DOCUMENT RESUME

ED 261 017 SP 026 233

AUTHOR	Larkin, Anthony I.; Keeves, John P.
TITLE	The Class Size Question: A Study at Different Levels of Analysis. ACER Research Monograph No. 26.
INSTITUTION	Australian Council for Educational Research, Hawthorn.
REPORT NO	ISBN-0-85563-390-5
PUB DATE	84
NOTE	125p.
AVAILABLE FROM	Australian Council for Educational Research, P.O. Box 210, Hawthorn, Victoria ($9.00).
PUB TYPE	Information Analyses (070) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Ability Grouping; *Academic Achievement; *Class Size; Elementary Secondary Education; Student Attitudes; *Teacher Student Ratio; *Teaching Methods

ABSTRACT

The purpose of this investigation was to examine the ways in which class size affected other facets of the educational environment of the classroom. The study focused on the commonly found positive relationship between class size and achievement. The most plausible explanation of the evidence seems to involve the effects of grouping more able students in larger classes, but the findings also indicated achievement gains beyond those expected soley from a consideration of differences in achievement levels. It is clear that an increased understanding of these features of the classroom is a necessary step towards sorting out the relationships between class size, teacher activities and student motivation, all of which appear as central themes in the class size question. (JD)

THE CLASS SIZE QUESTION: A study at different levels of analysis

Anthony I. Larkin and John P. Keeves

ACER RESEARCH MONOGRAPH NO. 26

THE CLASS SIZE QUESTION: A STUDY AT DIFFERENT LEVELS OF ANALYSIS

Anthony I. Larkin
and
John P. Keeves

Australian Council for Educational Research
Hawthorn, Victoria
1984

Published by
The Australian Council for Educational Research
Radford House, Frederick Street, Hawthorn, Victoria 3122

Printed and bound by
Allanby Press Printers Pty Limited
1a Crescent Road, Camberwell, Victoria, 3124

Cover design by Louise Coutts

National Library of Australia Cataloguing-in-Publication data.

Larkin, A.J. (Anthony Ignatius), 1955-.
The class size question.

ISBN 0 85563 390 5.

1. Class size. I. Keeves, J.P. (John Philip), 1924-. II. Australian Council for Educational Research. III. Title. (Series: ACER research monograph; no.26).

371.2'51

CONTENTS

LIST OF TABLES

8

LIST OF FIGURES

PREFACE

This investigation had its origins in the Staffing and Resources Study that was undertaken during the years 1979-1982 by the Australian Council for Educational Research at the request of the Australian Education Council. Initially some members of the Steering Committee for the study and some members of the staff of the ACER were hopeful that a similar study to the one presented in this report might be undertaken as part of the Staffing and Resources Study. However, a variety of factors prevented such an investigation being carried out during the life of the Staffing and Resources Study. Nevertheless, with a substantial body of data available which had been collected a decade or more before, it was resolved that when an opportunity arose, the data should be analysed to determine whether certain issues associated with 'The Class Size Question' might be explored in an Australian setting. Consequently this report is presented as a contribution not only to the development of a greater understanding of matters relevant to the debate on class size, but also to the research methodology concerned with how investigations of such complex issues might proceed, as well as advancing results that have clear implications for policy and practice, even if they are not explicitly stated in this report.

The data examined in this study were collected 15 years ago in a school setting that has changed in significant ways during the intervening years. Therefore, the findings should not be read and interpreted as applying to the schools that exist at the present time within that school system. Nevertheless, it is believed that the propositions advanced in Chapter 10 are likely to have a generality that extends beyond the specific time and the specific location where the data were collected. Indeed, the continued presence of positive relationships between class size and educational achievement which are being reported from studies conducted in natural classroom settings in Australia and in other parts of the world suggest that the phenomena being studied in this report still exist in our schools in spite of the findings to the contrary which have been reported from experimental studies into the concomitants of class size.

In conclusion, we would like to thank the many principals, teachers and students, both known and unknown, in the secondary schools of Canberra who assisted in so many ways with this project during 1969. We hope that they will be interested in the findings reported and the issues of both a substantive and a methodological kind which we have addressed.

The problems of undertaking and interpreting the results of non-experimental research studies are many, and we hope that those who assisted with this investigation will find in the report a contribution towards a better understanding of how the environment of the classroom influences educational outcomes.

John P. Keeves
September 1984

CHAPTER 1

THE CLASS SIZE QUESTION

Research has been used to examine many contentious educational issues. Unlike much research in other areas, both supporters and opponents of a particular educational policy or practice have frequently been able to look towards research and find support for their beliefs. The disparity between the findings of results on even quite simple educational issues has made precise conclusions and the consequent implementation of theory into practice a very tentative proposition. Educators are often trapped between several sets of contradictory research findings or, alternately, are faced with a set of inconclusive results which they had hoped could be used as the basis of a policy decision.

One such issue is that of class size which is currently perceived to be important by large sections of the community. Teachers commonly see it as the key indicator of the quality of their work environment which influences job satisfaction and their perceptions of job effectiveness. To parents it is a very obvious measure of their child's classroom environment. Parents may also be interested in the availability of books or aids and the physical state of the classroom, which are other overt indicators of classroom quality, but class size is a statistic readily available from each child and an immediate and easily perceived gauge of the classroom quality.

Furthermore, for educational policy makers and politicians, the class size question reaches into the heart of budgeting and the funding of schools. It is accepted that approximately 80 to 85 per cent of the recurrent resources provided for education in schools are spent on teachers' salaries and since the number of teachers required to staff the schools varies inversely with the average class size, then any change in average class size will have immense consequences for the required level of educational funding. For example, to reduce the average class size of a school or school system from 25 to 24, approximately four per cent more teachers would need to be employed. Based upon 1982 figures for teachers' salaries, this would cost the Victorian Education Department an additional $44 million each year. The class size issue does not only affect teachers, students and parents, but also clearly influences the priorities that a government sets. As such, it has an effect on all who use the services provided by governments, since increased support for education is likely to be at the expense of some other areas of governmental responsibility.

Despite the concern and interest in the class size question, it remains a controversial area of research, since the findings from studies in many parts of the world have been extremely diverse. Studies in natural school settings (for example, Husén, 1967; Peaker, 1967; Comber and Keeves, 1973) have sometimes found large classes to be associated with superior student outcomes, while other experimental studies have found smaller classes to be more effective (see Glass and Smith, 1978) or have produced inconclusive results.

It is hard to believe that, if other things were made equal, the addition of several more students to a classroom group would raise the average level of achievement of the class. Nevertheless, surveys undertaken in a wide range of settings have commonly reported moderate positive correlations between class size and achievement. However, there could be favourable circumstances associated with large classes which might explain results which differ from conventional beliefs.

These contradictory findings are generally unpalatable to teachers who are convinced that small classes are superior teaching environments for both teachers and students. On one side of the debate, teacher organizations in recent years have been strongly advocating the reduction of class sizes. The strength of their commitment to this issue has been demonstrated by their loss of income through strike action. On the other side, educational administrators, although perhaps not wishing to increase class sizes, are trying to maintain present class size levels in order to contain spending on education, and to use any resources available for what they perceive to be areas of greater need.

It has been perplexing that these different groups could maintain such strong opinions, either based upon intuition or frugality, when the educational research, viewed as a whole on this question, was so inconclusive. The need to clarify the issues associated with class size was recognized by Glass and Smith (1978), who employed the technique of meta-analysis to tease out a definite conclusion from the maze of contradictory findings that had been reported from previous experimental research. Their most quoted finding is a graph relating class size to achievement (measured in percentile ranks). The graph, shown in Figure 1.1, displays the inverse relationship that teachers have been describing for years without consistent support from research. A similar graph was produced when attitudinal outcomes were related to class size, but with a steeper gradient (Smith and Glass, 1979). Their results were widely accepted, although not universally, perhaps because it was the conclusion most intuitively expected. Although Glass and Smith's quantitative summarization of the research could be seen as a definitive statement upon the class size question, the actual interpretations of their graphs have been many and various.

Despite the clear inverse relationship between class size and achievement, the extent of achievement gains associated with reductions in class size for typical classroom situations was quite small. We again refer to the Victorian Education Department example given earlier. Using the Glass and Smith graph, decreasing class size from 25 to 24 would produce an achievement gain of less than one percentile rank. Furthermore, if class sizes were decreased from 25 to 20, the resultant achievement gains would be approximately two percentile ranks while the increase in teacher salary

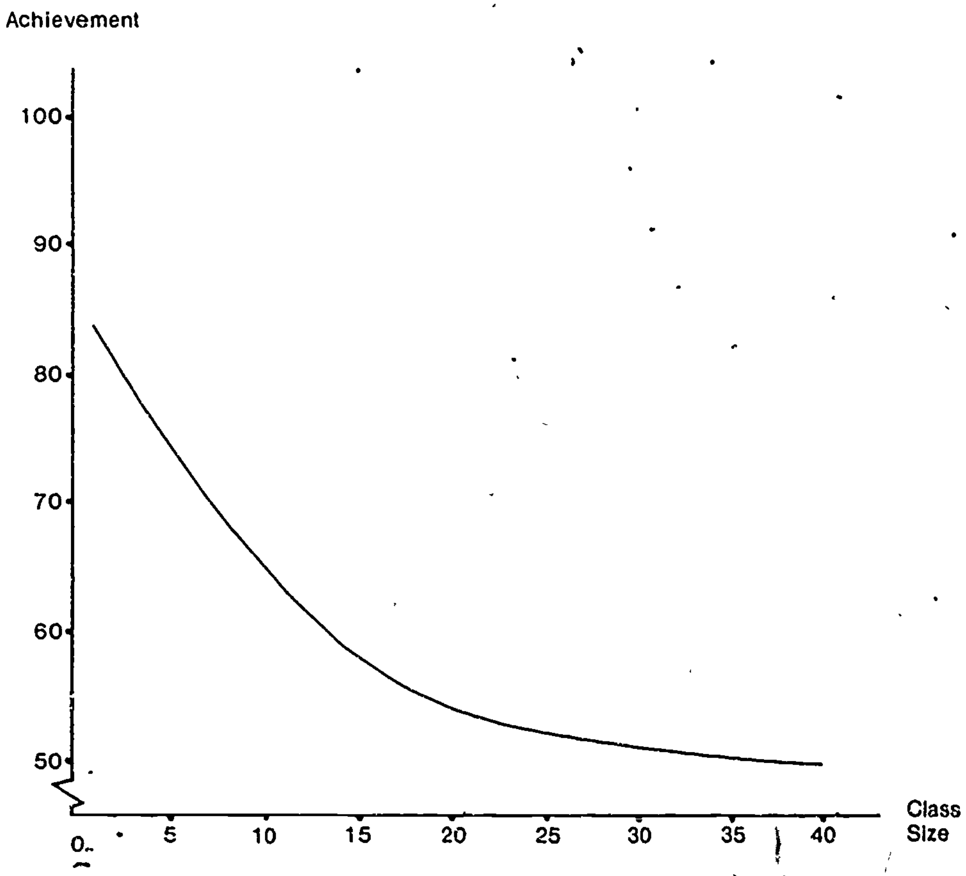

Figure 1.1 Graph Relating Class Size to Achievement Measured in Percentile Ranks From: Glass and Smith (1978)

costs would be nearly $220 million, again using 1982 figures. The obvious question posed by these figures is the relative importance that society places on an expenditure of $220 million or an estimated achievement gain of two percentile ranks in student achievement.

The problem of obtaining good value from the available resources is just as relevant in education as in commerce or industry. This difficulty was recognized by Karmel (1981) when he wrote:

> In schools, class size has become a sacred cow, and pressures for reductions in pupil/teacher ratios have continued in spite of great improvements over the past decade. It may well be that a more effective use of resources would require a trade-off between classroom teachers and special teachers to assist disadvantaged groups or ancillary staff of various kinds, or even a trade-off between primary and secondary teachers. The latest wisdom is that, although very small class size is an effective pedagogic device, small changes in the class sizes which commonly obtain do not produce significant effects (Glass and Smith, 1978). If this is correct, there may be a strong case for allowing the size of most classes to rise a litle so that special groups of children may be taught in very small classes. (Karmel, 1981:27)

This interpretation of the Glass and Smith findings would probably be opposed by teacher unions who could see in the research evidence a further endorsement of their desire for a reduction in the size of all classes.

One obvious problem with converting the Glass and Smith results into actual classroom policy is the global nature of their findings. Their graphs represent an amalgam of class size research so that conclusions specific to a particular instructional situation are no longer possible. Furthermore, their findings are purely descriptive and not analytical in the sense that while they noted that smaller classes had superior achievement levels, they could not provide an explanation as to why smaller classes had produced the higher achievement levels.

As their findings did not isolate which factors associated with smaller classes were responsible for the achievement gains, it is important to direct research towards answering this question. The problem is not just to establish that smaller classes are better, but to ask what things make smaller classes better. Thus, it is necessary to determine which student activities and which teaching practices are evident in smaller classes but absent in larger classes. Perhaps prematurely, we may advance some of the questions that could help to identify these differences. Are students more attentive in smaller classes? Are more experienced teachers allocated to smaller classes? Do students interact with their teachers more in smaller classes? As well as identifying which behaviours differ with class size, it is also essential to determine which of these behaviours might be responsible for differences in achievement levels. The examination of classroom practices should enable us to identify the areas where smaller classes are more effective.

To do this, we have sought to identify those classroom practices, teaching behaviours and teacher characteristics which are significantly different for small classes when compared to larger classes and then attempt to establish which of these factors are associated with superior achievement levels. Regression analysis is the most appropriate technique to examine such relationships and to control for socio-economic level and prior achievement, which have often been neglected as factors influencing the allocation of students to classes in previous non-experimental studies (see Linn and Werts, 1969). After such an analysis, we hope that we can give a more informed answer to the question 'Why are smaller classes better?'

The study of the class size question using regression analysis also raises technical issues concerned with the choice of an appropriate unit of analysis (Burstein, 1980). The class is the obvious unit of analysis since class size is a classroom measurement. However, class size can also be viewed as a variable which each student in the class experiences in the sense that it influences the student's environment. Hence, it is also possible to view the student as an effective unit of analysis. Although many researchers assume that one particular unit of analysis is appropriate and then legitimate their choice, there is some justification for using both units and hence some analyses should be conducted at each level. Classroom practices readily lend themselves to the use of the class and not the student as the unit of analysis. For example, a variable measuring the number of questions asked by students during class-time could not be viewed as a student variable as each student in the class would ask a different number of questions. Consequently, one measurement would not be equally appropriate to all students in the classroom in the analysis of data at the between student level, but would be appropriate as a characteristic of the classroom environment in an analysis at the between classes level.

It is also possible to conduct the analysis at the student within-class level. This might seem irrelevant since the class size measures would have no variation at this level of analysis. However, it is still possible at this level of analysis to examine relationships between the student variables which were found to be associated with class size. It is therefore desirable to consider the effect of within-class variation in one variable upon the within-class variation in achievement or attitudes. Hence, comparisons can be made between the findings at all three levels of analysis, either between classes, between students, or between students within-classes. Any differences in the findings of the separate regression analyses can be informative in the sense that they can provide a further insight into the manner in which a particular practice contributes to achievement. In fact, Burstein (in press) has suggested that knowledge of the unit of analysis is essential in the interpretation of a result.

In summary, we have two main aims in this study. The primary aim is to illuminate the relationship between class size and achievement. We seek to explain why Glass and Smith found that small classes were better by an examination of data relating to classroom practices. The secondary aim is to tease out meaningful relationships at three different levels of analysis. The findings should provide a greater understanding of an important issue for educational policy and practice as well as enabling comment to be made about the appropriateness of the different strategies and levels of analsysis in investigations in this complex field of research.

CHAPTER 2

A REVIEW OF CLASS SIZE RESEARCH

The research into the effects of class size upon educational outcomes has been confounded by uncertainty and inconsistency. Class size can be defined and quantified in many different ways. Definitions of class size differ in terms of how closely they actually reflect a student's or a teacher's classroom experience. Some definitions refer to staffing ratios such as the student-teacher ratio instead of actual measures of the number of students in a classroom. Furthermore, several instructors might be present at one time in a conventional classroom, thus reducing the effective instructional group size for the student. Nevertheless, group size is probably a sound measure of what the student experiences in a classroom since it relates to the amount of the teacher that a student shares. This ratio of students to instructors in a classroom is now the most widely accepted measure of class size, although earlier research used alternative definitions.

Some Issues in Class Size Research

Most class size studies have concerned themselves with the relationship between the number of students in the classroom and their achievement levels in different subjects. In a search for previous investigations in preparation for their meta-analysis study, Glass and Smith (1978) found over 300 documents which pertained to the issue of class size. Although many of these documents were unusable due to a lack of adequate data, there still remained a great body of information concerned with class size. Two recent comprehensive reviews of class size related issues are indicative of the texture of class size research throughout this century. Lafleur, Sumner and Witton (1974) in Australia and Ryan and Greenfield (1975) in Canada have provided a good coverage of the literature, including the extensive international surveys conducted by the International Association for the Evaluation of Educational Achievement in science (Comber and Keeves, 1973) and reading (Thorndike, 1973).

Glass and Smith (1978) noted that the research on class size and its effects upon achievement has gone through four stages. These were the pre-experimental era (1900-1920), the primitive experimental era (1920-1940), the large-group technology based era (1950-1970) and the individualization era (1970-). Class size was not a research issue during the 1940s. However, boundaries of these eras are not rigidly fixed, since studies of one type have occurred in eras largely devoted to another type. While the emphasis in each era probably reflects the educational emphasis of the times, the technology based era produced some massive empirical studies which were used to form national educational policies (Coleman, 1966; Peaker, 1967). Some of these studies surveyed tens of thousands of students, expecting that large studies using recently available computing power could unravel the mysteries of the classroom and the school. More recent research has shown a concern for establishing the benefits of individualization, and experiments have been conducted with radically smaller class sizes in an attempt to extract more conclusive results.

Another important difference between studies is concerned with the assignment of students and teachers to groups. The methods of assignment can be classified as 'random', 'matched', 'repeated measures' or 'uncontrolled'. These classifications are important in describing the degree of experimental control exercised in a study. The meaning of 'random' is clear. 'Matched' means there has been an attempt to equate smaller and larger classes by assigning students to classes using pretest achievement levels. 'Repeated measures' refers to using the same students or teachers in both small and large class situations. 'Uncontrolled' is obvious, but nevertheless important because in many natural classroom situations class size is strongly related to factors which influence achievement and other outcomes.

A further problem in class size research is that many important factors which are themselves related to class size are frequently overlooked. The omission of these factors from consideration provides reasons as to why some findings should be questioned. Many critics of class size research have noted the use of well-controlled measures of class size and achievement but also have noted the absence of any controls upon instructional variables. It is possible for teachers to adjust their teaching practices to the size of the class group such that the research is confounded by an interaction between teaching methods and class size. This suggests that there is an immense range of additional compounding factors which are probably impossible to examine completely or to control. Nevertheless, the majority of studies that have been carried out have controlled only for student ability, using either achievement or intelligence test scores.

Having recognized some of the issues and difficulties associated with definitions, the differing emphases of certain studies, and certain omissions which occur throughout the studies, it is appropriate to turn to the actual findings of the research. To provide a more orderly overview of the major findings we shall present summaries based upon subject taught and age or grade level.

Class Size and Achievement

Two subjects, reading or English and mathematics have been the most popular subjects for the investigation of relationships between class size and achievement. It is also within these two subject areas that virtually all research at the primary school level has been conducted. Science has also been a popular subject for research, but usually at the secondary level. To some extent, it is in these subjects that controlling for prior achievement is most necessary, since they are subjects where knowledge is cumulative and what students learn during one year is influenced by their earlier learning.

The findings of the studies of the effects of class size upon student achievement in reading or English can be summarized as follows.

1 With only a few exceptions, the evidence in favour of small classes has been restricted to reading in primary grades. These findings are typified by Balow (1969) who obtained slight achievement gains from smaller classes at Years 1 through 3. Flinker (1972) found no significant difference at Year 7 for reading. These findings would seem to suggest that an emphasis towards smaller classes was more important at the lower primary school level. The exceptions included studies by Little et al. (1971) and Davie et al. (1972) who found that children in larger classes made better than average progress in reading at infant levels.

