

CHAPTER THREE

THE RESEARCH DESIGN

Introduction

Chapter One discussed the research and the reports upon which the Mathematics K-6 (NSW Department of Education, 1989) syllabus was based. The reports were found to identify and address many of the deficiencies that existed in Mathematics programs during the first half of the twentieth century, but they were also found to neglect the influence of affective variables in the learning of Mathematics. Similarly, much of the research upon which the syllabus was based was also found to overlook the role of affect in student learning. This neglect led the researcher to review the literature on affect and the affective domain.

Chapter Two's review of the Mathematics Education literature on affect, revealed that much has been written about attitudes and beliefs but very little has been written about the emotions, especially positive emotions. Studies of emotion that have been undertaken have been dominated by research into stress, panic and the related issues of Mathematics anxiety and test anxiety.

In an attempt to understand the reason for the dominance of studies of unpleasant emotions, the researcher reviewed the literature on the theories of emotion. The review found that the early theories concentrated on the emotions that produced visceral responses and that those responses were usually related to negative emotions. The review of the literature also found that the theories of emotion proposed by Mandler (1984;1989) and Ortony et al. (1988) were relevant to Mathematics.

From the review of all the relevant literature the researcher concluded that affective variables do play a role in the learning process and that research into the full range of emotions has been grossly neglected. These conclusions prompted the formulation of a major research question:

Can the emotional states of children in a Mathematics environment be described?

To facilitate the inquiry into emotional states two subsidiary research questions were composed:

**Can descriptors of emotional states be developed?
Can indicators of emotional states be devised?**

Chapter Three will describe the design of the investigation into emotional states. In order to do this effectively the chapter has been divided into four sections.

Section one is a general section about the use of qualitative research in the study of affect. It begins by explaining why qualitative research is recommended for investigations into affective issues in Mathematics, particularly in relation to the study of the emotions. Section one also discusses some of the data collection methods that have been used successfully in qualitative research into affect, and it discusses triangulation, the main process by which the validity of a qualitative study can be strengthened.

The second half of section one begins with an explanation for the absence of studies of emotion in the Mathematics Education literature. It continues by giving the rationale for the researcher's investigation into emotional states and an overview of the data collection methods used in the study. The reasons for the choice of data sources are also found within that section of the chapter. Part one concludes with a description of the strategies that were used to analyse the research data.

The second part of the chapter discusses the procedures that were used in the research. To enhance the comparability and translatability of the study, part two provides an in-depth description of the teacher, the subjects, the Mathematics program in which the subjects participated, and the data collection methods. Two strategies used to control for bias and subjectivity (respondent validation and triangulation) are described in this section, as well as two limitations of the research.

Part three of Chapter Three describes the data analysis process. The method of coding advocated by Miles and Huberman (1984) was used as to analyse the data. Coding resulted in categories being formed. The categories were inspected for patterns and the patterns became the themes that emerged from the research. Identification of the categories and the themes enabled conclusions to be drawn from the study.

Chapter Three concludes with an evaluation of the research design. The evaluation section discusses the strategies that were used to strengthen the validity and reliability of the study and notes potential threats to those constructs.

QUALITATIVE RESEARCH

Why use the Qualitative Research Method in the Study of Affect?

Today there is a growing number of researchers (for example, Mehan, 1978) who claim that qualitative research makes a valuable contribution to investigations carried out in educational settings. Romberg and Carpenter (1986) give two main reasons for the increased popularity and use of qualitative research. First, such methods overcome the limited information obtained from studies conducted in clinical settings. Second, investigations that use qualitative methods provide a way of explaining how and why changes occur in students' conceptions.

McLeod (1994) has pointed out that the majority of studies of affective issues in Mathematics have involved the use of questionnaires and quantitative research methods. While McLeod (1994) agreed that those techniques have provided a great deal of useful information, he claimed that such research methods should be complemented by alternative data collection techniques such as those found in qualitative research, because they are better suited to capturing the complexity of affective issues. By using qualitative research and the multiple research methods it advocates, McLeod (1992) claimed that our knowledge of affective variables would be increased.

The following section describes three of the many data collection methods available for use in qualitative research. By providing examples of the successful use of each method it is possible to justify their inclusion in the current study of emotional states.

Data Collection Methods Used in Qualitative Studies of Affect

In the Mathematics Education literature on affect, qualitative research is increasingly being viewed as the logical alternative to the traditional research approach because of its ability to provide a detailed description of students' affective responses (McLeod, 1994).

Questionnaires and rating scales have been used to collect data in both qualitative and quantitative research. In Mathematics, questionnaires and rating scales have been used in qualitative investigations into attitudes and beliefs (FitzSimons, 1994), emotions (Clarke, 1987; Clarke et al., 1990), stress (Sarason, 1984), anxiety (Richardson and Suinn, 1972) and test anxiety (Kass and Fish, 1991).

Videotaping is a data collection method that is increasingly being used in qualitative research. Videotaping was used successfully in Leder's (1990) phenomenological study of children's ideas about Mathematics, and in Thompson and Thompson's (1989) study of affect and problem solving.

Interviews have been used in many qualitative studies of affect. The interview was the main data collection technique used by Buxton (1981) in his study of panic, by Blatchford (1992) in his study of children's attitudes to schoolwork, and by Byrd (1982) in her study of anxiety. Thompson and Thompson (1989) incorporated the interview into their investigation into affect and problem solving, while Lester et al. (1989) used the interview technique in their study of beliefs.

Cohen and Manion (1989) discuss four kinds of interviews that may be used in research: the structured interview, the unstructured interview, the non-directive interview and the focused interview. According to Cohen and Manion (1989), as the amount of structure of an interview decreases the flexibility and freedom of the interviewer to investigate responses increases. However, the more unstructured the interview becomes the less comparable the data becomes across time and space. Cohen and Manion (1989) warn that the use of any interview technique as a data collection method increases the risk of bias entering the data and therefore decreases a study's comparability.

To overcome the threat to validity caused by the use of the interview, or any single data collection method, qualitative researchers incorporate a variety of data sources into their research design. Validity may further be strengthened by the use of a strategy called triangulation. This process will be described below.

Use of Triangulation in Qualitative Research

The value of the process of triangulation lies in its ability to strengthen the validity of a study.

