CHAPTER 4

THE RESEARCH: OPERATIONALISATION, ADMINISTRATION

AND DATA ANALYSIS

4.1 INTRODUCTION

In Chapter 1 the chronology of the research and data collection procedures were given in outline. Briefly, these entailed: (i) the administering of questionnaires to groups of student teachers, lecturers, and practising teachers in 1976, and (ii) in 1978, sending out questionnaires to ex-students in the schools and administering questionnaires to student teachers in two colleges. Below, details of the operationalization and administration of the research as they relate to the research objectives enumerated in Chapter 2 are given, along with information on the respondent groups and an account of the statistical procedures employed.

4.2 THE RESPONDENT GROUPS

4.2.1 SUBJECTS USED IN THE MITCHELL STUDY

The subjects used in the 1976 replication were two cohorts of student teachers in one institution, their lecturers, and supervising teachers in the college's co-operating schools. The college was chosen on the basis of convenience to the investigator who worked in the teacher education programme as a lecturer and was able to gain access to the required groups. Limitations in the resources available to the investigator (e.g. the study was personally financed), the fact of replication, and the improbability of and logistical difficulties associated with gaining access to students in other institutions precluded any other course. Details of the groups are given below:

1. The student teachers:

a) <u>The 2nd semester group</u> was chosen because these students were close to the beginning of training and, in addition, had been in college long enough (one semester) to be able to ascribe views to the lecturers in the teacher education course. As well, they had not yet been on practice teaching though there had been some contact as a student teacher with schools through demonstration lessons and some involvement as teacher-aides. All of the cohort participated excluding those who for any reason (e.g. repeating a semester) had not entered the programme the previous semester. The group thus obtained numbered 98 subjects. All role norm inventories were returned virtually fully completed by this group.

b) <u>The 6th semester group</u>, being in the last half (semester) of their 3rd and final year were close to the end of their training and had completed all practice teaching. Again, those who had not started with the cohort were excluded leaving a final group of 97 subjects. Again, all questionnaire material and role norm inventories were returned and usable with minimal missing information.

2. <u>The lecturer group</u> was composed of all lecturers whose major involvement was with the teacher education course. Excluded thus were those lecturers who were wholly or primarily concerned with the college's two other teaching programmes, even though some such lecturers might occasionally be involved in the teacher education programme in some capacity such as, for instance, a guest lecturer. The number of lecturers thus obtained as subjects was 45. There was a hundred percent return of role norm inventories and all were fully completed. 3. The teacher group: the aim here was to obtain responses from teachers in primary schools used by the college for practice teaching and allied work. The target population was that group of teachers who in any such capacity had supervised or substantially helped supervise a student teacher during the period when the student teacher respondents had been in the schools for practice teaching, teacher-aide work, or observation. That is, the population in mind was that group of primary teachers with whom the students had predominantly interacted over the course of their training. Given the virtual certainty that during this period some teachers would have been transferred from or otherwise moved out of schools that were in the college's 'catchment' area, it was not possible to determine precisely how large this population was. On the basis of student placement in schools an estimate was made by the Schools Liaison Office of the college that 800 questionnaires Of these, 98 were returned unused. should be sent out. Of the 702 questionnaires taken by the schools, 655 were subsequently returned. Two of these could not be used, leaving a final 'sample' Overwhelmingly, the teacher returns of 653 primary teachers. were fully completed.

It should be noted here that the 2nd semester cohort of 1976 numbering 98 <u>became</u> the 6th semester cohort of 77 subjects in 1978. The 6th semester cohort of 1976 were tracked into the school system becoming a group of 91 teacher respondents in 1978.

4.2.2 SUBJECTS USED IN THE CASTLE HILL STUDY

Role Norm Inventory data from two Castle Hill College of Education respondent groups were obtained in 1978. The two groups were the entire 2nd and 6th semester cohorts of that year. That is, identical groups to those in the Mitchell study were used. The 2nd semester cohort numbered 74 while the 6th semester group consisted of 50 respondents. All inventories were returned with virtually no missing data from both groups.

4.3 INSTRUMENTS USED IN THE STUDY

Following is a list of the instruments used in the study together with details of when they were used and with whom. Each instrument will then be discussed more fully:

1. The Teacher Training Project questionnaire (T.T.P.Q.) containing the Foskett Role Norm Inventory, the Sieber and Wilder Teaching Style Inventory, questions concerning role models, commitment to teaching and opinions about college life (including courses), and other background information variables, was the major question-It was completed in its entirety by the 1976 6th naire used. The 2nd semester group of 1976 completed only semester group. the Role Norm Inventory when in 2nd semester and the whole questionaire in 1978 when in 6th semester. Though for the purposes of this study the R.N.I. data only were required from this group when in 6th semester, the remainder of the questionnaire was completed with a view to ongoing research.

Lecturer and teacher groups completed only the first section of the Role Norm Inventory in 1976 - the section entitled <u>Role Norm</u> <u>Inventory One</u> concerning respondents' <u>own norms</u> for the position of primary teacher. Second and sixth semester students from Castle Hill College of Education completed the whole questionnaire in 1978 - again with a view to ongoing research since the role norm inventory data only were needed for the present study.

2. Also in 1978 the students of the 1976 6th semester cohort who had been in the schools for two years were sent a <u>follow-</u> <u>up questionnaire</u> which attempted to assess aspects of students' adjustment to teaching. It contained questions concerning satisfaction with teaching, expressed commitment to teaching, and so on. This group was also sent the first of the Role Norm Inventories (R.N.I.1.) in order to estimate whether two years in teaching had effected any shift in students' <u>own norms</u> for the primary teacher role.

Before proceeding it is important to stress that while it was thought necessary to gather all of the data yielded by these instruments, it was always intended that only those aspects of the data directly relevant to answering the questions posed at the conclusion of the literature review would be used. Thus it is, for example, that numerous intrinsically interesting comparisons could have been made between responses given in the teacher training Project Questionnaire and those given in the follow-up questionnaire, but if these were not directly related to the research objectives of the study, they were omitted from consideration in the analyses that follow. It should also be added that some information was gathered at the same time for internal college purposes only, as part of an ongoing evaluation of the College's Teacher Education

Programme. The instruments are, of course, reproduced in their entirety in Appendix I with care taken to make clear those questions solely or primarily used for purposes other than the concerns of the present study.

Since this study is part of wider, ongoing research it should also be added that other procedures not reported here in any detail were used to complement the data gathered by questionnaire so that a more comprehensive picture might be built up of the various aspects of occupational socialisation being probed. Thus it is, for example, that a random sample of the 1976 cohort of 6th semester students were interviewed when they returned to college for graduation three months after entering teaching; the principal concern was to explore the initial impact of teaching amongst these students in the light of initial findings from the role Also, students were encouraged to write norm inventory data. in about what they saw as relevant teaching experiences. While it was not intended that this information would be systematically analysed for reporting in this study, it was thought that such data might occasionally be used to add further substance to the discussion of the results recorded herein. The transcript of part of the interview reported in Chapter 3 with the female student who resigned soon after entering teaching is one example of how such material has been used. For the most part, however, such data were gathered primarily for internal college purposes. Table 4.1 summarises the data collected for the investigation reported Details of other data collected for purposes of ongoing here. research are, in the interests of clarity, excluded from Table 4.1.

TABLE 4.1

SUMMARY OF DATA COLLECTION

			IN	STRUMENTS	
RESPONDENT GROUPS	NO. IN GROUP	ROLE NORM INVENTORIES 1 2 3 4	TEACHING STYLE INVENTORY	T.T.P.Q: BACKGROUND INFORMATION (etc.)	FOLLOW-UP SURVEY
(6th (Mitchell: 1976)	67	* * *	*	*	
<pre>Gth (Mitchell: 1976 - when in schools in 1978)</pre>	91	*			*
(2nd (Mitchell: 1976)	98	* * *			
, 6th (Mitchell: 1978)	77	* * *			
Lecturers (1976)	45	*			
Teachers (1976)	653	*			
6th (Castle Hill: 1978)	50	* * *			
2nd (Castle Hill: 1978)	74	* * * *			

Respondent groups bracketed are the same groups at different points in time. 1. N.B.

An asterisk signifies completion of the particular instrument designated. 2.

4.3.1 THE ROLE NORM INVENTORY

The major instrument used in this study was a role norm inventory for the position of primary school teacher developed and first used by Professor John Foskett of the Centre for the Advanced Study of Educational Administration at the University of Oregon, in the United States (Foskett, 1967a). The inventory consists of 45 role norm items divided into 4 different role sectors dealing with different teacher role relationships as follows:

ROLE SECTOR 1: ACTING TOWARD PUPILS (15 role norms) ROLE SECTOR 2: ACTING TOWARD COLLEAGUES (10 role norms) ROLE SECTOR 3: ACTING TOWARD PARENTS (10 role norms) ROLE SECTOR 4: ACTING TOWARD COMMUNITY (10 role norms)

Before discussing details of the construction of the instrument and its subsequent use by Foskett (1967a, 1969), McNamara (1972), and in the present study, the critical issues of validity and reliability need to be raised insofar as they relate to the use of the inventory both here and in previous studies.

Firstly, perhaps, it should be said that the Foskett inventory makes no claims to be an attitude scale. It is not meaningful, for example, to sum a respondent's 'score' on the inventory. Nor does the inventory make claims to unidimensionality for instance. As mentioned in Chapter 2 (p. 71), while an attitude scale typically measures some hypothetical construct which provides a conceptual 'bridge' between a persisting psychological state and an orientation to the social world (cf Newcombe, 1964: 41) a role norm inventory such as the Foskett schedule attempts to assess somewhat more directly an individual's propensity for behaviour in a specified social setting.

As was also mentioned (p. 71), the construction and validation of attitude scales have generally been characterised by Commenting upon this in the context of his own use rigour. of the Foskett inventory, McNamara (1972: 44) made the point that it was 'scarcely possible for a more behaviouristically orientated study (such as his) to approach the same standards of operationalisation and statistical analysis of the better attitudinal studies'. In the case of the Foskett inventory for example, the obtaining of the kinds of standard reliability estimates normally given for attitude scales was either logistically not possible in both the McNamara study and in this present study, or not approp-It was not possible, for instance, to obtain test-retest riate. estimates or to use a parallel form of the inventory, and it was not meaningful with such an inventory to obtain split-half reliability estimates.

Concerning these crucial matters of validity and reliability, Anastasi (1968: 481-2) has noted that with both attitude measurement and opinion surveys there are practical difficulties in obtaining validity criterial data such as ratings by close acquaintances, and comparisons with membership in contrasted groups. She observes, too, that data on reliability are meagre for opinion surveys. Anastasi (1968: 482) comments that the problems encountered are basically the same as those met in the construction and administration of other types of psychological tests. She

lists the major difficulties as centring upon:

- (i) The proper formulation of questions to avoid ambiguity, suggestion and other error
- (ii) adquate sampling of the population in respect of size and representativeness; and
- (iii) control of the conditions under which the survey is conducted, such as ensuring anonymity of replies and reducing the influence of interviewer characteristics.