2 When the evidence favoured smaller classes at higher year levels, the findings often applied to below-ability students. Woodson (1968) observed that class size made no difference to average or high ability students but that low ability students had superior results in smaller classes.

3 Longitudinal studies would appear to indicate that smaller classes had a cumulative positive effect upon achievement. The findings of Furno and Collins (1967) supported this conclusion. They investigated both regular and special classes over a five-year period in a study that controlled rigorously for many other variables.

4 At the secondary level, the results have either favoured larger classes or shown no significant differences. Ryan and Greenfield (1975) reported several studies where larger classes produced superior reading levels, while the work of Thorndike (1973) has suggested that reducing class size would have a limited impact upon student performance.

Overall, these findings suggest that the advantages of small classes are restricted to those students who are the most vulnerable. Smaller classes appear to aid the very young or those students of low ability. Again, it should be noted that nearly all the studies referred to have failed to control for differences in the teaching methods used by teachers.

Similarly, in studies of the effects of class size upon achievement in mathematics larger classes have been found to be as effective as smaller classes. These general conclusions are supported by several large surveys including those of Coleman (1966) and Husén (1967). Coleman found that the student-teacher ratio was not related to achievement for any group under any conditions. Husén, in an enormous cross-national study, was more tentative about the findings of the investigation and suggested a more detailed examination of what occurs in classrooms. Other studies have, however, obtained opposing results favouring smaller classes, so that the findings still remain uncertain and the issues unresolved.

The work of Moody, Bausell and Jenkins (1972) with respect to specific mathematical objectives found that class size strongly influenced achievement. Classes of size 1 produced superior performance levels to classes of sizes 2 and 5, which in turn produced significantly superior performance levels to classes of size 23. It should be noted that the studies quoted are only a sample of those available and that the effects of class size are well documented but the conclusions are equivocal. Again, it should be remembered that the comparisons were made without reference to teaching methods.

As mentioned previously, studies dealing with other subject areas, notably science, were usually conducted at the secondary school level. The overall pattern of the results suggested that class size was not a major factor influencing achievement at this level of schooling. The major studies concerning science achievement included those of Comber and Keeves (1973) who found that students only performed better in smaller classes in two of the 19 countries studied at the 10-year-old level. Rosier (1973) found that larger classes obtained better science results in Australian schools, but he also noted that larger classes had a higher proportion of students with an academic background. In another Australian study, Keeves (1972) found that the larger the class, the greater the gain in achievement during a year. It was suggested that this could be explained by the observation that: '. . . in larger classes a different standard of control was exercised, with a higher level of industry and more effective teaching.' (Keeves, 1972:210). However, detailed analyses to support these claims were not carried out. Only a small number of studies has endorsed the establishment of small classes for science at the secondary school level.

The findings relating class size to achievement have not been very conclusive. Although there is a lack of evidence supporting large classes, much of the evidence draws the conclusion of 'no difference'. The only areas of education where smaller classes have received general support from a large portion of the research were at the lower primary level and for remedial education where students require the individual attention only available from a teacher in a small class. We could be tempted to conclude that smaller classes have a beneficial effect upon achievement for younger students but that this effect diminishes greatly by the secondary school level. This statement is supported by the results for reading and English where the most significant differences were up to Year 3 and for science classes, usually at only the secondary level, where significant differences between small and large classes were uncommon. As such, a review of previous research does little to clarify the global relationship between class size and achievement except in a few specific situations.

Class Size and Student Attitudes

In addition to achievement outcomes, variations in class size could also influence a range of affective outcomes including attitudes to school and specific subjects, self esteem, teacher satisfaction and academic motivation. Also, variations in class size could affect the opportunities that teachers have had for doing different things in the classroom. Many teachers might not avail themselves of these opportunities. Nevertheless, differing class sizes would seem to influence the workload, the morale and the perceptions of teachers. Furthermore, both a teacher's and a student's satisfaction with school and a favourable affective climate in the classroom must be considered to be desirable outcomes in themselves.

The body of research into the influence of class size upon affective outcomes for students has been more decisive than have been the findings for achievement. The majority of studies in this area have pointed towards the superiority of small classes as a means of enhancing affective outcomes. This result is supported by the following typical studies. Walberg (1969) reported that class size influenced the learning climate as viewed by students and that class size was positively related to student perceptions of formality and diversity but negatively related to perceptions of intimacy and democracy. Welch (1971) could not predict an optimum class size but he was able to show that small classes displayed elements of cohesiveness, satisfaction, goal directedness and democracy. Keeves (1972), in the same study as cited earlier, found a negative relationship between class size and attitudes to science. Lindbloom (1970) reported that individualization, group activities, interpersonal regard and creativity all decreased as class size increased, especially at the primary level, but also for students over 15 years of age. However, Haskell (1964) reported that attitudes towards mathematics were not significantly related to class size. These results would seem to indicate a commonly observed, although not universal, inverse relationship between class size and affective outcomes. This general trend suggests that a reduction in class size would be beneficial to both affective outcomes and the teaching process itself.

Class Size, Teacher Attitudes and Teaching Practices

Another area of class size research is that of teachers' attitudes and job satisfaction. Although this does not directly relate to student outcomes, it must be regarded as a critical aspect of the class size research. Studies in this area include the investigation undertaken by the National Education Association (1969) in the United States which found that teachers ranked class size as their second largest source of problems. A survey by the Queensland Teachers Union (1972) also found that class size did affect workload, and consequently, teachers' job satisfaction.

As well as dealing with achievement and affective outcomes, and the attitudes of teachers, class size research has also examined the more relevant issue of classroom practices. The usual means of measuring classroom practices is for an observer to sit in the classroom and record the events that occur on an observation schedule. The findings have differed greatly between studies partly because different observation schedules have been used. Ryan and Greenfield (1975) have summarized the findings from such studies, many of which were conducted at Teachers College, Columbia University.

In particular, Vincent (1967) developed a measure called 'Indicators of Quality' dealing with teaching-learning procedures. Four main categories were noted - individualization, interpersonal regard, creative expression or divergency of thinking, and group activities. The scale was intended to measure attempts to accommodate individual differences in student growth, generate behaviour associated with warmth and respect among students and teachers, encourage the expression of intelligence in different ways, and facilitate group interaction to aid learning and improve social skills. In the following year, Vincent (1968) applied this scale to assess the quality of instruction in over 4000 classrooms at both the primary and secondary levels. At the secondary level, the quality of instruction deteriorated when classes rose above 16 students, while at the primary level, three distinct levels of instruction were observed. Classes below 16 students scored well on the measure, classes up to 25 students did not rate as well, and classes over 25 students received the poorest quality of instruction ratings.

Olson (1970; 1971) applied a similar measure in a larger study and confirmed that smaller classes scored better on the 'Indicator of Quality' scale, but he also noted that the subject being taught and the style of activity were more important factors in determining the quality of instruction. McCluskey (1978) warned that the data did not indicate that class size by itself governed quality. The type of educational activity appeared to be the determining factor. Some activities, such as discussion, consistently received higher quality scores than others, such as lecturing, irrespective of class size. Alternatively, Shapson et al. (1978) found that class size led to very few changes in the functioning of the classroom. No significant change in the level of student participation was noted. As class size increased, the amount of individualized interaction between students and teachers decreased because the teacher's time was distributed more sparsely amongst the students in the class.

Another significant observation was concerned with the number of teachers who failed to exploit the alleged opportunities available in smaller classes. The high rate of mass-oriented instruction and the uniformity of instruction for all students in small classes was cited as evidence. It seems that reducing class size does not lead to dramatic changes in teaching behaviour and that the teacher's own style is a more important determinant of classroom activities than class size. It has also been suggested that teachers have developed skills and techniques applicable to classes of 20 students or more and have not been taught how to optimise the opportunities provided by a small class situation. Despite these criticisms, the general tenor of the research relates smaller classes to increased individualization.

Summary of Research Findings

In summary, class size research is very inconclusive. Those studies which have concentrated upon academic achievement as a criterion were generally divided in their conclusions except for several select areas, notably very young and remedial students where smaller classes were found to be beneficial. Other studies involving educational process variables have generally supported class size reductions. Several important educational outcomes including attitudes, values and psychological growth measures were generally found to improve in smaller classes. In addition, teachers appeared to be virtually unanimous in the belief that they were more effective in smaller classes.

The Work of Glass and Smith

It was towards this large body of research, which they described as 'overly selective and insufficiently quantitative' that Glass and Smith (1978) directed themselves. Their specific aim was to draw a general conclusion which would incorporate all the findings that Porwell (1978) labelled as contradictory and inconclusive in a review of the available studies.

Initially, Glass and his colleagues directed themselves towards the academic achievement issue. They isolated approximately 80 studies which considered the class size-achievement relationship. Then they used these studies to form comparisons between classes of different sizes such that one study could provide many comparisons if it contained groups of several sizes. Seven hundred comparisons were available for examination. A standardized difference between the achievement levels of the two classes in the comparison was formed. Using regression analysis techniques, the 700 comparisons were combined into a single curve relating class size and achievement. The curve implied that in terms of achievement, there were 33 percentile ranks between the level of achievement of an individually taught student and that of a student taught in a class of 40.

When forming their comparisons they noted the grade level, the subject taught, the ability of students and the level of experimental control exercised in allocating teachers or students in the original study. Neither the grade level, the subject taught nor the ability of the students affected the basic relationship, although smaller classes were found to be slightly more beneficial at the secondary level. The only factor to alter the curve significantly was the level of experimental control in placing students or teachers in small or large classes. About 100 of the comparisons came from well-controlled studies. The curve obtained using poorly-controlled studies where no control over students or teachers was evident provided an inverse relationship again, but the relationship was much weaker than for the controlled studies. Glass and Smith concluded that more was learnt in smaller classes, regardless of the circumstances.

In the following year, Smith and Glass (1979) applied the same techniques to non-achievement outcomes including classroom processes, assessment of learning environments, student attitudes and behaviour as well as teacher satisfaction. About 70 studies provided over 300 comparisons. Since the integration of a wide range of affective outcomes on a single scale might have seemed unusual, they also performed separate analyses with affective outcomes for students, affective outcomes for teachers and instructional environment effects. Using all the instructional and attitude measurements, they again constructed a single inverse curve between class size and non-achievement outcomes. The effect of reducing class size was more pronounced for non-achievement outcomes than for achievement outcomes. The difference in non-achievement outcomes between a student in a class of 1 and one in a class of 40 was 46 percentile ranks. In all three specific categories, the same inverse relationship was exhibited although smaller classes were most influential upon teacher effects and least influential upon instructional effects. The improvement in non-achievement outcomes arising from decreasing class size was most evident at the primary level, contrary to the achievement outcome result. Also, in contrast to achievement outcomes, uncontrolled studies were more supportive of smaller classes than were well-controlled studies.

From these two studies, Glass and Smith concluded that class size influenced the classroom environment, and student and teacher attitudes. Perhaps independently, or perhaps because of these relationships, smaller classes were also associated with improved achievement. Although the scope of the research and the simplicity and general acceptability of the findings have given these studies considerable appeal, several conclusions would appear to contradict some previous findings. In the earlier research, one instance where smaller classes apparently led to achievement gains was for younger children, yet Glass and Smith found that smaller classes had a larger impact upon achievement at the secondary level. Why was the inverse relationship noted for younger students, yet missed for older students when, according to Glass and Smith's findings, it should have been more obvious? Like some earlier research, they established a relationship between certain classroom practices and class size but did not suggest the consequences for achievement outcomes that would flow from the varying incidence of certain practices.

The implications of Glass and Smith's work are widespread in terms of the direction that educational policies might move. Consequently, the work has attracted criticism. The most extensive critique of their work was conducted by the Educational Research

Service (1980), an 'independent, non-profit' American group, strongly supported by funds from school administrators. Their criticisms included the following points.

1 The shape of the graphs was influenced by a significant proportion of comparisons involving atypical class sizes below five students.

2 The 100 well-controlled comparisons came from only 14 studies such that the data base was not as extensive as suggested. Furthermore, only six of their 14 studies dealt with typical school situations.

3 The methods used hid the distinctions made in specific class size studies since statistically diverse data were combined.

Despite these criticisms, the simple manner in which Glass and Smith presented their findings meant that their results were more accessible and appealing than an enormous collection of inconclusive research findings that had previously served as a summary of class size research. Furthermore, their findings have given substance and quantity to the beliefs of most teachers. This is something that the previous research was unable to do.

Glass et al. (1982) answered their critics in several areas. They repeated their analyses with classes of one student omitted. This was done since it was claimed that very small classes biased the curve. The shape of the curve remained unchanged. They defended the small data base by emphasizing that the poorly controlled studies also supported smaller classes although not as strongly as the well-controlled studies. Also, they denied that their methods hid distinctions between different findings but claimed that by systematically classifying the findings, the literature had been clarified, not fragmented.

The problem is now to propose the direction in which class size research should head as a result of Glass and Smith's work. We do not want to return to another cycle of inconclusive studies. In one sense, Glass and Smith may be seen to have put the issue to rest provided their results are correct. Obviously, well-controlled studies are needed to replicate their findings. In addition, the question must now turn from 'Are small classes better?' to 'Why are smaller classes better?' To answer this question, research workers must return to the classroom.

Class size could influence what goes on in the classroom, what teachers do, how they handle students, how attentive students are, what activities students participate in and how students behave. These differences in classroom practices, in turn, could influence outcome measures like student achievement and attitudes. We have seen that reductions in class size provide opportunities for individualization, but to what extent and in what ways do teachers exploit these changes? Furthermore, which of these changes are worthwhile in terms of achievement and attitude gains? Are all the facets of individualization that are possible in smaller classes beneficial or are only certain techniques advantageous in smaller classes? These questions all relate to the idea that certain behaviours can occur in smaller classes which can lead to outcome gains. It appears that class size research must return to the classroom again, examine classroom practices to understand further the mechanisms operating in smaller classes in the hope of identifying the causal thread that runs between smaller classes and gains in educational outcomes.

CHAPTER 3

DESCRIPTION OF DATA

The effects of class size upon classroom practices and achievement and attitudinal outcomes have been examined in this report using the data collected in the study Educational Environment and Student Achievement (Keeves, 1972). Keeves' study provided suitable data since the general aim was to investigate the relationships between various measures of the student's home, classroom and peer group and the student's performance at school. To this end, measures of the student's achievement and attitudes were recorded as well as measures of the classroom environment, including class size and a wide range of teaching behaviours and classroom practices. Keeves' study has provided a set of data which were relevant to our proposed inquiry into class size and classroom process relationships.

The details of the study and the scales and procedures used in the measurement of all variables have been recorded elsewhere (Keeves, 1972; 1974a; 1974b). Thus, it is only necessary to give a brief description of the investigation. The study was conducted during 1969 in the Australian Capital Territory (ACT) and concentrated upon students who were entering Year 7 for the commencement of their secondary schooling. Year 7 was chosen as the year level for investigation as the beginning of the secondary school represented a marked change in subject content and the students' learning experiences. The study observed students in 72 classrooms and for the investigation of relationships with achievement and attitudinal outcomes, complete data were available on a total of 1986 students. It should be noted that a smaller simple random sample of 231 students was used in a more intensive investigation into home background. The 72 classes and 1986 students were drawn from a total population of 76 classrooms and 2348 students in 15 schools. The subpopulation on which complete data were obtained thus represented almost all Year 7 students in the ACT during 1969, but because of the unique nature of the ACT population, should not be considered to be fully typical of the entire Australian Year 7 population. The classrooms used in the investigation covered a range of class sizes from 15 to 45 and the distribution of class sizes is presented in Table 3.1 and displayed graphically on Figure 3.1. This represented a wider variation in class size than would be available in actual classrooms at the present time.

In 1969, all the students took an initial science test at the beginning of the secondary school year. At the end of the same year, the students were again tested in mathematics and science, but they were also asked to complete a general information booklet in order to obtain information on the student's home background, and an attitude questionnaire concerned with attitudes to both science and mathematics, liking of school and school learning, academic motivation and self-regard. To gain information on the

Table 3.1 The Distribution of Class Sizes: Canberra Year 7, 1969

Grouped Data	Cell counts					N	Comments and summary statistics
15-19	15	16	17	18	19		Lower extreme: 15
	1	1	1	-	2	5	
20-24	20	21	22	23	24		
	1	-	1	-	-	2	
25-29	25	26	27	28	29		Lower hinge: 29.6
	-	1	3	1	3	8	
30-34	30	31	32	33	34		Mean: 32.7
	7	5	3	2	10	27	Median: 33.4
35-39	35	36	37	38	39		Mode: 34
	5	6	3	6	2	22	
40-44	40	41	42	43	44		Upper hinge: 37.0
	5	1	-	-	1	7	
45-49	45						Upper extreme: 45
	1					1	
Total number						72	Standard deviation: 6.3

classroom environment, observations in science and mathematics classrooms and interviews with teachers were carried out during the second and third terms of the year. A preliminary program of classroom observation was conducted during the previous year to provide practice in the use of the classroom observation schedule so that more reliable and valid data could be collected. Furthermore, Keeves (1972) reported that the tests for science and mathematics as well as the attitude questionnaire were field tested and found to have satisfactory levels of reliability. A complete list of the variables employed in the investigation and a description of the scales and methods of measurement for each variable are provided in Appendix I. Reliability estimates, wherever appropriate, are reported in Appendix II.

The measures of the classroom learning environment were divided into the structural and process dimensions. The structural dimension included measures of teacher's age, sex, training and experience. For the school, the structural dimension included measures of class size and time allocated to instruction. For the students, the structural dimension measured factors like the socio-cultural background of the home, the level of ethnicity and the time spent on homework. As such, the structural dimension recorded the characteristics of the teacher, the school and the student. These measures are relevant to our study since they can be used to address questions like 'Which teachers are allocated to small classes?' or 'Do socio-culturally disadvantaged students receive instruction in small classes?'

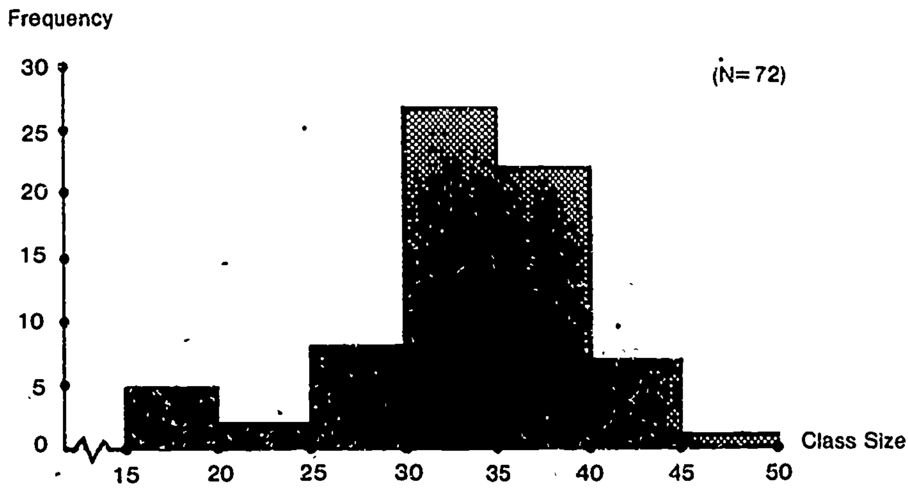

Figure 3.1 Histogram of Class Size Distribution: Canberra Year 7, 1969

For the process dimension, Keeves used previous research and, in particular, the review by Rosenshine (1971), to help identify seven different areas into which classroom practices could be grouped. The areas were as listed below.

1 Achievement press. This area involved the emphasis that the teacher placed upon study and achievement, and typical variables included measures of the regularity of testing and the checking of the students' work.

2 Independent study and inquiry. This area was concerned with independent work being done by students and the amount of choice they had concerning their work. These measures were intended to record the level of individualization occurring in the classroom. The findings of previous research suggested that individualized instruction should be more prevalent in smaller classes.

3 Work habits and order. These variables recorded the teacher's efforts towards developing a systematic and businesslike approach by students towards their work such that the classroom became an orderly working environment.

4 Warmth and affiliation. The specific processes which were equated with teacher warmth included the incidence of supportive statements to students and the occurrence of sympathetic laughter in the classroom. Previous research has suggested that smaller classes would show greater degrees of teacher warmth since these classes were noted as being more friendly and cohesive.