Cohen and Manion (1989) define triangulation as the use of two or more methods of data collection in the study of some aspect of human behaviour. Denzin (1970) identifies two main types of methodological triangulation: within methods and between methods.

Triangulation within methods concerns the replication of a study as a check on reliability and theory confirmation. Triangulation between methods involves the use of more than one method of data collection in the pursuit of a given objective. The between methods approach embraces the notion of convergence between independent measures of the same objective.

(Cohen and Manion 1989, p. 274)

In addition to between methods and within methods triangulation, Denzin (1970) extends his concept of methodological triangulation to include time triangulation, space triangulation, combined levels of triangulation, theoretical triangulation and investigator triangulation. Investigator triangulation refers to the use of one or more observers or participants in a research setting. According to Cohen and Manion (1989), the use of two or more observers independently can lead to more valid and reliable data. Cross-sectional and longitudinal studies are two research methods belonging under the heading of time triangulation. These types of studies attempt to control for the effects of social change and process (Cohen and Manion, 1989).

In qualitative research, data collection methods that contain a high degree of contrast are recommended so that maximum confidence can be gained from the data (Cohen and Manion, 1989). Reliance on one method of gathering data may distort the researcher's picture of the context she/he is studying. Using triangulation in an investigation provides the researcher with greater confidence that the data being collected is an accurate report of the context and not simply the product of the data collection method being employed. However, even when triangulation has been used, the establishment of a study's validity still remains the greatest problem faced by researchers (Cohen and Manion, 1989).

McCormick and James (1983) claim that by using the study's subjects to authenticate collected data, a partial solution will be found to the problem of establishing the validity of the research. McCormick and James (1983) call this approach 'respondent validation' while Lincoln and Guba (1985) call it 'member-checking'.

Credibility is a trustworthiness criterion that is satisfied when source respondents (like people who provide the information) agree to honour the reconstruction; that fact should also satisfy the consumer.

(Lincoln and Guba 1985, p. 329)

To strengthen the validity of the study into emotional states, respondent validation, between methods triangulation, investigator triangulation and time triangulation have been incorporated into the research design. These validation methods will be discussed in more detail in the Procedure section of the chapter.

The next part of section one explains the reasons for the absence of studies of emotion in the Mathematics Education literature.

Why the Scarcity of Studies of Emotion in the Mathematics Literature?

According to McLeod (1992), studies of emotional states have occupied a minor place in Mathematics research for three main reasons. First, the nature of emotions makes them unsuitable for measurement by the traditional instrument, the questionnaire. Questionnaires have been found to be appropriate for research in which stable factors have been measured (McLeod, 1992). However, since emotions are situation- and time-specific (Lester et al., 1989) and are therefore not considered to be stable, the questionnaire has been of limited use in their measurement (McLeod, 1992). McLeod (1992) has suggested that careful observations and detailed interviews would be more useful than questionnaires in the analysis of emotional states. Second, in past decades, Mathematics research has been more interested in studying measurable products than underlying processes. Consequently, most research into affect in Mathematics has investigated attitudes and beliefs rather than emotions. Third, a theoretical framework for interpreting the role of the emotions in the learning of Mathematics has not been created. The absence of a theoretical framework makes grounded theory (a qualitative research methodology in which theory is generated) an ideal research method to use in the study of emotions.

To compensate for the absence of studies in which a range of emotions were explored, the researcher attempted to investigate the emotional states that children experienced during Mathematics lessons. The last half of this section gives specific reasons for the researcher's choice of subject area, methodology, data collection sources and data analysis strategies.

The Emotional States that Children Experience in a Mathematics Environment: Choice of subject area, research type, data collection methods and data analysis strategies

In the Mathematics Education literature on affect one can find many studies that have investigated Mathematics anxiety and test anxiety (See Tobias, 1976; Buxton, 1981; Byrd, 1982; Hackett, 1985; Hembre, 1990; Kass and Fish 1991). However, only a few reports of strong emotional reactions to Mathematics (either positive or negative) have appeared in the research literature (McLeod, 1992). When they have appeared, the emphasis of the studies has been on cognitive rather than affective issues (McLeod, 1992). No study has systematically attempted to analyse the full range of emotions that children experience during Mathematics lessons.

It is for this reason that the major research question and the two subsidiary questions have been formulated:

Can the emotional states of children in a Mathematics environment be described?

Can descriptors of emotional states be developed?

Can indicators of emotional states be devised?

The qualitative research method was chosen for the study into the emotions for two reasons. First, the review of the literature on affect indicated that the qualitative approach was the most appropriate research method to use in the study of affective issues. Second, in the Mathematics Education literature there was a scarcity of studies of emotions and emotional states upon which the researcher could model her research design.

Multiple data collection methods that contained a high degree of contrast were chosen for the research. This was done for two reasons. First, to gain a complete understanding of the emotional states that children experienced during Mathematics lessons. Second, to enable the between methods triangulation strategy to be used to strengthen the validity of the study.

McLeod's (1992) recommendation that observations and interviews be included in studies of the emotions was considered when choosing the data collection methods for the research. Consequently, the following data collection sources were used in the research: three administrations of a modified version of the Clarke (1987) self-assessment response sheet, individual interviews and follow-up interviews, group interviews, a class discussion, children's journals, videotapes of Mathematics lessons, researcher observations and a parental survey.

Data analysis was conducted concurrently with data collection. The data was coded and categorised as it was analysed. As additional data was analysed it was compared to data already collected. Hypothesis generation and inspection of the data for patterns or relationships began with the analysis of the initial data. It then underwent continuous refinement by being fed back into the process of category coding. Patterns and relationships that emerged from the data were used to form the conclusions of the study. When necessary they were investigated further through the development of additional data collection methods. Glaser and Strauss (1967) called this research approach the constant comparative method. Today it is also known as the grounded research method (Goetz and LeCompte, 1984).

In the next part of Chapter Three detailed information is given about the procedures that were used in the study of emotional states.

THE PROCEDURE

The Procedure section of the chapter has been divided into four parts. Parts one, two and three provide background information about the researcher's position of teacher-researcher, the subjects used in the study of emotional states and the Mathematics program in which the subjects participated.

The fourth part of section two provides a detailed description of each of the eight data collection methods used in the research.

The final part of section two gives an overview of the strategies that were used to strengthen the validity and reliability of the study. A more detailed description of the strategies will appear in the last section of Chapter Three, the Evaluation of the Research Design.