In respect of this last requirement, details are given later in this chapter (see section 4.4) of the administration of the research. These will demonstrate that control of conditions under which respondents completed the instruments was rigorous in attempting to reduce the influence of interviewer characteristics, to forge good rapport (cf Oppenheim, 1976: 72), to assure participants of anonymity, and the like.

Concerning points (i) and (ii) above, the careful procedures followed by Foskett in developing, piloting and testing the instrument are given immediately below. Also outlined are relevant other procedures followed by McNamara (1972) and the author in their subsequent uses of the Foskett inventory. These procedures indicate the efforts made in this and previous studies to accommodate the kinds of difficulties listed by Anastasi. As detailed by Foskett (1967a: 12) specific criteria guided the selection of the role norms. Excluded were:

- (i) Statements concerning <u>attributes</u> of occupants of the position;
- (ii) statements pertaining to the <u>functions</u> or <u>goals</u> of the position;
- (iii) statements so <u>broad</u> or <u>general</u> that a specific form of behaviour could not be identified;
- (iv) statements that were vague or ambiguous; and
- (v) statements that were so <u>technical</u> or <u>'professional'</u> that they would be meaningless to the lay populations used in the Foskett studies.

In formulating the role norm statements an effort was made to state all role norms in terms of specific and explicit rules of behaviour. As described by Foskett (1967a, 1967b)¹ the first step in the development of the inventory was to compile an extensive list of role norms from a variety of sources - the literature and previous studies, suggestions by teachers outside of the community in which the research was conducted, and suggestions by the research team. The resulting list of 'several hundred role norm statements' (Foskett, 1967a: 12) was cut to 55 by applying the above-mentioned criteria of exclusion. There followed a series of field tests in which a number of changes were made in wording and a further 10 items deleted leaving a final

^{1.} The two studies are reported in companion monographs. They are <u>The Normative World of the Elementary School Teacher</u> (Foskett, 1967a) and <u>The Normative World of the Elementary</u> <u>School Principal</u> (Foskett, 1967b).

inventory of 45 items. According to Foskett (1967b: 13), items were eliminated in the field testing phase 'on the basis of ambiguity, lack of familiarity by lay subjects with the behaviour in question, and difficulty in securing reliable responses.' ². Foskett warns however that given even these procedures - procedures in which some effort was made to select norms that would be representative in the judgement of experienced practitioners and investigators - there was no way of knowing, finally, just how representative the 45 selected norms were of the total universe of norms associated with the position of primary teacher (cf. Foskett, 1967b: 13, 18-19).

Following the construction, field-testing, and subsequent modification of the instrument, Foskett used it first with a Pacific Coast Community in the United States (Foskett, 1967a). The focal population for the study was all of the elementary school teachers in the 34 schools of the district.

He obtained a 92 per cent completion rate of usable schedules. Additionally, three lay populations were included in the study. These were selected by means of area probability samples which yielded an 81 per cent rate of usable schedules. Interviewing of teachers was carried out by research team members by means of group interviews, while a public research firm conducted face-to-face interviews of members of the three lay samples.

^{2.} In an endeavour to obtain more detailed information about these procedures, the author wrote twice (in 1976 and 1977) to Professor Foskett but received no reply. A colleague of the author who visited Oregon University subsequently reported back that Professor Foskett had died some time prior to the letters being written; this colleague could find nothing more than is reported in Foskett's monographs.

Later, the study was replicated in two additional and differing communities (Foskett, 1969). <u>Prior to its subsequent use by</u> <u>McNamara in the United Kingdom (McNamara, 1972) and, in Australia</u> by the present investigator, the inventory had thus been carefully <u>compiled</u>, scrupulously piloted and field tested, and used in <u>a series of empirical studies</u>.

In the present study a version of the Foskett inventory anglicised by McNamara (1972) in a study of the socialization of female student teachers in the United Kingdom was trialled initially with 84 5th semester (1976) students who were asked to recommend changes in terminology appropriate to their milieu as they completed the entire inventory. In anglicising the American version of the instrument, McNamara (1972: 79) had asked thirty experienced teachers and college of education lecturers taking post-graduate degrees to make similar recommendations. The result of this process was a version of the inventory which was clearly comprehensible to the 84 students asked to trial it, with only minimal changes in terminology needed in McNamara's English version (for example, 'public house' was replaced by 'pub').

In using the Foskett inventory to assess socialization processes in a cross-sectional study of student teachers, McNamara found it necessary to use a rating system and factor analysis to facilitate examination and description of regular patterns within the data he obtained. Concerning his use of factor analysis (a principal components analysis of each role sector) McNamara 1972: 107-108) explained:

It must be emphasised that the purpose of the factor analysis is to provide an organising framework for the discussion of the data. The forty five items of the R.N.I. are not based upon some theoretical model of the teacher's role. It is therefore necessary, in order (a) to reduce the idiosyncracies of measurement error for each item, and (b) to describe the data more parsimoniously, to develop some framework for the analysis of the data. This could, of course, be done ex post facto on the basis of the researcher's intuition. However. it is preferable to use a more 'objective' procedure which removes the decision making process from the individual researcher. Therefore, a factor analytical model was employed ... Nevertheless it must be remembered that factors produced using a mathematical factorial model must have some substantive validity; they must make 'sense' within the context of the process of socialization within the college of education. Consequently, factors which appear to be arithmetical artifacts and which do not have substantive validity are precluded from further discussion.

McNamara classified the responses of each year group in his study by role sector since, conceptually, this was how the inventory was organised and presented, and subjected them to principal Factors were rotated according to the Varicomponents analysis. max procedure developed by Kaiser (1959). The number of factors extracted was determined by Kaiser's criterion. That is, only factors having latent roots greater than one were considered as common factors. This method, according to Child (1970: 43-4), is particularly suitable for principal components designs, the tendency being to extract a conservative number of factors where the number of variables analysed is less than 20. The stringent criterion McNamara used for deciding the significance of factor loadings was to consider only those factors having loadings of at least 0.30 (cf. Child, 1970: 45-6). In choosing to do

so, McNamara (1972: 107) argued that since methods employed to determine the statistical significance of factor loadings must be treated with caution, it was necessary to err on the side of conservatism, particularly when attempting to ascribe 'psychological' or 'sociological' sense to a factor. Further, where an item in the inventory showed a significant loading in more than one factor, it was included only in the description of the factor for which it had the highest loading. This was done to avoid according undue significance in the analysis to particular items (McNamara, 1972: 107).

To help in naming the factors that were yielded by the principal components analysis, and to facilitate examination of data patterns, McNamara firstly asked a group of twenty 'experienced educationists' to rate the forty five items of the role norm inventory according to whether a positive ('definitely should') response on each item would reflect the view of a 'traditional' or 'progressive' teacher. A 'don't know' category was also provided so as not to force a choice. While acknowledging the relative crudity of such a simplistic stereotyping of teachers, McNamara (1972: 103) cited numerous empirical studies (e.g. Barker Lunn, 1970; Entwistle <u>et al.</u>, 1971) to support such an approach.

The rating procedure gave clear 'traditional' ratings for items 1, 2, 3, 4, 6, 7, 12, 18, 27, 36 and 41 of the inventory, and clear 'progressive' labels for items 5, 8, 9, 11, 13, 15, 20, 28, 32, 33 and 34. The author noted that the distinction between 'traditional' and 'progressive' was most marked for items making up the role sector Acting Toward Children (McNamara, 1972: 106).

In summary, following the careful compilation, piloting, field-testing and use of the inventory in a series of studies by Foskett, the instrument was modified for use in the United Kingdom by McNamara who employed the ratings of experienced educators to assist in naming, and give some validation to the factors derived from a principal components analysis. In the present study and in the author's previous study (Sinclair, 1975) the Foskett inventory was again thoroughly trialled and minor refinements made to accommodate differences in the Australan educational setting. The rating procedure and factor analysis used by McNamara were repeated to provide a more elegant means of describing data patterns. Both of these measures produced very similar results to those obtained by McNamara thus enabling the descriptive framework he used to be retained in this study, with some modifications. Details of the factor analysis are given in section 4.5.7. Details of the rating procedure are given in Appendix 15 which presents both McNamara's results and those obtained in this investigation.

In accord with the research objectives listed at the conclusion of Chapter 2 each student teacher subject completed <u>four</u> copies of the inventory, while lecturer and teacher respondents completed one copy. This was as follows:

1. All respondents completed Role Norm Inventory One (R.N.I.1). The inventory attempted to ascertain how respondents thought a primary teacher <u>ought</u> to behave. That is it tried to find out what were respondents' <u>own norms</u> for the primary teacher role and was introduced thus:

1. continued

We are trying to find out how you think a primary teacher <u>ought</u> to behave. That is, given good working conditions and a sympathetic head-teacher and colleagues, how do you think that a teacher should behave?

There were five response categories for each role norm statement. They were:

- (1) definitely should (or 'will' for Inventory 2)
- (2) preferably should (will)
- (3) may or may not
- (4) preferably should not (will not)
- (5) definitely should not (will not).

Correspondingly, responses were weighted on a Likert-type 1 to 5 scale.

- Student teacher groups completed the same inventory from
 3 other standpoints. These were:
 - (1) <u>Role Norm Inventory Two (R.N.I.2)</u>, which asked students how they thought they <u>would</u> in fact behave when they began teaching, bearing in mind their experiences to date and what they knew about teaching. That is, R.N.I.2 attempted to discover what students' <u>own expectations</u> (or anticipations) for the primary teacher role were.

- 2. continued
 - (2) <u>Role Norm Inventory Three (R.N.I.3)</u>, which asked students how they thought their college lecturers would expect them to behave in the classroom. That is, this inventory aimed to find out the <u>norms attributed</u> to lecturers by students. What was meant by 'lecturers' was made clear in the rubric accompanying the inventory: 'Clearly different tutors would express different points of view, but insofar as it is possible, record your general impression from the majority of lecturers you come across.'
 - (3) <u>Role Norm Inventory Four (R.N.I.4)</u>, which asked respondents how, generally, they thought established members of the <u>teaching profession</u> would expect their colleagues (that is, practising primary teachers) to behave. It was made clear as to what was meant by 'teachers' in the statement prefacing the inventory: "Again, different teachers will have differing points of view, but drawing on your experiences in schools and with teachers, what would your general impression be of the majority?"