5 Stimulation for learning. Items that were typical of a stimulating classroom environment included measures of the variety of instructional materials and the diversity of teaching activities. Previous research had suggested that smaller classes provided increased opportunities for a wider range of activities.

6 Use of language. These variables assessed the presence of different forms of communication between the teacher and the students in the class.

7 Academic guidance. The facets of academic guidance which appeared as process variables in the classroom included the amount of feedback given to the students concerning their work, the time spent on actual learning, the time devoted to home study, and the time spent upon formal revision. These variables were intended to assess the types and amount of interaction between teachers and students with respect to their classroom work.

It was considered that the areas into which the classroom processes were divided covered adequately the range and types of behaviours and activities that occurred in the classrooms. This study has sought to detect if differences in class size led to varying incidence of these practices.

Hence, Keeves' data provided information on both mathematics and science teachers, and mathematics and science classrooms. Furthermore, it provided an extensive description of the activities occurring in the classrooms. The observation schedules for both types of classes were very similar with minor variations which allowed for differences in the activities which took place in these subjects. The observation schedules were field tested before the observational work commenced. In addition, following analyses of the data at the conclusion of the classroom observation phase, items of low reliability or a high degree of skewness were deleted. The remaining process variables that have been used were considered to have satisfactory reliability. The above summary of the process variables is intended only to provide examples of the variables relevant to each area and to indicate the general flavour of teacher and student behaviours that were recorded.

The data available from Keeves' study provided an opportunity to investigate the many facets of the class size question in relation to educational outcomes. The presence of a relationship between teacher characteristics and class size would indicate the types of teachers who were allocated to smaller classes. The characteristics of the students in the classes would indicate the types of students who were allocated to smaller or larger classes. Recognizing that remedial classes were usually small, it could be anticipated that less able students would be assigned more frequently to smaller classes. The variation in certain classroom process variables as class size changed would assist in providing answers to such questions as the level of individualization in smaller classes, the amount of teacher-student interaction in classes, the level of support provided by teachers, the use of educational materials, and the type and frequency of assessment procedures. All these measures would help to identify the differences in student and teacher behaviour between large and small classes. By undertaking analyses to examine these practices and these effects the reasons for the alleged superiority of smaller classes should be better understood and explained.

CHAPTER 4

CLASS SIZE RELATED MEASURES

The first step in the study of class size which involved measures of the classroom environment and educational outcomes was to identify those variables, either from the structural or process dimension which varied as class size changed. Since the majority of the variables were relevant at the classroom level, the class was considered as the appropriate unit of analysis. This was because the majority of measures referred to behaviours by teachers or activities occurring in the classroom, both of which were relevant to the class as a whole. Other measures referred to characteristics of the teacher or the classroom (e.g. materials available) which were again relevant to the entire class, not just to a particular student. The choice of the class as the unit of analysis is consistent with several major studies noted by Burstein (1980) although there is still considerable debate concerning the best unit of analysis for a given situation. Some of these issues will be discussed later when comparisons are made between the results obtained from the use of different levels of analysis.

To identify the variables which were strongly related to class size, either classroom measures (if appropriate) or class averages for those variables collected at the student level have been used. The latter variables refer either to measures of the home background or to the achievement and attitude measures collected from students through tests and questionnaires and the results have been aggregated to form class averages. Because of the large number of measures, a sifting procedure was developed to determine which variables were related to class size. The product moment correlation coefficient between each variable and class size was calculated and all variables with a statistically significant correlation at the 10 per cent level were noted. Although the data provided population measures, and although statistical significance is only relevant where a sample has been drawn, the level of statistical significance associated with 72 classrooms was considered suitable to screen the list of variables and to identify those related to class size. The correlation coefficients for all variables examined in this investigation in this way have been recorded in Appendix II. In this chapter, apart from the correlation coefficients for the aggregated variables reported in Table 4.1 all correlation coefficients referred to in the discussion have been recorded in parentheses in the text.

A further check was conducted to establish whether the relationship between the variable and class size was either linear or curvilinear. This was done by dividing the 72 classes into six groups ranging from very small to very large classes and then undertaking an analysis of variance for each variable with respect to class size. This was considered necessary because any variable found to be related to class size would be used in

Table 4.1 Correlation Coefficients between Aggregated Variables and Class Size Remaining after Screening

Number of classrooms = 72	Mathematics	Science
Positively related to class size:[a]		
Father's occupation	0.35	0.35
Student's intended occupation	0.57	0.57
Number of hours of homework per week	0.40	0.39
Student's expected level of education	0.60	0.60
Academic motivation	0.25	0.25
Attitude to school	0.30	0.29
Science prior achievement	-	0.57
Science achievement	-	0.61
Mathematics achievement	0.59	-
Participation in maths/science activities (NS)	0.17	0.18
Attitude to science (NS)	-	0.17
Attitude to mathematics (NS)	0.17	-
Negatively related to class size:		
Ethnicity of home	-0.22	-0.20
Participation in pop culture activities	-0.31	-0.31

[a] All correlations are significant at the 10 per cent level for 72 classrooms except where indicated by (NS).

subsequent analyses using regression procedures, and the existence of curvilinear relationships could confound the analyses.

Table 4.1 records the variables that were found to be related to class size using both correlation and analysis of variance procedures. It should be noted that mathematics variables were related to mathematics class size and similarly for science variables, while the remaining variables were related to the class sizes for both subjects. All these variables are class averages since the data were collected from the students and were aggregated to the classroom level before the correlations were calculated. The last three variables in the list (indicated by (NS)) were related positively, but not significantly to class size. Nevertheless, they were included in subsequent analyses because of their believed importance for achievement as suggested by the research previously reported by Keeves (1972).

While recognizing that the screening analysis was using correlation techniques only to refine the variable list, the initial impressions given by this table were perplexing. The results suggested that larger classes were associated with superior achievement and stronger affective outcomes. This was contrary to much of the previous research. These results might be explained by the relationships reported for both science prior achievement and father's occupation, which were both positively correlated with class size. This suggested that more able students were being placed in larger classes. It was considered that only from a regression analysis where either prior achievement or socio-economic status could be allowed for would a more accurate assessment of the effects of class size be possible. The observation that attitudes to the subjects of mathematics and science were not significantly related to class size, while many other attitudes were, was perhaps an early indication that an increase in class size might change attitudes from being favourable to less favourable.

The correlations between the structural variables and class size for science and mathematics classes are recorded in Tables A.4 and A.5 respectively in Appendix II. For teacher characteristics, only one variable was related significantly to class size. In mathematics, there was a positive relationship between class size and the teacher's membership of a Mathematical Association (0.24). This result would appear to be trivial unless it was indicative of the fact that more competent teachers were assigned to larger classes. A lack of relationships between teacher characteristics and class size is disappointing but not unexpected as teacher characteristics have a long history (Gage, 1963; Rosenshine, 1971) of being poor predictors in classroom research. No science teacher variable was related to class size.

For the school or classroom variables, the only variables from the structural dimension which were related positively to class size were the time on homework (Science: 0.44; Mathematics: 0.43) and total time studying science and mathematics (Science: 0.37; Mathematics: 0.36). Negatively related to class size in mathematics classes were the proportion of students from foreign language homes (-0.23), the number of regular teachers (-0.34) and the number of teachers for the year (-0.31). Clearly larger classes were given more homework, but this could have been because they contained more able students, as noted earlier, or this might have been due to other factors including class size. The finding that students from foreign language homes were more evident in smaller classes is consistent with the degree of ethnicity result found earlier, but it may be another consequence of less able students being assigned to smaller classes. The result that more teachers shared smaller classes on a regular basis is harder to explain. Perhaps teachers shared the responsibility of remedial classes which were usually smaller than normal classes. Again, these results could be better explained using regression analysis since the effects of prior achievement could be controlled statistically.

To consider the process variables more readily, the variables related to class size will be treated under the seven categories originally employed by Keeves (1972). Since the teaching practices occurring in mathematics and science classes varied considerably, so too would the processes which were related to class size in each subject. Therefore, separate reference is made to mathematics and science classes wherever necessary. The correlations between the process variables and class size for mathematics classes are recorded in Tables A.7 and A.9 and for science classes in Tables A.6 and A.8 in Appendix II.

1 Achievement press. For both mathematics and science, positive relationships were found between class size and various forms of assessment, in particular, emphasis upon the satisfactory completion of homework (Science: 0.22; Mathematics: 0.31) and the frequency of reports being sent to parents (Science: 0.26; Mathematics: 0.26). In mathematics, the range of assessment procedures used was also related positively (0.24) to class size. For both subjects there were no achievement press variables negatively related to class size. It would appear that larger classes had specific forms of assessment in each subject, either extended answer questions (0.26) in mathematics or short answer (0.29) and multiple choice questions (0.25) in science. Moreover, in larger classes, homework was more strictly monitored and official reporting more frequent. It would be interesting to know if these behaviours were a consequence of class size or the students' prior ability levels.

2 Independent study and inquiry. Only one variable in this area was related to class size. For mathematics, the frequency of invitations to students to participate in an investigation or inquiry (0.21) increased with class size. The absence of any negative correlates was surprising since previous research had indicated that smaller classes provided increased opportunities for individualized instruction. Furthermore, the original study (Keeves, 1972) included measures to detect individualized teaching styles. These measured the diversity of teaching methods as well as the frequency with which students were encouraged to act with autonomy. Clearly there was very little evidence to suggest that the teachers were exploiting any differences in class size in spite of the fact that there was a considerable range of class sizes reported in the study. In addition, those instances of independent study and inquiry that were found to vary with class size were more prevalent in larger classes and not smaller classes as might have been expected.

3 Work habits and order. For science classes, as class size increased there was greater usage of a pupil note book (0.32) and the teacher used less rebuke (-0.39) of any type to maintain order. The use of a pupil note book could well be a teaching strategy to maintain order in large classes. For mathematics classes, five work habit measures were related to class size. In larger classes, teachers emphasized the correct recording of homework (0.31), used one textbook extensively (0.56), and asked students to consider their work habits (0.20) more carefully. The time spent writing (-0.35) and the amount of wasted time (-0.36) were less in larger classes. The greater number of significant correlates for mathematics, when compared to science, suggested that science teachers emphasized work habits independently of class size while mathematics teachers in larger classes concentrated upon homework, good work habits and not wasting time. It is difficult at this stage to detect if these findings were a consequence of class size or student ability, but it was consistent with Keeves' (1972) observation that teachers of

larger classes sought a higher level of industry, and appeared to teach more effectively. It would seem that larger mathematics classes were more orderly and work oriented.

4 Warmth and affiliation. These variables assessed the warmth and level of affiliation in the classroom environment and the degree of encouragement shown by the teacher. For mathematics, teachers in larger classes gave more praise (0.3?) and less rebuke (-0.28) and were more supportive (0.29) of their students. This was contrary to the research which suggested that smaller classes would be more friendly and more supportive. However, these results might be related to prior achievement. Interestingly, all forms of laughter were more prevalent in larger classes (Science: 0.19; Mathematics: 0.32). This seems to suggest another explanation as to why larger classes were related positively to many variables examined in this study. The presence of more students in larger classes might simply have led to more interaction occurring because a larger number of students could ask more questions, generate more laughter or prompt more supportive statements. This seems to imply that more behaviours were observed as a direct product of more students being present. Alternatively, it was possible that mathematics classes might need to exceed a certain critical size to develop a coherence and an atmosphere of lively interaction.

For science, teachers seemed to reduce both other rewards (-0.21) and rebuke (-0.39) and make less positive support statements (-0.17) as class size increased. It appears that less interaction associated with warmth and affiliation occurred as class size increased. The nature of the relationships between warmth and class size differed greatly between mathematics and science classes. This difference is best exemplified by the number of positive support statements which increased with class size in mathematics (0.29) but decreased with class size in science (-0.17).

5 Stimulation for learning. This domain sought to assess the diversity of activities in which the students and teachers participated. It was hoped that it would record the increased opportunities for using different teaching materials and teaching styles that the research had suggested were available once class size was reduced. For science, larger classes used a second textbook more frequently (0.31). For mathematics, larger classes saw more television (0.42) and had a greater number of activity changes (0.19). Such a small number of variables which were correlated significantly with class size would seem to suggest that teachers were not varying their instructional methods greatly as class sizes changed.

6 Use of language. In this area of teaching practices an attempt was made to identify the different types of oral interaction occurring in classrooms. Only one variable was found to be significantly related to class size. In mathematics classes, more time was spent upon question and answer sessions as class size increased (0.40). One plausible explanation has already been considered. This result might be a direct

consequence of class size or simply a product of more students needing more time to ask more questions. Alternatively, it might be that in classes greater than a certain critical size it was possible to conduct more successful question and answer sessions because of the range of views present.

7 Academic guidance. This area referred to the activities that teachers employed to guide their students in how to use their time and how to tackle their work. For mathematics, the amount of homework set (0.41) and discussed in class (0.42) the number of questions asked by both students (0.26) and teachers (0.35) and the number of invitations to students to participate in academic work (0.19) were all related positively to class size. The frequency of revision homework (-0.21) and the number of teacher-student contacts (-0.45) decreased as class size increased. The variables mentioned here reiterate some earlier findings. Homework was again taken more seriously in larger classes. There was more questioning in larger classes but the questioning was distinct from teacher-student contacts which decreased with increased class size. This would seem to indicate that larger classes were more structured with direct questioning between teachers and students while smaller classes were less formal. For science, a distinct set of variables were found to be of interest. Large classes spent more time on revision (0.26), but like mathematics classes, there was less contact between teachers and students (-0.41). The academic guidance category included many measures of homework and specific teaching behaviours. It was interesting to find that many of these practices were not related to class size.

As well as noting those practices which varied with class size, it was also interesting to observe that many of the variables which the previous research had predicted would vary with class size did not appear to do so. This latter list included the amount of laboratory or small group work in science. In addition, the amount of feedback given to students and the amount of mass oriented instruction did not vary with class size. Many of these teaching practices were believed to vary with class size yet such variation was not detected in this investigation. Although the effects of prior achievement were uncertain, prior achievement certainly influenced the allocation of students to classes. Consequently, it could be claimed that teachers adjusted their teaching styles on the basis of the achievement level of the class, not the class size. This could be true for some variables, but many of the variables, particularly those relating to increased individualization, should have been more frequent in both less able and less numerous classes if the previous research findings were sustained. Since this was not found to be true, the effects of prior achievement would not appear to explain adequately the results for the exclusion of class size as an important factor.

Although the number of variables which were related to class size was smaller than expected, a significant set of variables was provided for further examination by the screening process reported in this chapter. The effects of these variables and class size upon achievement outcomes could only be examined using regression analysis. The use of this strategy of analysis was necessary to control for the strong relationship between prior achievement and class size which has been reported above. Only in this way could an investigation of the effects of the differing mechanisms occurring within large and small classes be undertaken effectively.

CHAPTER 5

CLASS SIZE AND CLASSROOM PROCESSES

The Simple Causal Model

In Chapter 4, the home background, structural and process dimension variables which were related to class size were identified, and it was possible to consider which of these variables might give rise to changes in either achievement or affective outcomes. Furthermore, it was argued that to examine the relationship between class size and educational outcomes, it was necessary that the influence of prior achievement should be controlled.

In this chapter, we shall consider the class as the unit of analysis. The variables under examination may be classified as either teacher or classroom characteristics. Classroom characteristics is a general label which includes the structural, process and attitudinal measures associated with a particular classroom. This broad area includes the use of educational materials, the practices and behaviours of teachers and students and the attitudes that students have towards themselves and their work.

Teacher characteristics exist before a class is formed whereas classroom characteristics are a product of the combination of a particular teacher and a particular class. This distinction is important in the construction of a causal model to guide the analyses that must be carried out since the paths should indicate the directions of the causal and temporal relationships between the variables.

An examination of teacher characteristics and their effects upon achievement has been shown to be largely unprofitable in previous research. Incorporating class size as an additional variable in this situation would be unlikely to change the nature of these findings since very few teacher variables have been found in this investigation to be significantly related to class size. Those that were found were either not linearly related to class size or not worthy of further consideration due to their largely trivial nature (e.g. membership of the Mathematical Association). The classroom characteristics were, however, of greater interest, and a number of these variables were related to class size. A causal model relating class size, classroom characteristics and achievement outcomes was developed from a consideration of previous research discussed in earlier chapters. It is presented in the path diagram in Figure 5.1.

The prior effects label was used in preference to prior achievement since a pretest measure was only available in science. For mathematics, the most suitable prior effects measure was an index of social background. Its predictive power would clearly not be as strong as prior achievement but its use was necessary in the absence of other measures. For the sake of providing comparisons between science and mathematics classes, the

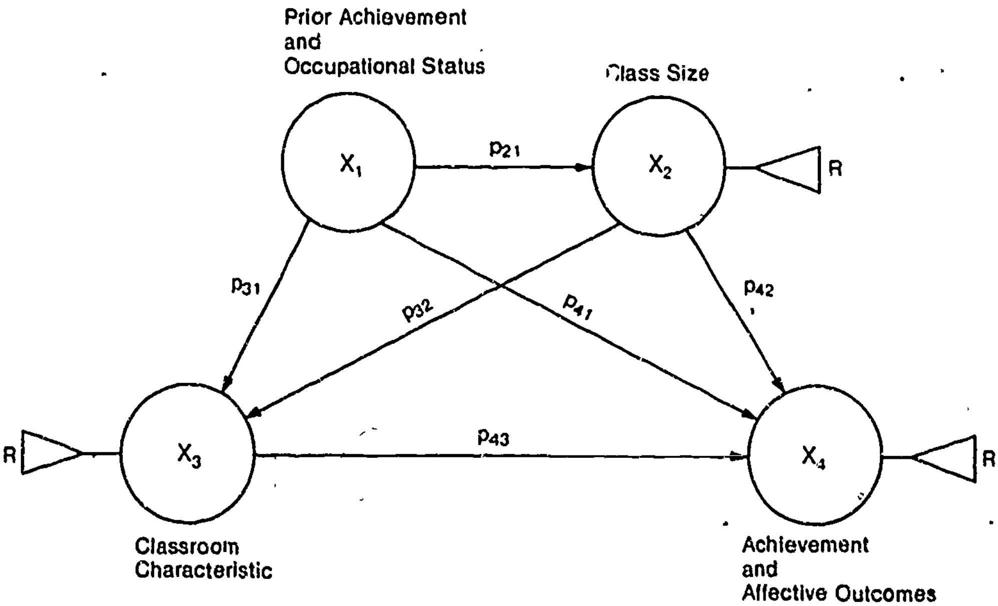

Figure 5.1 Causal Model Relating Class Size, Classroom Characteristics and Educational Outcomes

causal model has also been developed with father's occupational status as a prior effect in science. Four relevant educational outcomes were available for consideration: achievement in mathematics and achievement in science, attitudes towards mathematics and attitudes towards science. All these measures were class averages, derived from achievement tests or attitude questionnaires which were answered by the students under survey. Between prior effects and the outcomes were the mediating effects of class size and those classroom variables which were related to class size.

To facilitate an easier presentation and discussion of the causal model, the findings for each combination of subject, prior effect and outcome are given separately. Achievement outcomes will be presented first. To assess the importance of any relationship associated with the causal model, a path coefficient of at least 0.10 was chosen to indicate a substantial relationship between two measures. Any path coefficient below this value was considered trivial as it would explain less than approximately one per cent of the variance of the dependent variable. The establishment of such a criterion was necessary since a population and not a sample was under survey.

The results for the six separate regression models for varying classroom characteristics are given in Tables 5.1 to 5.6. The path coefficients presented in the tables were obtained using the SPSS Regression program (Nie et al., 1970).

Science Achievement Controlled for Prior Achievement

The path coefficients obtained from the analysis of science achievement controlled for prior achievement are recorded in Table 5.1 for all classroom variables related to class size. These path coefficients are given in parentheses where appropriate in the discussion that follows. The path coefficient p_{21} is not recorded in Table 5.1 since it is a simple correlation coefficient between prior achievement and class size. For science classes, this correlation coefficient was 0.56 for prior achievement and 0.36 for father's occupational status. For mathematics classes, the correlation coefficient between father's occupational status and class size was 0.37.

