CHAPTER THREE

RESEARCH DESIGN AND METHODOLOGY

INTRODUCTION

> If you want to know how people understand their world and their life, why not talk with them?
>
> (Kvale 1996, p. 1)

The quote above encapsulates the perspective taken by this study, which was that exploring students' perceptions of their worlds through conversation can reveal the personal significance of the options they are considering, and the decisions they make. It was argued in the previous chapter that only a limited understanding of the influences on students' science enrolment deliberations can be gained from analysis of enrolment statistics or self-completed questionnaires. The argument was advanced that a study based around an interpretative research strategy had the potential to reveal more of the texture and finer detail of science proficient students' decisions regarding enrolment in senior science subjects.

However, from another perspective Kvale's question 'why not talk with them?' can also be interpreted as a challenge to justify this approach. It is important, therefore, to detail the strategies used to gather and analyse these perceptions. This study involved the design and implementation of a complex research methodology, an overview of which is illustrated in Figure 3.1. The structure of this chapter follows that of the overview.

DETERMINING AN APPROPRIATE RESEARCH METHODOLOGY

The research methodology was informed by a number of important propositions based on the literature. First, it is apparent that students' science enrolment decisions are somehow influenced by aspects of their sociocultural domains (Ainley et al. 1994; Fullarton & Ainley 2000; Woolnough 1994). Second, different responses to science of two categories of science proficient student can also be attributed to aspects of their backgrounds (Costa 1995). A third assumption is that personal conceptions about science are socially constructed (Cobern 1998b; Solomon 1987) from what individuals perceive to be the conceptions of science held within their social environments (Duit & Treagust 1995).

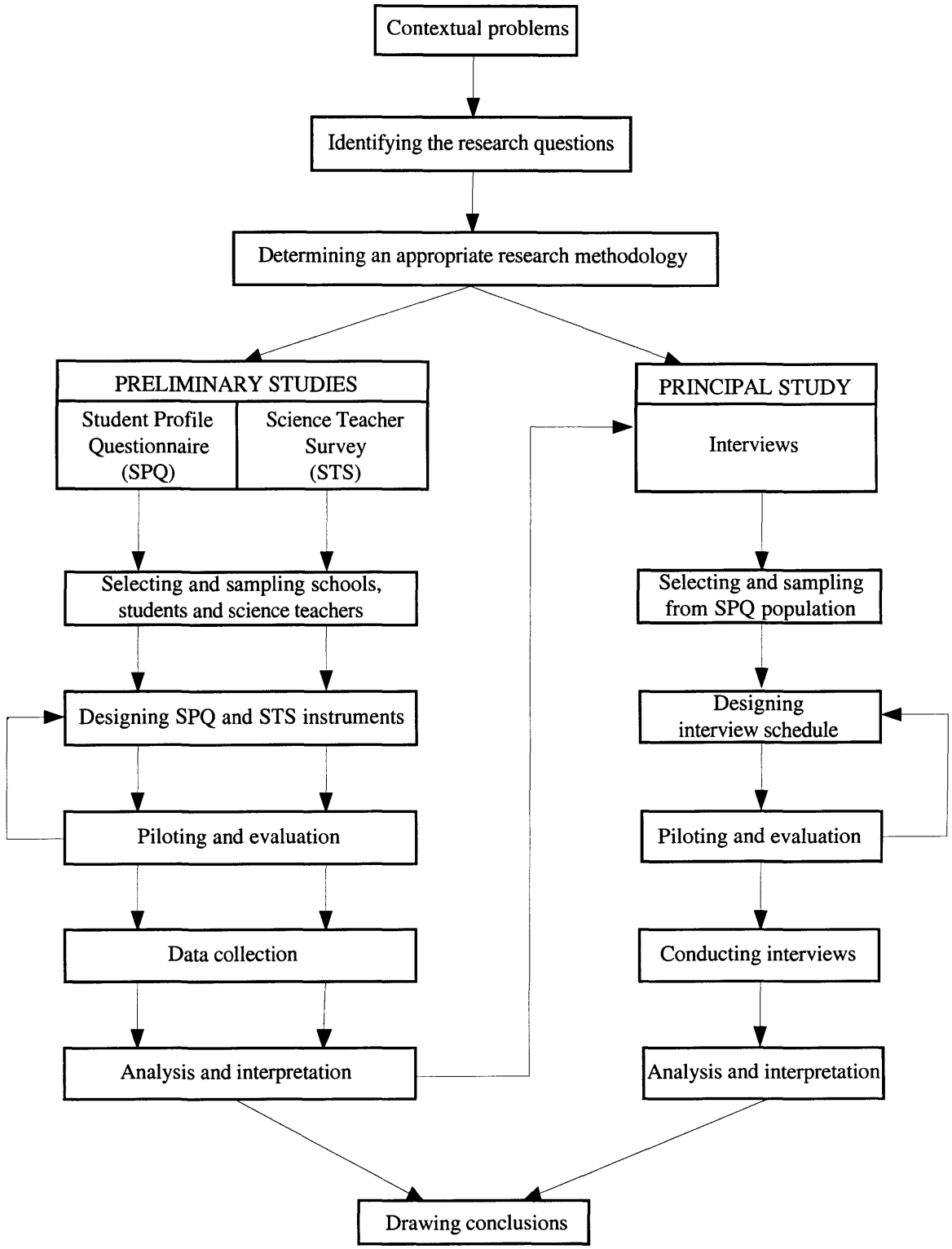

Figure 3.1 An overview of the methodology

Given these propositions, it was apparent that further investigation of science proficient students' decisions about school science demanded a principal research strategy better suited to the exploration of individual perceptions, cultural values and practices, and personal significance. Such a strategy required depth, flexibility and an openness to intuitive exploration, all of which are characteristics of qualitative research. This conclusion is supported by Eisner (1991, p. 35) who maintains that qualitative researchers are interested in looking beyond behaviours to matters of motive and the quality of experience undergone by those in the situation studied. His contention is that such strategies of inquiry provide a 'thick description' (Geertz 1973, in Eisner 1991, p. 35) of how individuals construct meaning.

Clearly, the research questions sought to look beyond the behaviour of choosing or not choosing particular science subjects, to the webs of significance which motivated that behaviour. There is a strong link between the meaning attached to concepts by individuals, and the behaviours of those individuals (Burns 1996; Eisner 1991). Given the evidence supporting the proposition that students' conceptions of science are socially constructed, it was decided to structure the methodology around a research strategy, in this case, in-depth open-ended interviewing, which had arisen out of a paradigm which assumes that reality is socially constructed (Burns 1996). However, in addition to this principal research strategy, it was recognised that complementary strategies would be required in order to help focus the interviews, and to provide a context for the findings. Thus, two preliminary studies, the Student Profile Questionnaire (SPQ) and the Science Teacher Survey (STS), were conducted prior to the main, interpretive, study.

Briefly, the SPQ performed as both a selection tool for identifying prospective interviewees, and as an exploratory relationship study (Cohen & Manion 1994) which would provide a profile of the backgrounds and consultation patterns of science proficient students. In terms of the iceberg metaphor, the SPQ was designed to take a single snapshot of many icebergs over a wide area. The purpose of the STS was also twofold. Since the other data generated in this study were limited to students' perceptions, it was decided that an alternative perspective would provide a valuable counterpoint. Science teachers were considered close enough to the student experience to have some appreciation of individual perspectives, but also to be well situated to report on enrolment patterns within particular schools over time. Extending the iceberg metaphor further, science teachers were conceptualised as relatively fixed observers, who see the surface features of many icebergs moving past over time.

Guided by the literature, and the analysis of these preliminary studies, the interview process sought depth and personal interpretation. While focusing on fewer individuals than the SPQ, the interviews explored the interplay of deeper influences and the effects these had on the surface phenomena. The interview process was conceptualised as probing not only the submerged features of the iceberg, but also the interactions between it and the surrounding environment. This multi-method approach thus involved both data triangulation and methodological triangulation (Denzin 1988). The contribution of triangulation to the trustworthiness of the study is discussed in the section on research integrity. First, however, each of the research strategies introduced above is described in detail.

FIRST PRELIMINARY STUDY: THE STUDENT PROFILE QUESTIONNAIRE (SPQ)

The primary aim of the SPQ (Appendix C) was to generate data on the backgrounds and subject choices of a substantial population of science proficient students. These data were used in both the quantitative and qualitative modes of inquiry, for three purposes. First, they aided the selection of prospective interview candidates according to school, gender and enrolment decision. Second, the SPQ data provided a statistical profile of the population, thereby establishing a context for the findings which emerged from subsequent interviews. Third, the data made administration of the interviews more efficient, since responses from a student's SPQ provided focus points for the interview, thereby eliminating the need for preliminary or redundant questions.

In addition to providing data on backgrounds and subject choice, the SPQ surveyed students regarding the sources upon which they relied when seeking advice about their senior subjects. As noted in the literature review, this was necessary since no assumptions could be made from existing studies about the consultation patterns of science proficient students.