The Teacher

The researcher who carried out the investigation into emotional states was the class's teacher and therefore was a participant observer in the study.

The teacher was the co-ordinator of the school's Mathematics committee and a strong advocate of the Mathematics K-6 (NSW Department of Education, 1989) syllabus. In

the year prior to the research, the teacher participated in an evaluation of the school's Mathematics program. Findings from the children's section of the evaluation, and a review of the literature of the affective domain, prompted the teacher to question the role of the emotions in the formulation of students' attitudes to Mathematics.

The Subjects

At the beginning of the year the class was allocated to the teacher-researcher by the principal of the school. A restriction of time due to work commitments resulted in the teacher-researcher using her class as the sample in the investigation. Any threats to the reliability of the study that might have been caused by the teacher-researcher's relationship with her students would have been overcome by the advantage of using a natural setting for the research, and by the fact that the students' association with the teacher-researcher was also a natural one.

Bailey (1978) has pointed out that there are a number of advantages of using the technique of participant observation in a natural setting. Two of those advantages are relevant to the study of emotional states. First, observation studies are superior to experiments and surveys when data are being collected on non-verbal behaviour. Second, the investigator is able to discern ongoing behaviour as it occurs and is able to make appropriate notes about its salient features.

The children chosen for the study were an intact, composite third/fourth grade class of nine and ten year olds in a southern New South Wales state primary school. While the class was an intact group the children were randomly assigned to it by the school's Executive Teacher and the Assistant Principal.

There were thirty-one students in the class: seven third graders (three boys and four girls) and twenty-four fourth graders (ten boys and fourteen girls). The children were predominantly white, middle-class and of Anglo-Saxon origin. One grade four boy was of Aboriginal descent. Twenty-seven of the children lived in town while four came from outlying farms. Four of the children (one grade three boy, two grade four boys and one grade four girl) had parents in the Defence Force and were enrolled at the school at the beginning of the year. All of the Defence Force children came to New South Wales from different states and territories. In addition to the Defence Force children another grade four girl was enrolled at the school at the beginning of the year. She also came from interstate.

Five of the grade three children and four of the grade four children were taught by the teacher-researcher when she had a composite kindergarten/first class in 1990. Moreover, three of the grade four children were in the teacher-researcher's class the year prior to the research. Consequently, the teacher-researcher was well acquainted with twelve of the thirty-one children in the sample and had developed what she considered to be a strong and trusting relationship with them.

The Class's Mathematics Program

During the research period the class's Mathematics program was comprised of activities from stage two of the Mathematics K-6 (NSW Department of Education, 1989) syllabus. The lessons in which the children participated were not specially prepared for the study. They were the normal sequence of lessons that had been programmed by the teacher-researcher from the syllabus.

Resources and assessment methods used with the class also came from Mathematics K-6 (NSW Department of Education, 1989) syllabus. The assessment strategies used by the teacher-researcher were: pencil and paper tests, oral tests, anecdotal notes, conferences with the teacher, teacher observations and inspection of written and constructed work.

For the months during which the research occurred the teacher-researcher gave the class two types of tests: oral tests and written tests. Written tests were conducted every second week. The grades from the tests were recorded in the teacher-researcher's assessment book. When the pencil and paper tests had been returned to the children they were taken home, signed by the parents, returned to school for the teacher-researcher to sight then given back to the children to keep.

The oral tests were randomly given during the school week. The children recorded their answers to the teacher-researcher's questions in their Mathematics exercise books. The questions in the oral tests were read out quickly and were used to assess the children's instant recall of Mathematical facts, such as the number of days in a year or the product of six and three. Marks from the oral tests were not always recorded in the teacher-researcher's assessment book.

Anecdotal notes, teacher observations and conferencing with the teacher occurred on a needs basis. Inspection of written and construction work was undertaken when the children had completed the task on which they were working.

It was the policy of the school in which the research was conducted that its practising teachers should use the content, teaching strategies and assessment methods prescribed by the Mathematics K-6 (NSW Department of Education, 1989) syllabus.

DATA COLLECTION METHODS USED IN THE STUDY OF EMOTIONAL STATES

The following section discusses in detail the eight data collection methods that were used in the research.

Strongly contrasting sources of data and three different types of data were incorporated into the study for three reasons. First, to strengthen the validity and reliability of the research. Second, to thoroughly explore the range of emotions that children experienced during Mathematics lessons. Third, to enable the optimal number of indicators and descriptors of emotional states to be developed.

It was anticipated that data from a l eight sources would be needed to address the major research question:

Can the emotional states of children in a Mathematics environment be described?

Indicators of emotional states were primarily expected to be found in the visual and oral data while descriptors of emotion were expected to come from the written and oral data.

The first five data collection methods were sequentially planned. Data obtained from the first administration of the modified Clarke (1987) self-response sheet was used to structure the questions for the individual interviews. The responses from the individual interviews were used to plan the questions for the group interviews. Data from the group interviews were used to construct the questions for the class discussion.

Findings obtained from all the data collection sources were compared with findings from the literature.

The Questionnaire

In the investigation into emotional states the researcher used a modified version of the Clarke (1987) self-assessment response sheet (See Appendix A).

The Clarke (1987) self-response sheet was developed to give Victorian highschool teachers a means of effectively monitoring student learning in Mathematics, particularly in the area of student affect. Even though the questionnaire had not been normed or validated, its use in the IMPACT Project (Clarke, 1985) established its accuracy as a measurement instrument.

The first and last parts of the original Clarke (1987) self-assessment response sheet asked the children to describe Mathematics topics that were causing them concern. The middle section of the questionnaire gave the students the opportunity to describe the feelings they experienced when they participated in Mathematics lessons (See Appendix A for a copy of the original questionnaire).

The questionnaire used in the study of emotional states included some additional questions not employed on the original Clarke (1987) self-assessment response sheet. The additional questions were not related to the study of emotions, but had been added to the questionnaire so that the instrument could be used as a part of the school's evaluation of its Mathematics program in the year prior to the research. Only the responses from two questions on the modified Clarke (1987) self-assessment response sheet were included as data in the research: those from the question pertaining to the liking and disliking of Mathematics and those from the question about the feelings experienced during Mathematics lessons.