Prior achievement had an effect upon many classroom characteristics to the extent that the influence of class size on these characteristics was reduced when prior achievement was controlled. Those affective measures that were still related to class size included the student's occupational aspirations (0.18) and educational aspirations (0.13) and academic motivation (0.13). It is of considerable interest that these aspirations were greater in larger classes even after controlling for prior achievement. In addition, attitudes to science (-0.14) decreased as class size increased after adjusting for prior achievement. Many of the process variables remained related to class size. The most interesting results involving teaching practices were that larger classes continued to spend more time reviewing work (0.42) but received less rebuke (-0.22).

Prior achievement had a predictably strong effect on science achievement with a path coefficient commonly over 0.90. The size of the path coefficient arose from the use of class averages. Class size had a positive effect upon achievement for a majority of the classroom variables although the strength of the relationship was marginal. Only a small number of classroom characteristics influenced achievement. Again, the student's occupational aspirations (0.13) and educational aspirations (0.22) were related to achievement. Only one classroom practice was important. The amount of positive support given to students (0.10) had a significant path coefficient with achievement. The absence of any other process variables would seem to indicate that the activities pursued by teachers in classes of differing sizes did not have a recognizable influence upon achievement.

A positive relationship between class size and achievement is contrary to the findings of Glass and Smith (1978). It could be argued that the model has ignored some variable which could explain more of the variation in science achievement. However, the absence of more substantial relationships between the classroom processes and achievement was possibly due to the strong explanatory and predictive power of prior achievement. It is possible that this measure has accounted for so much of the variation in achievement that the other variables have been overshadowed. For this reason, it is doubtful if the addition of other variables would change these results greatly. The

Table 5.1 Science Achievement Controlled for Prior Achievement

N = 72 Classroom characteristic	Classroom characteristic regressed on: Prior achievement P_{31}[a]	Classroom characteristic regressed on: Class size P_{32}	Science achievement regressed on: Prior achievement P_{41}	Science achievement regressed on: Class size P_{42}	Science achievement regressed on: Classroom characteristic P_{43}
Ethnicity of home	-0.07	-0.17	0.90	0.09	-0.06
Occupational aspirations	0.69	0.18	0.81	0.07	0.13
Number of hours homework per week	0.56	0.08	0.91	0.10	-0.02
Educational aspirations	0.82	0.13	0.73	0.07	0.22
Pop culture activities	-0.53	-0.02	0.89	0.09	-0.03
Maths/science activities	0.40	-0.05	0.88	0.10	0.07
Academic motivation	0.22	0.13	0.89	0.09	0.06
Like school	0.43	0.05	0.88	0.09	0.06
Like science	0.54	-0.14	0.86	0.11	0.08
Use of short answer tests	0.21	0.17	0.92	0.11	-0.06
Use of multiple choice tests	-0.22	0.37	0.92	0.08	0.05
Assessments involving student choice	0.33	-0.09	0.92	0.09	-0.04
Homework reprimand given	0.09	0.19	0.90	0.10	0.00
Frequency reports sent home	-0.34	0.46	0.93	0.07	0.06
Use of textbook B	-0.32	0.48	0.32	0.07	0.05
Use of pupil notebook	0.37	0.11	0.91	0.10	-0.01
Prop'n from foreign language homes	-0.47	0.03	0.87	0.10	-0.07
Time on science homework	0.03	0.42	0.90	0.10	-0.01
Total time on science	0.21	0.25	0.91	0.11	0.04
Time on all homework	0.56	0.06	0.91	0.10	-0.02
Teacher reviews work	-0.30	0.42	0.91	0.09	0.01
Teacher contacts students	-0.40	-0.18	0.89	0.09	-0.04
Other reward	-0.26	-0.06	0.90	0.09	-0.02
Deliberate rebuke	-0.19	-0.20	0.91	0.10	0.02
Casual rebuke	-0.36	-0.12	0.91	0.10	0.01
Other punishment	-0.15	-0.22	0.90	0.10	0.00
Total rebuke	-0.31	-0.22	0.91	0.10	-0.02
Positive support	-0.20	-0.05	0.93	0.10	0.10

[a] Path coefficients greater than 0.10 have been underlined.

effects of class size upon a student's aspirations for his or her career and education and their subsequent effects upon achievement suggest that, even after the effects of prior achievement have been allowed for, larger classes must be reinforcing these aspirations.

Science Achievement Controlled for Occupational Status

The findings using occupational status as a prior effect are similar to those reported above, and the results for science achievement controlled for occupational status are presented in Table 5.2. Among the most important effects of class size was the raising of the occupational aspirations (0.34) and educational aspirations (0.34) of students although most of the other classroom characteristics maintained substantial relationships with class size in the present of the variable controlling for home background. Fathers' occupational status had a strong influence (approximately 0.65) upon science achievement, but the path coefficient was not as great as with prior achievement. Class size continued to have a strong effect upon achievement (approximately 0.36). The path coefficient was much greater than when prior achievement was the controlling variable and its increase is viewed as a consequence of occupational status explaining less of the variation in achievement outcomes. This result adds strength to the earlier observation that the positive path coefficient between class size and achievement might be explained away if other suitable predictors could be incorporated into the model.

The attitude measures were important in predicting achievement outcomes. Students' occupational aspirations (0.34) and educational expectations (0.70), attitudes to science (0.19) and school (0.15) and their level of participation in mathematics and science activities (0.14) all had positive path coefficients with science achievement. The students' level of participation in pop culture activities had a negative path coefficient (-0.18). In fact, the students' expected educational level (0.70) was a better predictor of science achievement than was occupational status and when it was included in the model, the path coefficients from both occupational status (0.15) and class size (0.13) to science achievement decreased markedly.

The few process variables which were significantly related to science achievement were interesting. It appeared that other punishment (0.13) increased achievement while other reward (-0.10) reduced achievement levels. It seems that in science classes, although class size does influence some classroom practices, they in turn fail to influence achievement. A more important determinant of achievement is the extent to which favourable attitudes to the subject and to learning can be generated. To some extent, larger classes seem to increase student aspirations. The positive relationship between class size and achievement is unexpected from the review of previous research. Although our findings are internally consistent, there is a suggestion that our model is incomplete in that the effects attributed to class size might possibly be due to other unidentified variables or to the operation of a different causal model.

Table 5.2 Science Achievement Controlled for Occupational Status

N = 72 Classroom characteristic	Classroom characteristic regressed on: Occupational status P_{31}[a]	Classroom characteristic regressed on: Class size P_{32}	Science achievement regressed on: Occupational status P_{41}	Science achievement regressed on: Class size P_{42}	Science achievement regressed on: Classroom characteristic P_{43}
Ethnicity of home	-0.10	-0.17	0.65	0.36	-0.04
Occupational aspirations	0.64	0.34	0.43	0.25	0.34
Number of hours homework per week	0.56	0.19	0.63	0.36	0.04
Educational aspirations	0.71	0.34	0.15	0.13	0.70
Pop culture activities	-0.36	-0.19	0.59	0.33	-0.18
Maths/science activities	0.32	0.06	0.61	0.36	0.14
Academic motivation	0.28	0.15	0.64	0.36	0.04
Like school	0.32	0.17	0.60	0.34	0.15
Like science	0.42	0.02	0.58	0.36	0.19
Use of short answer tests	0.15	0.23	0.65	0.37	-0.01
Use of multiple choice tests	-0.14	0.30	0.65	0.37	-0.01
Assessments involving student choice	0.19	0.02	0.64	0.37	0.06
Homework reprimand given	0.00	0.24	0.65	0.35	0.06
Frequency reports sent home	0.01	0.26	0.65	0.41	-0.18
Use of textbook B	-0.11	0.34	0.64	0.40	-0.11
Use of pupil notebook	0.16	0.26	0.63	0.33	0.14
Prop'n from foreign language homes	-0.48	-0.06	0.61	0.36	-0.09
Time on science homework	0.12	0.40	0.66	0.40	-0.07
Total time on science	0.13	0.33	0.65	0.36	0.03
Time on all homework	0.55	0.17	0.63	0.36	0.03
Teacher reviews work	-0.14	0.31	0.64	0.40	-0.11
Teacher contacts student	-0.40	-0.25	0.62	0.35	-0.07
Other reward	-0.16	-0.15	0.64	0.35	-0.10
Casual rebuke	-0.36	-0.19	0.64	0.36	-0.02
Deliberate rebuke	-0.27	-0.21	0.67	0.38	0.06
Other punishment	-0.34	-0.18	0.69	0.39	0.13
Total rebuke	-0.39	-0.25	0.67	0.38	0.06
Positive support	-0.05	-0.15	0.65	0.37	0.01

[a] Path coefficients greater than 0.10 have been underlined.

Mathematics Achievement Controlled for Occupational Status

The findings from the previous chapter suggested that process variables may be more important in mathematics classes than in science classes. This was tested with the model presented in Figure 5.1 and the results of the analyses for mathematics achievement controlled for occupational status are presented in Table 5.3. Substantial relationships, with path coefficients greater than 0.10, were more numerous than for science classes with class size influencing all but two of the chosen classroom characteristics when occupational status was controlled, and furthermore, some of these different behaviours appeared to influence achievement outcomes.

As with science, occupational status (approximately 0.65) and class size (approximately 0.33) had a strong influence upon mathematics achievement. Attitude measures were again important in predicting achievement outcomes. The students' educational aspirations (0.67) and occupational aspirations (0.27) and attitude towards mathematics (0.23) and school (0.12) had positive path coefficients, while participation in pop culture activities (-0.14) had a negative relationship. As for science classes, the students' expected education level was a better predictor of achievement than occupational status. Several other classroom variables were significant. The use of extended answer tests (0.16) and a broad range of assessments (0.16) increased achievement. In addition, the time spent on mathematics (0.19), the time spent writing (0.12), the number of invitations for students to participate (0.19) and inquire (0.11) into academic work, and consideration of work habits (0.12) all enhanced achievement, while the number of regular teachers (-0.15), and number of teachers throughout the year (-0.10), the time spent on unclassified activities (-0.14) and the number of rebuke statements (-0.15) to students all had substantial negative path coefficients.

While attitudes were found to be important in mathematics classes, the model also contained many significant process variables. The increased fruitfulness of mathematics classes for an investigation of process variables has made the absence of a pretest measure more regrettable. The strong path coefficient between class size and achievement (approximately 0.33) would probably be weakened if a prior achievement measure were available. Nevertheless, all three achievement situations that have been considered have indicated a positive relationship between class size and achievement, even when prior achievement or occupational status was incorporated into the model.

In general, the results of the analysis of the causal model for achievement outcomes could be summarized in a few observations. First, the effects of class size upon process variables were more pronounced in mathematics classes and many of the processes produced changes in achievement. For both science and mathematics classes, certain attitudes led to clear gains in achievement. In particular, larger classes would appear to generate greater aspirations for the future and these hopes appeared to be

Table 5.3 Mathematics Achievement Controlled for Occupational Status

N = 72	Classroom characteristic regressed on		Mathematics achievement regressed on		
Classroom characteristic	Occupational status P_{31}[a]	Class size P_{32}	Occupational status P_{41}	Class size P_{42}	Classroom characteristic P_{43}
Ethnicity of home	-0.11	-0.17	0.66	0.34	-0.01
Occupational aspirations	0.66	0.33	0.49	0.26	0.27
Number of hours homework per week	0.55	0.19	0.65	0.34	0.02
Educational aspirations	0.71	0.33	0.18	0.12	0.67
Pop culture activities	-0.33	-0.19	0.62	0.32	-0.14
Maths/science activities	0.33	0.05	0.65	0.34	0.05
Academic motivation	0.30	0.14	0.66	0.34	0.03
Like school	0.35	0.16	0.62	0.32	0.12
Like mathematics	0.01	0.14	0.66	0.31	0.23
Use of extended answer tests	0.37	0.13	0.61	0.32	0.16
Range of assessment	0.34	0.12	0.61	0.32	0.16
Frequency of revision homework	-0.04	-0.19	0.66	0.33	-0.05
Frequency homework set	0.14	0.34	0.65	0.31	0.09
Written record of homework	0.14	0.20	0.65	0.34	0.03
Record homework in notebook	0.29	0.21	0.64	0.33	0.07
Homework discussed	0.08	0.39	0.66	0.33	0.04
Homework must be completed	0.28	0.23	0.66	0.34	0.01
Homework work habits score	0.20	0.29	0.65	0.32	0.07
Homework completed score	0.23	0.22	0.65	0.33	0.04
Frequency reports sent home	0.01	0.26	0.66	0.39	-0.20
Use of textbook A	0.29	0.42	0.66	0.34	0.02
Use of printed workbook	-0.11	-0.29	0.66	0.33	-0.04
Use of TV	0.03	0.41	0.67	0.37	-0.08
Prop'n from foreign language homes	-0.48	-0.06	0.63	0.34	-0.07
Time on maths homework	0.24	0.36	0.65	0.33	0.05
Total time on maths	0.19	0.28	0.63	0.29	0.19
Total time on homework	0.55	0.17	0.65	0.34	0.02
Number of teachers in year	-0.23	-0.21	0.64	0.32	-0.10

[a] Path coefficients greater than 0.10 are underlined.

(continued)

Table 5.3 Mathematics Achievement Controlled for Occupational Status (continued)

N = 72 Classroom characteristic	Classroom characteristic regressed on: Occupational status P_{31}[a]	Classroom characteristic regressed on: Class size P_{32}	Mathematics achievement regressed on: Occupational status P_{41}	Mathematics achievement regressed on: Class size P_{42}	Mathematics achievement regressed on: Classroom characteristic P_{43}
Number of regular teachers	-0.19	-0.27	0.64	0.30	-0.15
Time on question and answer session	-0.06	0.41	0.66	0.39	-0.11
Time students write	-0.03	-0.30	0.67	0.38	0.12
Unclassified time	0.04	-0.40	0.67	0.29	-0.14
Number of changes in activity	-0.25	0.26	0.67	0.34	0.02
Teacher contacts student	-0.13	-0.38	0.66	0.33	-0.04
Teacher asks question	0.01	0.32	0.66	0.33	0.03
Student asks question	0.15	0.20	0.66	0.34	0.04
Invitation to participate	-0.09	0.22	0.68	0.30	0.19
Invitation to inquire	0.12	0.17	0.65	0.33	0.11
Consider work habits	-0.21	0.28	0.69	0.31	0.12
Casual praise	-0.31	0.41	0.65	0.36	-0.05
Total praise	-0.32	0.42	0.65	0.37	-0.06
Other punishment	0.04	-0.17	0.67	0.31	-0.18
Deliberate rebuke	0.01	-0.28	0.67	0.30	-0.15
Total rebuke	-0.21	-0.20	0.63	0.31	-0.15
Positive support	-0.16	0.32	0.66	0.36	-0.05
Laughter with	0.03	0.03	0.66	0.34	0.02
Laughter at	0.20	0.23	0.67	0.35	-0.01

[a] Path coefficients greater than 0.10 are underlined.

Table 5.4 Attitudes to Science Controlled for Prior Achievement

N = 72 Classroom characteristic	Science attitudes regressed on		
	Prior achievement P41[a]	Class size P42	Classroom characteristic P43
Ethnicity of home	0.55	-0.11	0.16
Occupational aspirations	0.29	-0.20	0.37
Number of hours homework per week	0.42	-0.15	0.21
Educational aspirations	0.16	-0.20	0.47
Pop culture activities	0.45	-0.14	-0.16
Maths/science activities	0.35	-0.11	0.48
Academic motivation	0.46	-0.18	0.38
Like school	0.34	-0.16	0.46
Use of short answer tests	0.54	-0.13	0.09
Use of multiple choice tests	0.56	-0.17	0.10
Assessments involving student choice	0.53	-0.13	0.02
Homework reprimand given	0.53	-0.16	0.11
Frequency reports sent home	0.58	-0.19	0.12
Use of textbook B	0.57	-0.19	0.11
Use of pupil notebook	0.53	-0.14	0.02
Prop'n from foreign language homes	0.59	-0.14	0.11
Time on science homework	0.54	-0.13	-0.02
Total time on science	0.54	-0.13	0.00
Time on all homework	0.44	-0.15	0.18
Teacher reviews work	0.52	-0.10	-0.07
Teacher contacts student	0.58	-0.12	0.11
Other reward	0.60	-0.12	0.22
Casual rebuke	0.59	-0.12	0.14
Deliberate rebuke	0.54	-0.14	0.00
Other punishment	0.55	-0.12	0.06
Total rebuke	0.56	-0.12	0.07
Positive support	0.60	-0.12	0.28

[a] Path coefficients greater than 0.10 are underlined.

reflected in achievement gains. It would seem that by collecting able students in larger classes their ambitions were raised above the normal, either by competition, general classroom climate or in some unknown way, and this influenced their achievement.

Attitudes to Science Controlled for Prior Achievement

The path coefficients obtained from the analysis of the model for attitudes to science controlled for prior achievement are presented in Table 5.4. The effects of prior achievement and class size upon the classroom characteristics were the same as for achievement outcomes, so it is only necessary to discuss the effects of prior achievement, class size and classroom characteristics upon attitudes to science. Prior achievement had a substantial positive influence upon attitudes to science (approximately 0.55), while class size had a substantial negative path coefficient (approximately -0.15) with attitudes to science for all classroom characteristics. The negative path coefficient between class size and attitudes to science is an expected

Table 5.5 Attitudes to Science Controlled for Occupational Status

N = 72 Classroom characteristic	Occupational status P_{41}[a]	Class size P_{42}	Classroom characteristic P_{43}
	Science attitudes regressed on		
Ethnicity of home	0.43	0.05	0.17
Occupational aspirations	0.11	-0.14	0.48
Number of hours homework per week	0.28	-0.03	0.24
Educational aspirations	0.04	-0.20	0.65
Pop culture activities	0.33	-0.02	-0.23
Maths/science activities	0.25	-0.01	0.50
Academic motivation	0.31	-0.04	0.37
Like school	0.26	-0.06	0.49
Use of short answer tests	0.41	0.01	0.02
Use of multiple choice tests	0.42	0.00	0.06
Assessments involving student choice	0.40	0.02	0.07
Homework reprimand given	0.41	-0.01	0.14
Frequency reports sent home	0.41	0.03	-0.03
Use of textbook B	0.41	0.01	0.01
Use of pupil notebook	0.39	-0.01	0.10
Prop'n from foreign language homes	0.47	0.03	0.11
Time on science homework	0.42	0.04	-0.06
Total time on science	0.41	0.01	0.04
Time on all homework	0.30	-0.02	0.21
Teacher revises work	0.39	0.06	-0.14
Teacher contacts student	0.46	0.05	0.11
Other reward	0.44	0.04	0.17
Casual rebuke	0.46	0.04	0.14
Deliberate rebuke	0.42	0.03	0.03
Other punishment	0.46	0.04	0.15
Total rebuke	0.46	0.05	0.11
Positive support	0.42	0.05	0.21

[a] Path coefficients greater than 0.10 are underlined.

result from the research of Smith and Glass (1979) and confirms that increased class size does damage students' attitudes to that particular subject.

In addition, all the home background and attitudinal measures had a substantial influence upon attitudes to science. This was not surprising since the attitudes measured would be components of a more general attitude. Many of the structural and process measures considered in the model also influenced attitudes to science. Among the strongest predictors were the use of other rewards (0.22), the amount of positive support (0.28) and the time spent on homework (0.18). Unlike achievement outcomes, it appeared that there were some process variables that were good predictors of affective outcomes in science classes.

Attitudes to Science Controlled for Occupational Status

In Table 5.5 the path coefficients for the model with attitudes to science controlled for occupational status are recorded. Just as for achievement outcomes, occupational status was not as strong a predictor of attitudes towards science as was prior achievement. The path coefficient between occupational status and attitude to science fluctuated markedly depending upon the predictive power of specific classroom characteristics, although it was almost always substantial. The presence of suppressor effects was suspected. The relationship between class size and attitudes towards science was usually negligible, but again it varied according to the classroom variable being considered. The same set of attitude variables as listed for achievement outcomes continued to have substantial path coefficients.

When the attitude measures showed larger path coefficients, the predictive power of occupational status decreased and the path coefficient between class size and attitude to science changed from positive to negative values. Many of the other classroom characteristics had a substantial influence upon attitudes to science. The main predictors were again the use of other reward (0.17), the amount of positive support (0.21) and the time spent on homework (0.21).