Selection and Sampling

The process of selecting students for the SPQ was primarily one of choosing schools, since the potential population for the SPQ consisted of *all* science proficient Year 10 students in those schools. This section therefore addresses the criteria used to decide on the number, type and location of the schools involved in the study. Nine schools were originally approached to participate in the research. Seven of these were targeted as the study schools, while another was designated the pilot school. The ninth was chosen as a stand-by in the event that one of the other schools should be unable to participate in the study. This possibility was reduced by informal approaches made to the science coordinators at prospective schools to gauge their levels of interest in the study and willingness to be involved. All science coordinators expressed interest from the outset, as did science teachers more generally.

The decision to use seven study schools was based upon both conceptual and logistical considerations (LeCompte & Preissle 1993). Conceptually, the decision was guided by the requirement to make comparisons between students in three choice categories: those enrolling in physical science, those enrolling in biology/other science and those taking no science. Based upon an approximation of two students per choice category per school, and an interview population similar to Costa's (1995) 43 students, this resulted in a target of seven schools.

In terms of logistics, it was noted in Chapter One that the scope of the study was limited by the short time window available for data collection. Calculations regarding the size of SPQ and interview populations and, consequently, the number of schools involved, took into account the time required for coordinating the pilot studies and surveys, the need for preliminary analysis of the survey data before undertaking interviews, and the time required for these interviews. Furthermore, it was intended that data be collected as soon as possible following students' enrolment decisions, so that their consultation and deliberation processes

would be fresh in their minds and not influenced by subsequent experiences in the courses they had chosen.

The precise population size of science proficient students in these schools was not regarded as crucial, although since 'A' and 'B' grades were normally awarded to the top 30 per cent of students, calculations based on discussions with science coordinators pointed to an SPQ population of about 220 students. As noted previously, however, one of the schools was forced to withdraw at a stage in the study when it was too late to involve the stand-by school. This reduced the SPQ population to 196 and the interview cohort to 37.

The decision to select schools representative of the major categories in NSW was made in order to enhance trustworthiness, in accord with the principles of selection outlined by LeCompte and Preissle (1993) and Warwick and Lininger (1975). Furthermore, the major distinctions in NSW high school education reflected some of the variables discussed in the literature review as possibly being associated with different science enrolment patterns (Ainley et al. 1994; Young, Fraser & Woolnough 1997). These variables were location (urban/regional), education system (government/non-government) and school type (single-sex/coeducational). The proportions represented by these major categories, and the situation of the study schools within this classification, are illustrated in Figure 3.2.

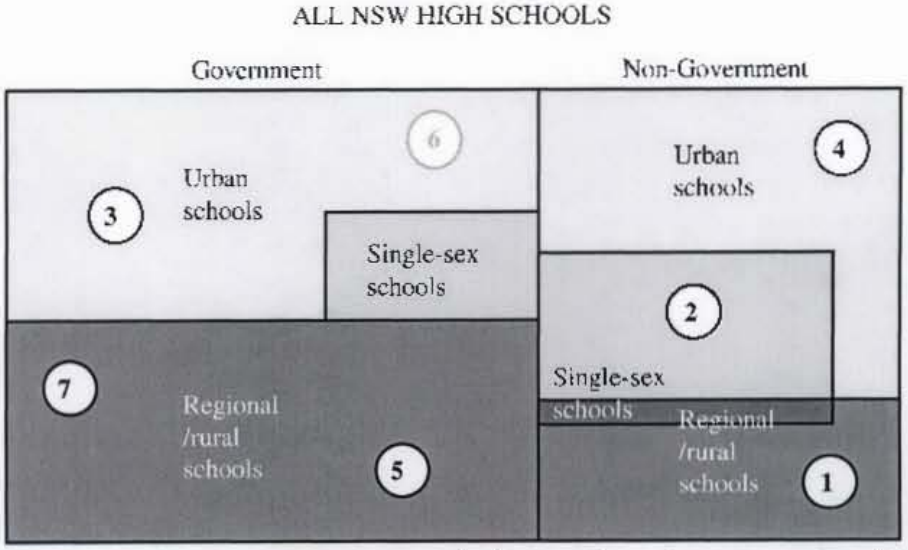

Figure 3.2 Schools involved in the study. The diagram shows the representation of the original seven study schools within the major categories of NSW high school. (School '6' later withdrew). The size of each area within the diagram is proportional to the number of schools in that category (Data provided by: NSW Department of Education & Training; Association of Independent Schools; Catholic Education Commission).

One category of NSW high school not represented was that of government selective schools. These schools select students on the basis of high academic ability, and claim to provide an educationally enriched environment by grouping talented students together, concentrating school resources and using specialised teaching methods (Department of Education and Training [DET] 2000a). Consideration was given to including a selective school in the study because of the high proportion of science proficient students which might be found in such a

school. However, the small proportion which these schools represent (23 out of 770, or less than 3 per cent of high schools statewide) mitigated against such an inclusion. Furthermore, there was some doubt regarding the comparisons which could be made between students who had experienced the 'enriched' science teaching and learning environments of selected schools, and those from other schools in the study.

Once the required categories of school were determined, the possibilities were narrowed to those schools large enough to ensure a reasonable number of science proficient students within each of the three choice categories. There is a polarity in NSW in terms of population density, with the urban areas - three major cities - clustered together on the coast, and the remainder of the state dotted with a few regional centres of between 20 000 and 50 000, as well as many smaller towns. The study schools were chosen from one urban, and two regional, centres. The regional centres were 600 km and 700 km, respectively, from the urban area. Schools in small rural towns did not have the Year 10 population sizes required for the study.

The five largest government and non-government coeducational high schools in the regional centres were approached, and agreed to participate. In the city, the four largest schools representative of each type (single-sex non-government, coeducational non-government and coeducational government) found within twenty kilometres of the CBD also agreed to participate in the study. It is argued that the criterion-based selection method (LeCompte & Preissle 1993) employed represented the best compromise between the conceptual requirements of the study and the time and cost constraints. Random sampling, which is often a requisite for generalisability in quantitative studies (Kerlinger 1973), may not have produced a study population representative of the main categories of NSW high school. This point is expanded upon later in the chapter.

Designing the SPQ

The format of the questionnaire was guided by the recommendations of Burns (1996), Leedy (1993), Warwick and Lininger (1975) and Wiersma (1991) with respect to structure, presentation and parsimony. The general arrangement of the questions followed the funnel, or 'reversed-pyramid', model (Lin 1976), proceeding from less taxing to more thought provoking questions.

Questions 1 to 8 of the SPQ sought demographic data. These required little effort to complete and were placed so as to lead the respondent well into the questionnaire, thereby building up momentum and encouraging continuation (Burns 1996; Leedy 1993). Questions asking the respondent's date of birth and gender were followed by three others concerning the ethnic origins of the student and their parents, as determined by countries of birth, years in Australia and languages spoken at home. In line with the studies by Ainley et al. (1994) and Myhill, Herriman and Mulligan (1994), ethnic identity was determined both by parents' countries of origin and the language most commonly spoken at home. Having both parents born in the same country, or the common use of the language of one parent's native country in the home, identified the ethnic background of that student.

Question 5, regarding brothers and sisters of the respondent, was concerned with whether the student had access to older siblings who had experience in the process and pitfalls of subject choice. Question 7, which asked about the respondent's religion, was felt to be relevant in view of the possible choice of one of the religious studies subjects (which may then limit choice of other subjects) as well as offering a potential point of interest in interviews with respect to ontological or philosophical perspectives. Questions 9 and 10 requested details of subjects and levels chosen in Years 10 and 11.

Question 11 was in a different format, incorporating scaled responses corresponding to the degree to which a student relied on advice provided by a number of significant individuals. These included the student's mother, father and best friend, as well as senior students, careers counsellors, science teachers, and any other source of advice. The format and categories were similar to those used by Warton (1997). This item had two purposes. First, it identified individuals who had the greatest direct influence on the decisions of each student, so that discussions in the ensuing interviews could then focus on these individuals. The prior identification of influential others was intended to save interview time. Second, the format of the responses was designed so that basic statistical analyses of the ratings could be used to provide an overview of the sources of advice upon which science proficient students relied. Statistically significant differences in ratings of individuals by students in different choice categories were of particular interest to the study. The use of the term 'reliance' indicated that the study was not interested so much in the quantity or soundness of the advice, but in students' perceptions of the capacity of that advice to influence their decisions. This construct therefore included a relationship dimension which could be explored during interviews.

Questions 12 and 13 were included in order to cross reference the schools' identification of these students as science proficient. It was thought there was a possibility that students' perceptions of their academic achievement in science may differ from the impressions of their teachers, given that grades were awarded according to the judgment of teachers, rather than in line with purely objective criteria. For example, a student may have been given a grade 'A' while seeing themselves as having below average academic ability in science. The format of the questions also allowed statistical comparisons to be made between students' self-efficacy ratings and background variables or enrolment decisions.

The final item on the SPQ gave students the opportunity to declare their willingness to be interviewed in the main study, thus identifying the potential interview cohort. Originally the SPQ was designed to include a number of other items including personal interests and attitudes to school science. However, in the final draft it was thought that a more parsimonious version would be better received by science teachers in the sample schools, in terms of minimising completion time and the necessity for teachers to clarify questions for respondents. There was also the possibility that students would be reluctant to respond openly to attitudinal questions knowing that their science teachers would be collecting the responses. This reticence was confirmed in the pilot study.