At this point some comment is necessary about the fact that these respondents were given four formats of the Foskett inventory in the order described above. It is possible that such a procedure is likely to generate interaction effects amongst formats. If the <u>primary</u> intent of the research had been to provide a relatively comprehensive description of students' perceptions of how lecturers and teachers viewed the teacher role the problem of possible interaction effects would have made necessary some modifications in the administration of the inventories (for example, the completion of each schedule at different times, by each student, or in different sequences by different students). However, it needs to be stressed that the principal focus of the research was to ascertain the relative influence of lecturers and teachers on student teachers' developing As stressed in the review of literature role conceptions. impinging on the much-discussed practice/theory 'gap' in teaching, the prime emphasis in this study was on exploring aspects of the process of identification of student teachers with the college (through its lecturers) and/or the school (through practising teachers): whether and how students changed over training in respect of whom they saw as influenc-The point of having the student teacher responding them. ents do the four versions of the inventory together was that they would indeed make direct comparisons between themselves and their lecturers, themselves and teachers, and between lecturers and teachers.* It was considered that a single session wherein each respondent completed the four inventories would best facilitate such comparisons, the more especially so in view of the logistical difficulties of getting large numbers of respondents back individually to complete four inventories on separate occasions. To eliminate the possibility of order effects as a source of difference between respondent groups, all groups completing the four inventories did so in the same order.

^{*} It perhaps should be noted that both Foskett (1967a, 1967b, 1969) and McNamara (1972) used the instrument in this way, i.e., to ensure respondents made direct comparisons.

As well as permitting an overall comparison between the original study (Sinclair, 1975) and the present investigation, the gathering of these data enabled numerous of the research objectives to be met:

1. Responses to Role Norm Inventory One provided the basis for determining (a) the prevailing climate of opinion about important teacher role relationships amongst student teachers and their significant others, (b) differences between student groups near the beginning and those near the completion of training, (c) differences between students and lecturers, (d) differences between students and teachers, (e) differences between lecturers and teachers, and (f) degrees of within- and between-group consensus and conflict about teacher role relationships.

2. Responses to Role Norm Inventory Two and comparisons between these and the responses to R.N.I.1. permitted determination of (a) students' expectations for their future behaviour as teachers, (b) differences in expectations between the student groups, (c) levels of idealism in these groups, and (d) consensus on role expectations, and conflict between norms and expectations.

3. Responses to Role Norm Inventory Three and comparisons between these and the responses to R.N.I.1. gave measures of (a) differences amongst and between the student groups in their perceptions of lecturers, (b) the relative closeness in viewpoint of the students to what they perceived to be the views of lecturers - that is, the 'impress' of lecturers, (c) the accuracy of students' perceptions of lecturers and (d) consensus on views attributed to lecturers and conflict between students' norms and these attributed views.

4. Reponses to Role Norm Inventory Four and comparisons between these and R.N.I.'s 1 and 3 gave the means of ascertaining (a) differences amongst and between student groups in their perceptions of teachers, (b) the relative closeness of the student groups to what they saw were the views of teachers, (c) the accuracy of these perceptions, (d) students' views of lecturers vis à vis teachers, (e) the accuracy of students' views of lecturers vis à vis teachers, (f) differences between lecturers' and teachers' actual views, (g) consensus on views attributed to teachers, and (h) conflict between students' views, their attributed views and the actual views of significant others.

4.3.2 THE TEACHING STYLE INVENTORY

In an endeavour to elucidate further the perceived 'influence', if any, of significant others on students' role conceptions and behaviour, use was made of a relatively short instrument, the Teaching Style Inventory, developed by Sieber and Wilder (1967) in the United States. In developing the inventory the authors argued that the role behaviour of teachers can be seen as reflecting instrumental processes, that is, those processes concerning the daily behaviour of teachers and pupils, rather than the terminal goals or eventual outputs of education (Sieber and Wilder, 1967: 303). Such day-to-day teacher-pupil interaction involving

different preferences for the various instructional practices which had been the subject of considerable community debate in the post-Sputnik era, was an important area of potential conflict between parents and educators according to Sieber and Wilder (1967: 303). Since the inventory concentrated upon the very sector of the teacher role - that of teacher-pupil interaction - which, in this investigator's original study had proved to be potentially the most conflict-inducing, and since the debate over 'progressive' and 'traditional' teaching styles has intensified in the 1970's³, the Sieber and Wilder instrument appeared to be particularly appropriate for this additional, exploratory phase of the research.

The Teaching Style Inventory operationalises a more general orientation to the teaching role than does the Foskett role-norm schedule. When they developed the instrument Sieber and Wilder pointed out that the difficulty of measuring classroom behaviour and the consequent dearth of empirical evidence on what teachers actually do in classrooms had resulted in the deriving of ideal constructs concerning teaching styles 'from "philosophies" of teaching, from controversies over progressive versus traditional education, and from the everyday discourse of practitioners and parents' (1967: 304). They detailed the development of their own instrument as follows:

Two especially important aspects of the teaching role are widely discussed in the literature: (1) the extent to which subject matter is emphasised, and (2) the extent to which adult authority is exercised. By dichotomizing and combining these two dimensions, we obtain four distinct styles of teaching:

^{3.} See, for example, Bennett, N. <u>Teaching Styles and Pupil</u> <u>Progress</u>, Open Books, London, 1976.

Emphasis on Subject Matter

		High	Low
Relations Between Teacher	Adult-centred (Authoritarian)	CONTROL ORIENTED	CONTENT ORIENTED
and Child	Child-centred (Permissive)	DISCOVERY- ORIENTED	SYMPATHY- ORIENTED

The four styles singled out for study are not exhaustive of the popular conceptions of teaching and are not wholly accurate reflections of behaviour patterns, but they do represent some of the most common images that are held of teaching at the elementary and secondary levels.

(Sieber and Wilder, 1967: 304-5)

In using the inventory, the version anglicised by McNamara (1972) in his study of the occupational socialization of female student teachers was preferred once again, with only one further minor terminological modification made, appropriate to the Australian context: in the questions asked of respondents, 'Her Majesty's Inspector or County Organiser' was replaced by 'Inspector'.

Once more, a number of lecturers and teachers, and the previously mentioned group of 84 5th semester (1976) student teachers not involved in the research were asked to comment on the presentation and wording of the inventory before it was finally used with the 6th semester group. After the above-mentioned change no problems for respondents were envisaged by any of the group asked to vet the presentation of the inventory and accompanying questions.

The four teaching styles corresponding to the four cells in the Sieber and Wilder typology were presented as suggested by the authors under the following rubric:

Although teachers have to concern themselves with many different things in their jobs, some teachers emphasise certain things more than others. Below are descriptions of four types of teachers who emphasise different things: Then followed a description of the four teaching styles:

- Teacher is most concerned with maintaining discipline, seeing that pupils work hard, and teaching them to follow directions.
- Teacher feels that he (she) should cover the syllabus thoroughly, that pupils should know their subject matter well and that their progress should be tested regularly.
- Teacher stresses making the class interesting and encourages pupils to be creative and to think things out for themselves.
- 4. Teacher thinks that it is important that he (she) should be friendly and well liked by pupils and be able to understand and help with their problems.

The descriptions corresponded with, respectively, the control-, content-, discovery-, and sympathy-oriented styles that constitute the typology. Following McNamara (1972) 9 questions were then asked of respondents as presented below:

Remembering that these are not complete descriptions of a type of teacher but rather where a teacher places most emphasis, please answer by placing the appropriate number in the box:

- A. Which would you emphasise given free choice?
- B. Which did you emphasise on your last school practice?
- C. Which do you think your last school practice college supervisor wished you to emphasise?
- D. Which do you think your class teacher on your last school practice wished you to emphasise?

- E. Which do you think that you will emphasise when you begin teaching?
- F. If, when you begin teaching, an authority such as an Inspector came into your classroom, which would you emphasise?
- G. Looking back, think of the teacher who, in your opinion, was your best primary school teacher. Which would he or she emphasise?
- H. And which would your best secondary school teacher emphasise?
- I. And which would your best college lecturer emphasise?

It was hoped that answers to these questions might shed some light upon any modification of role expectations and behaviour during the final practice teaching session before the students entered the profession, upon the likely effects on role behaviour of the sorts of authority figures that loom large in a neophyte teacher's probationary years, and upon the possible influence of teacher and lecturer role models on students' preferred teaching styles. That is, this information would enable other research objectives – as specified at the conclusion of chapter 2 – to be met.

4.3.3 THE BACKGROUND INFORMATION SECTION OF THE TEACHER TRAINING PROJECT QUESTIONNAIRE

The background information section of the questionnaire was designed to yield details of respondents' sex, year group, residence, expressed and 'felt' commitment to teaching, views on the perceived influence of potential role models, degree of integration with campus life, and views on aspects of college training. It was intended that some of this information, such as that concerning integration with campus life and views on training, would provide useful data on aspects of the college context. Details of these questions are therefore given in Chapter 5 describing the institutional setting. Apart from the obvious questions about sex, year group and so on, the remaining items dealt with commitment to teaching. These are discussed below.

4.3.3.1 COMMITMENT TO TEACHING: OPERATIONALISATION

Given the conceptual inadequacies of uni- or two-dimensional definitions of commitment mentioned in the review of literature, a number of measures of commitment were employed in an effort to discriminate between students who were clearly more highly committed to teaching and clearly less committed than others. These measures were drawn from the work of Mason (1961) in his study of the beginning teacher for the United States Office of Education, Walker (1962, 1967) in his research on the occupational plans of finalyear student teachers and beginning teachers in Australia, and McNamara (1972). Once more, the operationalisation process involved scrutiny by the previously mentioned 84 5th semester student teachers to clear up ambiguities and recommend changes in wording. Again, however, such changes were minimal and of a minor nature. The relevant questions are presented below, as operationalised, with brief explanatory remarks:

The first question was one of three devised to tap what McNamara (1972: 63) in a searching analysis of the concept of commitment as applied to teaching, calls the 'temporal dimension' of commitment, deriving from Dreeben's (1970: 180) notion of consistency of behaviour over a period of time. This temporal dimension, by definition, reflects past, present and future aspects of career choice. Thus past commitment was operationalised in the question below by reference to that period in time when a <u>firm</u> career decision was taken:

- * At what stage in your school career did you make a FIRM decision to become a teacher?
 - (1) While in Primary School?
 - (2) In Secondary School before Fourth Form?
 - (3) In 6th Form but before the exams?
 - (4) After matriculating/receiving good results?
 - (5) After the offer of a Teachers College scholarship.
 - (6) Other (specify)

Similarly, <u>present</u> commitment and <u>future</u> commitment levels were intended to be elicited by the two questions immediately below. The first of these was devised to reflect a retrospective degree of expressed satisfaction with the choice of teaching as a career at the end of the students' college training, while the second asked about their future career plans -

- * If you had to make the choice again would you:
 - (1) Certainly go into teaching.
 - (2) Probably go into teaching.
 - (3) (be) Uncertain.
 - (4) Probably not go into teaching.
 - (5) Certainly not go into teaching.