Attitudes to Mathematics Controlled for Occupational Status

The relationships for attitudes to mathematics when occupational status was used as the control were surprising and the path coefficients have been recorded in Table 5.6. The direct path coefficient between these two measures was not strong with most values in the vicinity of zero although different classroom characteristics generated fluctuations. Class size sometimes had a substantial positive path coefficient with attitudes to mathematics but again this varied according to the classroom variable under consideration.

All the attitude measures were substantially related to attitudes to mathematics. The direction of the relationship was positive except for the students' participation in pop culture activities (-0.31). As noted for the corresponding analysis of science attitudes, as the path coefficients between certain attitudes and the attitudes to mathematics measure increased, the path coefficient between occupational status and attitudes to mathematics became negative and the path coefficients from class size either decreased or became negative. Many of the process measures possessed substantial path coefficients. The main predictors of attitudes to mathematics included the amount of laughter (laughter with: 0.27; laughter at: 0.31), the number of invitations for students to participate or inquire (0.24) and the number of work habit comments (0.29). The use of a major textbook damaged attitudes (-0.24).

In general, the results from the analysis of the causal model for affective outcomes could be summarized as follows. First, the effect of class size upon attitudes appears to be hidden unless prior achievement is included as a predictor to account for much of the variation in attitudes and thus reveal a negative relationship. Even after controlling for prior effects, most of the attitude measures collected in the original study appeared to be substantially related to attitudes to these two subjects. This is not surprising since all these measures could well be components of a general attitude dimension. Finally, classroom practices seemed to have less effect upon achievement than upon attitudes such that achievement appeared to be independent of most of the teacher behaviours that were related to class size.

One disappointing feature of the analyses of the causal models for all outcomes was the small number of classroom characteristics which substantially affected achievement outcomes. The use of only individual classroom variables is a valid criticism of the model and the subsequent analyses. It might have been more appropriate to have combined several classroom process variables into one compound variate with the hope of producing a stronger classroom effect. This was considered but the construction of such a compound variate was rejected since it would have required the combination of conceptually different variables, including time measures, frequency counts and interview schedule results. It was decided that for the analyses reported in this investigation it would be desirable to simplify the issues of interpretation so that comparisons across levels of analysis would be more meaningful. Furthermore, the current list of variables did contain some combined measures for homework, assessment and the use of teaching aids. These variables represented compound measures of a specific range of classroom activities and some of these variables were examined in the causal model.

Another difficulty with the analyses conducted, particularly when a prior achievement measure was absent, was to explain enough variation in achievement so that the effect of class size would be validly observed. To overcome this difficulty, it would be necessary to identify other variables which might be introduced into the model. These new variables should explain variation in achievement, distinct from that already explained by prior achievement or more especially by occupational status. To identify these additional variables, the best clue is given in the regression models for science and mathematics achievement already constructed. In most of these models, a collection of attitude measures possessed substantial path coefficients when either prior achievement or occupational status was controlled. The general attitude measures in the model included the students' occupational and educational aspirations, attitudes to school, academic motivation and level of participation in mathematics or science activities. These five attitude measures provided an important additional dimension to the causal model.

A Causal Model Incorporating Attitudes

The usefulness of the attitude measures as a supplement to the predictive power of prior achievement or occupational status was indicated by the analyses with the simple causal model. Class size had a substantial effect upon aspirational measures and academic

Table 5.6 Attitudes to Mathematics Controlled for Occupational Status

N = 72	Mathematics attitudes regressed on		
Classroom characteristic	Occupational status P41[a]	Class size P42	Classroom characteristic P43
Ethnicity of home	0.03	0.16	0.10
Occupational aspirations	-0.25	0.01	0.40
Number of hours homework per week	-0.05	0.12	0.12
Educational aspirations	-0.50	-0.09	0.72
Pop culture activities	-0.09	0.08	-0.31
Maths/science activities	-0.07	0.13	0.26
Academic motivation	-0.15	0.07	0.54
Like school	-0.16	0.06	0.50
Use of extended answer tests	-0.01	0.14	0.06
Range of assessment	0.01	0.14	0.02
Frequency of revision homework	0 01	0 14	-0 03
Frequency of homework set	0.02	0.15	0 03
Written record of homework	0.02	0.15	-0.02
Record homework in notebook	-0.07	0.08	0.30
Homework discussed	0.01	0.14	0.00
Homework must be completed	-0.02	0.09	0.16
Homework work habits score	-0.01	0.09	0.15
Homework completed score	-0.03	0.08	0.21
Frequency reports sent home	0.02	0.19	-0.25
Use of textbook A	0.09	0.23	-0.24
Use of printed workbook	0.03	0.14	0.05
Use of TV	0.02	0.19	-0.16
Prop'n from foreign language homes	0.11	0.14	0.18
Time on maths homework	0.02	0.13	-0.01
Total time on maths	-0.01	0.09	0.13
Total time on homework	-0.09	0.10	0.19
Number of teachers in year	0.00	0.10	-0.10
Number of regular teachers	0.01	0.11	-0.05
Time on question and answer session	0.01	0.19	-0.16
Time students write	0.02	0.16	0.09
Unclassified time	0.02	0.10	-0.06
Number of changes in activity	0.02	0.12	0.01
Teacher contacts student	0.04	0.19	0.17
Teacher asks question	0.02	0.08	0.14
Student asks question	0.00	0.09	0.17
Invitation to participate	0.04	0.07	0.24
Invitation to inquire	-0.01	0.09	0.24
Consider work habits	0.08	0.05	0.29
Casual praise	0.06	0.07	0.13
Total praise	0.07	0.06	0.15
Other punishment	0.01	0.15	0.14
Deliberate rebuke	0.02	0.13	0.03
Total rebuke	0.05	0.15	0.14
Positive support	0.01	0.14	-0.05
Laughter with	0.01	0.05	0.27
Laughter at	-0.04	0.05	0.31

[a] Path coefficients greater than 0.10 are underlined.

motivation in all three models. All the attitude measures that were discussed (except for academic motivation) had a substantial effect upon achievement even after controlling for the influence of prior effect measures. In order to develop a causal model which incorporated the additional effect of attitudes, it must be recalled that all attitude measures were obtained towards the end of student's year in Year 7. Hence, these attitude measures do not represent a prior effect but should be viewed as a component of the environment of the classroom. As such, the students' attitudes could be influenced by class size. Moreover, the interaction between attitudes and the processes which occur in the classroom are complex. The activities that teachers pursue can influence student appreciation of the subject and hence their attitudes. Alternatively, teachers may perceive the presence of favourable attitudes and the students' enthusiasm and modify their behaviour to match a particular class. In view of the fact that the effects of classroom characteristics upon attitudes to school subjects have already been considered, it may be more profitable to assess the effects of the presence of favourable and unfavourable attitudes upon the behaviours and practices that the teachers employed.

The second causal model relating class size, classroom behaviours, attitudes and academic achievement that has been discussed in the previous paragraphs is given in Figure 5.2. Since attitude measures are included in the model, it is necessary to restrict the analyses to a consideration of achievement outcomes only. Since all five attitude measures were obtained from questionnaires, it was considered possible to develop a general attitude measure from the five variables using factor analysis. The results of this analysis have been reported in Appendix III. While the use of a general attitude measure might seem theoretically preferable, some investigatory runs with the measure indicated that it would not be as productive as the analyses using the specific attitude measures since it failed to provide an understanding of the relevance and importance of different facets of the students' attitudes. The causal model was examined for science classes using only prior achievement as the prior effects variable. The model included only the one process variable which had been found to have a substantial effect upon science achievement using the first model. The results of the examination of the causal model given in Figure 5.2 are presented in Table 5.7.

1 Students' occupational aspirations. The inclusion of this attitude appeared to make little difference to the effect of class size. The most interesting finding was that the amount of support given by teachers decreased as prior achievement increased (-0.44), as attitudes increased (-0.35), and as class size grew (-0.12). It appeared that teachers did not waste their support on classes of high-achieving, well-motivated students. Prior achievement (0.86), students' occupational aspirations (0.10), and the amount of positive support (0.10) were all substantial predictors of science achievement, while the effect of class size upon achievement was still positive (0.08), but no longer substantial, after adjustment for the factors included in the model.

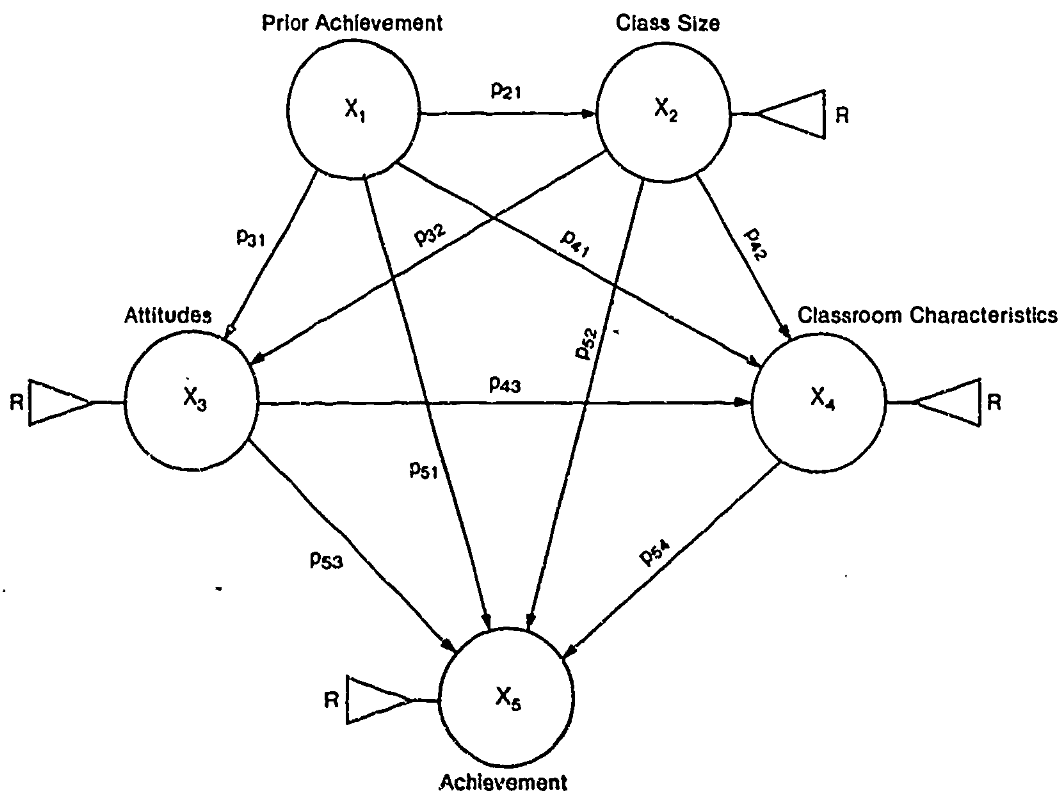

Figure 5.2 Causal Model at the between Classes Level of Analysis Including Attitudes, Class Size and Classroom Characteristics

2 Students' educational aspirations. The inclusion of this attitude into the causal model is very perplexing, for although the effect of class size upon positive support did change slightly (-0.17), the effect of prior achievement upon positive support changed markedly (-0.90). It seems that prior achievement and the students' educational aspirations are confounded by each other's influence and that the attitude measure was having an extreme effect upon positive support (0.85). The change in the path coefficient suggested that educational aspirations have acted as a suppressor variable and hence altered the predictive power of the other variables. The effect of all four predictors upon academic achievement was similar to the previous attitude measure except that educational aspirations (0.14) were superior to occupational aspirations as a predictor of achievement. The effects of positive support (0.08) and class size (0.08) were still positive, but not substantial.

3 Mathematics/science activities, academic motivation and attitudes to school. The effect of class size upon positive support was unchanged by the addition of any of these

Table 5.7 Path Coefficients for the Analysis of the Causal Model at the between Classes Level of Analysis

N = 72 Positive support regressed on		Prior achievement p_{41}[a]	Class size p_{42}	Attitude p_{43}
Occupational aspirations		-0.44	-0.12	-0.35
Educational aspirations		-0.90	-0.17	0.85
Academic motivation		-0.20	-0.06	0.01
Like school		-0.23	-0.06	0.06
Maths/science activities		-0.28	-0.05	0.19
Science achievement regressed on	**Prior achievement p_{51}**	**Class size p_{52}**	**Attitude p_{53}**	**Positive support p_{54}**
Occupational aspirations	0.86	0.08	0.10	0.10
Educational aspirations	0.80	0.08	0.14	0.08
Academic motivation	0.91	0.09	0.06	0.11
Like school	0.90	0.10	0.05	0.11
Maths/science activities	0.91	0.10	0.05	0.10

[a] Path coefficients greater than 0.10 are underlined.

three variables. Furthermore, none of these variables was as strong a predictor of science achievement as were the first two attitude variables considered above, and hence the effect of class size upon achievement remained substantial (approximately 0.10). The amount of positive support given by teachers maintained a substantial path coefficient (approximately 0.11) with science achievement in each case.

The addition of attitude measures to the causal model was interesting if the students' occupational and educational aspirations were considered. These measures did not greatly influence the relationship between class size and positive support, or the relationship between positive support and science achievement, but influenced instead, the relationship between class size and science achievement. The attitude measures help to explain a greater proportion of variation in achievement and hence the effect of class size was reduced, occasionally to non-substantial levels. Despite the weakening of the effect of class size in comparison to the previous model to the point where it was barely substantial, the relationship was still consistently positive. This remains a result contrary to the extensive body of previous research reviewed.

The relative effect of the students' occupational and educational aspiration measures when compared to those of academic motivation, participation in mathematics and science activities and attitudes to school can be matched to certain results from the factor analysis of these five attitudes. Although all five measures were aligned on the first principal factor, presumably some general attitude factor (see Appendix III), the two aspirational measures were assigned negative values for the second principal factor while the other three attitudes were assigned positive values. This appeared to indicate a distinction between the aspirational measures and the other attitudes, although it would seem difficult to find an appropriate label for the second factor.

Summary and Conclusion

A perplexing result is the consistent positive relationship between class size and achievement in both subjects. The use of a prior achievement measure went part of the way t reducing the strength of this link, but it has remained substantial in the presence of controls with achievement, occupational status, and attitudes. The inclusion of additional predictors may further reduce its influence, but its persistence as a predictor of academic achievement in all models does suggest that the result has some strength. It appears that large classes, even after controlling for the presence of high performing students, were able to generate additional achievement gains. Furthermore, these achievement gains went beyond the effects of the process dimension which might have suggested that teachers of larger classes employed superior teaching styles. It would appear that the achievement gains possible in larger classes were an indirect consequence of certain properties or characteristics of the students who formed the classes, for they would appear to interact with each other and amplify achievement beyond its initial level. This phenomenon would seem to be a product of the way students are chosen for larger classes, and the attitudes of students in these classes, for they have produced an enhancing effect.

Apart from achievement scores, it is hard to identify those student characteristics which influence placement in particular classes. The high correlation between prior achievement and class size does indicate that School Principals or Subject Coordinators allow the size of high performing classes to creep up. The students in these classes must be sufficiently well motivated to support their position. Although no measurements of attitudes at the commencement of the year were available, there appears to be some interaction between achievement, class size and attitudes. We have already considered the effect of class size upon attitudes, but it seemed that being well motivated and able influenced the size of the class to which a student was allocated. Therefore, there is a two-way interaction between class size and attitudes, and this interaction should be incorporated into a regression model. Such a model, appropriate to our present considerations, has been examined in Chapter 8. It was hoped that it would provide further insight into how students were allocated to classes by identifying other factors used in placing students. Furthermore, these factors might explain how the assemblage of able students was able to produce enhanced achievement gains.

In conclusion, we can note that class size did affect a range of classroom activities, particularly in mathematics, but that the number of variables so affected was a small percentage of the total number of variables investigated in the study. However, the incidence of certain activities was not entirely consistent with the scenarios of small and large classes painted in debate and research. It appears that class size has a slight positive effect upon achievement in both subjects although the reasons for this result are still hidden in the way that large and small classes are formed and develop. As such, classroom process variables were not found to be as important a determinant of achievement as was originally supposed, for it seems that the attitudinal interactions among the students that form a class are a more important factor than the teaching and learning activities that occur within the classroom.

Although the relationship that has been found between class size and achievement is contrary to the research summary provided by Glass and Smith (1978), an examination of the attitude measures provided further understanding of how groups of able students seemed to enhance their achievement levels. It should be noted that although our findings disagree with Glass and Smith, they are consistent with several major non-experimental studies noted in the review of previous research. Moreover, throughout the analyses attempts have been made to control for some of the criticisms that were levelled against studies where measures were not taken to compensate for prior achievement differences. Furthermore, an examination has been carried out into mechanisms operating in large and small classes using a wide range of structural and process variables. The small number of interesting results from this list of variables confirms the observation made by Ryan and Greenfield (1975) that teachers were not exploiting the opportunities available in smaller classes. Our findings suggest that the ability of the class, rather than the size of the class, is a more important factor influencing the use of various teaching styles.

Finally, we found some support for the claim that large classes damaged attitudes. In science classes, where prior achievement measures were available, a substantial negative path coefficient between class size and attitudes to science was observed after controlling for other relevant factors in the causal model. This finding represented support for the claim that class size can damage student attitudes to a subject. It was also found that many classroom processes had an effect upon the students' attitudes towards that subject. As such, this investigation has contradicted many of the research findings from the achievement domain, but supported the findings from the affective domain.

CHAPTER 6

THE EFFECT OF CLASS SIZE AT THE STUDENT LEVEL

The causal model that was developed and discussed in the previous chapter concentrated upon class size, classroom processes and educational outcomes with the class as the unit of analysis. Although the choice of the level of analysis was well justified by the character of the original measurements, the class does not need to be the sole unit of analysis. In fact, the use of a single unit of analysis has been criticized. As Rogosa (1978) explained:

> . . . no one level is uniquely responsible for the delivery and response to educational programs . . . confining substantive questions to any one level of analysis is unlikely to be a productive research strategy. (Rogosa, 1978:83)

As a response to this comment, it is desirable to conduct further analyses using the student as the unit of analysis.

The use of class initially as the unit of analysis seemed a natural choice since the primary aim of the study was concerned with the class size issue. However, it has been observed (Burstein, 1980) that by aggregating student test results to the classroom level, the relationships between these test results and other variables are strengthened by a compounding of prior achievement and other, often unmeasured, background characteristics. Moreover, the increased correlations that generally result in educational studies from the aggregation of data reduce the capacity to identify the effects of teacher-student interactions. Some evidence of the way that aggregation magnifies an effect can be seen in the very strong relationship between prior achievement and final achievement which was displayed in the set of analyses presented in Chapter 5. Furthermore, the aggregated measure may not represent the original meaning of the variable when it was recorded at the individual level. Alternatively, Blalock (1964) has suggested that by aggregating data, a more pure measure of the variable was possible since the 'nuisance' effect of individual differences would be lessened in the formation of classes.

Brophy and Good (1974) have argued that the nature of the relationship between the teacher and the student demanded the use of the student as the unit of analysis. Teacher behaviour is often directed towards individual students and not the entire class, and in return, individual differences between students influence teacher behaviour. Furthermore, even teacher behaviours directed to an entire class are received and responded to in different ways by individual students. These arguments would sugest that the student would be a more appropriate unit of analysis in the examination of the effects of teacher behaviours and classroom practices upon the achievement outcomes of students.

However, the nature of the data in the original study does not enable us to reproduce readily the findings of Chapter 5 at the between students level of analysis, for although individual measures from tests, questionnaires and attitude scales are available, the same is not true for the classroom measures. Measures of teacher behaviour and classroom practices were obtained at the class level. Individual students may experience the behaviours and practices to varying degrees, according to ability, concentration or otherwise, but assessment was not made of the impact of behaviours and practices upon the individual students being considered in this study. Therefore, it was not possible to duplicate the previous chapter's analyses without disaggregating the data collected at the classroom level. The complication with using disaggregated data arises because, in general, teacher behaviours can not be attributed equally to all the students in the class.