Finally, the introduction to the questionnaire was worded in a clear, open and inviting way so as to maximise the reliability of responses and the numbers of returns. Dillman (1978)

emphasised the importance of introducing a survey in such a manner that the respondent recognises that they and their opinions are being regarded positively by another person, that genuine appreciation is expressed for their time and effort and that they are being consulted on an issue of importance.

Piloting and Evaluating the SPQ

Thorough trialling of the instruments was crucial to the trustworthiness of the study (Rosier 1988; Warwick & Lininger 1975). Wiersma (1991) describes the purpose of a pilot as being to refine the instrument and identify misunderstandings, ambiguities and useless or inadequate items.

The questionnaire was piloted at a coeducational high school that was not involved in the main study. The pilot group consisted of five female and four male students who were proficient in science and representative of the three choice categories: physical science, biology/other science and no science. These criteria anticipated the characteristics of the eventual study population. Prior to writing, the 'Instructions to Students' were read aloud, and students were then asked whether they had any questions. The instruction process and the completion of the SPQ were timed and shown to take no more than twelve minutes, including distribution and collection of papers. This period was regarded as being short enough to allay any reservations which teachers' and parents' may have had about loss of lesson time, an important consideration since permission from these sources was required for participation in the study .

Following completion of the pilot SPQ there was a discussion with the students on the format of the instrument and their interpretations of the wording. The students were quite forthcoming and a number of minor modifications to the instrument and Instruction Sheet were made following their suggestions. These were all concerned with matters of style and clarity of wording. Of particular interest were students' interpretations of Question 11, which asked them to rate the extent to which they had relied upon the subject advice of others. When asked for their interpretations of this question, and of the term 'reliance' in particular, all students gave responses which, directly or indirectly, included some reference to the significance of their relationship with the advisor. For example, in the debriefing, the question of why students had relied on a particular source of advice resulted in some description of the relationship between the student and their advisor and, more importantly, the capacity of that person to influence the decision. As intended, 'reliance' was not seen to be solely a measurement of the degree of consultation, nor of the amount of advice offered, although these were also facets of the responses. The validity of the construct 'reliance on the advice of' was further enhanced by seeking clarification of students' responses to this question in subsequent interviews (Interview Question 9).

Of concern to some students was the possibility that they could be identified by their responses in the thesis or in any subsequent publications. The mechanisms for maintaining anonymity were then explained and deemed to be satisfactory. In general, however, no substantial alterations to the SPQ were considered necessary on the basis of the pilot study. Furthermore, all students expressed a willingness to participate in the interviews.

Collecting and Analysing the SPQ Data

The Student Profile Questionnaire was distributed to the study schools in the final term of 1998. The survey packs consisted of parental permission notes, plain language and consent forms, the SPQ instrument, and instructions for its completion (see Appendix C). Initially, the science teachers informed students about the study and distributed the Parental Permission forms. Once the permission slips were returned, science teachers allocated class time for completion of the questionnaire. As part of the regular communication maintained with science coordinators, I contacted each following the administration of the SPQ. No difficulties were reported.

In all, 332 SPQs were returned to the university. However, 115 of these were discarded as they had been completed by students who had achieved a science grade lower than 'B'. This eventuality had been anticipated since, in most of the schools, science classes were not academically streamed. In the interests of internal validity and school convenience it had been decided earlier that all students in mixed ability classes should complete the questionnaire, though returns from students not considered to be science proficient would be ignored. This decision is discussed in more detail in the section on Research Integrity.

A thorough cleaning of the remaining SPQs resulted in a further 21 returns being discarded, as they had either been incorrectly completed (n=2) or did not include the names of the students (n=19). The proper identification of students was important, since SPQ returns were checked against lists of 'A' and 'B' grade students provided by each school in order to identify science proficient students. Unidentified questionnaires could not, therefore, be assumed to have come from such students. Ultimately, the number of eligible SPQ returns was 196. A profile of this population is shown in Table 3.1.

Table 3.1 Overview of the SPQ population.

SPQ Population n=196											
Regional n=100 (51%)						Urban n=96 (49%)					
School 1 (NG/coed) n=35 (17.9%)		School 5 (G/coed) n=34 (17.3%)		School 7 (G/coed) n=31 (15.8%)		School 2 (NG/ss) n=49 (25%)		School 3 (G/coed) n=27 (13.8%)		School 4 (NG/coed) n=20 (10.2%)	
f	m	f	m	f	m	f	m	f	m	f	m
n=25	n=10	n=18	n=16	n=20	n=11	n=49	n=0	n=12	n=15	n=5	n=15
12.8%	5.1%	9.2%	8.2%	10.2%	5.6%	25%	0%	6.1%	7.7%	2.6%	7.7%

Key: G=government school; NG=non-government school; coed=coeducational; ss=single sex; f=female; m=male. Percentages are of the total population.

The higher proportion of female students overall (129, compared to 67 males) was primarily due to the inclusion of School 2, which was a single-sex girls school. This ratio had implications for the statistical comparisons, and in some cases School 2 data were excluded from calculations. Even ignoring this school, however, the higher representation of females in

the remaining SPQ population reflected the fact that, statewide, more females than males achieved an 'A' or 'B' grade for science in their School Certificates in 1998 (BoS 1999a).

Three data sets were compiled from the SPQ responses. The first consisted of background data, as outlined earlier. Non-numerical variables such as location, gender and ethnic background were allocated a nominal value and entered into a spreadsheet document. The second data set contained the subject enrolment details which enabled the SPQ population to be divided into those who were continuing with science study and those who were not. As noted above, the science group was subdivided further, acknowledging the distinct differences between physical science students and those taking biology/other science, in terms of their motivations and background factors (Ainley et al 1994; Fullarton & Ainley 2000; Woolnough 1994). The resulting choice categories are detailed in Table 3.2.

Table 3.2 Choice categories of the SPQ population, based upon their science enrolment decisions.

Choice Category	Courses chosen	No. students
Physical science *(physci)*	Physics only Chemistry only Physics and chemistry Physics, chemistry & biology 3 or 4 unit science	80
Biology/ other science *(biother)*	Biology only Marine Science only Biology and Marine Science General Science Science for Life	50
No science *(nosci)*	No science	39
Cross category *(crosscat)*	Physics and biology Chemistry and biology	27

Students whose subject profiles contained more units of physical than biological science were included in the physical science choice category. Those with more units of biological than physical science, or who had chosen non-specialist courses such as Science for Life or General Science, were included in the biology/other science category. However, there were 27 students who chose an equal number of units from each category. The difficulty of allocating such students to choice categories had been encountered in other research attempting to distinguish between physical and biology/other science students (Ainley, Jones & Navaratnam 1990), and was resolved in that research by resorting to further classification with reference to the amount of mathematics being taken with the science courses. This complexity was avoided in the present study by the decision to establish a fourth category for these students. With regard to statistical analysis of SPQ data, while statistical comparisons were generally made between students in the first three choice categories, comparisons were

also made between all four choice categories to determine the effects on contingency tables of the cross category students. As the effect of including these students was either inconsequential, or blurred significant differences between the other three choice categories, cross category students were only included in bivariate comparisons between science and non-science students.

The third set of data related to students' perceptions of the degree to which they relied upon the advice of others in their deliberations. The responses to the Likert items were given ordinal values and recorded in a spreadsheet document. Explorations of the SPQ data were conducted via three series of crosstabulations between the data sets, using SPSS (see Figure 3.3).

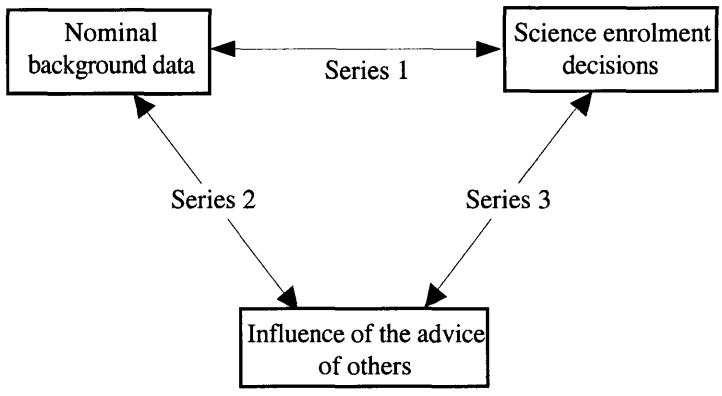

Figure 3.3 Analysis of the SPQ data. Three series of crosstabulations were conducted to search for indicative patterns in the data.

Each crosstabulation produced a contingency table displaying the number of cases falling into each combination of the categories of two or more categorical variables (SPSS 1995). This mode of analysis suited the exploratory nature of the preliminary studies, and the requirement for relatively simple bivariate statistics which would indicate the existence of statistically significant relationships between background characteristics, enrolment decisions and sources of advice. These relationships could then be explored during the interview stage. Furthermore, the non-interval nature of much of the data precluded the use of more complex analyses. The decision was also influenced by Kerlinger's (1973, p. 176) recommendation that a researcher should never use a complex analysis when a simpler one will accomplish the analytical task.