The modified Clarke (1987) questionnaire was considered to be an appropriate starting point for investigating the major research question because it provided children with the opportunity to describe the positive and negative emotions they experienced during Mathematics lessons.

The modified Clarke (1987) self-assessment response sheet was administered three times during the research: at the beginning of the study, then twice more at two monthly intervals. The emotion words that each child circled on the first administration of the questionnaire were discussed with her/him during her/his individual interview.

While the first administration of the questionnaire was used as the starting point for the research, the second and third administrations were used to strengthen the internal validity of the study. This was achieved in two ways. First, by using the three administrations as a time triangulation strategy. Second, by determining the trustworthiness of the responses of each child by comparing them across administrations.

As well as strengthening the validity and reliability of the study, it was hoped that the three administrations of the questionnaire would provide the researcher with a wide range of words to describe the emotions that the children experienced.

During each administration of the questionnaire the children were told that their responses would be confidential and seen by no-one except the researcher. In an attempt to encourage the children to give an honest account of their feelings the questionnaires were covered as they were being completed.

The individual interviews commenced after the first administration of the questionnaire. No interviews were scheduled after the second and third administrations of the questionnaire.

Semi-structured Individual Interviews

The review of the literature on the affective domain and the Mathematics Education literature on affect raised three issues that the researcher hoped to investigate during the course of her study. They were: physical reactions, Mathematics anxiety and Mathematics test anxiety. The interview was considered to be the most appropriate method for collecting data on the issues raised in the literature, and for more fully investigating the responses gained from the modified Clarke (1987) self-assessment response sheet (A format sheet for the individual interviews is included in Appendix B).

In the individual interviews the researcher continued her investigation into the major research question:

Can the emotional states of children in a Mathematics environment be described?

The interviews also began to address the research questions:

**Can indicators of emotional states be developed?
Can descriptors of emotional states be devised?**

The semi-structured interview format was chosen for the individual interviews so that the interviewer would have some flexibility and freedom to explore responses (Cohen and Manion, 1989). The content, sequence and wording of the questions were carefully selected in order to minimise the amount of researcher bias that might be reflected in them (Cohen and Manion, 1989).

The individual interviews were used to verify the data gained from the first administration of the modified Clarke (1987) questionnaire. Responses to the questions on the self-response sheet were checked with each child during her/his interview.

Each interview was audiotaped and transcribed on the day that it was held. A follow-up interview was held the day after the initial interview. The follow-up interview, which was also audiotaped and transcribed, was carried out in order to verify the responses made during the original interview. During the follow-up interview each child had the opportunity to make changes to the data if she/he considered that they had been incorrectly recorded by the researcher. This strategy (respondent validation) was used to verify the informant's responses, and was an additional way of strengthening the internal validity of the research.

Group Interviews

When the follow-up individual interviews had been completed, the group interviews were commenced (See Appendix C for the interview questions). Each group's interview was audiotaped and transcribed.

The group interviews continued to address the major research question and the subsidiary questions:

Can the emotional state of children in a Mathematics environment be described?

Can indicators of emotional states be developed?

Can descriptors of emotional states be devised?

During the group interviews the researcher hoped to explore an issue that had been discussed in the Mathematics Education literature: that of children rarely expressing overt signs of enthusiasm for Mathematics (Thompson and Thompson, 1989; McLeod, 1992).

To minimise the amount of disorganisation that would occur in the Mathematics lessons in which the interviews were to be held, the groups that were interviewed were those in which the children were arranged within the classroom. In the class there were four groups of six children and one group of seven children. One of the groups was made up of grade three children only. Another group had one grade three child and five grade four children in it. The two remaining groups contained grade four children.

The children were interviewed in the hatroom adjacent to the classroom. That area was chosen for three reasons. First, it was hoped that the quietness of the hatroom would allow the children to speak honestly and openly about their feelings, and that they would do so free from the fear of being laughed at by the other children in the class. Second, the hatroom provided the children with an environment in which confidentiality and trust could be established. Third, it was thought that the amount of embarrassment the children might experience in discussing the questions would be minimised by having them withdrawn from the classroom setting.

To establish the trustworthiness of the data gained from the group interviews, each child's statements were compared to those obtained from her/his individual interview, and to the responses given during the first administration of the questionnaire. This strategy strengthened the internal validity of the research.

Follow-up group interviews were not scheduled as the whole class discussion was considered to be the follow-up to the group interviews.

The Class Discussion

Issues that arose from the group interviews, or themes that had begun to emerge from the data were more fully investigated during the class discussion. If no issues or themes had appeared in the data prior to the scheduling of the class discussion then this data collection method would have been abandoned.

The large number of informants involved in the class discussion meant that it needed to be the most structured of the three verbal data collection methods used in the research. The class discussion was videotaped and audiotaped. The audiotape was transcribed. (See Appendix D for the questions asked during the class discussion.)

Journals

Each child kept a Mathematics journal for the duration of the research. All entries made in the journals were dated. One of the tasks that the children completed in their journals was to describe their thoughts and the way they felt, before and after they participated in a written test. The children were also asked to estimate the mark they thought they would get in the test. Completing that task enabled detailed data to be gained about children's emotions and thoughts during written tests.

In the last two months of the research the children were asked to record their thoughts and feelings about the Mathematics activities *and* tests in which they participated. By gathering data about both types of tasks it was thought that the study would obtain a complete representation of the emotions that children experienced during Mathematics lessons. This might not have been achieved if data had only been obtained from either the activity or test situation.

The journals continued to provide data that addressed the major research question and the research question pertaining to indicators of emotional states. Data from the journals also contributed to the research question:

Can descriptors of emotional states be devised?

Extracts from the children's journals have been used to support or refute the findings from the research. As expected, the children's written work contained spelling mistakes, grammatical errors and word omissions. Where the statements from the journals have been quoted in the research, they have been presented as they were written, unless the researcher considered that the meaning of the piece of writing was unclear. Where there was ambiguity, a translation of the word, sentence or sentences has been placed near the original statement.

Videotaping of Mathematics Activities

Mathematics lessons were mainly videotaped to assist the researcher in her development of indicators of emotional states. While watching the videotapes the researcher noted the indicators of emotion expressed by each child. The frequency of the indicators was also recorded.

Each child was taped working with the group of children with whom she/he sat. This was done to minimise the amount of movement and reorganisation needed in the classroom during the lessons in which the videotaping occurred.