In addition to the restrictions necessary because of level of analysis considerations, the model was limited to science classes. The poor predictive power of the father's occupational status as a sole prior effects measure suggested that it did not act as an adequate control in the analyses of the mathematics data. Hence, any results for mathematics classes would remain inconclusive. However, incorporating both prior achievement and occupational status into one model was seen as improving the model's level of specification. Teacher behaviour and process variables were removed from the model for two reasons. One difficulty was the conceptual problem of using disaggregated data, while the other was the general failure of these variables, particularly in science classes, to influence the relationship between class size and achievement outcomes. However, an apparent contradiction was the use of class size, obviously a class measure, in an individual level model. This discrepancy was justified by noting that each student experienced the same class size, while the other classroom measures which were shown to be substantial in science classrooms were not so evenly distributed to all students in a class group.

The purpose in developing a new model was to identify further the effect of class size upon science achievement by considering its influence in a different context. By examining the influence of class size upon the individual's performance as distinct from the performance of the entire class, the intention was to further isolate the way in which class size regulated science achievement. The resultant causal model is presented in Figure 6.1.

The model was examined with the same five attitude measures that were employed in the previous chapter. They were the student's occupational and educational aspirations, liking for school, academic motivation and participation in mathematics and science activities. These five measures were selected because of their repeated interaction with class size and academic achievement.

All path coefficients were obtained using the SPSS Regression program (Nie et al. 1970). The same criterion and nomenclature as used previously were again used to

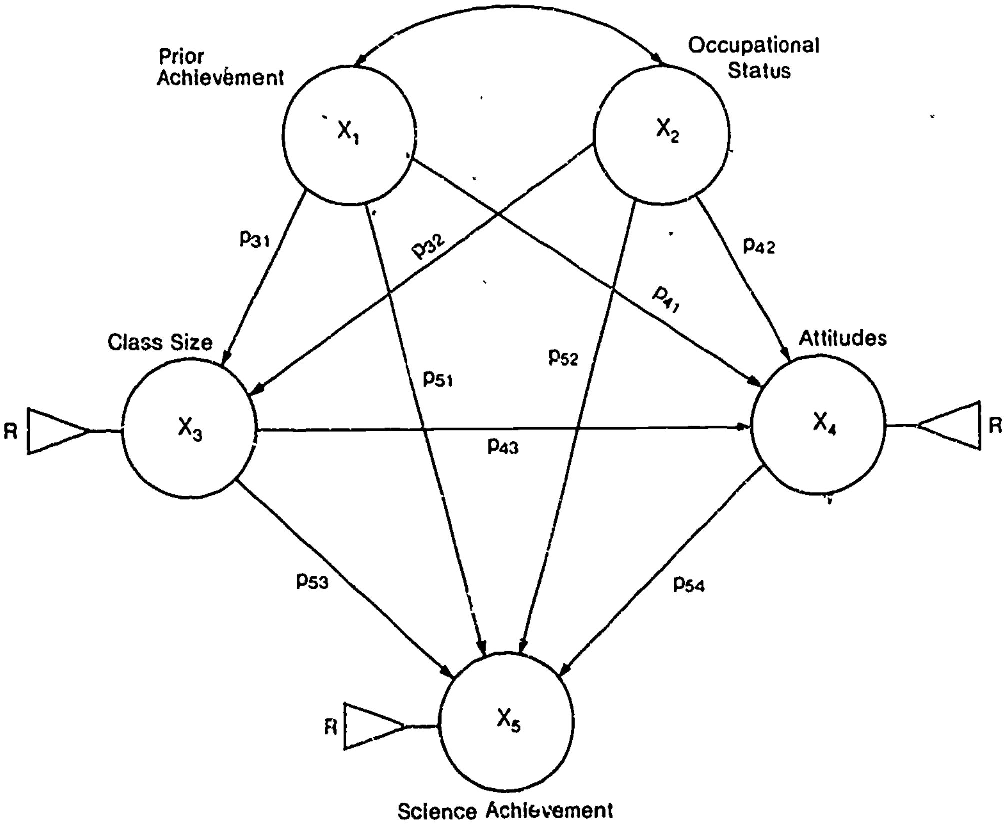

Figure 6.1 Causal Model at the between Students Level of Analysis Including Class Size and Attitudes

indicate when a path coefficient represented a substantial relationship between two measures. In Table 6.1, the results from the causal model for each attitude are presented.

The most interesting results in Table 6.1 were the positive influences of class size upon aspirational measures, while class size did not appear to affect greatly the other attitude measures. Neither of the prior effects measures (prior achievement or occupational status) influenced consistently attitudes in all areas. Furthermore, both prior achievement (approximately 0.65) and class size (approximately 0.13) were substantial predictors of science achievement regardless of the attitude measure being considered. Although prior achievement was the strongest predictor, class size was a better predictor than all other measures except for educational aspirations (0.17). In addition, class size had a small indirect effect upon achievement through the aspirational measures. The path coefficient between class size and science achievement was larger when the student, not the class, was used as the unit of analysis.

Table 6.1 Path Coefficients for the Analysis of the Causal Model at the between Students Level of Analysis

N = 1986 Class size regressed on	Prior achievement p_{31}[a]	Occupational status p_{32}		
Class size	0.20	0.04		
Attitude measures regressed on	**Prior achievement p_{41}**	**Occupational status p_{42}**	**Class size p_{43}**	
Occupational aspirations	0.17	0.22	0.11	
Educational aspirations	0.33	0.27	0.14	
Like school	0.11	0.09	0.04	
Academic motivation	0.03	0.10	0.04	
Maths/science activities	0.09	0.05	-0.02	
Science achievement regressed on	**Prior achievement p_{51}**	**Occupational status p_{52}**	**Class size p_{53}**	**Attitude p_{54}**
Occupational aspirations	0.65	0.10	0.13	0.09
Educational aspirations	0.61	0.08	0.11	0.17
Like school	0.66	0.11	0.13	0.10
Academic motivation	0.67	0.11	0.13	0.09
Maths/science activities	0.66	0.12	0.14	0.08

a Path coefficients greater than 0.10 are underlined.

Effect upon Attitudes and Practices

Superior prior achievement levels led to an improvement in aspirations and liking for school. Not surprising, the strongest relationship was with the student's educational aspirations. Clearly, students who had done well at school expected to continue with their studies. Occupational status related positively with aspirations and academic motivation. Interestingly, occupational status was a stronger predictor of academic motivation than prior achievement, suggesting that home background rather than academic success at school was more important in providing general motivation for students towards their studies.

The effect of class size upon attitudes is more interesting. While larger classes appeared to enhance a student's occupational and educational aspirations, they appeared to have little effect upon the other three measures. This distinction between the findings for the two aspirational measures and the other three measures has already been noted in Chapter 5.

Effect upon Science Achievement

All four antecedents had substantial path coefficients with science achievement. Prior achievement always had the greatest effect on science achievement. In four of the five cases, class size had the second largest path coefficient to science achievement, followed by occupational status and attitudes. When the attitude measure was educational aspirations, it had a stronger effect on achievement than did class size, with a path coefficient approximately double that of the other attitude measures.

The importance of these results when compared to the equivalent results obtained when class was the unit of analysis is difficult to assess. Differences in the models do not enable direct comparisons. Prior effects were stronger when the class was the unit of analysis, but there were two prior effect measures present in the latter analyses so that the extent of the difference could not be directly determined. Furthermore, the effect of aggregation in strengthening relationships has already been noted. The path coefficients between the attitudes and science achievement were similar at both levels of analysis although liking for school had a stronger effect when the individual level was examined.

Occupational status had a weaker effect in the presence of the prior achievement measure, but more important in terms of the issues for this study, were the changes in the effects of class size. At the class level, class size had a barely substantial effect on science achievement with the magnitude of the path coefficient depending upon the combination of attitude and classroom characteristic being considered. However, at the individual level, class size had a consistently substantial effect on science achievement, and as has already been noted, often had a stronger effect than attitudes or occupational status. It is possible that the class size effect was mediated through teaching practice, but since teaching behaviours had such weak relationships with science achievement, this appeared unlikely.

The reasons as to why class size should appear to be more effective in increasing science achievement at the individual level as distinct from the class level are difficult to provide. In the previous chapter we discussed the possibility of larger classes having a certain climate due to the assemblage of more able students. If this were so, it would be envisaged that analyses conducted at the class level could be more successful at detecting this effect. The results contradict this suggestion. The explanation is more likely to lie with a consideration of the effects of aggregating the data to the class level. The strong path coefficient between prior achievement and achievement would appear to be accounting for so much of the variance in science achievement that the path coefficients for the other measures were being reduced. Although it is difficult to quantify the relative influence of class size upon achievement at the two levels of analysis, a positive relationship between class size and achievement has been confirmed. Even if the results obtained in this study continue to disagree with the Glass and Smith findings from a review of experimental studies, a certain confidence in the results is possible due to their consistency at the two levels of analysis.

CHAPTER 7

COMPARISONS BETWEEN DIFFERENT LEVELS OF ANALYSIS

The importance and implications of choosing a suitable level of analysis for an investigation has already been discussed briefly in previous chapters. However, the issues and their consequences are more complex. The increased importance of the choice of an appropriate level of analysis is indicated by the number of research workers who have considered it necessary to address themselves to this issue during the past decade. They include Cronbach and Webb (1975) and Barr and Dreeben (1977), although a more extensive summary of the work related to the level of analysis issue has been given by Burstein (1980).

Among the points emphasized by Burstein was the possibility of a variable having different meanings depending upon the unit of analysis under consideration and the difficulties associated with drawing meaningful conclusions when an examination of the data was conducted at more than one level of analysis. A simple illustration of his first point can be made with the occupational status variable that has been used in the causal models that have been examined in this study. At the individual level, this variable is an indicator of home background, while if the variable is aggregated to the classroom level, the same variable may become a measure of the type of community that the classroom services. The second point refers to the appropriateness of extrapolating a result obtained from a classroom study to the individual students in those classes.

The causal model developed in Chapter 6 was a response to the problems associated with the construction of appropriate models. Its aim was to assess the effect of class size upon science achievement at the individual, as distinct from the classroom, level. The use of a class size measure in an individually based model was justified in the sense that all students experienced the effects of the class size, and that this effect was equal for all students in a class. As such, the purpose of the model was to be relevant to the current issues as well as contain measures suitable at the individual level of analysis.

Reporting of analyses conducted at both the individual and school level has been undertaken by Comber and Keeves (1973) and Peaker (1975). Both reports were associated with the cross-national Six Subject Study undertaken by the International Association for the Evaluation of Educational Achievement. Just as for the different levels of analysis reported in Chapters 4, 5 and 6, these studies obtained results which differed slightly between levels of analysis. However, no regular pattern was discernible and therefore it was not possible to draw conclusions for one level of analysis in terms of the results obtained at the other level of analysis. The difficulties associated with cross-level inference were again illustrated. The possibility of drawing conclusions was further complicated by the use of dissimilar models.

Student within class is another level of analysis that has also been employed and reported (Peaker, 1967; Burstein, 1980; Keeves and Lewis, 1983). This type of analysis is often described as the 'frog pond effect' and it is used to examine the relative standing of a student within a class. The importance of a student's relative standing can be illustrated in several ways. For example, a student's performance may improve in a weaker class because the student's self-image is enhanced by being a 'big frog in a little pond'. Furthermore, the attention that students receive from their teachers and peers is a function of their relative position in the class. Other illustrations of the importance of the within-class effect involve the nature of instruction received by students. For example, if the knowledge required to answer a question on an achievement test has been presented to the students, then all students should be able to answer the item correctly. Obviously, this does not happen. Some students respond to being taught while others do not, either due to inattention, lack of ability to comprehend the instruction or for other reasons. It is clear that while all students in the class receive the same instruction, students within the class might benefit in unequal amounts. As such, classroom characteristics would not be evenly shared throughout the class although each student might have the same opportunity to receive them.

The nature of the student within class level of analysis can be further elaborated since the student's relative standing within a class is an integral part of the performance of the individual student. If the variable X_{ij} denotes the performance of student j in class i, and X_i denotes the mean performance of class i, then these measures can be related by the expression:

$$X_{ij} = X_i + (X_{ij} - X_i)$$

such that the performance of the student may be broken down into a class component and a within class component. Thus, the relative performance of a student within a class is an integral part of the student's performance although it is interrelated to the class to which the student belongs.

Comparisons at all three levels of analysis (i.e. student, class and student within class) have been carried out on Comber and Keeves' (1973) American data using regression techniques and were discussed by Burstein (1980). The study sought to assess the effects of ability, sex, father's occupation, number of books in the home, years of science instruction, amount of science instruction per week and the use of discovery science teaching methods upon science achievement. The path coefficients obtained from the three levels of analysis were equivalent in both pattern and relative magnitude. The similarity of the results at different levels of analysis made more detailed interpretations of the effects of unit of analysis upon the causal model very difficult since differences between the levels of analysis were needed to isolate particular effects.

Another study where results were obtained using different levels of analysis but where the results were not internally consistent was also discussed by Burstein (1980). In that study, correlation coefficients were obtained between the rate of student success, as a measure of the degree of difficulty of the work presented to them, and reading and mathematics achievement, which were adjusted for prior achievement. At the student within class level, high success rates were positively correlated with achievement. At the between classes level, many of the correlation coefficients were negative. This result was interpreted to mean that teachers who assigned easier material to their students and hence provided their students with a greater proportion of successful learning experiences tended to have lower class achievement. At the student level, the results were similar to those obtained at the student within class level. This example indicates how different levels of analysis may be used to increase the understanding of an educational situation by providing additional insight into the reasons for changes in student achievement.

To extend the idea of comparing the three separate units of analysis mentioned above to the present situation, several modifications were needed before a suitable causal model could be developed. The most obvious difficulty was associated with class size because there was no variability in this measure at the student within class level since its value would always be zero. This was regrettable since the investigation was primarily concerned with class size. Consequently to provide a direct comparison between the three levels of analysis, class size had to be omitted from the model and hence the discussion must concentrate solely upon comparing the effects of the other antecedent measures at the various levels of analysis.

Furthermore, the between classes model used in Chapter 5 and the between students model used in Chapter 6 are also not directly applicable since they included class size. The between classes model also used classroom characteristics while the between students model incorporated both prior achievement and occupational status as prior effects. To produce a causal model which could be applied in the same form to all three levels of analysis, the models used previously had to undergo further modification.

The improved predictive power of the model with two prior effect measures together with the unsuitability of any model without a prior achievement measure suggested that it was worthwhile restricting the comparisons to science classes only. The strength of the attitude measures as predictors of science achievement indicated that they were worthy of continued inclusion in the model. However, since the five attitude measures used in the two previous models were chosen for their importance in conjunction with class size, it seemed appropriate to extend the model to include the complete list of attitude and practice measures available at the student level since class size was no longer under consideration. The attitude and practice measures included occupational and educational aspirations, amount of homework per week, level of

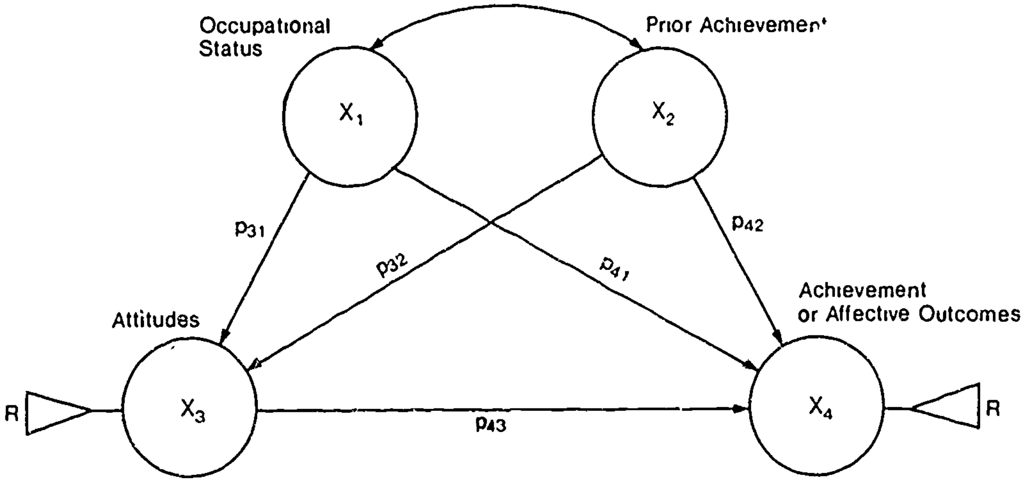

Figure 7.1 Causal Model Appropriate to the between Classes, between Students and between Students within Classes Levels of Analysis

participation in both pop culture or mathematics and science activities, academic motivation, liking of school and self regard. In addition, since attitudes were an integral part of the model, it was decided that the outcomes measures should include both science achievement and attitudes to science.

A causal model which incorporates prior achievement, occupational status, attitudes and either science achievement or attitudes to science is presented in Figure 7.1. Data were available for all variables at each of the three levels of analysis. Just as before, class measures were obtained from 72 classes, individual measures from 1986 students and students within classes measures by constructing the difference between the student measures and the appropriate class measure for the 1986 students. The results of the regression analyses conducted at each of the three levels of analysis are presented in Table 7.1.

It should be noted that the data were not weighted to correct for unequal size of class groups as is sometimes considered desirable if classes have been sampled. In this study, an almost complete population was under survey and naturally occurring class groups were used in the analyses that are reported. All path coefficients were obtained using the SPSS Regression program (Nie et al., 1970). The same notation as used in previous chapters is again used to indicate if a path coefficient represented a substantial relationship.

Level of Analysis: Between Classes

1 Outcome - Science Achievement. The antecedent measures of occupational status and prior achievement both enhanced the student's occupational aspirations (0.34 and

0.52) and educational aspirations (0.33 and 0.64), the amount of homework done each week (0.36 and 0.33) as well as the level of participation in mathematics and science activities (0.11 and 0.28). In addition, occupational status had a positive effect upon academic motivation (0.25) and self regard (0.38) while prior achievement positively influenced liking for school (0.43). Classes with high prior achievement levels contained students who participated less in pop culture activities (-0.58). In the presence of prior achievement, occupational status was not a good predictor of science achievement. Similarly, occupational aspirations (0.15) and educational aspirations (0.26) were the only attitude measures with substantial path coefficients to science achievement. The absence of other substantial results was partly due to the strong relationship between prior achievement and science achievement at the class level. This relationship was characterized by path coefficients generally in excess of 0.8.

2 Outcome - Science Attitudes. As for achievement outcomes, occupational status was not a substantial predictor of attitudes to science except in one case. Prior achievement and all the attitude and practice measures had a substantial effect upon attitudes to science. The behaviour of the model when occupational and educational aspirations were considered was unusual. The normally weak but positive relationship between occupational status and attitudes to science changed to a weak negative relationship and the path coefficients between prior achievement and attitudes to science decreased markedly, suggesting that these two measures might be causing suppressor effects.

Level of Analysis: Between Students

1 Outcome - Science Achievement. At this level of analysis, both occupational status and prior achievement influenced occupational (0.22 and 0.19) and educational aspirations (0.28 and 0.36). In addition, occupational status positively influenced academic motivation (0.10) and the amount of homework done per week (0.12). Prior achievement influenced liking for school (0.12) and self regard (0.13). Also, the more able students spent less time on pop culture activities (-0.16). At this level of analysis there were fewer substantial path coefficients than at the between classes level, but otherwise the findings were similar to those at that level except for the very weak influence of prior achievement upon mathematics and science activities at the between students level (0.09) and the contrasting effects of occupational status and prior achievement upon self regard. At the class level, self regard appeared to result from the overall home background status of the class (0.38), while at the individual level, the student's own academic performance (0.13) appeared to be the most influential factor. The importance of these two effects will be considered in comparison with the findings from the between students within classes analysis.