Each contingency table within a series of crosstabulations produced indicators of association, including patterns within the residuals, which are the differences between actual and expected frequencies. The significance of any associations was determined by reference to the Pearson and Likelihood Ratio chi-square statistics. These statistics indicated whether the null hypothesis, that is, the assumption that the row and column variables were independent, could be rejected. Significance levels of 0.05 or less (Everitt 1977; Hair et al. 1998) were interpreted as indicating a high probability of association, and thereby as reliably supporting rejection of

the null hypothesis. However, due to the exploratory nature of the SPQ preliminary study, significance levels up to 0.1 were considered suggestive of an association worthy of further investigation in the interview stage, despite the increased chance of a Type 1 error, as long as the statistics were accompanied by frequencies and residuals which also suggested an association (Cooksey 1997; Everitt 1977; Kerlinger 1973).

A final consideration in the interpretation of contingency tables concerned the range of cell values. Conventionally, statistics resulting from crosstabulations have been regarded as unreliable in cases where any cell has an expected value less than 1.0, or where more than 20 per cent of cells have expected values less than 5.0 (SPSS 1995). However, some studies indicate that this condition is probably too stringent and can be relaxed (Everitt 1977; Norusis 1994). Thus, while the present study generally complied with the conventional limitations, crosstabulations which resulted in cell values lower than these limits were not discounted out of hand, particularly where the patterns of residuals and the significance levels were indicative of a potentially interesting relationship. Where such cases are reported in later chapters, however, the caveat regarding expected values is noted.

It is apparent from the description above that interpretation of the contingency tables did not always adhere to the most stringent of guidelines required by some quantitative researchers. However, the more flexible levels of tolerance observed by the present study were appropriate given that findings from these crosstabulations were not intended to be stand-alone conclusions, but regarded rather as indicators of relationships which deserve further exploration in the interview process. As Hair et al. (1998) maintain, the context of the study should dictate the margins of error which are acceptable.

While each series of crosstabulations was conducted using all variables within the relevant data sets, only a few of these crosstabulations produced results which were significant or relevant to the study. Because the SPQ data were primarily meant to prompt further inquiry in the interviews, findings from crosstabulations are only reported in the thesis where they supported, qualified or conflicted with findings from the STS or interview analyses. Consequently, contingency tables, significance levels and graphs from the SPQ crosstabulations are not presented and discussed in a single chapter, but incorporated into relevant arguments in various chapters.

With regard to analysing students' ratings of their reliance on the advice of others (Q. 11), consideration was given to converting the ordinal data into interval measurements through Rasch analysis (Wright & Masters 1982), so that the significance of differences between ratings could be determined. Ultimately, however, this approach was not used, for two reasons. First, Rasch analysis of ordinal Likert values demands that ratings refer to a single variable, or unidimensional construct (Snyder & Sheehan 1992), where, for example, items can be ordered from lowest to highest, or easiest to hardest. The construct investigated by the present study, 'perceived capacity of others to influence decisions', was not considered to be unidimensional, since students' ratings of the advice of others depended upon a number of variables, including their willingness to consult others, the opportunities for consultation, students' relationships with different individuals, and the quantity and qualitiy of advice

offered. With so many variables, the order of ratings would not necessarily reflect identifiable characteristics of the rated individuals, or of the students.

Second, the study was less interested in the order in which students rated these individuals, than in significant ($p<0.05$) or suggestive ($p<0.1$) differences which might be found between the ratings of students in different choice categories, since such differences might provide clues regarding influences on students' enrolment decisions which could be followed up in the interviews. Hence, there was little necessity for rating order to be an interval scale measurement.

SECOND PRELIMINARY STUDY: THE SCIENCE TEACHER SURVEY (STS)

As mentioned previously, the literature contributed few insights regarding teachers' perceptions of the influences on students' science enrolment decisions, with the notable exception of Woolnough (1993). In view of this paucity of research, and in order to examine students' enrolment deliberations from a different vantage point, it was decided to survey science teachers at the sample schools.

The Science Teacher Survey (Appendix D) looked for two types of information. First, it sought background data on science teachers at the participating schools, with particular regard to the breadth and extent of their teaching experience. Second, the STS garnered teachers' views as to why some high achieving students choose not to take senior science, and whether any change in the numbers of such students over time had been detected.

Selecting the Science Teachers

It was determined from the outset that only science teachers from the sample schools would be invited to participate, so as to reflect similar school, regional and systemic variables to those of the students in the study. In this way the data generated by these teachers were seen to have greater credibility when referenced to the student data than would have otherwise been the case. A disadvantage of this decision, however, was that the total number of science teachers in the sample schools was not sufficient to allow some otherwise germane analytical processes to be used, thus limiting the significance of the data and reducing the possibility of them being generalisable. Despite this, it was felt that the disadvantages of surveying teachers outside of the sample schools outweighed any statistical advantages of doing so.

Designing the STS

The design rationale of the STS was consistent with that which guided the SPQ. Primarily, the survey was required to be brief, uncomplicated (Cohen & Manion 1994) and undemanding in terms of completion time so as to maximise the number of returns (Warwick & Lininger 1975). The instrument itself (Appendix D) asked, firstly, for nominal data and details concerning teaching experience in order to build up a profile. The focus of the survey was then reintroduced and teachers asked what they considered to be the main motivations behind the decisions by some high achieving students to forgo senior science study. The

format encouraged open, extended responses. The final questions asked teachers about their perceptions of changes in the proportions of science proficient students enrolling in science in their school. The purpose of this question was to determine whether the teachers saw the phenomenon as localised, perhaps due to school factors, or more pervasive. The response format of this item also encouraged an extended answer.

Evaluating the STS

A draft of the STS was reviewed by a panel of experts, consisting of two senior science teachers at the high school involved in piloting the SPQ, two senior university lecturers in science education, and a doctoral candidate with fourteen years experience as a high school science teacher. The draft was passed in with only minor corrections and no recommendations for modification. The use of a panel of experts was consistent with the recommendations of Warwick and Lininger (1975, p. 43).

Collecting and Analysing the STS Data

The STS questionnaires were mailed to science coordinators for distribution to their teachers. Each form included a Plain Language Statement and Consent form (see Appendix D). Teachers in the six schools responded very positively considering that it was a mail survey which encouraged extended responses and that it was imposed on them at a very hectic time of the school year. A total of 24 valid responses were returned, representing a response rate of approximately 70 per cent.

This questionnaire generated both nominal data and extended responses. Because of the function of the STS and the size of the sample, sophisticated analytical tools were not required. The nominal data were coded and entered into a spreadsheet document to create a profile (see Appendix E). The extended responses, including perceptions of students' motivations, were entered into the computer application QSR NUD*IST and indexed according to categories of response. NUD*IST, which stands for Non-numerical Unstructured Data Indexing, Searching and Theorising (QSR 1997), was an ideal tool for storing and analysing qualitative data. As the NUD*IST application was primarily used to analyse the interview responses, the procedures involved are detailed later in the section describing that analysis.

Findings from analysis of the teachers' responses are discussed in Chapter Four, and compared with students' own explanations for their enrolment decisions. This comparison also provided points of interest for the interview process.

THE PRINCIPAL STUDY: INTERVIEWS

Referring again to the iceberg metaphor (Figure 2. 3), the interview process was designed to explore below the topographical features, such as subject profiles and rationales, to deeper influences on students' perceptions and behaviours. The interview was regarded as the ideal research tool for revealing how participants conceived of their worlds and how they explained these conceptions (Schatzman and Strauss 1973), and for delving into the motivations of respondents and their reasons for responding as they do (Kerlinger 1973).

Specifically, the interview process performed three main functions. First, it developed in more detail the profiles of selected students and their worlds, providing depth and texture to the background data collected by the SPQ. Students were able to furnish accounts of the significant others and interpersonal dynamics existing within their worlds; that is, what Kvale (1996, p. 5) refers to as 'the life world of the interviewee.'

Second, the interviews revealed the respondents' various perceptions of, and interactions with, science within the home, peer, mass media and school worlds, as discussed in Chapter Two. Of particular interest were students' experiences with school science and the contrast which may exist between the ways science was conceptualised in school and within their other worlds.

The interview's third purpose was to encourage respondents to relate their experiences of the process of subject choice, including their personal priorities, influential others and the extent and sources of advice. The portrayals of different worlds, experiences with science and deliberations about enrolments could then be used to explore the degree to which cultural congruence was influential in these students' decisions about enrolling in senior science.

Selection and Sampling

The interviews comprised the core of this study, and so design issues and rationales for selection and sampling were regarded as crucial, particularly in terms of credibility. The intended size of the interview population, as discussed earlier in the chapter, was calculated on an average of six students from each of seven schools. Selection and sampling of interviewees within each school was made primarily on the basis of students' willingness to participate and on their representation within the three choice categories.

However, within each school SPQ population there were a number of students representative of each choice category who were willing to be interviewed. Therefore, a process combining criterion-based selection (LeCompte & Preissle 1993) and probability sampling was used to identify the interview population. With regard to the criteria, a reasonable balance between males and females within each choice category was desirable, given the gender differences in enrolment motivations noted by Kelly (1988). Second, it was important to the study to include some students who had a non-English speaking background, given the associations between this characteristic and patterns of science enrolment (Ainley et al. 1994). Third, it was originally intended that fourteen students from each choice category would be interviewed. However, with the withdrawal of one school, the proportions were revised in the remaining schools. Rather than evenly reducing the number of interviewees in each category, it was considered more important to focus on physical science and non-science students, given the greater concern about declines in physical science enrolments. This decision accounts for the lower number of biology/other science students.