To accustom the children to having the videocamera operating while they were working, the machine had been used many times in lessons prior to the commencement of the research. By using the videocamera with the children in earlier lessons, the researcher thought that the children's actions during the research period would be natural and spontaneous. Making the children accustomed to working in front of the videocamera also reduced the likelihood of the Hawthorne Effect influencing the validity of the study.

Eight consecutive thirty minute Mathematics lessons were videotaped. Each tape showed the date on which it was made. This was done so that data from the videotapes could be used to verify data from other sources.

The tasks in which the children participated were those that had been programmed by the teacher-researcher for the weeks during which the research was to occur. The tasks came from the Number and Space strands of the Mathematics K-6 (NSW Department of Education, 1989) syllabus. No activities from the Measurement strand of the syllabus were recorded as the researcher believed that the movement of the children around the room during those lessons could have resulted in the videocamera being bumped, or the recorded image being distorted.

The taped activities were:

- * one written test
- * two oral tests
- * construction work using plastic shapes that clipped together
- * a number game involving a partner
- * individual stencils that practised multiplication tables
- * bookwork and blackboard work that practised division algorithms. (See Appendix F for a more detailed account of the tasks that were videotaped)

All of the lessons that were videotaped came from the normal sequence of Mathematics work programmed for the class. No lessons were specially prepared for the tapes.

Researcher Observations

Observations of each child engaged in Mathematics tasks occurred when the videotaping had been completed. Each observation period lasted between five and ten minutes and was made during the class's normal Mathematics lessons.

The researcher observations included the children working on stencils, blackboard algorithms and group activities. Observations were also made during one written and one oral test.

The researcher observations recorded indicators of emotional states and they supplemented the visual data obtained from the videotapes.

Parent Survey

The final data collection method used in the study was a parent survey (See Appendix E). During the third last week of the research, a short survey was sent home to the parents of the children in the class. The survey was planned to be a way of verifying the data obtained from the questionnaire, the individual and group interviews, the class discussion and the journals.

Strategies Used to Control for Bias and Subjectivity in the Investigation into Children's Emotional States

The following section provides an overview of the strategies that were used by the researcher to strengthen the validity and reliability of the study. A more detailed description of the strategies appears in section of the chapter entitled Evaluation of the Research Design.

The researcher used four strategies to strengthen the internal validity of the research. The strategies were respondent validation and three forms of triangulation: time triangulation, between methods triangulation and inter-observer agreement.

Multiple data collection methods were incorporated into the research design to strengthen the validity and reliability of the study and to enable the use of the between methods triangulation strategy. The between methods triangulation approach was used to provide a rigorous control for researcher bias.

Between methods triangulation helped determine the consistency of the written and verbal data from the multiple administrations of the questionnaire, the children's journals, the individual and group interviews and the class discussion. The videotapes and researcher's observations were mainly used as a source of comparison for the visual data collected during the study.

The use of one independent observer was another triangulation method incorporated into the research design. The employment of an independent observer to assist with the data analysis, was an attempt to minimize the amount of data contamination caused by researcher bias. Use of an independent observer also served to strengthen the validity of the research.

Multiple administrations of the modified Clarke (1987) self-assessment questionnaire were a time triangulation strategy that was used to enhance the validity of the data. This

longitudinal study of responses to the questionnaire took into account the effects of change over time.

Respondent validation was a fourth strategy that was employed by the researcher to strengthen the validity of the research. Respondent validation was used twice during the research: in the individual interviews and follow-up interviews.

Limitations of the Research

While the research had been designed to provide the most accurate representation possible of the phenomena it was investigating, there were two minor and unavoidable difficulties that were encountered in the planning of the study. These difficulties could have influenced the data that was obtained and consequently, the conclusions of the research.

A significant problem faced by the researcher was that of developing the reliability of the coding system to be used in the analysis of the data. In an attempt to achieve this with the data from the videotapes, the researcher and the independent observer used an intercoding reliability measure.

According to Miles and Huberman (1984), intercoder reliability refers to the ability of two or more researchers to achieve the same or similar coding results from a set of responses. One method of doing this is to express the number of agreements as a percentage of the number of coded items.

$$\text{Intercoder reliability} = \frac{\text{number of agreed codings}}{\text{number of possible codings}} \times 100$$

The use of the planned reliability measures (participant observation, peer examination of the data, mechanically recorded data and an intercoder reliability measure) helped to ensure that the reliability of the study was established and maintained.

Limited resources was the second difficulty that was encountered in the design of the research. An absence of funding for the study, coupled with time restrictions due to work commitments, meant that the researcher had to use a convenience sample in the research. However, as noted earlier, this had the advantage of the research occurring in a natural setting. The use of a natural setting reduced the possibility of data

contamination. This might have occurred if the data had been obtained from an artificial environment.

Lack of funding also forced a limitation on the number of researchers and videocameras available for the study.

The strategies (described earlier in this chapter) that were incorporated into the research to enhance its reliability and validity, were considered capable of overcoming the limitations of the research design. Those strategies ensured that reliable and valid data was gained from the study, and consequently, that accurate conclusions were drawn.

The following section will describe the methods by which the data from the study was analysed.

DATA ANALYSIS PLAN

The data analysis section describes the strategies the researcher used to analyse the data that was collected from the research.

The main research methodology used in the study of children's emotional states was grounded theory. Grounded theory is a general methodology for developing theory that is grounded in data systematically gathered and analysed (Strauss and Corbin, 1994). In this methodology, theory may be generated (Strauss and Corbin, 1994). Theory evolves during the actual research. It does this through a continuous interplay between the data analysis and the data collection. A central feature of this analytic approach is a "general method of [constant] comparative analysis" (Glaser and Strauss 1967, p. vii); hence the approach is often referred to as the constant comparative method. By using grounded theory in the study of children's emotional states it was hoped that theory about the emotions would be generated.

A number of strategies may be employed by researchers to assist them in their use of the constant comparative method. Miles and Huberman (1984) have described seven such strategies: contact summary sheets, codes and coding, pattern coding, memoing, site analysis meetings, interim site summaries and data accounting sheets. Codes and coding, and pattern coding, were the strategies used to analyse the data in the research into children's emotions.