- * When you think about your future career which of the statements below best expresses your future plans:
 - (1) I expect to stay in schools until retirement.
 - (2) I expect to continue in the field of education until retirement but hope to move eventually from school teaching to another area of education.
 - (3) I expect to leave teaching in order to have a family but would hope to return to teaching later.
 - (4) I expect to leave teaching in order to have a family and will not return to teaching.
 - (5) I expect either never to teach or to leave teaching for some other occupation.

The next question was as follows:

- * (a) Did you at any time make an application to go
 to University but were unable to obtain a place?
 (1) Yes
 (2) No
 - (b) If you obtained a place at University but turned it down in favour of College put a (3) in the box.

It was hoped that in making a decision about the overall level of commitment characterising the responses of a given student, answers to this question might assist by revealing those students who were enthusiastic enough about the prospect of teaching to prefer a College place to the more prestigious University scholarship, and those by contrast for whom a teachers' scholarship at College was secondary to a University place. Next, the students were asked to rate themselves by placing a cross on a seven point scale in terms of how closely they thought they fitted the description at each end:

* A person who sees teaching - - - - A person who is dedias just a job, a means of providing money.
* A person who is dedicated to teaching and who thinks that teaching will be one of the most important things in life.

Finally, each student was briefly interviewed by the investigator about the reasons for choosing to become a teacher, any special factors which influenced this choice and the level of commitment to teaching felt by each respondent. At the conclusion of this loosely structured session which lasted about five minutes on average, the student was asked to capture the essence of the discussion on paper:

* In the space below could you outline in a sentence or two the reasons why you decided to become a teacher. What were the particular factors which influenced you in your decision. Could you also give some indication of the degree to which you feel committed to teaching as a career.

While this was being done the interviewer made notes about the discussion. These notes and the students' written comments provided the basis on which students were rated as high, moderate or low according to the degree to which they <u>felt</u> committed to teaching.

As well as being relatively convenient for the purposes of data collection, it was intended that these questions would provide a reasonable basis for determining differing degrees of commitment to teaching amongst the students of the 1976 cohort. No attempt was made to combine the various commitment criteria in any mathematically precise fashion to arrive at an overall commitment rating for each student. Rather, judgement had to be exercised by the investigator, but since the objective was to do no more than ascertain which student sub-groupings were clearly more highly committed or clearly less committed than the others, this was considered a defensible procedure. Where there was doubt about whether a student should be assigned to a high or low commitment category that student was usually given a 'moderate' overall The process of allocating subjects to one commitment rating. category on the basis of a number of criteria led the investigator to conclude that commitment may not be the unitary phenomenon that it has been operationalised as being in much research, but that it might be more accurate to conceive of various kinds of commit-For example, it is surely possible for someone to feel ment. strongly committed to teaching while in the job (in the sense of dedication to pupils, colleagues and so on) but only plan to stay in teaching a relatively short time (say, a few years).

While, in general, students' responses to the questionnaire items on commitment were in accord with how committed they felt to teaching as this latter emerged in the brief interview sessions, there were not a few instances of marked discrepancies. For example,

student 034 (M) made it clear in his interview that he would do as conscientious a job as he could while he was in the teaching service but that he would leave immediately that he was 'called by God'. He was asked if he had any idea as to when this was likely to occur and replied that he did not know but that he would recogise the moment when it came and did not expect to be teaching for too many years.

Thus it was that with some students the interview modified judgements that would have been made on the basis of questionnaire items alone. In the case of student 034 (M), his questionnaire responses signified an otherwise reasonably committed student. This experience led the investigator to believe that matters as potentially complex as commitment to a profession may often be more amenable to probing via interview techniques than by use of questionnaires.

4.3.4 THE FOLLOW-UP SURVEY

Two years after entering teaching the 97 subjects comprising the 6th semester cohort of 1976 were sent a copy of Role Norm Inventory 1 to complete together with a relatively short questionnaire aimed primarily at attempting to determine the ex-students' current levels of satisfaction with and expressed commitment to teaching after they had been in the job sufficient time to have settled into teaching properly and to have at least begun to cope with the problems that appear to beset the probationary year (see references in the literature review chapter). The responses obtained from this follow-up survey provided (i) a measure of any shift in students' ideal role concepts two years after graduating, and (ii) one basis for assessing the degree to which the role images carried into teaching might be associated with subsequent levels of satisfaction with and commitment to the profession - that is, students' adjustment to teaching. The particular problem explored concerned whether or not student differences in the degree of perceived role congruence with their lecturers and teachers at the end point of training were associated with later differences in the degree of expressed satisfaction with and commitment to teaching. Since, as stated, the first year of teaching has been widely considered as likely to be troublesome for most entrants to teaching, it was decided to wait until the second year had passed in order to maximise the possibility that dissatisfaction and expressed commitment levels reported would not merely reflect a beginner's problems but might, conceivably, represent a more enduring state of conflict.

As well as incorporating questions intended to yield direct comparisons with questions asked of the cohort before entering teaching, the follow-up questionnaire also encouraged open-ended responses at various points to enable the investigator to go some way toward putting 'flesh' on the bones of the bare numbers which may not always do justice to the complexity of a subject's response. The questionnaire is reproduced in Appendix 2a. Briefly, it asked questions concerning:

 (i) the number of schools appointed to and the reasons for any transfers, since this was considered a possible source of dissatisfaction, especially in the probationary years where transfers for bureaucratic purposes have been a commonplace in New South Wales;

- (ii) <u>respondents' estimates of their present level of satis-</u> <u>faction with teaching</u>; the question used was taken from the Wiseman and Start (1965) follow-up study of teachers in the United Kingdom, to which was added a request that subjects list their <u>chief sources of satisfaction</u> <u>and dissatisfaction</u>;
- (iii) the future career plans of respondents; these were questions developed originally by Mason (1961) for the Beginning Teacher Study conducted by the United States Office of Education, and used in Australia subsequently by Walker (1967) in a longitudinal study of occupational commitment amongst teachers;
- (iv) retrospective judgements of the training received at college including an assessment of the general competence of lecturers to train future teachers; and
- (v) respondents' retrospective judgements as to the positive and negative influences of various significant others on the way in which the teacher role was viewed.

While each of these measures are of intrinsic interest and were indeed, in some degree, taken for internal college purposes, they will only be elaborated upon subsequently insofar as they enabled the investigator to compose a relatively comprehensive picture of satisfaction and commitment current in the 1976 cohort two years after entering teaching, and to relate these to levels of perceived role congruence and conflict with lecturers and teachers evident upon leaving college. The relatively small number of subjects involved precluded other than the basic analyses which follow in the results section but, in view of the associative intent and tentative nature of the exploration, these were considered adequate.

Before being sent out, the questionnaire was piloted with a small group of 6th semester students and local teachers who were asked to note and report any difficulties associated with completing the questions. Their recommendations led to several changes in wording and layout in the final version. For example, question 3 concerning future career plans was originally intended to be the same question as that asked of the 1976 cohort prior to leaving College. However, on the recommendation of some of the pilot respondents a further response category was added to distinguish between those who saw teaching as a lifetime career and those who were committed in the 'forseeable future' but not necessarily until retirement. The questionnaire is given in full in Appendix 2a.

4.4 ADMINISTRATION OF THE RESEARCH

Prior to administering the Teacher Training Project Questionnaire (or part thereof), all students were first contacted by note through the college's internal mail system to outline the broad purposes of the research and to invite participation. In the note were specified the time and place of a more detailed, personal explanation of the study by the investigator. For the 2nd semester group this took place in a compulsory mass lecture time slot while for the 6th semester group the follow-up contact was made in tutorial groups since no mass lecture time existed for the group.

Subsequently, the students came individually or in small groups to the investigator's office where they were taken step-bystep through the instrument so that any matters needing clarifying could be attended to before each subject took away the questionnaire It was explained to the students that since to complete it. some follow-up was envisaged, each questionnaire was numbered so that later it could be related to names on a list kept and seen only by the investigator who was known by all of the students involved (through his teaching on numerous semester courses and through schools liaison work) and, apparently, trusted. Assurances of confidentiality were personally given to each respondent. It was stressed that the focus of the research was not on the individual but the group. That the investigator's assurances were accepted is reflected in the fact that all available students in both cohorts returned usable questionnaires, that overwhelmingly these were fully completed and that the investigator received many unsolicited comments to the effect that the students found the research proposal interesting and that it was about time someone sought their views.

Moreover, many students opted to complete their questionnaires either on the spot in the investigator's office or in one of the teaching rooms a few metres away. It was emphasised that questionnaires must be completed alone. On average, it took the 6th semester students about forty five minutes to complete the entire questionnaire while it took 2nd semester respondents about twenty minutes to complete the role norm inventories.

These procedures were time-consuming, taking several weeks before all completed questionnaires were finally collected. The whole approach however was informed by an attempt to create a confidential but personalised atmosphere in order to maximise the prospects of a full and honest response from each participant. Throughout, the investigator alone administered the research - from the initial note requesting participation, to the collection of the last questionnaire - and was always on hand to respond to students' queries and thus to ensure uniformity of interpretation where necessary. It was hoped that such an approach would help offset criticisms made about the use of questionnaires to investigate respondents' views (see for example Shipman, 1967a).

The 45 respondents in the lecturer group were all contacted personally, either directly or by internal College telephone to explain the purposes of the research and request participation. This latter was readily agreed to by all respondents. The role norm inventory was subsequently explained to each subject face-toface and was either completed there and then or returned through the College's internal mail system.

As explained previously the 653 teacher respondents were contacted through the College's Schools Liaison Office. A letter and accompanying role norm inventories were sent out with practice teaching materials to each school Principal or Headmistress. The letter explained briefly the nature and purposes of the study, requested participation and gave details of how the inventories should be completed. It was asked that the inventories be returned with various practice teaching forms.
The 6th semester, 1976 students who were tracked in 1978 two years after they had entered teaching, were located through files kept by the New South Wales Department of Education Western Area Office in Bathurst. Before leaving College each student had been told that an effort would be made to contact him or her within two years of entering teaching. To this end a letter was sent in June 1978 to each of the 97 members of the 1976 cohort reminding them of the research project, requesting their further participation, assuring confidentiality and enclosing the follow-up questionnaire and a copy of Role Norm Inventory One dealing with respondents' ideal role conceptions. After a reasonable time had elapsed during which over a 70% return was achieved, a second (handwritten) letter was sent out to those who had not responded. This resulted in returns from most of the rest who did take part in this followup phase. Finally, telephone calls were made to four respondents who were known to be in schools but who had not responded to the two letters. All were apologetic assuring the investigator that they had intended to participate.

These procedures yielded a 94% rate of return (91 out of the original 97 respondents). The remaining 6 respondents could not be traced. What information could be gleaned about them suggested that had left teaching for a variety of reasons (such as to go overseas).