Within these parameters, it was also desirable to include interesting and atypical cases, as recommended by LeCompte and Preissle (1993, p. 70), in order to uncover a wider range of perspectives (Glaser & Strauss 1967; Patton 1990; Taylor & Bogdan 1998). Thus, the

interview cohort eventually included some students who had altered their original choices from chemistry to no science, another who had changed from no science to physics; females who had chosen six units of science; high 'A' grade science students who had chosen no science; and 'B' grade students who had chosen four units of physical science. The implications of the selection and sampling methods for the credibility of the study is discussed later in the chapter.

In those situations where several students within a school fitted the necessary criteria, random sampling was employed using the numerical order of identity codes, which had been applied at random to returned SPQs earlier in the study. Profiles of the resulting interview cohort, with the pseudonyms given each student, are found in Appendix F.

Designing the Interview Schedule

The interview schedule (Appendix G) was informed by the theoretical model and the research questions. The questions were designed to explore students' perceptions of the influences within their worlds, in the context of deliberating about senior science enrolment. The ways in which particular questions addressed the relationships and constructs within the model are illustrated below in Figure 3.4.

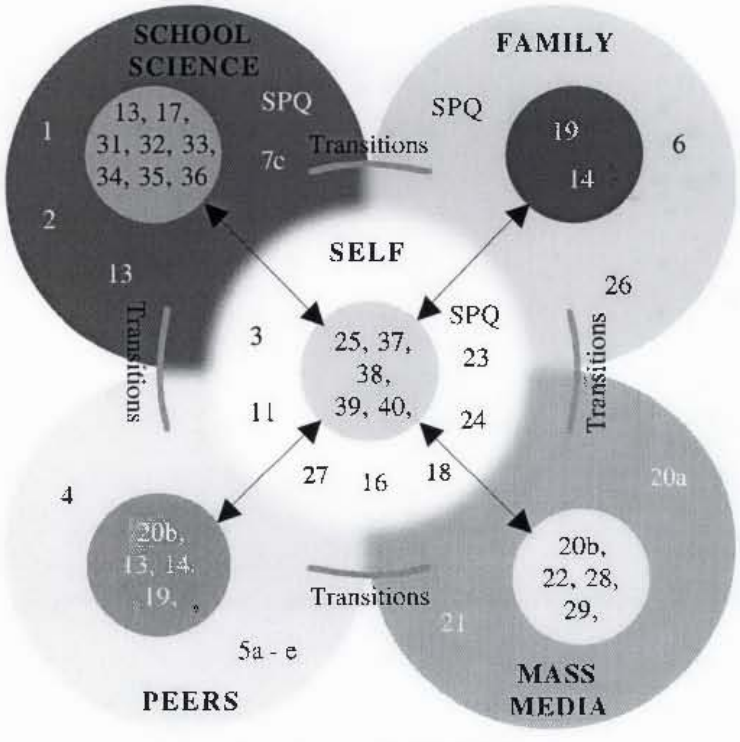

Figure 3.4 The relationship between the Interview Schedule items (see Appendix G) and the multiple worlds model. The smaller circles indicate questions which addressed conceptions of science within the worlds, as in Figure 2.6.

Situating several items in a single region of the diagram illustrates how a number of questions approached a particular issue from different perspectives, in the search for rich and varied information (Kvale 1996). However, while many of the items addressed specific worlds, as shown in the diagram, there were a number of non-specific questions, which are listed separately. The reason for not specifying a world in every question is discussed later in this section.

As the interview schedule was constructed specifically for this study, there were a number of fundamental design considerations. The hierarchy of decisions and the design parameters of the schedule were guided by the recommendations of several sources in the literature. While some of these were quite prescriptive (Kerlinger 1973; Tuckman 1978), the final design owed more to the flexibility and variety in format recommended by LeCompte and Preissle (1993) and Fontana and Frey (1994).

The first decision concerned whether the interviews should be structured, semi-structured or unstructured (Burns 1996; Cohen & Manion 1994; Fontana & Frey 1994; Kvale 1996). A semi-structured format was considered to be the most suitable arrangement, as this offered a good compromise between structure and flexibility. Kvale (1996, p. 129) describes the semi-structured interview as following a sequence of themes and questions, while at the same time being open to changes of sequence and form which allow the interviewer to follow up responses and further explore the stories told by respondents. This format also encourages the researcher to be more natural and responsive than would a tightly structured schedule (LeCompte & Preissle 1993), while following a sequence enables fair comparisons to be made between students' responses.

The literature presented a range of apparently dichotomous design options, such as open and closed questions, direct or indirect questions, general or specific, fact or opinion, and so on (Cohen & Manion 1994; Kerlinger 1973; Patton 1990; Tuckman 1978; Wiersma 1991). The design followed the advice of Pelto and Pelto (1978) and Schatzman and Strauss (1973) in not treating such categories as mutually exclusive, but rather seeing the options as a smorgasbord of useful alternatives which may be appropriated for different, or complementary, purposes within the same instrument.

In relation to open or closed questions, or more accurately, questions designed for open and closed responses (LeCompte & Preissle 1993) both types were suited to different parts of the schedule. While open response questions were not always necessary, Cohen and Manion (1994, p. 277) have pointed out that such questions can lead the researcher along an unexpectedly fruitful path. The schedule also contained a composite of direct and indirect questions. While the use of the former might seem the most rational and economical path, Tuckman (1978, p. 198) has noted that some respondents shy away from such an approach. The 'rule of thumb' used in this present study was that the more personal the focus, the more indirect the question. The risk, of course, was that some questions might be so oblique as to generate data of little use. In such instances, the questions were asked again in a more focused manner. This approach, however, resulted in an increased number of questions and

therefore a greater requirement of time. As with all design decisions, concessions needed to be made to time and other limitations.

In the quest for credible responses, consideration was given to whether the wording of certain questions could result in guarded or biased responses. Questions focusing on specific worlds, for example, on family or peers, had the potential to cause disquiet or suspicion (Cicourel 1964; Cohen & Manion 1994). Consequently, it was decided at an early stage not to specify particular worlds in a number of the questions probing perceptions of cultural features (see Figure 3.4). It was thought that this approach would avoid 'leading' the interviewee, and that responses to such questions would provide some indication of the relative prominence of influential agents in the mind of the student. A disadvantage of this position, again pointed out by Tuckman (1978, p. 198) is that the design required a number of indirect questions in order to lead honestly, if circuitously, to the information.

The schedule also contained a balance of factual and opinion responses (Tuckman 1978), requiring the respondents to alternately draw from shallow and deep reflection. That is, students could answer some questions without needing to deliberate over their responses, while other questions required them to articulate concepts and experiences that they may not have previously considered at length. In fact, during the debriefing phase, several interviewees noted that the questions required them to think more deeply about their decision making processes than they had done when choosing their subjects. As a counterpoint, some items (Qs. 16, 20, 28) were worded in a simple, breezy manner, temporarily changing the tone of the interview. However, they were not simply 'padding questions' (Lin 1976) since they generated relevant data, as well as providing respite from the more probing questions.

Piloting and Evaluating the Interview Schedule

The interview procedure was trialled using a subset of the group which piloted the SPQ. Two male and two female students, representative of the physical science, biology/other science and no science choice categories, were interviewed individually under conditions anticipated in the sample schools. The students acted firstly as interviewees, and later as evaluators, being asked for their opinions on question comprehension, wording and relevance at the end of each page of the schedule. All of the interviews were tape recorded and timed.

Outcomes from the interview pilot study

Overall, few modifications were made in the light of the pilot study. No problems emerged regarding the explanatory covering letter or permission note, and all students agreed to be tape recorded. It was noted that students appeared to ignore the tape recorder once the interview was underway. The interviews took between 46 and 53 minutes each to complete.

No student recommended modifications to the schedule as a result of the evaluations. In listening to the recordings later, however, some ambiguity in students' interpretations of the term 'science' was noticed. In asking about attitudes to science, two students understood the question to pertain to school science while the others interpreted it as concerning science more generally. In considering how to clarify this interpretation, it was decided that there was some benefit in leaving the term undefined and without context in early questions, in order to

appreciate how different students' conceptualised 'science'. In other questions, however, particularly those addressing students' classroom experiences, it was necessary to use the specific term 'school science'.

Following analysis of the pilot transcripts, two additions were made to the schedule. The first, Q. 29a, investigated students' awareness of science careers. This was included in response to comments by two students indicating a very limited appreciation of the variety of possible science careers. The second item (Q. 36) was included to provide students with an opportunity to reflect on what they considered the most significant influences on their science enrolment decisions, in the light of their interview responses up to that point.