According to Miles and Huberman (1984), a code is an abbreviation or symbol applied to a segment of words - most often a sentence or paragraph of transcribed field notes - in order to classify the words. Codes are categories. They usually derive from research questions, hypotheses, key concepts or important themes. Codes are retrieval and organising devices that allow the analyst to quickly spot, pull out, then cluster all the segments relating to the particular question, hypothesis, concept or theme. Clustering sets the stage for analysis (Miles and Huberman, 1984).

Miles and Huberman (1984) stated that there were three levels of coding: descriptive, interpretive and explanatory. Descriptive codes attribute a class of phenomena to a segment of text. They do not attempt to interpret it. Interpretive codes describe the reason for a phenomena occurring. Explanatory codes are an inferential interpretation of data that fit emerging themes, that is, they attempt to explain the relationships found in the data (Miles and Huberman, 1984).

The three types of codes discussed by Miles and Huberman (1984) were used in the coding of the data from the research. Descriptive coding was used in the analysis of the questionnaire and the children's journals. Both descriptive and interpretive coding were used in the analysis of the individual interviews, the group interview and the class discussion. Explanatory coding occurred during the analysis of the teacher's observations, the video data, the group interviews, the children's journals and the parent survey.

Coding occurred continuously throughout the research. The coding that occurred during the analysis was grounded in the data that had already been collected.

Categories that were formed during the analysis of the data were changed if the researcher considered it necessary. Categories were dropped if there was insufficient data to support them or created if incoming data did not fit with the categories already in existence. If the data analysis had not been done efficiently then significant data could unwittingly have been discarded. However, the use of multiple data collection methods prevented that from happening.

When the initial coding was complete, pattern coding was used to identify emerging themes. Pattern codes are explanatory or inferential codes that identify an emerging theme, pattern or explanation for an occurrence. Pattern codes pull a lot of material together into meaningful units of analysis. Pattern coding is a way of grouping the code summaries into a smaller number of interrelated themes or constructs (Miles and

Huberman, 1984). As the themes in the research emerged it was found that one of them, 'body language', needed further investigation. Additional data collection methods were devised to explore the theme. See Chapter Four for an explanation of the theme and the data collection methods used to obtain data on it.

Coding and pattern coding were simple and efficient strategies used to organise the large amount of data that were collected during the study. The use of a coding system enabled the researcher to identify emerging themes early in the study and to investigate them in as much detail as they warranted.

During the analysis of the data the researcher placed her analysed data and other relevant information in a loose-leaf folder. The contents of the folder included: a list of children who were sitting at each group of tables, data on each child from each of the data collection methods, the original research questions, the procedural plan for the collection of data (revised when necessary), summary sheets of the sample's responses to the three administrations of the Clarke (1987) self-response sheet (including numbers and tallies), descriptors and indicators of the emotional states as they emerged, propositions that were developed, reflective remarks, summaries of the data from the individual and group interviews and the class discussion, emerging hypotheses and models that were devised from the data. Using a loose-leaf folder made it possible to regroup and rearrange data and other items as required and to add or delete data when necessary. The videotapes, audiotapes and other original material were stored in a filing cabinet.

The following section of Chapter Three describes the strategies that were incorporated into the research so that it could obtain maximum reliability and validity. The establishment of the study's reliability and validity enables its conclusions to be generalised.

THE EVALUATION OF THE RESEARCH DESIGN

This part of the chapter discusses the strategies that were used to strengthen the reliability and validity of the research, and consequently, the quality of the research design.

Part four begins with a short, general discussion of the need for qualitative research to guard against threats to reliability and validity. The section then proceeds by discussing the measures taken to enhance the reliability and validity of the study of emotional states. The methods employed by the researcher were those suggested by Goetz and LeCompte (1984). When applicable, further limitations in the research design are noted.

Reliability and Validity in Qualitative Research

The results of ethnographic research are regarded by many scholars as unreliable and lacking validity and generalisability (Goetz and LeCompte, 1984). In an attempt to overcome such criticisms, ethnographic researchers, like their quantitative counterparts, deliberately employ strategies that will strengthen the reliability and validity of their investigations. The strategies used in qualitative studies are aimed at minimising the amount of subjectivity that enters the data through observer and informant bias.

Reliability

Goetz and LeCompte (1984) have described two types of reliability that influence the credibility of a qualitative study: external reliability and internal reliability. The authors have pointed out that both facets of reliability need to be addressed.

Internal reliability refers to the degree to which other researchers, given a set of previously generated constructs, would match them with data in the same way as did the original researcher. External reliability addresses the issue of whether independent researchers would discover the same phenomena or generate the same constructs in the same or similar settings (Goetz and LeCompte, 1984).

Internal Reliability

Crucial to internal reliability is inter-rater or inter-observer reliability, or the extent to which the sets of meanings held by multiple observers are sufficiently congruent that they describe and arrive at inferences about phenomena in the same way (Goetz and LeCompte, 1984).

Five strategies may be used by ethnographers to reduce threats to internal reliability. The strategies are: the use of low-inference descriptors, multiple researchers, participant

researchers, peer examination of data and mechanically recorded data. (Goetz and LeCompte, 1984). Four of the strategies discussed by Goetz and LeCompte (1984) were incorporated into the study of children's emotions. They were: low-inference descriptors, a participant researcher, peer examination of data and mechanically recorded data. The absence of multiple researchers at the research site posed a threat to the internal reliability of the study. However, the use of an independent observer compensated for this limitation.

Written and Verbal Data using Low-Inference Descriptors

Whenever possible, verbatim accounts of the participant's conversations and direct quotations from the researcher's fieldnotes have been used to illustrate children's emotional states and their descriptors and indicators. Verbatim accounts have also been used to verify emerging themes. Quotes included in the research were chosen for their low level of ambiguity and idiosyncrasy.

Participant Researcher

The researcher's position of teacher of the sample made her a participant observer in the study. Her affiliation with the research site, its operation and its occupants would have resulted in trustworthy data being collected.

Peer Examination

One independent observer was used to assist the researcher in her analysis of the data from the videotapes. The independent observer was a female undergraduate student who was enrolled in a Bachelor of Social Science degree at the university attended by the researcher.

Mechanically Recorded Data

Audiotapes and videotapes were used in the research.

The audiotapes were used to provide a complete record of the individual and follow-up individual interviews, the group interviews and the class discussion. The tapes complemented the rough fieldnotes made at the research site. Expanded notes were made from the tapes and were placed in a file on each child.