The Castle Hill data were gathered by Dr. Marie Kelliher, a long-standing personal friend and former colleague of the investigator. It was through Dr. Kelliher that access to the Castle Hill College was negotiated. The investigator visited the College

a number of times prior to the administration of the Teacher Training Project Questionnaire, the role norm inventories sections of which were needed for the research reported here. During these visits Dr. Kelliher was taken carefully through the relevant administrative procedures and the purposes of the research were explained.

It was decided not to confine the research to the collection of the role norm inventory data required for this study but to administer the whole Teacher Training Project Questionnaire to Castle Hill respondents with an eye to ongoing research. Accordingly the Questionnaire was administered by Dr. Kelliher to the entire 2nd and 6th semester cohorts of students in 1978 in a compulsory mass lecture time slot. Again, while confidentiality assurances were given by Dr. Kelliher who knew all of the students involved through her teaching and practice teaching contacts with them, questionnaires were numbered (as previously explained in respect of the Mitchell data) so that respondents could later be tracked.

4.5 STATISTICAL PROCEDURES

4.5.1 INTRODUCTION

Before describing the statistical procedures used, it should be pointed out that the five-point Likert scale used for the Foskett Role Norm Inventory generates data that, strictly speaking, are sub-interval in nature. That is, no assumption can really be made regarding equal intervals between each of the five response categories to which numerical values were assigned, the categories ranging from 'definitely should/will' (category 1) to 'definitely should not/will not' (category 5). It has been argued (e.g. by Stevens, 1951, 1968; Wilson, 1971; Siegel, 1956) that unless a researcher can establish that his data exhibit interval scale properties, then the application of powerful parametric statistical procedures is inappopriate and other so-called non-parametric statistical procedures designed specifically for ordinal or nominal data should be used. Despite the controversy that this purist 'weak measurement' orientation has provoked (see for example Acock and Martin, 1974), it is a commonplace in educational investigations such as this for researchers to use the more powerful parametric techniques on data which seldom achieve interval or ratio scale properties.

Apart from the fact that if the strictures of what Acock and Martin (1974: 427) call the 'Methodological Purism' approach were adhered to, only a small proportion of the data collected by social and behavioural scientists could utilise such analyses (cf. Harris, 1975: 226), it has been argued (e.g. by Labovitz, 1970; Abelson and Tukey, 1970; Petrinovich, 1966; Anderson, 1961; Hermann and Braskamp, 1970; Harris, 1975) that most such data are, in fact, amenable to parametric statistical treatment since there is strong mathematical and empirical evidence for the robustness of statistical procedures under violation of the assumptions of normality and homogeneity of variance. Further, as Winkler and Hays (1971: 843) observe, hypotheses tested by order methods seldom equate with those tested by parametric techniques in that, whereas the latter usually test the hypotheses of equal population means, order methods essentially test the hypotheses of identical population distributions, leaving the possibility that the kinds

of true differences among populations that such tests fail to detect may be quite unknown. Moreover, order and other non-parametric techniques lack power relative to parametric counterparts, so that with equivalent n's and alpha levels, the use of the former increases the probability of Type II errors (accepting the null hypothesis when it is false). A further disadvantage of order techniques is that because only ordinal features of the raw numerical data are used not all of the information in the data is used.

Summing up the arguments brought to bear upon an issue that is not easily resolved, Kerlinger (1973: 288) says:

In brief, in most cases in education and psychology, it is probably safer - and usually more effective to use parametric tests rather than non-parametric tests,

and adds that, in a definitive article on the whole controversy, Anderson (1961: 315) concluded that 'parametric procedures are the standard tools of psychological statistics, although nonparametric procedures are useful minor techniques.' Granted, therefore, that parametric statistical procedures are widely used with data such as that derived from responses to the role norm inventories in this study, it was thought that their use could be justified here if, as in the foregoing discussion, the appropriate cautionary note had been sounded and due care subsequently taken in interpreting the data.

The problems of data interpretation inevitably raised the question of statistical significance testing. In a withering attack on the use of traditional statistical methods in educational research, Carver (1978) makes a plea for the abandonment of what he calls 'the corrupt scientific method' involving the interjection of the null hypothesis into the research process, a method characterised by Novick and Jackson (1974: 245) as 'giving a misleading answer to a question nobody is asking'. Carver points out that, despite assuming in research that two sample groups both represent the same population, we cannot be sure in practice that the null hypothesis is true and that, if we could, there would be no point in testing for statistical significance at all. A result that is designated 'statistically significant' then, is one where, given acceptance of the null hypothesis, the probability is low that one would get the particular result obtained. It is erroneous to reverse this by interpreting the obtained p value as relating to the probability that the null hypotheses is true as Carver rather colourfully explains:

What is the probability of obtaining a dead person (label this part D) given that the person was hanged (label this part H); that is, in symbol form, what is p(D/H)? Obviously, it will be very high, perhaps .97 or higher. Now, let us reverse the question. What is the probability that a person has been hanged (H) given that the person is dead (D); that is, what is p(H/D)? This time the probability will undoubtedly be very low, perhaps .01 or lower. No one would be likely to make the mistake of substituting the first estimate (.97) for the second (0.01); that is, to accept .97 as the probability that a person has been hanged given that the person is dead. Even though this seems an unlikely mistake, it is exactly the kind of mistake that is made with interpretations of statistical significance testing - by analogy, calculated estimates of p(H/D) are interpreted as if they were estimates of p(D/H), when they are clearly not the same.

(Carver, 1978: 384, 5)

That is, the p value in statistical significance testing is the probability of getting the obtained difference given the null hypothesis is true and not the probability that the null hypothesis is true given the obtained difference. Equally erroneous interpretations are those seeing statistical significance as the probability of getting the same result in a replication of the research, and that the complement of p (i.e., .95 for a p of 0.5) is the probability that the research hypothesis is true, reflecting the degree of validity of the results.

Carver warns also against interpreting trivial results as 'significant' when they simply reflect what would rarely happen when randomly sampling from the same population using large sample sizes. That a statistically significant difference is not necessarily a difference of importance, and vice versa, has been pointed out before (e.g., by Tyler, 1931: 115; Gold, 1969: 42). In this respect Hargreaves (1972: 84), in a comment relevant to the particular concerns of this thesis, warns:

Many researchers give what may be a spurious importance to their findings by demonstrating discrepancies between ... expectations which are statistically significant. This is not enough. The researcher needs to demonstrate the psychological significance of these discrepancies \ldots I suspect that such discrepancies may need to be quite marked before they provoke a reaction of psychological or sociolog-Small differences in expectical significance. ations may have a significance which is only statist-Sometimes, too, researchers do not ascertain ical. that the areas in which they reveal discrepancies in expectation are \ldots central and important aspects of the role involved. The demonstration of discrepancies in expectation concerning peripheral aspects of a role may be without significance for the persons involved or for the social systems of which they are part.

Furthermore, as Rozeboom (1960: 416-28) has argued, whether the evidence is statistically significant or not, if it is in accord with the research hypothesis it should increase the degree of belief in that hypothesis. However, where data are not statistically significant it has been common practice to accept the null hypothesis as the most reasonable explanation despite rival hypotheses that warrant investigation.

It was against this background then, and in response to Stevens' (1968) question, 'Can no one recognize a decisive result without a significance test?' that the initial data analyses, at least, were carried out using techniques that did not utilise statistical significance testing. These methods are described below. Following these analyses various statistical significance tests were used despite Carver's plea for their abandonment, but care has been taken to avoid the interpretive fallacies he exposes. The tests have been regarded as complementing the initial analyses and, where trends appear to have emerged, differences that were not significantly different statistically have not been excluded from consideration automatically, on the grounds that there are a priori reasons for believing the null hypothesis is generally false (Bakan, 1966: 423ff). Given that this is so, the whole position can perhaps be summarised as follows: if the null hypothesis is actually true, as is assumed when statistical significance tests are used, then any difference between means is a real difference, and a judgement must be made as to its likely importance; recourse to statistical significance testing as a formal and objective way of ascertaining the statistical awareness of a result should aid (rather than determine) such a judgement (cf. Winch and Campbell, 1969: 140). Understandably pessimistic about the chances of eliminating statistical significance testing from educational research, Carver (1978: 397) argues for, at very least, the abandonment of the 'corrupt' scientific method, the addition of 'statistically' in front of 'significant' and the incorporation of effect-size measures where possible to indicate the significance - as against statistical significance - of a particular result. An

effort has been made in presenting the results of this study therefore, to try to reflect some of this thinking, as well as that of Cronbach and Snow (1977), who urge that the basic descriptive statistics should always be shown.

4.5.2 DESCRIPTIVE MEASURES

To begin with, the responses of all groups under all role norm inventory conditions were punched onto computer cards, checked, and mean response scores, standard deviations, and frequency distributions for each of the 45 variables for all groups separately were obtained using the BMD 2PD programme. At the same time, a simple category count programme was written to find the total number of responses in each of the five response categories for any group, or combination of groups, for any combination of variables such as those making up a particular role sector.

From the mean response scores, mean differences per role norm over a role sector or the total role were calculated for all within-group comparisons (that is, for example, a group's own norms versus their expectations) and between-group comparisons. These were done with a desk calculator so that the researcher could immerse himself in the data. The simple mean difference per role norm statistic, obtained by calculating the average difference between mean response scores over all relevant items in a comparison, was employed by Foskett (1967a, 1967b, 1969) throughout the reporting of the results of his Pacific Coast investigations and was found to be useful indeed in this study also, for as well as being in accord with the thrust of Carver's (1978) advice to researchers, it enabled the investigator to discern early the broad trends in the data prior to statistical significance testing. In the presentation of results that follows, this simple measure is used to complement the analyses and to summarise data trends.

4.5.3 CONSENSUS MEASURES

The essential problem in attempting to ascertain the degrees of consensus in the data was to find measures that could be used for within- and between-group comparisons and also accommodate both single expectations and sets of expectations such as those constituting a role sector. The measurement of consensus has, in past research, proven to be complex. For example, Gross et al. (1958), in their major study of School Superintendents, used a number of different estimates of variance for single role expectations but ran into difficulties when trying to determine between-group consensus with sets of expectations (see Gross et al., 1958, pp. 104-115). The use of a statistic such as, for example, gamma (e.g., see Rogers, 1969) also is problematic in that it can provide misleading pictures of data (Kohout, 1974). To avoid these difficulties, it was decided therefore to use two separate measures of consensus.

The first measure was one developed by Leik (1966) and subsequently used by Foskett (1967a, 1967b, 1969). It is a measure of ordinal consensus based on the degree of dispersion of responses over the five categories for each role norm statement and derived from the cumulative relative frequency distribution of responses. The statistic reflects the degree to which responses cluster along a continuum from 'definitely should' through to 'definitely should not'. The theoretical range of the statistic is 2.0, from a score of + 1.0 where all responses are in one response category (i.e. perfect consensus), through 0.0 where, akin to random entry, 20 per cent of the responses are in each response category (i.e. no agreement), to - 1.0 where 50 per cent of the responses are in each extreme category (i.e. perfect dissensus). The measure is named the Agreement Score (Foskett, 1967a) and reflects a concept of consensus as lack of dispersion.

