after the training workshop, contradiction which may arise, how they resolved these contradiction and the general feeling of the activity/session. The data was sorted out according to relevant categories to fit in into the various sections of the findings and discussion. The triangulation process of all the data would be discussed further in the later part of this chapter.

3.8.2 Quantitative Data

The data collected through interviews (during the second phase of the study) was also quantified for easy analysis. After the qualitative data was recorded and transcribed, common themes were located. Then descriptive and conceptual categories were coded where the deductive method of coding was primarily utilised. According to the deductive method, categories would be chosen from the researcher's theoretical knowledge and the data would be searched for relevant passages (Miles & Huberman, 1984).

During the coding process, concepts or categories that clustered together were linked. The coded data was then transferred to data summary sheets which record the frequency and type of response to each item by the various groups in the study (King, Morris & Ftiz-Gibbon, 1987). This form of data was useful in the triangulation process of the findings to form a clearly picture of the phenomenon in study.

3.8.3 Triangulation

Triangulation of both the quantitative and qualitative data obtained would help the researcher to strengthen the validity of the research findings. More importantly, it

enhanced the researcher's own understanding of mental processes of the participants' while assessing and evaluating writing and how the formulated conceptual model could later be used in equipping beginning teachers with the necessary knowledge and skills to be proficient ESL teachers. Based on related literature, the researcher concludes that triangulation of the various data in the study will build stability and confidence in how data is being interpreted as it will illuminates problems and raises new questions (Freeman, 1998).

Actually, the quantitative data (derived from qualitative data) would give both breadth and depth to the study as it would be able to provide certain information related to the participants or their general attributes. On the other hand, the qualitative data, which were collected through verbal protocol analysis and interviews, would also give further depth to the study as the researcher would be able to assess the cognitive domain of the participant to gain insight of their attributes that affect their performance in assessing and giving feedback on students' writing. The data collection and data analysis procedure are summarised in the Table 3.8 below.

Data collection	Data analysis
Phase 1 (Jun 2006) (Modelling of the ER mental model) Stage 1 Verbal protocol analysis - briefing - video-taping	- transcribing recorded verbal protocol - analysing using coding schemes – knowledge states, conceptual operators - identifying differences and similarities - identifying lines of reasoning - Categorising and comparing comments - Formulating the ER mental model
Stage 2 Field notes and observation - debriefing after the verbal protocol session - audio-recording	- transcribing recorded interviews - categorising interviews data into themes - triangulating data with verbal protocol data
Phase 2 (Jun 2008) (Trialling of the ER mental model)	
Stage 1 Interview teachers (two weeks after they attended the workshop)	- Transcribing recorded interviews - Categorising interview data into themes - Evaluating effectiveness of training workshop

Table 3.8 Summary of data collection and data analysis procedures

3.8.4 Reporting the Findings

Since the analysis involved 50 video transcripts, verbal reports and observation/field notes, the researchers had to select certain video transcripts, verbal reports and observation/field notes to be presented in the following chapter to show the participants' cognitive strategies in assessing and giving feedback on students' writing and participants' perception of usefulness of the training workshop on assessing writing which was designed based on the constructed mental model of the expert raters.

Careful selection of certain participants' video transcripts, verbal report and observation/field notes also allows a better understanding of the participants'

cognitive processes and individual perceptions of particular situations pertaining to the contexts of the study.

In reporting the findings in phase one of the study, the participants' cognitive processes pertaining to the problem-solving task of assessing and giving feedbacks to students' writing are classified and described. The quantitative data derived from the analysis of the verbal protocol data of the expert and novice raters is compared and presented in tabulated form to show frequency of occurrence, especially of knowledge states and conceptual operators. This is supported by the related segments of their think-aloud verbal reports so that a clear picture of their cognitive processes can be drawn and compared between the two groups of raters.

The raters' lines of reasoning were derived from this analysis and these were later used to construct the raters' mental model to represent their respective cognitive strategy in assessing and giving feedbacks to students' writing. In phase two of the study, the findings are reported in the form of verbal report segments of their perceptions toward the training workshop and usefulness of the constructed ER mental model in helping the novice acquire the experts' necessary skills in assessing and giving feedback on students' writing.

3.9 Chapter Summary

This chapter started with an overview in section 3.1. It was followed by describing the design of the study in section 3.2. The research conceptual framework and phases of the study was illustrated and presented in section 3.3 to provide 'scaffolding' to the study as a whole. Section 3.4 mentioned about the sampling method, in particular the participants in both phases of the study and explained the rational for choosing them.

Section 3.5 touched on the instrument used in this study and highlighted the process of validation of the instrument.

Section 3.6 reported on the pilot study which was carried out to trial out the research data collection procedure and to investigate the feasibility of implementing the data analysis procedure. Procedures involved in this study, namely, gaining access to the study sites, data collection procedures, ethical procedure, credibility of the study and the training workshop were discussed in section 3.7. Section 3.8 provided an explanation on how the data of this study were analysed and to be reported. Like in the previous chapters, this chapter also ends with a summary of the chapter under section 3.9. The findings of the study will be presented and discussed in the following chapter.