The videotape data was collected for three reasons. First, to provide visual data in the investigation of the major research question. Second, to assist in the development of indicators and descriptors of emotional states. Third, to substantiate and extend the data from the researcher's observations, the questionnaire, the individual and group interviews and the class discussion.

External Reliability

According to Goetz and LeCompte (1984) no study can attain perfect external reliability. However, ethnographers can enhance the external reliability of their data by recognising and addressing five major issues that arise in research: researcher status and position, informant choices, social situations and conditions, analytic constructs and premises and methods of data collection and analysis (Goetz and LeCompte, 1984). The following section shows how the researcher has addressed each of these problems.

Researcher Status and Position

One of the main threats to the external reliability of the research came from the fact that the researcher was a teacher and therefore represented the relatively narrow perspective of a teacher (Goetz and LeCompte, 1984). This perspective could only be accurately replicated by researchers who assumed a comparable role. Studies which might in the future investigate the emotional states of students, would therefore need to be considered as supplemental or comparative rather than replicative (Goetz and LeCompte, 1984).

Furthermore, the researcher was the teacher of the class being studied and therefore held a position of significance in the research group. The researcher's position of power in the group could have influenced her perceptions in the analysis of the data and thus caused her to draw false conclusions from the study. However, to use another person or another setting would have made the situation less natural.

Informant Choices

Limited resources of both time and money meant that the sample chosen for the research was selected for convenience. Consequently, the children used in the study were the thirty-one students in the teacher-researcher's class. The type of sample and its small size could be considered to pose a threat to the external reliability of the study. However, as the sample was an intact group used in a natural setting, any threat to the external

reliability of the study that was imposed upon it by a limitation of resources, would have been overcome.

Social Situations and Conditions

The school in which the sample was found contained nine classes. In addition to the researcher's composite third/fourth grade class there was one straight grade three class. The teacher-researcher's class was the only one that contained grade four children.

The curricula used by the sample were those prescribed by the New South Wales Department of Education. The Mathematics K-6 (NSW Department of Education, 1989) syllabus was one of the set curricula.

The external reliability of the research would have been strengthened if the study had been conducted at multiple sites over a period of years.

Analytic Constructs and Premises

To facilitate the replication of the study, the constructs, definitions, and units of analysis have been clearly delineated in the Procedure section of this chapter, and in the literature review in Chapter Two.

Methods of Data Collection and Analysis

Multiple data collection methods that contained a high degree of contrast resulted in written, verbal and visual data being obtained from the study.

The written data came from the three administrations of the modified version of the Clarke (1987) self-assessment response sheet, the children's journals and the parental survey. Verbal data was obtained from the individual interviews and follow-up interviews, the group interviews and the class discussion. The individual and group interviews were audiotaped and transcribed while the class discussion was videotaped, audiotaped and transcribed. The researcher's observations and the videotapes provided the visual data.

The constant comparative methodology was used in the research. Data was analysed by coding. The coding resulted in categories being formed. The categories that were formed were inspected for patterns or themes. From the patterns conclusions were

drawn. Hypotheses that evolved during each step of the analysis were tested against data collected from the various sources. This method of data analysis continued until the researcher was satisfied that the conclusions that had been drawn, were an accurate and valid representation of the phenomena being described.

Validity

Validity is concerned with the accuracy of scientific findings. Establishing validity requires determining the extent to which conclusions effectively represent empirical reality, and the extent to which constructs devised by researchers represent or measure the categories of human experience (Goetz and LeCompte, 1984). There are two types of validity that researchers need to consider when conducting a qualitative investigation: internal and external validity.

Internal Validity

Internal validity refers to the degree to which scientific observations and measurements are authentic representations of some reality (Goetz and LeCompte, 1984). Threats to internal validity include history and maturation, observer effects, selection and regression, mortality, and spurious conclusions (Goetz and LeCompte, 1984). The researcher has addressed each of these issues in her design.

History and Maturation

Changes that occur in the overall social scene are comparable to what experimenters designate as history, while changes that involve progressive development in individuals are considered to be maturation (Goetz and LeCompte, 1984).

By using the constant comparative method and the between methods triangulation strategy, the researcher was able to compare the verbal, written and visual data for consistency. This strategy enabled any minor changes that occurred at the research site to be assimilated into the study without them affecting the study's findings.

As the research was not longitudinal, effects due to maturation were minimal.

Observer Effects

All the children in the class were used as informants in the investigation. Therefore the view of reality that was obtained from the study was the widest possible for that particular population.

The data from the research is able to be accepted with a high degree of confidence for two reasons. First, the researcher had taught many of the children in the sample in previous years and had developed a good rapport with them. Second, the study took place over a six month period.

The researcher's rapport with the informants, and the lengthy period of time in which the research occurred also overcame two additional problems that may otherwise have effected the internal validity of the study: that of informants wanting to present their ideal self to the researcher, and the possibility of the informants telling the researcher what they thought she wanted to be told (Goetz and LeCompte, 1984). Those problems, however, may still have occurred with the children who were new to the class. In order to minimise the threat to validity caused by observer effects of that type, a number of procedures were incorporated into the research design. They were: multiple data collection methods, the use of an independent observer in the analysis of the video data, and the comparison method of data analysis. By presenting data to the informants in the interviews and follow-up interviews, distortions that occurred in the data were revealed and were able to be corrected.

The use of respondent validation ensured that the categories formed after the follow-up interviews were meaningful and honest representations of the way in which the participants experienced reality (Goetz and LeCompte, 1984).

Between methods and time triangulation strategies were included in the research design to minimise any threat to the internal validity of the research that might have been caused by the researcher's influence on the research site as a participant observer, and by the researcher's inexperience in using a coding system to analyse data.

While the researcher's position of participant observer was not expected to exert an undue influence on the research findings, her personal relationship with each of the children in the sample could be construed as a possible threat to the internal validity of the research.

The development of close teacher-pupil relationships may have resulted in data contamination due to researcher bias. Alternatively, the development of those

relationships could have produced highly trustworthy data. It is therefore possible that the positive influence of the teacher-pupil relationships offset any negative effects caused by the relationships. If that occurred then the internal validity of the study would not have been jeopardised.