Since, to the writer's knowledge, this useful measure has not been used in Australian research, a brief description of it, based on Leik's (1966) more detailed explication of the rationale for the procedure, is provided in Appendix 4. Despite advantages such as relative ease of calculation, the fact that it has ratio properties, makes no assumptions about intervals, and seems preferable in most circumstances to a similar measure of Interval Consensus (see Leik, 1966, 88), the measure has the drawback that there seems no way to derive sampling variability of the index, and so significance tests have not been developed for it and confidence intervals not constructed.

Since, despite its usefulness as a measure of consensus, the statistical significance of between-group comparisons could not be ascertained using Leik's measure of ordinal consensus, a measure of dispersion was sought with a known sampling distribution that allowed tests of significance to be performed to determine the

probability of an obtained result occurring when two multivariate profiles were compared. The statistic chosen was the Mahalanobis Distance (MD^2) which is an inferential measure of the amount of separation between two sets of scores. Essentially, the calculation of MD^2 requires the use of Pythagorean distance (D^2) measures, with the scores used in the calculations first being transformed so that they are standardized and uncorrelated (Kerr, 1978: 312). This process, according to Cronbach and Gleser (1953: 467), equates to 'what would be obtained if one factored the correlation matrix into k orthogonal factors, computed the person's scores on these components, and then applied the D^2 formula to measure similarity.' The transformation of the scores allows the MD^2 statistic to be referred to the F distribution for determination of statistical significance. Its advantage as a measure of between group (interposition) consensus is thus obvious but it has limitations as a within group (intraposition) estimate of dispersion as Kerr (1978: 312) observes: 'Since it derives k orthogonal factors, one parallel to each of the original k variates, and since it assigns equal weight to each factor, some factors composed almost entirely of error variance will receive relatively heavy weighting.' Moreover, though stable as a between-group estimate of dispersion over two or more trials with the same group, this is not so for the within-group situation (Cronbach and Gleser, 1953). For these reasons the Mahalanobis Distance was used initially to signify where differences between groups were statistically significant and Leik's useful Measure of Ordinal Consensus for all subsequent comparisons where the statistical significance of between group differences for whole sets of expectations (role sectors) had been established using MD^2 .

4.5.4 THE MULTIVARIANCE PROGRAMME AND POST HOC ANALYSES

Because the role norm inventory was conceived of in terms of different role sectors, each made up of a cluster of norms, it was deemed inappropriate to conduct <u>as a first step</u> separate univariate analyses of variance for each role norm item. What was needed in the case of each of the four sectors of each inventory was an initial analysis pertaining to the entire set of role expectations (or role sector) followed by appropriate <u>post hoc</u> probing where an <u>overall</u> difference was found between the responses of the various groups to the entire cluster of items.

The multivariate analysis of variance programme 'Multivariance' (Finn, 1974) was used to carry out this first step. A separate analysis was done for each of the four role sectors. Thus for the role sector Acting Toward Pupils which consisted of a cluster of 15 separate role norms, all of the responses of each of the four groups to each of these variables were together analysed to test the hypothesis of equality of dependent mean vectors for the groups. Where a statistically significant multivariate F was found, the analysis proceeded in order to determine which of the role sector items contributed to the overall significance. The Multivariance computer programme gave statistical significance levels for each of these univariate F-tests.

Having determined those role norm items contributing to the overall significance, relevant <u>a posteriori</u> comparisons between group means were made for each such variable to ascertain precisely where the between-group differences lay. Specifically, the relevant comparisons were made as follows:

- 1. For each student-teacher group, their own norms were compared with their own expectations (i.e., Role Norm Inventories 1 vs 2), their own norms with those they attributed first, to lecturers (R.N.I.'s 1 vs 3), and then, to teachers (R.N.I.'s 1 vs 4), and finally the norms they attributed to lecturers with the norms they attributed to teachers (R.N.I.'s 3 vs 4).
- 2. The own norms of the student-teacher groups were compared with each other (i.e., R.N.I. 1 for 2nd semester students vs R.N.I. 1 for 6th semester students), then their own expectations (R.N.I. 2), and then their attributed norms (R.N.I.'s 3, 4).
- 3. The own and then the attributed norms of each studentteacher group were compared with the actual (i.e. own) norms of, firstly the lecturers, and then the teachers.
- The norms of the lecturers and teachers were compared with each other.

In making such comparisons a choice had to be made amongst the numerous methods commonly used for post hoc analysis of this kind. It was recognised that the problem of multiple comparisons in the analysis of variance remains a controversial topic with no simple solution. The various techniques most often used in behavioural research often differ in the way in which the rate of Type 1 errors (rejecting a true hypothesis) are dealt with. For instance, some use an error-rate per comparison, such a rate being defined as the probability that any one of the group of comparisons will be designated significant when H_o is true. However, such an error-rate seems justified only when the comparisons in question are statistically independent as in the case of orthogonal comparisons. On the other hand, some use an experimentwise error rate, defined as the probability that one or more Type 1 errors will be made in the comparisons derived from a given experiment. This more conservative approach is exemplified in statistical tests such as the Scheffe test.*

Although early in the history of the analysis of variance, Sir Ronald Fisher suggested using individual t-tests to make multiple comparisons after finding a significant F, and despite the fact that this procedure has often been used and continuesto be used (Glass and Stanley, 1970: 382), such a procedure was rejected for the reason that the t-test was derived for use in a situation in which only two random samples were drawn; violation of this assumption can result in a situation wherein the probability of rejecting a true hypothesis is likely to exceed the specified significance level. To illustrate, where the alpha level is set at, say, .05 in a multiple comparison situation following a oneway analysis of variance, one would expect to find a statistically significant difference by chance alone for every twenty such comparisons.

For these reasons it was decided to err on the conservative side in analysing the data and so the Scheffe test, which uses a single range value for all comparisons and is exact for unequal group sizes, was used for all relevant between-group comparisons.

*See Scheffe, H.A. The Analysis of Variance, Wiley, New York, 1959.

Though the test is the least powerful of all the procedures available, its use guards against the likelihood of attributing difference of an educationally significant nature to some results which intuition would show to be trivial, even though statistical significance may have been found using more powerful methods. In addition, the very rigorous .01 criterion of statistical significance was used for the most part to increase the probability that obtained differences were 'real' differences.

Additionally, to comply with Carver's (1978) suggestion that effect-size measures ought to be incorporated in results tables, omega squared (ω^2) was calculated (see, for example, Winkler and Hays, 1975: 765-7) for those role norm items where statistically significant differences were found so that these could be ranked in order of magnitude according to the proportion of variance accounted for in the obtained difference. Briefly, the larger the value of omega squared the greater the 'effect', since large values indicate that the differences between groups are big relative to within-group variability and the overlap between the scores in the different groups is small (cf. Lindman, 1974: 51). Ιt should be added that omega squared differs from 'F' in that whereas 'F' is a descriptive statistic merely giving a ratio of variances, and varying according to the size of N (the number of subjects/ observations), omega squared is a measure of what might be termed the practical significance of the differences and is unaffected by the size of N.

4.5.5 THE MEASUREMENT OF CHANGE

To assess change in the responses of the 2nd semester 1976 students who again completed the four inventories in 1978 when in 6th semester, use was made of the B.M.D. P2V Analysis of Variance and Covariance with Repeated Measures computer programme. Essentially, a repeated measures analysis allows advantage to be taken of the systematic behaviour of individuals and thus eliminate any part of the original error which could be ascribed to this (cf. Burroughs, 1975: 226). While in one-way or factorial analysis of variance there is no question of correlation between groups since the independence of groups and observations is a sine qua non of the design, in a repeated measures approach such as that used here, the same subjects are 'measured' more than once and are thus not independent. Because the variation due to individual differences is known to be substantial if, as in a repeated measures analysis, it can be isolated and extracted from the total variance, there is a resultant increase in precision because this source of variation can be subtracted from the total variance and thus a smaller error variance used in evaluating obtained differences (cf. Kerlinger, 1973: 272; Winer, 1971: 261).

Analysis proceeded, as detailed previously for the role norm inventory data on a sector by sector basis for each of the 4 inventories. Where change was signified, <u>post hoc</u> comparisons of means for individual items were made as before to ascertain where specific differences lay.

4.5.6 CROSSTABULAR ANALYSES

For the Teaching Style Inventory data, and the data yielded by the Follow-up Survey and Background Information section of the Teacher Training Project Questionnaire, cross-tabular analysis (together with associated statistical techniques such as chisquare) were used, as befitted the data. Briefly, a cross tabulation is a joint frequency distribution of cases classified according to two or more variables and is, according to Nie et al. (1975: 218) the most commonly used analytical tool in the The chi-square statistic, which is the most social sciences. frequently used measure of statistical significance with tables of cross-classified data, assists in determining whether a systematic relationship exists between two variables by comparing observed with expected cell frequencies. Measures of association used with chi-square such as the co-efficient of contingency and the phi co-efficient enable the strength of the relationship between variables to be estimated.

4.5.7 THE FACTOR ANALYSIS: PROCEDURE AND RESULTS 4.5.7.1 PROCEDURE

The large number of responses yielded by the role norm inventories necessitated the use of data reduction techniques to provide, if possible, a framework wherein more elegant generalisations could be made than in the 1975 study where statements about the data were based on items intuitively grouped by the researcher. In the original study for example, role norm items 3 (deprive a pupil of privileges as a form of punishment), 7 (use extra academic work as a form of punishment) and 12 (use corporal punishment as a disciplinary measure) were together taken to

indicate degrees of 'punitiveness' or 'authoritarianism' with respect to teacher-pupil relationships. Whilst such a procedure is defensible since the researcher's groupings and consequent analyses are open to inspection it was thought that in the interests of parsimony, more rigorous, objective procedures were desirable Therefore, the responses of all of the 1976 groups at this stage. taken together (2nd and 6th semester student teachers, lecturers and teachers) were subjected to a Principal Components Analysis, role sector by role sector. That is, the responses of all subjects to all 15 items for the role sector Acting Toward Pupils, for all Role Norm Inventory Conditions, constituted the raw data of the first analysis, this being followed by similar analysis of the 10 items of Role Sector 2 (Acting Towards Colleagues), the 10 items of Role Sector 3 (Acting Toward Parents) and the 10 items of Role Sector 4 (Acting Toward Community), making four analyses in all.

Principal Components analysis, like other factor analytic approaches, is a technique for reducing the dimensionality of a set of variables but, unlike the factors derived from a factor analysis, principal components are exact mathematical transformations of the original variables with each subject's score on a component being a linear combination of his scores on the original variables. The first principal component is that linear combination of the original variables which discriminates maximally among the subjects in a given sample. Likewise, each successive principal component of the total of p components extracted from the data accounts for as much of the residual variance in the original data as possible, with the scores on any component being uncorrelated with scores on the other components. One advantage of this is that duplication in interpretation of the way subjects' responses are affected by, or affect other sets of variables can be eliminated. Another is that, if needed, the scores derived for each subject from the principal components can be used in subsequent statistical analyses (such as multiple regression) with minimal, or no loss of information, depending upon the number of components retained.