Apart from these additions, no substantial changes were made to the schedule following the pilot. On the other hand, in reviewing the recordings, it was recognised that more attention needed to be paid to the interview process itself, including the pre-checking of levels on the tape recorder to optimise sound quality, and the need to concentrate more on the respondents and their answers and less on the recording apparatus and interview schedule. There was also a need to minimise my interruptions to the flow of the responses, and allow pauses in the conversations, since pauses sometimes encouraged students to follow up or clarify their initial responses to a question. Finally, I recognised that vernacular terms and neologisms would require annotation in the transcripts in order to be properly understood. References to non-verbal communication would also need to be made in transcripts, in order to communicate more fully the tone and context of responses.

To determine the quality of the data generated by the pilot interviews, the responses were evaluated using Kvale's 'Quality Criteria for an Interview' (1996, p. 145). While this was an idealised guide, it at least provided an itemised list for comparison and reflection. The evaluation confirmed the fruitfulness of the interview as a strategy in terms of the quantity and rich quality of the data. As Eisner (1991, p.182) commented, 'it is surprising how much people are willing to say to those whom they believe are really willing to listen.'

Collecting and Analysing the Interview Data

Conducting the interviews

Once the prospective interview participants for a given school were selected, I contacted the science coordinator and negotiated arrangements for conducting the interviews, which were typically completed over two or three days. Interviews generally took place in a library tutorial room, or an interview room in the administration area, though they were also conducted in classrooms, laboratories and, in one instance, a playground. This variation between interview environments had implications for the quality of the data, an issue which is discussed in the section on research integrity. Each interviewee was given a brief outline of the nature and purpose of the study in order to orientate and relax them (Kvale 1996; Tuckman 1978).

The initial interviews took between 50 minutes and an hour, though, following early analyses, some interviews took a little longer as I delved for more detail in particular areas to support or test emergent theory. This fine tuning of questions in response to early data is advocated by Kvale (1996) and Eisner (1991). All interviews were tape recorded with the consent of the

respondent. In addition, notes were recorded on individual interview schedules to provide back up for the recording, to record non-verbal communication and to register interesting ideas or new avenues for examination that arose during the interview. Upon completion of each interview, a short debriefing was conducted to smooth the transition from, in some cases, the intensity of the interview to the normality of the school day, and also to once again assuage any concerns as to confidentiality or the purpose of the study (Kvale 1996).

Treatment and Analysis of the Interview Data

Unlike the analytical strategies used for quantitative data, there is far less specification about analysis of non-numerical data, no *via regia* to arrive at essential meanings and deeper implications (Kvale 1996). Researchers need instead to believe that the meanings are there in the data, and that they can be teased out and presented in a way which respects the integrity of the individual's responses, while at the same time showing how these can also be part of a larger pattern. As a consequence, I needed to be conscious of both the assumptions of the theoretical model and the emergence of unanticipated patterns. That is, I employed both deductive and inductive methods interactively, as suggested by Miles and Huberman (1994). As well as ensuring the coherence of the study, the deductive approach had the advantage of focusing and reducing the data that could be included. Induction, on the other hand, enabled reexamination, modification and where necessary, the rebuilding of the original constructs (Sowden & Keeves 1988). Analytical methods involving the inductive approach also have greater credibility, since there is no temptation to ignore patterns which do not conform to hypotheses.

In the light of these advantages, analysis of the interview data was based upon the constant comparative method, originated by Glaser and Strauss (1967) and refined by Maykut and Morehouse (1994).

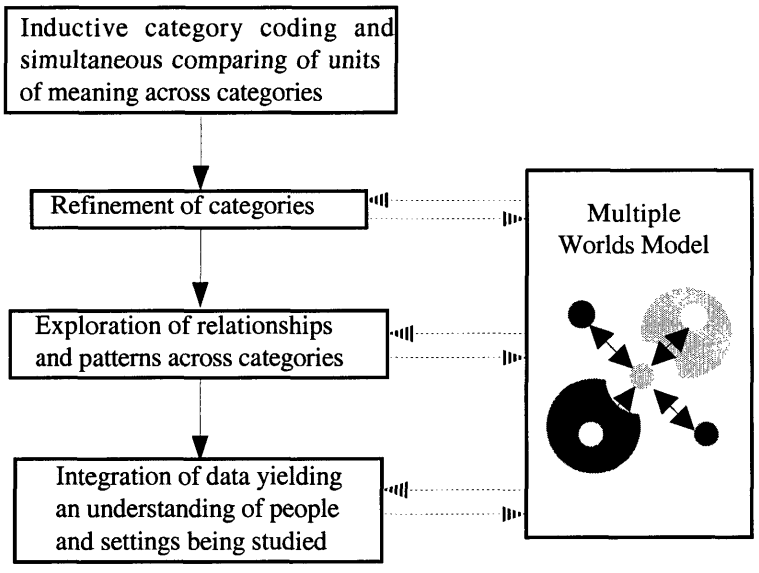

Figure 3.5 The analytical process. Analysis of the interview data followed the constant comparative method (Glaser & Strauss 1967; Maykut & Morehouse 1994) with reference at each stage of refinement to the multiple worlds model.

Using this method, each new unit of meaning, for example a particular student response, was compared to all other units and subsequently grouped (categorised and coded) with similar units of meaning. The flow diagram suggested by Maykut and Morehouse (1994) has been modified in Figure 3.5 by the addition of reference loops back to the model at each stage, indicating a dialogue between the induction process and the theory. Analysis within each step of the constant comparative method used many of the tactics suggested by Miles and Huberman (1994) for deriving meaning from qualitative data.

One aim of the study was to explore the degree to which cultural congruence, or incongruence, between different worlds had played a part in students' deliberations. In analysing the narratives, therefore, it was important to be sensitive to indications of resonance or dissonance, as evidence of possible congruence or incongruence (Festinger 1957, 1964; Misiti & Shrigley 1997; Rea-Ramirez & Clement 1998; Sears et al. 1988). Resonance in enrolment deliberations could be indicated by a student's sense of confidence regarding his or her choices, a clear articulation of motivations, descriptions of advice being reinforced by other sources, or by expressing satisfaction with choices. Signs of cognitive dissonance, on the other hand, could include descriptions of difficult or stressful subject deliberations, descriptions of contradictory or confusing advice, unclear motivations, references to competing imperatives, or expressions of regret regarding decisions.

The tool which facilitated analysis of the interview data was QSR NUD*IST, the computer application mentioned earlier when discussing STS responses. This powerful programme supports the processes of coding data, searching text, searching patterns of coding and theorizing about the data (QSR 1997, p. 2). The interview recordings were initially transcribed into a text format accessible by the programme. Notes were made of any non-verbal responses, such as laughing, sighing or pausing for thought, as well as contextual information, for example, cynicism, sarcasm, enthusiasm, reluctance or guardedness. This annotation was in line with Kvale's (1996, p. 179) recommendations that qualitative researchers be cognisant not only of what is said, but of what is not said. In addition, memoranda, or 'memos' (QSR 1997), were created for each interview document, summarising the initial impressions and interesting points which arose during interviews and transcription.

*The QSR NUD*IST project*

After being introduced as a NUD*IST document, each interview transcript was coded and stored as a series of text units which could be identified and retrieved at will. A text unit is the smallest unit of a data document handled by the NUD*IST application (QSR 1997) and generally consists of all the text between one 'hard return' and the next. In this study, each uninterrupted response, question or prompt was considered a single text unit. All text units were numbered for reference, and these numbers accompany each interview extract quoted in this thesis.

An index tree was constructed in order to organise both the raw data and the results which emerged from analysis. An abridged version of the final index tree is shown in Appendix H. The subtree labelled 'BASEDATA' indexed whole documents in appropriate categories,

according to gender, location, choice category, ethnic background, and so on, so that comparisons between categories, such as science and non-science students, could be made in relation to themes or patterns which arose from the analysis. The 'BASEDATA' subtree also indexed all responses to each of the schedule items.

Three other subtrees were constructed to store text units and the results of analysis. The structure accommodated both analyst-constructed typologies, which were based upon the multiple worlds model, and indigenous typologies which were created in the process of inductive analysis (Patton 1990). The subtree 'MULTIWORLDS', was structured around four major categories, or parent nodes (QSR 1997), corresponding to the four worlds. These nodes and their descendants stored all references to each world. Each parent node stored the accumulated data of all its descendants.

A third subtree, 'SELF', stored all references to students' perceptions of themselves and their own interests, attitudes and values. Specific references to decision making processes, however, were stored in a fourth subtree titled 'SUBJECTDELIB'. NUD*IST allowed multiple coding of the same text unit, meaning that, for example, a student's reference to her mother's advice about taking chemistry could be stored in each of the nodes 'Family', 'Parents', 'Mother', 'External Influence' and 'Chemistry'.

The fifth subtree, 'EMERGENT', stored the results of explorations and comparisons across categories. The major node structure here followed the conceptual framework in distinguishing between topographical features and underlying sociocultural influences. The cultural profiles of the worlds were further developed in terms of their structural, attitudinal and dynamic dimensions. NUD*IST allowed text searches and a range of index searches to investigate patterns, the results of which were stored in this subtree. Moreover, as part of the reflection process, node memos were produced to express ideas, suggestions and thoughts on the results of node intersections. Significant memos themselves became part of the node structure, providing system closure, a process by which the results of searches and analysis are fed back into the system (QSR 1997).