Table 7.1 Path Coefficients for the Analysis of the Causal Model at the between Classes, between Students and between Students within Classes - Levels of Analysis

Attitudes and practices regressed on	Between classes level (N = 72)		Between students level (N = 1986)		Between students within classes level (N = 1986)	
student attitudes/practices	Occupational status P31	Prior achievement P32	Occupational status P31	Prior achievement P32	Occupational status P31	Prior achievement P32
Occupational aspirations	0.34	0.52	0.22	0.19	0.16	0.10
Educational aspirations	0.33	0.64	0.28	0.36	0.21	0.22
No. of hours homework/week	0.36	0.33	0.12	0.09	0.04	0.00
Pop culture activities	0.05	-0.58	-0.03	-0.16	-0.02	-0.10
Maths/science activities	0.11	0.28	0.04	0.09	0.03	0.06
Academic motivation	0.25	0.09	0.10	0.04	0.07	0.02
Self regard	0.38	-0.03	0.04	0.13	0.02	0.09
Like school	0.03	0.43	0.09	0.12	0.07	0.07

Science achievement regressed on	Between classes level (N = 72)			Between students level (N = 1986)			Between students within classes level (N = 1986)		
student attitudes/practices	Occupational status P41	Prior achievement P42	Attitude P43	Occupational status P41	Prior achievement P42	Attitude P43	Occupational status P41	Prior achievement P42	Attitude P43
Occupational aspirations	0.02	0.83	0.15	0.10	0.68	0.11	0.06	0.58	0.08
Educational aspirations	-0.02	0.74	0.26	0.07	0.63	0.19	0.04	0.55	0.14
No. of hours homework/week	0.08	0.91	-0.03	0.12	0.69	0.05	0.07	0.59	0.03
Pop culture activities	0.07	0.89	-0.03	0.12	0.68	-0.07	0.07	0.58	-0.08
Maths/science activities	0.06	0.89	0.06	0.12	0.69	0.07	0.07	0.58	0.08
Academic motivation	0.05	0.90	0.07	0.12	0.69	0.10	0.07	0.58	0.11
Self regard	0.05	0.91	0.06	0.12	0.69	0.07	0.07	0.58	0.05
Like school	0.07	0.88	0.06	0.12	0.68	0.11	0.07	0.58	0.11
Science attitudes	P41	P42	P43	P41	P42	P42	P41	P42	P43
Occupational aspirations	-0.01	0.23	0.30	0.04	0.27	0.13	0.04	0.22	0.11
Educational aspirations	-0.04	0.14	0.40	0.00	0.20	0.25	0.02	0.20	0.18
No. of hours homework/week	0.02	0.33	0.19	0.05	0.28	0.14	0.05	0.24	0.12
Pop culture activities	0.10	0.30	-0.17	0.06	0.28	-0.06	0.05	0.23	-0.04
Maths/science activities	0.03	0.26	0.48	0.05	0.26	0.35	0.04	0.22	0.27
Academic motivation	0.00	0.36	0.36	0.02	0.27	0.42	0.03	0.23	0.33
Self regard	0.01	0.40	0.22	0.06	0.26	0.22	0.04	0.19	0.53
Like school	0.08	0.20	0.45	0.03	0.24	0.40	0.03	0.21	0.32

* Path coefficient greater than 0.10 are underlined.

Occupational aspirations (0.11), educational aspirations (0.19), academic motivation (0.10) and liking for school (0.11) all substantially influenced science achievement at this level of analysis. The effect of prior achievement was weaker at the individual level with most path coefficients being just below 0.7, while most of the path coefficients between occupational status and achievement were around the threshold value of 0.1. Although occupational and educational aspirations were among the strongest attitudinal predictors of science achievement, their relative strength was not as pronounced as at the class level.

In general terms, the path coefficients between occupational status or prior achievement and the attitude and practice measures were generally larger at the class level owing to the clustering of students within classes. Similarly, prior achievement was a stronger predictor of achievement at the class level. On the other hand, occupational status and some of the attitude measures were stronger predictors at the individual level although occupational and educational aspirations were an exception to this statement. The strong influence of these two measures, particularly at the class level, has already been noted at other stages of the study.

2 Outcome - Science Attitudes. Prior achievement was a consistent predictor of good attitudes to science and all the other attitude and practice measures substantially influenced the student's liking for science except for participation in pop culture activities. Occupational status was not a substantial predictor of attitudes to science. The suppressor effects observed with occupational and educational aspirations at the between classes level were not evident at this level although the path coefficients between prior achievement and occupational status and attitudes to science had their lowest value (0.20 and 0.00 respectively) when educational aspirations were incorporated into the model.

Level of Analysis: Between Students within Classes

1 Outcome - Science Achievement. The purpose of conducting the analysis at the student within class level was to detect if the relative level of occupational status, prior achievement or any of the other measures were important in determining achievement and affective outcomes. The relative level refers to the student's status or performance compared to the class to which the student belonged. At this level of analysis, both occupational status and prior achievement substantially enhanced occupational (0.16 and 0.10) and educational aspirations (0.21 and 0.22). In addition, prior achievement differences reduced the level of participation in pop culture activities (-0.10). Although these results were also obtained at the other two levels of analysis, there were less substantial results at the student within class level. Major differences are discussed after all the results are presented.

At the student within class level, prior achievement (approximately 0.58), educational aspirations (0.14), academic motivation (0.11) and liking for school (0.11) had a substantial influence upon science achievement. Prior achievement had a weaker effect on science achievement at this level of analysis with all path coefficients being below 0.6. Occupational status was not substantially related to science achievement.

2 Outcome - Science Attitudes. Prior achievement had a consistently strong effect on liking for science (approximately 0.22) while occupational status did not. Although self regard (0.53) had the greatest effect from among the attitude and practice measures, all but one of these measures had a substantial effect on attitudes towards science.

Summary and Discussion

As discussed in the introduction to this Chapter, the differences between the findings at the three levels of analysis were considered to indicate the manner in which certain variables acted upon other measures. Also, it was necessary to determine if different findings were possible using different levels of analysis. This was in response to the warning given in relation to drawing inferences across different levels of analysis. To facilitate a more orderly summary and discussion of these issues, the findings have been divided into three categories.

1 Effect of occupational status and prior achievement upon attitude and practice measures. Several discrepancies were noted during the initial presentation of the findings. The most notable was the relative influence of the two antecedent measures upon self regard at different levels of analysis. At the class level, occupational status enhanced self regard, while at the student level, prior achievement enhanced self regard. However, at the student within class level, neither prior effect had a substantial influence on self regard. It appears that students raise their self regard in two ways; by belonging to a high status class or by performing well academically as an individual. As such, self regard is a function of the status of the class but also of the individual performance of the student. Hence, being from a high status home but belonging to a low status class would not significantly enhance self regard unless supported by superior achievement levels.

The measure of the amount of homework done each week also depends upon the level of analysis. More able classes and high status classes do substantially more homework, high status students do more homework, but more able students in a particular class appeared to do no more homework than less able students in the same class. This results suggests that teachers set homework according to the ability and status of the class such that within a given class all students do much the same amount of homework.

Another inconsistency concerned the influence of the antecedent variables upon the attitude to school and academic motivation measures. At the student within class level, no substantial relationships were found, yet prior achievement enhanced liking for school, and occupational status improved academic motivation at the other two levels of analysis. It would appear that these two results represent absolute effects, not relative effects, in relation to the respective prior effect. The other results obtained were similar regardless of the level of analysis although minor differences were present. These included both prior effects substantially influencing participation in science and mathematics activities at only the class level.

2 Effect of occupational status, prior achievement and attitude and practice measures upon science achievement. Regardless of the unit of analysis, prior achievement was a consistent predictor of science achievement. Occupational status was only a substantial predictor at the individual level which suggested that relative occupational status was unimportant for science achievement while the effect at the class level was probably reduced by the strong prior achievement effect. Among the attitude and practice measures, occupational and educational aspirations were stronger predictors of science achievement at the class level, although educational aspirations possessed substantial path coefficients at all three levels of analysis and occupational aspirations had a substantial effect at both the class and student level. These results were consistent with the effect of compounding of aspirations that has already been noted in high ability classes and the detection of effects between prior achievement and aspirations and between aspirations and achievement at all levels of analysis thus supporting the proposition that the grouping of more able students enhanced aspirations beyond those levels predicted by achievement measures only.

Liking for school and academic motivation were the only other substantial predictors of science achievement at specific levels of analysis but the results were generally borderline. Overall, very few differences were detectable since many of the predictors failed to influence science achievement at any level of analysis.

3 Effect of occupational status, prior achievement and attitude and practice measures upon attitudes to science. It was with this set of path coefficients that the least variation in results occurred between the levels of analysis. Prior achievement had a consistently strong path coefficient to attitudes to science. Occupational status was not a substantial predictor of liking for science except for one borderline case. Most of the attitude and practice measures appeared to influence attitudes to science in a similar manner except for participation in pop culture activities which only damaged attitudes to science substantially at the class level.

Conclusion

The results and comparisons discussed on the previous pages mean that it is very difficult to comment on the relative merits and appropriateness of different levels of analysis. This is, in part, due to the similarity in the findings at all three levels of analysis. As such, the suitability of a particular level of analysis to the causal model cannot be determined from the analyses carried out. Evidence of some of the problems associated with different levels of analysis were observed. The strong relationship between prior achievement and science achievement at the between classes level could be construed to be a result of aggregation effects described by Burstein (1980). Alternatively, this strong effect could be interpreted as a purer measure of the relationship between prior achievement and science achievement without the 'nuisance' effects caused by individual differences that Blalock (1964) noted at the student level.

Although the comparison of different levels of analysis did not incorporate the primary issue of class size, it did provide assurance that the sets of analyses considered in the two previous chapters were suitable for an investigation of the class size question since the results obtained at different levels of analysis generally supported each other. More importantly, the similarity between the results at the different levels of analysis was tentative support that the measures used in the model were appropriate to all levels of analysis. Although startling differences between levels of analysis might have been more interesting, a consistency of results is nevertheless reassuring. Furthermore, some of the small differences that did occur between different levels of analysis have provided some insight into the more subtle mechanisms acting between measures and, even if only in a small way, has increased the understanding of the educational setting.

CHAPTER 8

RECIPROCAL CAUSAL MODEL WITH ATTITUDES AND CLASS SIZE

One of the major results presented in Chapters 5 and 6 was the consistent positive relationship between class size and achievement. This finding contradicted the work of Glass and Smith (1978). Although larger classes contained more able students, this result was obtained after controlling for prior achievement and father's occupational status and after adjustment for attitudes and, at the class level of analysis, the processes occurring within the classroom. It appeared that class size, other things being equal, was able to produce further achievement gains. Moreover, these gains were distinct from any influence of teaching practices and the effect of superior attitudes because these factors had been incorporated into the causal model.

It appeared that the gains in achievement possible in larger classes were more than just a direct consequence of larger classes containing more able or better motivated students, and therefore would seem to be dependent upon certain other characteristics of the students that formed the classes. It was suggested that the students in larger classes interacted with each other in some way to amplify achievement beyond the levels expected from either prior achievement or home background. The implication was that these interactions were a product of the way certain students were allocated to larger classes, and the manner in which these classes generated superior aspirational levels was leading to the reinforcement of achievement.

The criterion for allocating students to larger classes, apart from prior achievement levels, was uncertain. Clearly, the attitudes of students in larger classes was sufficiently good for them to maintain a positive approach to their work. Although no attitude measures were available at the commencement of the school year, favourable attitudes appeared to be necessary for a student to function well in a larger class. In addition, favourable attitudes seemed to be important for a student to be allocated to a larger class. The effect of class size upon attitudes has already been noted. In the presence of prior achievement controls, class size damaged attitudes to science, although class size had a significant positive effect upon both a student's occupational and educational aspirations. There would appear to be a two-way interactive relationship between class size and attitudes. The effect of class size upon attitudes is well documented, but simultaneously, the attitudes of students seems to be important in the formation of classes.

The absence of any prior attitude measures made assessment of the effects of attitudes upon the size of the class to which a student was allocated difficult. The use of the attitude measures obtained during Year 7 was not theoretically justifiable since the temporal order of measurement was not fully consistent with the causal sequence of the model. The suitability of an attitude measure obtained during the year for use as an antecedent measure of attitude would depend upon the stability of attitudes over the time period of the investigation. Since this time interval covered the important transitional period from primary to secondary school, the use of the same attitude measure appears to be even less defensible. For these reasons, it was not possible to construct a causal model to assess the effects of attitudes upon class size directly. Instead, the use of a reciprocal causal model was considered to be an appropriate way to approach the problem of interaction between attitudes and class size.

The distinction between a reciprocal or non-recursive model and a recursive causal model is that the former allows for a two-way interaction between two specified variables. The reciprocal model in this study recognizes that attitudes influence class size and, in turn, class size influences attitudes. The use of an interactive model does not dispel the difficulties associated with using attitude measures obtained during Year 7. However, instead of strongly asserting that attitudes measured after the classes have been formed had influenced class size, the inferences are slightly weaker in the sense that the reciprocal model simultaneously considers effects operating in both directions. Furthermore, the model would contain the indirect effects of prior achievement and occupational status upon class size through the attitude measures.

A non-recursive causal model incorporating the interaction between class size and attitudes as well as prior achievement and occupational status is presented in Figure 8.1. The model was developed from an example given by Hauser (1973) of an interactive model in a sociological setting. The model only considers science classes since no prior achievement measures were available for mathematics. In addition, five attitude measures were identified in Chapter 5 as influencing achievement after allowance had been made for differences in class size. The five attitudes were occupational and educational aspirations, academic motivation, participation in mathematics and science activities and liking for school. These five attitude measures were considered in the non-recursive causal model. Achievement outcomes were not included since this non-recursive model could be superimposed upon existing achievement models. The path coefficients between the variables in the model and science achievement would not change as a consequence of the introduction of a reciprocal link between attitudes and class size.

The crucial feature of the reciprocal model which permits the estimation of the effects of attitudes upon class size and vice versa is the assumption that father's occupational status only affects class size indirectly. That is, father's occupational status influences class size through its relationship with prior achievement which directly affects class size and also by an indirect influence through attitudes. The number of path coefficients to be found and the number of correlation coefficients available for substitution into the regression equations required the deletion of one

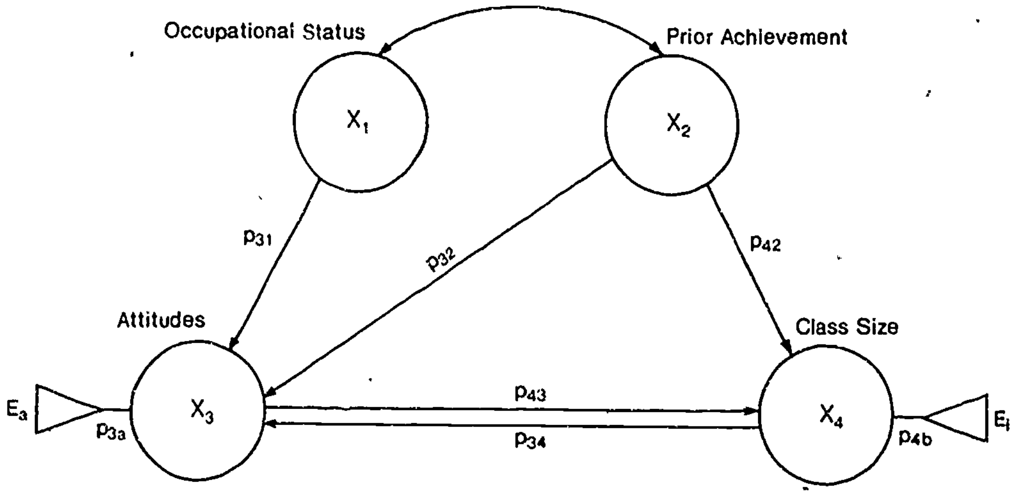

Figure 8.1 Reciprocal Causal Model Relating Attitudes and Class Size

causal path and the use of an instrumental variable formed as a consequence of the proposed deletion. The direct influence of occupational status upon class size appeared a highly plausible one to remove. Thus, father's occupational status acted in the revised model as an instrumental variable. The validity of omitting this path depends clearly upon the strength of the observed relationship between occupational status and class size. The model would not be valid if this observed relationship were too great.

Calculation of Path Coefficients in a Reciprocal Model

Because of the reciprocal effect, the path diagram no longer provides accurate guidance in writing regression equations and greater reliance must be placed upon the application of the basic theory of path analysis. The major equations for the model are:

$$X_4 = p_{42}\, X_2 + p_{43}\, X_3 + p_{4b}\, E_b \qquad (1)$$

$$X_3 = p_{31}\, X_1 + p_{32}\, X_2 + p_{34}\, X_4 + p_{3a}\, E_a \qquad (2)$$

It must be assumed that

$$r_{a1} = r_{a2} = r_{b1} = r_{b2} = r_{ab} = 0.$$

By considering equation (1), expressions for the correlations of x_4 can be written:

$$r_{41} = p_{42}\, r_{21} + p_{43}\, r_{31}$$

$$r_{42} = p_{42} + p_{43}\, r_{32}$$

These are not the normal symmetric equations of multiple regression analysis, but they can be solved for p_{42} and p_{43}. By considering equation (1) again, the following expressions can be obtained:

$r_{43} = p_{42} r_{23} + p_{43} + p_{4b} r_{b3}$, and

$r_{b3} = p_{34} r_{4b}$

whence,

$$r_{43} = p_{42} r_{23} + p_{43} + p_{4b} p_{34} r_{4b} \qquad (3)$$

Also, we find that

$$r_{44} = 1 = p_{42} r_{24} + p_{43} r_{34} + p_{4b} r_{b4} \qquad (4)$$

Rearranging equations (3) and (4), p_{34} is obtained as

$$p_{34} = \frac{r_{43} - (p_{42} r_{23} + p_{43})}{1 - (p_{42} r_{24} + p_{43} r_{34})}$$

By applying the basic theorem to equation (2) to find the correlations of v_3 with the other variables and by shifting terms involving p_{34} to the left hand side, the following equations are obtained:

$$r_{31} - p_{34} r_{41} = p_{31} + p_{32} r_{21} \qquad (5)$$

$$r_{32} - p_{34} r_{42} = p_{31} r_{21} + p_{32} \qquad (6)$$

Since p_{23} and all the correlation coefficients are known, equations (5) and (6) may be solved for p_{31} and p_{32}. Hence, all path coefficients in the model can be calculated.

The path coefficients of the non-recursive causal model for each of the five attitude measures at both the class and individual student level were estimated using the equations summarized above.

Level of Analysis: Between Classes

The correlation matrix between all the measures at the class level is presented in Table 8.1. The coefficients were obtained using the average of each variable for the 72 classes. The most notable feature of the correlation table was the large correlation between occupational status and class size. The magnitude of the link between these two measures (r = 0.36) suggested that the basic assumptions of the reciprocal model were violated at this level of analysis.

Table 8.1 Correlation Coefficients between Reciprocal Model Measures at the between Classes Level

N = 72	Occupational status	Prior achievement	Class size
Prior achievement	0.79		
Class size	0.36	0.56	
Occupational aspirations	0.76	0.79	0.57
Educational aspirations	0.83	0.90	0.59
Academic motivation	0.33	0.29	0.24
Maths/science activities	0.34	0.37	0.17
Like school	0.38	0.45	0.28

Level of Analysis: Between Students

The correlation matrix obtained from the measures at the individual level is presented in Table 8.2. The table was obtained by considering the measures for all 1986 students used in the study. The correlation coefficient between occupational status and class size was trivial ($r = 0.08$) and indicated that only a weak link existed between these measures at the between student level. Therefore, it was possible to use the reciprocal model with some confidence since the basic assumptions appeared to be satisfied.

The path coefficients obtained from the equations developed earlier in this chapter are presented in Table 8.3. The path coefficients are very similar to those in the ordinary causal model discussed in Chapter 6. Prior achievement influenced class size and all attitudes except academic motivation. Occupational status had an effect on all attitude measures except for participation in mathematics and science activities.