Principal components analysis then, can be used when it is desired to make a relatively straightforward transformation of a set of variables into a new set of composite variables that are orthogonal to each other, that account for a maximum of variance in the original data and that make no particular assumptions about the underlying structure of the data. In brief it is a technique used as an aid in describing the 'structure' underlying a system of variables (Harris, 1975: 174) and was chosen here because of its especial usefulness in summarizing in a smaller number of variables as much as possible of the information contained in the original variables.

Although theoretically n principal components are needed to account for all of the variance in n original variables, in practice often the most important components only are retained. The computer programme used carries out a varimax rotation and specifies the number of components extracted these being determined by Kaiser's criterion with only components (or factors) having eigenvalues greater than one being retained.

4.5.7.2 RESULTS: ROLE SECTOR 1 - ACTING TOWARD PUPILS

Table 4.2 gives the varimax rotated component analysis factor matrix for the role sector Acting Toward Pupils. Using Nunally's (1967) conservative rule-of-thumb approach, most interpretive emphasis was placed on variables loading .30 or more, i.e. those contributing approximately 10% to the variance in the factor. Using the Burt-Banks formula for determining factor loadings, those of .30 or greater are, generally, statistically significant at the .01 level (Child, 1970: 97). Such loadings are marked with an asterisk. However, since with a large sample size and 15 variables (in the case of Role Sector 1) this is a stringent criterion for significance⁴, variables with loadings approaching .30 were not <u>automatically</u> excluded when it came to interpreting the extracted factors.

In naming the factors the writer was able to draw upon the work of McNamara (1972) who in a study of the professional socialization of female student teachers in an English College of Education carried out the only other components analysis of the Foskett inventory known to the writer. Where there were marked similarities between the factor structure of the components found in this study and those reported by McNamara, the same or similar labels to those he chose were used if they were thought adequate for providing an acceptably accurate, broad descriptive framework for the data.

^{4.} In determining the significance of factor loadings a number of considerations are relevant and were used as guidelines here to interpret factors: (i) as the sample size increases, the smaller loading required to reach a given level of statistical significance, (ii) as the number of variables to be analyzed increases, the smaller the loading can be to reach statistical significance, and (iii) as the number of factors increases the larger the loading needed to achieve statistical significance. (See the discussion on these matters in, for example, Hair et al. 1979, pp. 234-237.)

TABLE 4.2

VARIMAX FACTORS FOR ROLE SECTOR 1, ACTING TOWARD PUPILS: 1976 GROUPS

		FACTORS				
ROLE NORM		1	2	3	4	5
1.	give homework regularly,	.58*	05	.14	.16	32*
2.	make and carefully follow detailed lesson plans.	.18	08	05	00	0.74*
3.	deprive a pupil of privileges as a form of punishment.	.58*	.25	13	.09	.07
4.	give pupils a great deal of rote learning in the basics.	.57*	.04	.03	16	16
5.	evaluate work of pupils on basis of individual improve- ment rather than by comparing with other children.	11	.15	34*	.50*	07
6.	give greater attention to the more able than to the less able students.	.25	04	29	0.67*	.15
7.	use extra academic work as one form of punishment.	.56*	07	.18	06	03
8.	experiment with new teaching techniques.	.08	09	23	.61*	0.03
9.	permit each pupil to follow his own educational inter- ests most of the time.	06	21	69*	.11	00
10.	smoke in situations where a pupil might see them.	.08	17	00	00	.63*
11.	devote most of their time to working with individual pupils or small groups.	11	03	75*	.07	03
12.	use corporal punishment as a disciplinary measure.	.60*	05	.02	16	.26
13.	encourage pupils to discuss various religious beliefs in the classroom.	02	79*	09	.07	.02
14.	(Teachers) express own polit- ical views in the classroom.	02	78*	08	12	.02
15.	encourage pupils to question the opinions held by the teacher.	03	48*	13	.44*	.27
	Eigenvalues	2.26	1.70	1.40	1.11	1.03
	Percentage Variance	15.07	11.33	9.36	7.38	6.88

In order of the magnitude of their contribution to the variance in the first principal component the variables with statistically significant loadings are as follows:

Variables	Loading	Description
12	.60	Use corporal punishment as a disciplinary measure.
3	.58	Deprive a pupil of privileges as a form of punishment.
1	.58	Give homework regularly.
4	.57	Give pupilsrote learningin the basics.
7	.56	Use extra academic work as punishment.

Three of these items (12, 3 and 7) describe teacher behaviour that could be labelled 'authoritarian' or 'punitive' while items 1 and 4, when considered with the other three variables convey a stereotype, perhaps, of the 'traditional' classroom teacher concerned with order, regularity and the basics. McNamara found an almost identical factor and named it 'traditional authoritarian behaviour' which seems accurately descriptive of the component's structure. In addition to the five loaded items above, item 6 (give greater attention to the more able than less able...) carried a .50 positive loading in McNamara's first factor and it will be noted from Table 4.2 that here also item 6, with a positive loading of .25, approaches the accepted .30 criterion. As a simple means of confirming the obvious similarities in the structure of the factors obtained here and by McNamara, the 15 variables of Role Sector 1 were in each case ranked from the role norm carrying the highest positive loading to that with the largest negative loading and a Spearman rank-order correlation coefficient calculated. A value of 0.84 was found, indicating the strong similarity (p < .01) in structure between the obtained factors. Add to this the fact that in both cases the same variables were the statistically significant positively loaded items contributing substantively to the variance in the factors, and it was decided to follow McNamara in naming the first factor 'traditional authoritarian behaviour'. This procedure was used throughout where existing factor names were used.

Variables	Loading	Description		
13	79	Encourage pupils to dis- cuss various religious beliefs.		
14	78	(Teachers) express own political views in the classroom.		
15	48	Encourage pupils to question opinions held by teacher.		

The structure of the second factor is as follows:

Again, McNamara's second factor was virtually identical, involving the same major contributing variables in the same order. As well, the overall relationship of all of the variables was similar as the obtained rank order correlation of 0.66 (p < .01) suggests. Since each of the variables relates strongly to the notion of an open, democratic classroom atmosphere where there is positive encouragement to express a viewpoint, McNamara's label 'freedom of expression' was accepted.

Variable Loading Descripti		Description
11	.75	Devote most time to indi- viduals/small groups
9	.69	Permit pupils follow their own educational interests
5	.34	Evaluate the work of pupils on the basis of individual improvement rather than by comparing

The variables contributing most to the third factor (or, principal component 3) were, in order of magnitude:

Each of these variables suggests a teaching orientation characterised by concern for the learner as an individual and so this factor was labelled 'pupil-centred teaching behaviour'. McNamara's third factor was similar in that items 9 and 11 carried the highest loadings and that item 5 was also a significant contributor, but role norms 8 (experiment with new teaching techniques) and 2 (make and carefully follow detailed lesson plans) were also included, leading him to use the broader label 'progressive child-centred behaviour' which suggests perhaps two components under one label. In fact, the analysis of the data of this present study did yield two such separate components – the third factor just discussed and the fourth factor discussed below.

Variable	Loading	Description		
8	.61	Experiment with new teach- ing techniques.		
5	.50	Evaluate work of pupils on basis of individual improvement		
15	.44	Encourage pupils to ques- tion the opinions held by the teacher.		
6	67	Give greater attention to the more able than to the less able pupils.		

The statistically significant variables making up this fourth factor were:

Before attempting to name this factor its bipolar nature should be considered. Harman (1970: 100) has pointed out that bipolar factors are not essentially different from others but are merely ones for which some of the variables have significant Such variables can be held to measure negative projections. the negative aspect of, or absence of the postulated characteristic. Thus if the positively loaded variables on a factor were interpreted as meaning, say, 'courage', a negatively loaded variable could be taken to signify 'cowardice' or 'lack of courage'. In the case of this fourth factor the positively loaded variables suggest elements of openness to new approaches (8), child-centredness (5) and a democratic approach to classroom interaction (15) while the negative loading on item 6 suggests a concomitant anti-elitist Bearing in mind the factor 'progressive child-centred stance. behaviour' obtained by McNamara which appears to be a combination of the third and fourth factors found here, it was thought that the combination of variables making up this fourth factor might

reasonably be taken to signify a 'progressive' orientation to teaching and so was called 'progressive teaching behaviour'. Whilst terms such as 'progressive' and 'traditional' are perhaps too open to interpretation to be entirely unambiguous, they are nonetheless widely used and very much part of the educational zeitgeist of the 1970's and '80's.^{5.} Moreover, it is stressed that in this study such labels are defined in terms of the variables that subsume them and are used primarily to describe the data more parsimoniously.

The variables contributing most to the fifth factor were as follows:

Variable	Loading	Description		
2	73	Make and carefully follow detailed lesson plans.		
1	32	Give homework regularly.		
10	.63	Smokewhere a pupil might see them.		

Again, this is a bipolar factor and is not easy to name, the variables together perhaps describing teaching behaviour that might be characterised by the term 'formality'. Accordingly, the factor was called 'formal teaching behaviour' despite the limitations of such a label given the incomplete evidence upon which it was based.

^{5.} See, for example, a work already mentioned - Bennett, N. <u>Teaching Styles and Pupil Progress</u>, Open Books, London, 1976. This book, which aroused considerable interest and generated some controversy, investigated the efficacy of so-called 'traditional' and 'progressive' approaches to teaching.

Summarizing the above, the fifteen variables of this role sector were reduced to five factors which provide a framework wherein the responses of all of the groups can be described with defensible accuracy using terms such as 'traditional', 'authoritarian', 'punitive', 'democratic', 'open', 'pupil-centred', 'progressive', and 'formal'.