The index tree shown in Appendix H presents only part of the ultimate configuration, which consisted of 316 nodes. The relationship between subtrees altered many times over the period of analysis, as new themes or relationships emerged. NUD*IST allowed nodes and whole subtrees to be shifted and reattached as ideas developed and were tested. Furthermore, while some nodes were constructed in the light of the conceptual and theoretical frameworks before analysis began, many more were created during the analysis.

The considerable time spent introducing, coding and categorising the data using NUD*IST was repaid many times over by the flexibility and power of the indexing system. Hunches or hypotheses could easily be investigated by commands to intersect or otherwise collate nodes, or by searching text selections. Most importantly, a number of substantial findings from this research were unexpected, and the patterns that gave rise to these findings would perhaps have been overlooked had a less thorough and flexible tool been used.

ISSUES OF RESEARCH INTEGRITY

The integrity of this thesis relies on the demonstration that it was conducted in accordance with acceptable research standards and practices. These range from consideration of ethics, such as gaining consent and preserving anonymity, to establishing the trustworthiness of results and conclusions. The steps taken to maintain the integrity of both the findings and the individuals involved in the study are detailed below.

Ethics of Data Collection and Reporting

The importance given to ethical considerations in research has greatly increased over the last decade (e.g. Cohen & Manion 1994; Taylor & Bogdan 1998). The principal considerations addressed in undertaking this project included gaining the consent of participants (and, in the case of minors, their parents or guardians), clearly communicating the aims and implications of the research, guaranteeing and maintaining anonymity and honestly representing the data (Bell 1991; Cohen & Manion 1994).

Because multiple data sources were used, the process of gaining consent was a complex and protracted one, as illustrated in Figure 3.6.

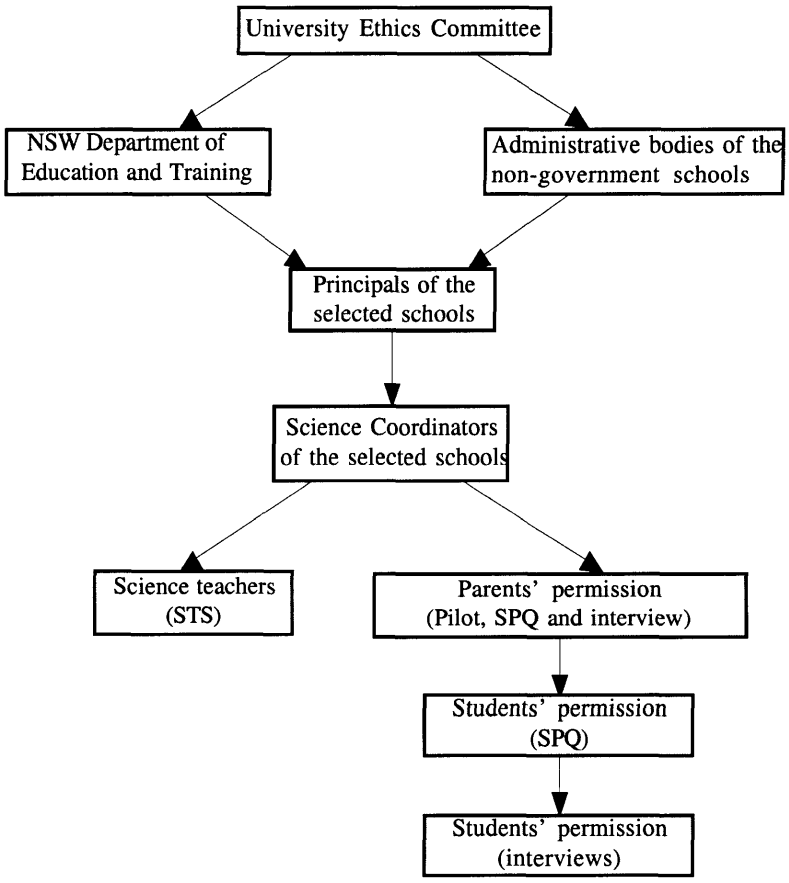

Figure 3.6 Permission Flow Chart. Consent was required from each of the parties above, in the order indicated by the arrows. See Appendix I for university and departmental correspondence giving consent for research.

At all stages, the data collection needed to comply, and be seen to comply, with ethical guidelines prescribed by both the University of New England Ethics Committee and the NSW Department of Education and Training. Plain Language Statements were provided to parents, students and science teachers and consent was obtained from all three parties. At the beginning of each interview, the aims and implications of the study were explained and students asked whether they had any questions. One copy of each of the consent forms was retained by the student while a second was kept on file.

Another ethical consideration was ensuring the anonymity of the respondents (Cohen & Manion 1994; Leedy 1993). For the SPQ, student and school identification were recorded on a master sheet against a coded number, as noted earlier. From this point on the scripts were identified by this number, and all other references to identity were removed and destroyed. At the beginning of each interview the student was again reminded of the confidentiality of their responses and the independence of the researcher in relation to the school or school system. Tapes, transcripts and reports of interviews, either electronic or hard copies, were labelled with the identity code and securely stored.

As an experienced teacher, I am familiar with the NSW Department of Education and Training policies on child protection. All efforts were made in interviews to avoid situations in which the respondent may have felt uncomfortable or which in any way transgressed departmental guidelines.

Establishing the Trustworthiness of the Findings

The value of any research study hinges upon its ability to demonstrate conformity to the criteria of integrity generally accepted by a particular paradigmatic community. Since the criteria vary between research models, studies such as this which are not situated entirely within any one paradigm face particular difficulties in rationalising the criteria used in one model with that used in another. Positivists, for example, require evidence of reliability, internal validity and external validity (Kerlinger 1973). However, in developing their own standards, qualitative researchers have looked beyond the terminology of positivist criteria to the outcome served by these requirements. That is, they have examined the 'end' to which reliability and validity are the means, and concluded that this end is the demonstration of trustworthiness (Ely 1994; Kincheloe & McLaren 1994; Lincoln & Guba 1985; Maykut & Morehouse 1994). Trustworthiness, they argue, should be seen as the overriding criterion for all research methodologies, and reliability and validity the specific standards for achieving this when using quantitative measurement. The trustworthiness of interpretive research, however, is established through demonstrating dependability, credibility and transferability (Lincoln & Guba 1985).

Dependability

In common with reliability, dependability is established if there is a high probability that another trained researcher with the same tools and sample would generate the same data. Determination of dependability is contingent upon the number of variables involved in the data collection and the degree to which these can be recreated. In the naturalistic settings of interpretive research this is problematic (LeCompte & Preissle 1993; Kerlinger 1973),

although in the present study, the interview environments mitigated this problem to some extent. Although I 'entered the field', in as much as the interviews were carried out in schools, the time and settings were not truly naturalistic. In effect, interviewer and interviewee met 'half-way', in an environment neither totally familiar nor totally unfamiliar to either party, but one non-threatening to both. For the students, although the interviews occurred in the broader context of school, it was a new situation. The neutrality of such an arrangement meant that the social and environmental variables were reduced to a room, a tape recorder, some furniture and two people. The advantage of this arrangement over a more naturalistic one, say the home, or a school playground, was that most of the situational variables are reproducible in any confirmative study. Furthermore, the use of individual interviews avoided the contextual influences that group discussions in classrooms, peer groups or families, have on responses (LeCompte & Preissle 1993; Ogbu 1974).

An advantage of the semi-structured interview schedule was that, while it allowed for some flexibility and encouraged the flow of thought, it did have a structure ensuring that all interviewees were asked the same principal questions. The checks and probes not only allowed development of responses, but ensured a confirmation of meaning between interviewer and interviewee (Taylor & Bogdan 1998). Another trained researcher using the same schedule with the same students should elicit generally similar responses. Nevertheless, there would be differences, in cases where narratives took an unexpected direction and the schedule was put aside for a short period, or in interviews where one further probe unearthed a vault of possibilities. Ironically, it is these differences which are the *raison d'etre* of qualitative work and the source of its greatest strength - credibility.

Reliability of the quantitative data

The trustworthiness of the study was also enhanced by the reliability of the SPQ data. Kerlinger (1973, p. 480) noted that problems with reliability of questionnaires can be minimised by piloting to eliminate ambiguities and inadequate wording, a process undertaken prior to data collection. Furthermore, the SPQ was accompanied by an Instruction sheet (Appendix C) to be read out to students by their science teachers, thereby optimising consistency in administration across schools. In addition, the elementary and non-threatening nature of the questions lessened the chance of unreliable responses (Warwick & Lininger 1975), as did the basic biographical character of much of the SPQ.

Credibility

In addition to dependability, trustworthiness rests upon the pillar of credibility (Altheide & Johnson 1994; LeCompte & Preissle 1993; Lincoln & Guba 1985). The credibility of a study can be undermined by bias insinuating itself at any stage, but particularly during the interview process. Bias, through omission or commission, is found in all research, since no form of representation includes the whole (Eisner 1991). Nevertheless, bias in research can be minimised. This study took the view that awareness of sources of bias (Cohen & Manion 1994; Guba & Lincoln 1994; Kitwood 1977; Robson 1993) was the first step in reducing its occurrence.