Selection and Regression

An intact group was chosen as the sample for the research.

The class that participated in the study was formed the previous year by the random assignment of eight year olds from the school's population and the assignment of all the school's nine year olds to the group. This had been done in accordance with the school's policy on class formation and reflected the New South Wales Department of Education's policy on the subject.

The method in which the sample was formed limits representativeness to similar situations.

Mortality

It was anticipated that the number of subjects in the study would remain constant for the duration of the research. However, if changes did occur to the sample's numbers during the investigation, then they would have been typical of changes that took place in any class population during one school year.

Spurious Conclusions

In an attempt to prevent the researcher from drawing false conclusions from the study, multiple research methods were incorporated into the research design. Multiple research methods enabled the researcher to use the between methods triangulation strategy. Use of that strategy strengthened the internal validity of the research and reduced the likelihood of the researcher drawing false conclusions.

A second strategy that was used to prevent the researcher from drawing false conclusions was that of retrospective data inspection. As the analysis proceeded, the researcher considered alternate explanations for the data being obtained. If data from one or more sources indicated that false conclusions were being drawn then the data was re-inspected. When the re-inspection showed that false conclusions had been made then

those conclusions were ignored and the data was re-analysed. New conclusions were only accepted as accurate when they had the support of all the data.

The researcher might have been at risk of drawing the wrong conclusions from the study if the following criteria had not been taken into consideration: the use of multiple data collection methods, the use of retrospective inspection of data, the length of time over which the research was conducted, and the rapport the researcher had developed with her students.

External Validity: Comparability and Translatability

External validity refers to the degree to which scientific observations and measurements can be compared legitimately across groups (Goetz and LeCompte, 1984). According to Goetz and LeCompte (1984) threats to the external validity of ethnographic findings consist of those effects that obstruct or reduce a study's comparability and translatability.

Comparability refers to the degree to which the components of a study are described and defined. When this is done accurately it enables other researchers to use the results of the investigation as a basis for comparison with studies that address related issues (Goetz and LeCompte, 1984). Translatability refers to the degree to which the researcher uses theoretical frames, definitions and research techniques that are accessible to, or understood by, other researchers in the field or related disciplines (Goetz and LeCompte, 1984).

To establish comparability and translatability the study must be shown to exhibit typicality, that is, the ability to be compared and contrasted along relevant dimensions with other phenomena (Goetz and LeCompte, 1984). This is best achieved by the use of more than one site for the study of a phenomena. For studies in which only one site is used, external validity is strengthened by providing a highly detailed description of the setting and the procedures used in the research (Goetz and LeCompte, 1984).

The small sample of the school population that was used in the study of emotional states may have led to the selection of a group that was not representative of the larger school population. Even though the class was typical in structure and formation to a composite class found in any New South Wales state school, and the Mathematics program for the class was solely based on the syllabus that was mandatory for all New South Wales schools, it was possible that the sample contained idiosyncrasies which made it different to the rest of the school population or to the system in which it was located. The

teacher's regular use of written tests as a means of evaluating student progress, and the small number of children in the third grade section of the class were two variables that may have made the sample unique. These two variables, and the fact that the sample was only one class at a single site, could have been seen to pose a threat to the external validity of the research.

In a quantitative sense the findings from the study of emotional states would not be able to be generalised because the research would be seen to lack external validity. However, using the comparability and translatability conception of external validity suggested by Goetz and LeCompte (1984) the research should be considered to be externally valid.

The research design provided a detailed description of the research methodology, the procedures employed in the study, the strategies used to analyse the data and the measures taken to enhance the study's reliability and validity. The provision of that information strengthens the comparability and translatability of the research and therefore enables its findings to be generalised. The threat to the study's external validity which might have come from the uniqueness of the sample, its idiosyncrasies and the single site setting, has no impact upon this research because it is not a quantitative study.

SUMMARY

Chapter Three began by discussing the significance of qualitative research to studies of affect and the emotions. This was followed by a discussion about the use of triangulation in qualitative research. The first section of the chapter continued with the researcher providing the rationale for the study of emotional states, her statement of the research questions, her reasons for choosing the qualitative method and the data collection sources and data analysis methods that were used in the research.

The Procedure section of the chapter described the constructs that were used in the study and provided an overview of the strategies employed to control for bias and subjectivity. The Procedure section provided an in-depth description of the multiple data collection methods used in the research. The data collection methods that were employed resulted in oral, written and visual data being collected.

The third section of the chapter, the data analysis section, described the main strategies used to analyse the data. The strategies that were discussed were coding, pattern coding and the retrospective inspection of data. The strategy of constantly comparing new data to that already obtained was given as the means by which conclusions were drawn from

the research. Section three also discussed two of the limitations of the research: the researcher's inexperience in using a coding strategy to analyse data, and a restricted amount of time and money available for use.

The final section of Chapter Three provided an evaluation of the research design. In this section the terms reliability and validity were defined, and their meaning in qualitative research was discussed. Possible threats to the internal and external reliability and the internal and external validity of the research were noted.

Precautions that were taken to guard against threats to the internal reliability of the study were: the use of low-inference descriptors, the use of a participant observer, peer examination of data and the use of mechanically recorded data. The main threat to the internal reliability of the research came from the absence of multiple researchers at the research site.

External reliability was addressed through the explanation of the researcher's status and position at the research site, the choice of informants, the description of the social situation and conditions in operation during the research, the analytical constructs and premises on which the research had been based and the various methods employed to collect data. The external reliability of the research was threatened by the type and size of the population used in the study, the use of a single site for the research and the teacher's emphasis on written tests. The researcher's status as a teacher was also considered to be a threat to the external reliability of the study because it represented the narrow perspective of a teacher.

Issues that might have effected the internal validity of the study were also addressed. They were identified as: history and maturation, observer effects, selection and regression, mortality and spurious conclusions. The researcher's relationship with the sample, her position within the group and her inexperience in using the coding method of data analysis were construed as possible threats to the internal validity of the research. Using Goetz and LeCompte's (1984) conception of external validity, no threats were found to the external validity of the study.

The following chapter will discuss the data that was collected during the research. This will be done by a systematic analysis of each of the eight data collection methods used in the study. Whenever possible, direct quotes are given to substantiate the researcher's analysis. Themes that emerged from the data will be discussed in the fifth chapter of the thesis, Emerging Themes.