4.5.7.3 RESULTS: ROLE SECTOR 2 - ACTING TOWARD COLLEAGUES

Table 4.3 shows the factor matrix for Role Sector 2 - Acting Toward Colleagues:

TABLE 4.3

VARIMAX FACTORS FOR ROLE SECTOR 2,

ACTING TOWARD COLLEAGUES: 1976 GROUPS

ROLE NORM		FACTORS			
		1	2	3	
16.	devote time outside regular teaching duties to school affairs, such as curriculum planning, without additional pay.	.16	.15	72*	
17.	take up active membership in a teachers' professional organisation.	.73*	.00	29	
18.	use surnames like 'Miss Smith' or 'Mr. Jones' when addressing other teachers in front of pupils.	.31*	. 56*	.02	
19.	include other teachers in their circle of close friends.	.53*	.05	02	
20.	continue to take further pro- fessional courses as long as they continue to teach.	.35*	07	47*	
21.	insist upon extra pay for duties, like coaching a team, that require extra time.	.22	30*	.64*	
22.	make or receive personal tele- phone calls while at school.	.04	72*	.10	
23.	discuss serious personal problems with the principal.	.15	43*	40*	
24.	join a teacher organization such as the N.S.W.T.F.	.76*	.03	.02	
25.	engage in part time work during term time.	01	62*	.26	
	Eigenvalues	2.11	1.54	1.07	
	Percentage Variance	21.12	15.37	10.66	

The first factor was labelled 'professional political activism' since the variables contributing most (24 and 17) clearly suggest this and since it can be surmised that the teacher who is thus active is likely to cultivate the friendship and support of his colleagues (19), to advocate that teachers keep abreast of professional developments (20), and perhaps even maintain a professional image in front of pupils in ways such as that detailed in item 18 (a minor contributing variable).

The second factor is described by the negatively loaded items, in order of magnitude, 22 (-.72) 25 (-.62) 23 (-.43) and 21 (-.30) and one positively loaded variable, 18 (.65), which together suggest a pattern of teacher behaviour along a 'professional' (in the sense of acceptable) - 'unprofessional' continuum with a particularly strong element of 'extra-curricular remuneration', to use the label McNamara employed for his second factor which was composed of negative loadings on variables 21, 22 and 25. Combining the two notions to incorporate the contributions of variables 18 and 23, this factor was called <u>extra curricular profess-</u> <u>ionalism</u>.

The third factor is also bipolar in structure, the two main contributing variables being items 16 with a positive loading of .72 and 21 with a negative loading of -.64. Items 20 and 23 with positive loadings of .47 and .40 respectively are the other significant contributors. In attempting to describe this factor, cognisance was also taken of two other items sufficiently close to the .30 criterion to be worth considering. These were variables 17 (take up active membership in a teachers' professional

organization) with a positive loading of .29, and 25 (engage in part-time work during term time) with a negative loading of .26. A very similar factor obtained by McNamara was named 'extra curricular altruism' because of a significant emphasis in some of the contributing variables on activities outside of school hours demanding some sacrifice by the teacher. This label was accepted here also. It should be noted that this factor appears to describe behaviour very much the reverse of that described by factor two.

In all, on the basis of the principal components analysis of all subjects' responses to this role sector, descriptors such as 'activist' (in the professional political sense), 'mercenary', 'dedicated' and 'altruistic' were derived for use in generalizing about any patterns in the data.

4.5.7.4 RESULTS: ROLE SECTOR 3 - ACTING TOWARD PARENTS

Table 4.4 gives the rotated matrix for the third role sector, that of 'Acting Toward Parents':

TABLE 4.4

VARIMAX FACTORS FOR ROLE SECTOR 3

ACTING TOWARD PARENTS: 1976 GROUPS

	ROLE NORM		FACTORS	
		1	2	3
26.	accept the judgement of parents when there is dis- agreement about the needs of the child.	16	.07	.67*
27.	insist that parents contact them only after obtaining permission from the principal.	.20	.02	.70*
28.	visit every pupil's home at the beginning of the school year.	34*	.35*	.09
29.	discuss with parents the child's scores on standardized attain- ment tests.	15	.74*	.16
30.	tell a parent the tested I.Q. of his child.	03	.80*	.11
31.	attend parent teacher assoc- iation meetings.	56	.05	.26
32.	encourage parents to visit the classroom at any time.	65*	.04	04
33.	contact parents whenever any problem arises about their children.	70*	.03	.03
34.	attempt to find out what in the home situation, may con- tribute to the misbehaviour of a pupil.	63*	01	20
35.	discuss freely with parents the weaknesses of other teachers.	.11	.51*	20
	Eigenvalues	2.05	1.47	1.05
	Percentage Variance	20.54	14.67	10.47

Three factors were extracted. Of these, the first is made up principally of items 33 (-.70), 32 (-.65), 34 (-.63), 31 (-.56) and 28 (0.34) which together deal with co-operation between teachers and parents in the education of the child. McNamara described a factor composed of the same items with very similar loadings 'teacher - parent communication'. Here the rather more specific label 'teacher parent co-operation' will be used.

The second factor for this role sector is described chiefly by items 30 and 29 with loadings of .80 and .74 respectively virtually identical to a factor named by McNamara 'reporting intellectual attainment'. Here, however, there are two other contributing variables - item 35 (discuss freely with parents the weaknesses of other teachers) with a loading of .51 and item 28 (visit every pupil's home..) with a .35 loading. This component therefore appears to be describing behaviours at the opposite end of a hypothetical scale of communication between teachers and parents described by factor one, namely, behaviours characterised by a desire to preserve professional <u>distance</u> from parents. Accordingly, the component was named 'teacher-parent distance'.

Principal component three, with significant loadings on variables 27 (.70) and 26 (.67) seemed somewhat similar to the second component in that distance between teachers and parents is also implied. Additionally though, these role norms suggest a recognition of the formal rights of parents, and situations where recourse to formality in relationships might be expected. The factor was therefore called 'teacher-parent formality'.

Summing up, analysis of the role sector dealing with teacherparent relationships yielded three factors which it was thought could be described without undue violation to the original data by such terms as 'co-operativeness', 'distance' and 'formality'. In general, these characterise the <u>communication</u> between teachers and parents.

4.5.7.5 RESULTS: ROLE SECTOR 4 - ACTING TOWARD COMMUNITY

Table 4.5 gives the rotated matrix for the role sector 'Acting Toward Community':

TABLE 4.5

VARIMAX FACTORS FOR ROLE SECTOR 4,

ACTING TOWARD COMMUNITY: 1976 GROUPS

ROLE NORM		FACTORS			
			2	3	
36.	exercise great caution in expressing views outside of classroom on controversial issues	.26	.01	68*	
37.	live within the neighbourhood of the school.	.64*	.01	.09	
38.	be active in at least one community youth group.	.67*	16	11	
39.	attend church regularly.	.55*	33	08	
40.	spend an eight-hour day at school.	.53*	.06	11	
41.	remember that a stricter stan- dard of conductapplies to them because they are teachers.	.43*	11	61	
42.	patronize locally-owned busin- esses and services.	.58*	.16	17	
43.	make political speeches.	.22	.12	.76*	
44.	serve alcoholic drinks	01	.92*	.03	
45.	visit a pub.	01	.91*	.12	
	Eigenvalues	2.52	1.75	1.17	
	Percentage Variance	25.19	17.48	11.7	

Since the factors yielded above are very similar to those McNamara obtained, his labels will be used. The first principal component is made up chiefly of items 38 (.67), 37 (.64), 42 (.58), 39 (.55) and 40 (.53), the first four of which imply some commitment to the community by the teacher and appear to constitute a positive public image for the teacher, while item 40 (spend an eight-hour day at school) is also congruent with the image of the committed community-conscious professional conveyed by the four other variables. The factor is therefore named <u>community</u> supportiveness.

McNamara argues that the second factor, consisting of very high loadings on variables 44 and 45 (in the case of his data .94 for each variable, while here .92 and .91, respectively) together with a small negative loading (in this case) of -.33 for item 39 (attend church regularly) refers not so much to a 'drinks' factor but to a pattern of more liberal and independent behaviour than that described by factor three (below). The undoubted tolerance of alcohol in contemporary Australian society makes this a plausible argument in that whether or not a teacher engages in the behaviours described in the two variables seems, generally, to be regarded as 'his own business'. McNamara's name for the factor 'independent community behaviour' will therefore be used.

The third principal component is described by item 43 with a positive loading of .76 and items 36 and 41 with negative loadings of .68 and .61 respectively and is labelled by McNamara 'correct community behaviour' - 'correct' presumably in the sense that teachers might, generally be expected to adopt a low profile and not actively seek controversy because they are so patently under public scrutiny. That is, there is a perceived expectation that teachers ought to behave with propriety in respect of such public matters.

The primary school teacher's relationship with the community as described by the ten items contained in Role Sector Four can therefore in summary be described by behaviour toward the community marked by 'supportiveness' and 'correctness' (or 'propriety') in most respects and 'independence' in others.

4.6 SUMMARY

In 1976 a replication of the author's original study (Sinclair, 1975) was carried out at Mitchell College of Advanced Education in New South Wales. Primarily, this involved administering a role norm inventory for the position of primary teacher divided into four role sectors (corresponding to important role relationships) to groups of 2nd and 6th semester student teachers, and to their lecturers and co-operating teachers. Each of these groups completed the inventory from the standpoint of their own ideal views about the role of teacher. Additionally, the student teacher respondents completed the inventory from the standpoints of their future expected behaviour in the role of primary teacher and the views they attributed to lecturers and teachers.

To provide a relatively concise framework within which to describe data trends, subjects' responses were factor analysed by role sector using a principal components computer programme. This reduced the 45 original variables to 14 factors from which descriptors were derived to summarise results.

Subsequently role sector-by-role sector analyses were carried out using firstly a multivariate analysis of variance programme followed by <u>post hoc</u> Scheffé tests on group means where the analysis
of variance revealed individual role norm items to be statistically significant contributors to the variance in the entire sector. Other, simpler measures such as the calculation of a mean difference per role norm for all relevant within- and between-group role sector comparisons were also used to facilitate analysis of data trends. Consensus was determined by using the Mahalanobis Distance (MD²) statistic in the first instance to ascertain between-group statistical significance, and then Leik's (1966) Measure of Ordinal Consensus to summarise both within- and between-group trends.

In 1978, when the 2nd semester students of 1976 were in 6th semester, they again completed the four role norm inventories so that changes in role perception could be assessed. This was done by means of an analysis of variance and covariance with repeated measures computer programme. The inventories were also completed by 2nd and 6th semester cohorts of student teachers from Castle Hill College of Education. The data thus gathered were analysed as for the 1976 Mitchell study.

Finally, the 1976 6th semester Mitchell cohort was the focus of exploratory follow-up research. Prior to leaving college they had completed - in addition to the role norm inventories - a Teaching Style Inventory which attempted to assess the 'impress' upon student teachers of significant others in both hypothetical and actual teaching situations, and questions relating to teaching commitment contained in the background information section of what was called the Teacher Training Project Questionnaire. The group was followed up in the schools in 1978, two years after leaving college and asked to complete the first role norm

217

inventory again and a questionnaire aimed at assessing aspects of adjustment to teaching, the intention being to relate this to relevant data collected from the group before leaving college. This information was analysed primarily by cross tabulation. Student sub-groupings pertinent to these analyses were identified by utilizing ranking procedures based on absolute differences between students' own and attributed norms.⁶.

Throughout, care was taken to provide adequate descriptive statistics and, where possible effect-size measures were incorporated in analysis. That is, in accord with suggestions made in the literature, an effort was made to provide a relatively comprehensive picture of the data.

6. These procedures are described in detail in Chapter 8.