Cohen and Manion (1994) summarised the most common sources of bias found in interviews. These include the influence of the attitudes and opinions of the interviewer, a tendency for the interviewer to see the respondent in his or her own image, or a tendency for the interviewer to seek answers that support preconceived notions. There may also be misunderstandings on the part of the respondent of what is being asked, or misperceptions on the part of the interviewer of what the respondent is saying. In order to reduce these forms of bias, efforts were made to phrase the questions in such a way that meanings were clear, and free from any insinuation particular responses might be expected or favoured. This was easier with the semi-structured interview schedule than would have been the case using a completely unstructured approach, since probes were used to confirm interpretations of questions, and the multiple perspective design of the schedule allowed a number of items to revisit important foci.

Other sources of bias could not be eliminated, and therefore need to be acknowledged. Despite the advantages of the interview over less dynamic research tools, it does include a confounding variable, the interviewer. Eisner (1991, p. 33) speaks positively of the 'interviewer as instrument', and Kvale (1996, p. 296) emphasises the interpersonal dynamics of 'Inter Views'. However, researchers must be aware that, as in quantum mechanics, the presence of the observer necessarily affects the outcome. This may manifest itself in a respondent's eagerness to please the interviewer, or perhaps as a vague antagonism which arises during the interview (Borg 1987). Such interpersonal vagarities may affect the responses, and there was a need to be sensitive to, and record any recognition of, interviewer effect.

Another potential source of bias was the effects of the interview settings on students and their responses. The possibility existed that a student's initial perception of the relationship between the interviewer and the science coordinator may have influenced the student's responses. It was important that students appreciate the impartiality of the interviewer with respect to the school and the science department. Furthermore, where schools provided an interview room in the administration building, for example, there may have been a tendency for the student to attach greater import to the interview, or to defer to the interviewer, taking their cues from what they interpreted as the esteem in which the research, or the researcher, was held by the school. In the one school where no interviewing facilities were provided and the interviews were conducted in the playground, perceptions may have been different. The degree to which the physical context of the interview influenced each respondent's attitude, manner and responses was difficult to gauge. Nevertheless, there was a need to be aware of the effects of the setting, and teacher-researcher interactions, on the respondents and to include such observations in the interview notes (LeCompte & Preissle 1993).

A number of authors (Cannell & Kahn 1968; Kvale 1996; Tuckman 1978) make mention in one form or another of a tendency for respondents to present aspects of themselves in a positive light. Consequently, it was important when asking questions to maintain, as much as possible, a neutral reaction to those responses with implications of social desirability. Furthermore, because many questions focused on school science, it was desirable to avoid giving the impression that there were negative implications to the student's decision to forgo

science, or that they or anyone influencing their decisions were in any way being judged. For example, positively worded comments and questions about students' decisions not to choose science subjects were used to allay any suspicion that deliberations about non-science subjects were somehow of lesser value or interest to the study.

The previous point raises the issue of the subjectivity of the data and its implications for credibility. It could be argued that a student's perceptions of the dynamics within their family world, for example, or of why they made a particular decision, may differ from those of their teachers or parents. It could further be argued that the data triangulation could have been extended to include parents' perspectives. However, the decision to focus primarily on the perceptions of the students themselves was made for three reasons. First, it was assumed (and later supported by the interview data) that the vast majority of students in this study made their own enrolment decisions, based upon their perceptions. Hence, the degree to which these perceptions varied from those of others within their different worlds was irrelevant to this study. As Levy et al. (1997, p. 48) argued in their investigation of students' perceptions, 'whether or not the students' perceptions were supported by external observers or the teachers themselves, they represent reality for the students.'

Second, students' perceptions have repeatedly been shown to be both reliable and valid (Levy et al. 1997), comparing well with those of experienced observers (Peck, Blattstein & Fox 1978; Stallings, Needels & Stayrook 1979). Indeed, Levy et al. (1997) make the point that students' perceptions of classroom activities are more often in agreement with those of observers than are teachers' perceptions. A third reason for focusing primarily on the perceptions of students was the need for research presenting their points of view. Despite all the literature on enrolment and career motivation, Stables (1996, pp. 133-134), for example, concludes that 'we do not yet have much on the students' perspectives on why they are following their chosen paths'.

Internal validity of the quantitative data

The internal validity of the SPQ data contributed to the overall credibility of the study. Internal validity was optimised by thorough piloting of the instrument, by constructing the scaled items with reference to a similar study (Warton 1997), and by standardising the data collection procedure. The large sample size (n=196) and the post-facto confirmation of students' interpretations of SPQ questions in subsequent interviews also enhanced internal validity. The specific problem of non-response, or volunteer-bias (Belson 1986) and its allied difficulties was avoided by having the students complete the questionnaires during class time, thereby maximising returns.

One potential threat to internal validity was the possible perception that the study was directed only towards science proficient students, and that those who did not see themselves as proficient, or those who were not intending to take further science courses, would not regard the study as relevant. To avoid this, science teachers were asked not to discuss the study with students in terms of science-proficiency, and to invite *all* continuing students to participate, regardless of their grades or subject choices. Thus, the classroom populations and settings were not modified for the administration of the SPQ, an important consideration in terms of

validity (Maykut & Morehouse 1994). This approach did, however, result in a large number of unusable returns, as noted previously.

Transferability

Transferability concerns the degree to which particular research findings are applicable to other contexts (Guba & Lincoln 1994; Kincheloe & McLaren 1994). Interpretations of what constitutes transferability within different paradigms vary significantly. In quantitative work, transferability is referred to as external validity (or generalisability) and is normally demonstrated by the representativeness of the sample (Kerlinger 1973). Consequently, the processes of probability sampling need to be rigorous. Even so, Kerlinger (1973, p. 324) makes the point that establishing external validity is not a matter of saying whether or not the findings can be generalised, but determining to whom and what they can be generalised.

The criterion-selection procedures used to select the schools in this study did not conform to the rigorous sampling protocols required in purely quantitative research, mainly because of the small number involved. The seven schools initially selected constituted slightly less than one per cent of the all high schools in NSW[1]. Despite their representativeness of the major school types, findings cannot therefore be generalized with any certainty, though Kvale (1996, p. 102) argues that the size of the sample does not necessarily preclude it from external validity either.

Nevertheless, positivist standards of generalisability were not a priority in this study. As Wiersma (1991, p. 6) noted, a research design which heightens internal validity may decrease external validity, and the credibility of this study was paramount. Interpretative notions of transferability were more applicable to this study, which was designed for depth and the interpretation of meaning. In interpretive studies the onus of transferability shifts away from the researcher to the reader (Kvale 1996). As Lincoln and Guba (1985, p. 316) put it:

> It is not the task [of the researcher] to provide an *index* of transferability; it is his or her responsibility to provide the data base which makes transferability judgments possible on the part of potential appliers.

Suggestions on enhancing transferability differ greatly within the qualitative literature, a feature symptomatic of the existing state of conceptual and semantic flux within the paradigm. Nevertheless, claims for transferability of this study rest upon commonly accepted guidelines, including the lucidity of its frameworks (Robson 1993), the thoroughness of the design (LeCompte & Preissle 1993) and the credibility of the sampling and data collection techniques (Kincheloe & McLaren 1994).

Trustworthiness through Triangulation

The multi-method approach taken by this study involved both data triangulation and methodological triangulation. Data triangulation uses two or more data sets to provide a more sophisticated representation of some phenomenon, thereby enhancing the credibility of the

[1] Based upon 1998 information supplied by the NSW Department of Education and Training, the Association of Independent Schools and the Catholic Education Commission

research (Guba & Lincoln 1994). Such a method does not necessarily seek to confirm or refute an impression gained from one set of data, but can be used to provide different perspectives, in which a divergence of perceptions is seen to be as valuable as convergence. In the present study, the STS and interview responses provided data triangulation with regard to students' motivations. The purpose was not to validate either perspective, but to appreciate similarities and differences in the perceptions of teachers and students.

Methodological triangulation is often used to help overcome the inherent weaknesses of single measurement instruments (Denzin 1988). In the present study, a triangulation of strategies was used to enhance the credibility and transferability (Guba & Lincoln 1994) of the SPQ scaled items regarding students' reliance on the advice of others. During the interviews, students were again asked about the significant others upon whom they had relied for subject advice. Their responses were compared with their SPQ ratings, thereby cross-checking the SPQ data and exploring in more depth what students had understood by the term 'relying upon the advice' of others. This in turn increased the construct validity of these questionnaire items.

SUMMARY

The description of the methodology used in this study has been extensive, particularly with regard to issues of research integrity. As noted in the introduction, such issues underpin the trustworthiness of any study, but particularly one using complementary strategies from different research paradigms.

The conceptual framework, which differentiated between the topographical features of science subject choice and the sociocultural influences on students' decisions, is reflected in the structure of the next four chapters. Chapter Four reports and discusses the SPQ findings concerning enrolment patterns and influential sources of advice. The opinions of science teachers on students' enrolment motivations are also presented and compared with the explanations provided by interview participants for their science enrolment decisions.

Chapters Five, Six and Seven present and discuss the findings relating to influences within students' multiple worlds. Each chapter begins with results and implications from the SPQ data, before turning to the interview narratives. These chapters provide insights into the sociocultural domain of science proficient students, and the influences within their multiple worlds on deliberations about enrolling in senior science courses.