For the reciprocal effects, all five attitudes had positive path coefficients towards class size, but the effect was weakest for the two aspirational measures. Although this statement is dependent upon the stability of the attitude measures, there was support for the claim that students with favourable attitudes were placed in larger classes. This effect was distinct from the influence of prior achievement levels. In reverse, class size

Table 8.2 Correlation Coefficients between Reciprocal Model Measures at the between Students Level

N = 1986	Occupational status	Prior achievement	Class size
Prior achievement	0.29		
Class size	0.08	0.20	
Occupational aspirations	0.28	0.25	0.16
Educational aspirations	0.38	0.44	0.23
Academic motivation	0.11	0.07	0.05
Maths/science activities	0.07	0.10	0.00
Like school	0.13	0.15	0.07

Table 8.3 Path Coefficients for the Reciprocal Causal Model Relating Attitudes and Class Size at the between Students Level[a]

N = 1986	Occupational status to attitudes P_{31}[a]	Prior achievement to attitudes P_{32}	Prior achievement to class size P_{42}
Occupational aspirations	0.22	0.18	0.18
Educational aspirations	0.27	0.34	0.17
Academic motivation	0.10	0.07	0.19
Maths/science activities	0.05	0.17	0.16
Like school	0.10	0.15	0.17
	Attitudes to class size P_{43}		**Class size to attitudes P_{34}**
Occupational aspirations	0.09		0.03
Educational aspirations	0.07		0.09
Academic motivation	0.19		-0.16
Maths/science activities	0.42		-0.45
Like school	0.20		-0.17

a Path coefficients greater than 0.10 are underlined.

did not have a substantial effect upon educational and occupational aspirations. However, students from larger classes had lower levels of academic motivation and liking for school and spent less time participating in mathematics and science activities. In relative terms, these results were similar to those presented in Chapter 6, but the detrimental effect of class size upon certain attitudes was more pronounced in the reciprocal model.

The path coefficients obtained from the causal model presented in Figure 8.1 can be combined with the results obtained from the model given in Figure 6.1 to form a complete occupational status, prior achievement, attitude, class size and science achievement model included reciprocal paths between attitudes and class size. This complete model is presented in Figure 8.2 and gives the most extensive picture of the class size issue, at least at the individual level since classroom processes were not considered in the model.

The results from the non-recursive model emphasise the differences between the aspirational measures and the other attitude measures. Larger classes enhanced aspirations slightly while they damaged other attitudes. As such, the student's aspirations appeared to profit marginally from the increased diversity and competition provided in large classes. There was thus some evidence that well-motivated, as distinct from more able students, were placed in larger classes. However, this result should be treated with some caution due to the lack of a clearly defined temporal sequence in the measurement of the attitude variables. The manner in which School Principals or Year

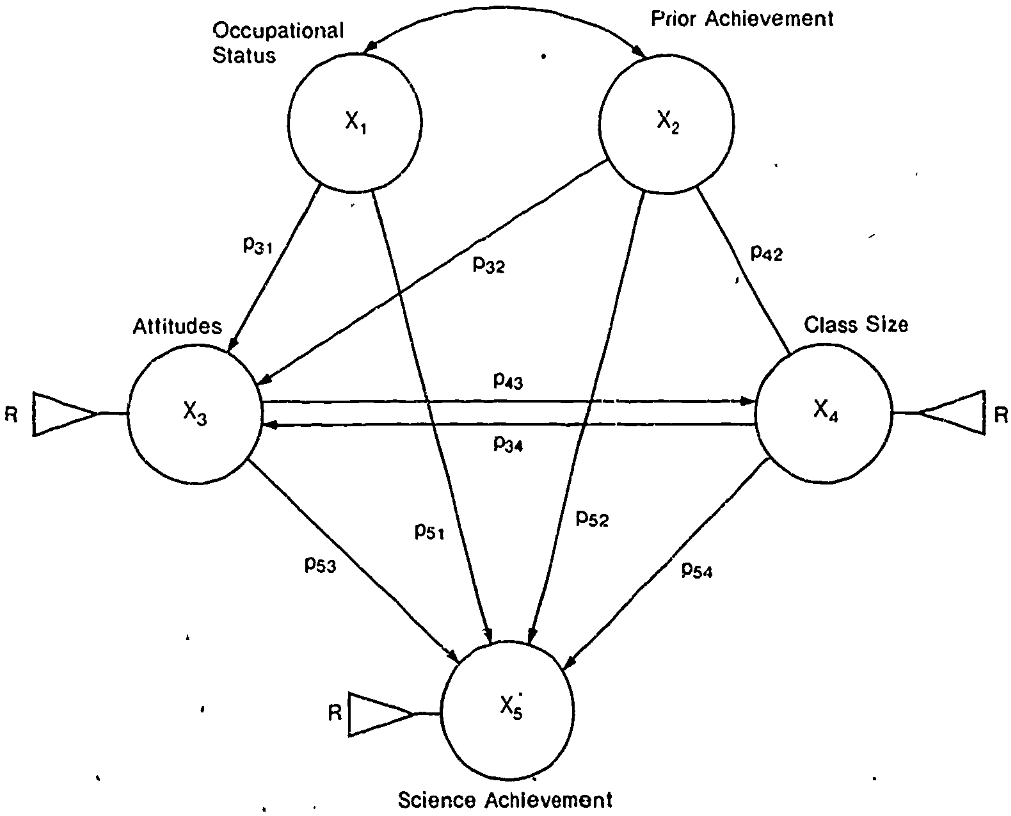

Figure 8.2 Reciprocal Causal Model with the Effects of Attitudes and Class Size upon Educational Outcomes

Co-ordinators perceived these attitudinal differences between students in the absence of formal measurements was also uncertain.

Summary and Conclusion

The result that remained inconsistent with the expected findings of the study was the positive relationship between class size and achievement. Most other findings have been either consistent with previous research or intuitively acceptable. The desire to explain this link has led to the inclusion of attitude measures into the model to assist in refining the link between class size and achievement. The attitude measures weakened the link but class size remained a significant predictor of achievement. Although additional predictors may further reduce the size of the path coefficient, the inclusion of other measures is hard to justify in terms of the nature of the model. The suggestion that larger classes containing more able students led to increased achievement gains has been confirmed in two regards by the reciprocal model. As has already been noted several times, larger classes enhance the student's aspirational levels and in turn, improved aspirations enhance achievement.

Alternatively, the possibility that students with favourable attitudes were allocated to larger classes would seem to suggest that these students were more capable of handling the classroom environment that they encountered in larger classes and therefore the effects of large classes might not detract from their achievement levels to the same extent as suggested by Glass and Smith (1978). All these suggestions are intended to support the earlier proposition that the assembling of more able students, usually into larger classes, had the effect of producing a class atmosphere of competition, stimulation and a general desire to learn which gave rise to a group effect beyond that of the ability of the individual students.

CHAPTER 9

MULTILEVEL ANALYSIS OF CLASS-SIZE DATA

In the previous chapters the data available from the Australian Capital Territory Year 7 population of students in the 15 schools in 1969 have been examined at three different levels of analysis: between classes, between students, and between students within classes. The findings from these analyses have shown that, while there were many similarities between the results obtained at the three levels, there were also some interesting differences. Thus it would appear that the issue is less one of choosing an appropriate unit of analysis for the data, but rather of the conceptualization of the problem and the identification of research questions that might be answered by using one or more levels of analysis. However, it is also possible that alternative approaches to the analysis of data might be found which would combine analyses at different levels in order to make full use of the different types of information that had been obtained from classrooms, students and from the students relative to the classroom group in which they were placed.

In searching for alternative strategies for the examination of data that had been collected at different levels, it was important to recognize (see Linn and Werts, 1969) that prior effects, including both home background characteristics and prior achievement, not only influenced very significantly final performance, but also teaching behaviours, classroom characteristics and the behaviours of the students both within and between classrooms. Unless appropriate allowance were made both for the influence of prior effects on final achievement and for the influence of prior effects on the treatment or mediating conditions in the classroom and thus on final performance, the analyses carried out would be of questionable value (see Keeves and Lewis, 1983). As argued in earlier chapters the use of some form of regression analysis, or an equivalent type of analysis such a path analysis is desirable, because the data were collected without randomly assigning students to classrooms and classrooms to treatments. That is, the data were derived from a natural setting and not generated through an experimental design. Burstein, Linn and Capell (1978) have suggested that the slopes obtained from the series of between students within classes regression analyses could be considered as outcomes in the multilevel analysis of such data. The origins of this approach are uncertain, but the idea of undertaking regression analyses in which the dependent variables are the regression coefficients that have been obtained from other regression analyses conducted at a lower level of analysis has been recognized for many years by social science research workers, econometricians, and statisticians. However, only recently has this strategy of analysis been discussed by research workers in education (see Cooley, Bond and Mao, 1981) and it does not appear to have been employed for the analysis of data obtained from school and classroom studies. A possible reason why it has not been used is the obvious difficulty likely to be encountered in interpreting the results obtained from such analyses. Furthermore, the degree of sophistication required of users in order to cope with the treatment of multilevel analysis has clearly made the use of such approaches very difficult for non-statisticians. Nevertheless, this commonly encountered problem in educational research where data are available at two levels of observation (for example, students and the classrooms or schools to which they belong) suggests that the use of the procedures of multilevel analysis should be investigated within the context of educational research.

The Framework of Multilevel Analysis

In this treatment of multilevel analysis the formulation advanced by Mason, Wong and Entwisle (1983) has been followed. The data available in this present study were collected at two levels of observation; at the student or micro level and at the higher classroom or macro level. It is assumed that student performance at the micro level depends on such factors as the prior effects of home background and prior achievement as well as student attitudes, and that the influence of these determinants at the micro level will vary systematically with the fixed effects context variable of class size. In this way student performance is seen to be influenced indirectly by class size.

In this statement we assume that there are the following variables:

one micro level response variable	- final achievement (Y)
three micro level regressors	- prior achievement (X_1)
	- father's occupation (X_2)
	- student attitudes (X_3), and
one macro level regressor	- class size (G_1).

The micro level equation may be stated:

$$Y_{ij} = b_{0j} + b_{1j}X_{1ij} + b_{2j}X_{2ij} + b_{3j}X_{3ij} + \varepsilon_{ij} \qquad (1)$$

where j = 1..., J for classrooms;

i = 1..., n_j for students within classrooms;

$N = \sum_{j=1}^{j=J} n_j$ for the total number of students; and

ε_{ij} = random error at the micro level.

At the macro level the equations may be stated

$$b_{0j} = c_{00} + c_{01}G_{1j} + \alpha_{0j} \qquad (2)$$

$$b_{1j} = c_{10} + c_{11}G_{1j} + \alpha_{1j} \tag{3}$$

$$b_{2j} = c_{20} + c_{21}G_{1j} + \alpha_{2j} \tag{4}$$

$$b_{3j} = c_{30} + c_{31}G_{1j} + \alpha_{3j} \tag{5}$$

where α_{kj} = random error at the macro level for $k = 0,1,2,3$.

These equations are written with the usual assumptions associated with the rank condition, and with the error terms at the micro level independent of the errors at the macro level. Equations (2) to (5) represent the effects of the contextual variable - class size (G_1) on the four parameters of the micro level model, and it is assumed that once the systematic component associated with class size has been removed from b_0, b_1, b_2, and b_3 their resulting variability is strictly random.

A single equation can be stated for the multilevel model by substituting equations (2) to (5) in equation (1).

$$\begin{aligned} Y_{ij} = {} & c_{00} + c_{01}G_{1j} + c_{10}X_{1ij} + c_{11}X_{1ij}G_{1j} \\ & + c_{20}X_{2ij} + c_{21}X_{2ij}G_{1j} + c_{30}X_{3ij} \\ & + c_{31}X_{3ij}G_{1j} + \varepsilon'_{ij} \end{aligned} \tag{6}$$

This equation has no unusual estimation or computation problems and analysis can proceed using ordinary least squares regression analysis. However, the analysis must be carried out with caution because Mason, Wong and Entwisle (1983) have presented a case where unsatisfactory results were obtained with such an analysis and where a Bayesian perspective yielded results which would otherwise have remained hidden.

Some Issues of Analysis

The analyses can be carried out at the micro level with one, two or three predictor variables included in the regression equations. The number of students in the classrooms for which complete data were held, with the exception of one classroom, ranged from 15 to 45. However, one classroom contained too few students for effective regression analysis with three variables and was deleted from the analyses at the micro level. Thus instead of the 72 classrooms included in the analyses in the earlier chapters there were only 71 classrooms which provided data for the subsequent analyses at the macro level. In the presentation that follows, the analyses at the micro level have been undertaken successively with prior achievement in science alone as the predictor variable, with prior achievement in science and father's occupation as the two predictor variables, and with prior achievement in science, father's occupation and educational aspirations as the three predictor variables in regression analyses with final achievement in science as the criterion variable. Educational aspirations was chosen as the most powerful of the measures of students' attitudes from the analyses presented in earlier chapters of this report, and as the variable, from among the five attitudinal variables, most likely to have a stable and substantial effect.

A question arises as to whether to use the standardized or the unstandardized (or metric) regression coefficients obtained from the analyses at the micro level as the criterion variables for the analyses at the macro level. It must be recognized that the standardized regression coefficients have been calculated making allowance for the variances and covariances of the variables included in the regression model, as well as the variances of the variables not included in the model but included under the error term. Thus these coefficients may be compared across the same sample but not between samples. However, the unstandardized regression coefficients remain relatively stable across different settings or samples and therefore can be used for the purposes of generalization across settings and samples (Pedhazur, 1982). In this study, it was proposed to compare the regression coefficients across 71 classroom settings and thus the unstandardized regression coefficients had to be employed.

Only final achievement in science has been used as the criterion variable in the analyses which follow. While a multilevel examination of the factors influencing science attitudes would have been of some value, it was the strong and persistent effect of class size on achievement in science that was seen to be of greatest interest. Not only was this relationship of interest as a substantive question, but also the strength and consistency of the relationships reported at all three levels of analysis suggested that the present situation in which the contextual variable of class size was found to be related to science achievement was an appropriate one in which to examine the usefulness of such analytical procedures.

Macro Level Regression Analyses of Class Size Data

The results of the macro level regression analyses of the effects of class size on the regression coefficients obtained at the micro level with three, two and one of the predictor variables included in the regression equation are presented in Table 9.1. Levels of significance are recorded at the 10 per cent, five per cent and one per cent levels, and both the standardized and unstandardized or metric coefficients are reported for each analysis of the set of 71 unstandardized or metric coefficients obtained by analyses at the micro level.

In the analysis of the three predictor model, with prior achievement in science, father's occupation and educational aspirations as the predictor variables and with final achievement in science as the criterion variable, it is observed that only the regression slopes for prior achievement are significantly related to class size. The standardized

Table 9.1 Macro Level Regression Analyses of Class Size Data

N = 71 Regression slope variables	F value $F_{(1,69)}$	Significance regression coefficient P	Standardized regression coefficient b_{k1}	Metric regression coefficient C_{k1}	Intercept C_{k0}
Three predictor model					
Intercepts	0.08	>0.1	0.033	0.034	12.64
Slopes for prior achievement	3.35	<0.1	0.215	0.014	0.59
Slopes for father's occupation	0.79	>0.1	0.107	0.019	-0.89
Slopes for educational aspirations	0.06	>0.1	-0.029	-0.004	0.72
Two predictor model					
Intercepts	0.34	>0.1	0.070	0.068	13.62
Slopes for prior achievement	6.46	<0.05	0.293	0.018	0.53
Slopes for father's occupation	0.28	>0.1	0.064	0.011	-0.76
One predictor model					
Intercepts	3.53	<0.1	0.220	0.170	9.09
Slopes for prior achievement	8.26	<0.01	0.327	0.021	0.50

regression coefficient of 0.215 is both substantial and significant, if significance testing is considered appropriate in the analysis of data that comprise almost a total population. In addition, it should be noted that while the standardized regression slope of 0.107 for father's occupation is not significant, and the F value is less than one, the slope might be considered substantial in terms of the criteria specified in earlier chapters of this report, namely, exceeding 0.1, since in excess of one per cent of variance is explained. Under these circumstances it would seem appropriate to eliminate the least significant variable, namely, educational aspirations, from the regression analyses at the micro level of between students within classrooms and repeat the analyses with only two predictor variables.

In the analyses of the two predictor model with educational aspirations excluded, the results obtained are similar to those for the three predictor model, except in so far as the relationships associated with father's occupation at the macro level are of reduced size and are no longer substantial. In addition, the relationships for prior achievement in science are significant at the five per cent level and the standardized regression coefficient has increased to 0.293. The evidence from these analyses indicates that the model would be improved by the deletion of a second variable, father's occupation, from the analysis at the micro level, thus reducing the model to a one predictor model in the analysis of the achievement in science of students within classrooms.

With only prior achievement in science as a predictor in the analyses at the micro level, it is found that there is a significant relationship between class size and the intercepts at the 10 per cent level with a substantial standardized regression coefficient of 0.220. In addition, there is a highly significant relationship between class size and the slopes of the regression lines for prior achievement in science at the one per cent level and with a standardized regression coefficient of 0.327. Nearly 10 per cent of the variance in the slopes of the regression lines is accounted for by differences in class size in this analysis at the macro level.

It is of interest to express these significant and substantial relationships in the single multilevel equation:

$$\hat{Y}_{ij} = 9.09 + 0.17G_j + 0.56X_{ij} + 0.02G_jX_{ij}$$

where $\hat{Y}_{ij}$ = fitted final achievement in science of student i in class j

G_j = size of class j, and

X_{ij} = prior achievement in science of student i in class j

This equation is highly informative since the positive coefficients for G_j and G_jX_{ij} indicate that not only is there an effect associated with class size, such that the larger the class the greater the student's level of final achievement, but that th is also an interaction effect such that for students of a higher level of initial achievement there is a greater advantage associated with a larger class size.

An Interpretation of Results

In discussing the results obtained from the analysis of the data for the one predictor model, it is noted that the residual effect for a class group associated with final achievement in science after allowance has been made for initial achievement in science is greater where the size of the class group is greater. Thus the final performance of students in larger classes is greater even after the students' scores have been adjusted for prior performance. These findings are consistent with results presented earlier in this report.

In addition, it is noted that the slopes of the between students within classes regression lines in the regression analysis of final achievement in science regressed on prior achievement in science are greater in larger classes and less in smaller classes. Thus in the period of one school year between when the initial and final achievement tests in science were administered, it was in the smaller classes where there was an effect associated with the attainment of more equal educational outcomes between the higher and lower performing students. Moreover, it follows that in the larger classes there was an effect associated with the attainment of a greater divergence in performance between the higher and lower performing students. In larger classes, the 'good get better and the poor get worse' relative to their classmates. It would be tempting to suggest that from these analyses of the slopes of the regression lines presented above, a strong case could be advanced for the establishment of smaller classes, because such classes would seemingly promote greater equality of outcomes, while larger classes would give rise to a greater inequality of outcomes. However, it must also be noted that in larger classes there is an overall gain in average level of performance of the students in the class group. Consequently, it would be the relative magnitudes of these two effects for a student, namely, class size and relative performance of the student within a class that would determine whether the student would fare better in a larger or smaller class group.

The multilevel analyses of this set of data with respect to the influence of class size, while permitting the effects of class size to be more accurately estimated, do not necessarily provide a greater understanding of the way in which class size worked to influence achievement in the science classes in the Australian Capital Territory at the Year 7 level in 1969. Beyond the fact that within smaller classes there is a trend towards greater equality of outcomes, and that within larger classes there is a trend towards a greater divergence in outcomes (and these are results that intuitively make sense), there is no further understanding provided as to why such effects might have been observed.

The relationship between father's occupation and final achievement, after allowance has been made for both prior achievement in science and attitudes of educational aspirations, are of less consequence. They are essentially similar in kind to those recorded for prior achievement and the trend in the results nevertheless is worthy of comment. Class size is positively related to the slopes of the regression lines for father's occupation as a variable in the between students within classes regression analyses. Thus in the larger classes it is those students from higher status homes who gain in performance during the school year and those from lower status homes who decline in performance relative to their classmates during the year. Moreover, in smaller classes there is greater equality in outcomes between students from high and low status homes. These trends would seem to suggest that smaller classes are associated with equalizing outcomes. However it must again be noted that associated with larger classes there is a gain in the average level of achievement in science of the class group.

Summary and Conclusion

The analyses reported above have been undertaken with some reservations for the strength of the statistical procedures employed. Although the analyses have not provided a further understanding of why class size is having the effects that it would appear to be having, they have yielded a more detailed account of how class size operates in relation to student achievement in science. In addition, the analyses have provided more accurate estimates than were available from previous analyses of the relationships between class size and achievement.

CHAPTER 10

CONCLUSION

The purpose of this study was to examine the ways in which class size affected other facets of the educational environment of the classroom. In particular, it was intended to assess the work of Glass et al. (1982) in a specific classroom setting and provide an explanation for their findings. In their work, meta-analysis techniques were used to summarize the diverse and often contradictory findings of class size research. Glass and his associates condensed their findings into a series of graphs relating class size with both achievement and affective outcomes. In the circumstances considered, they found increased class size to be detrimental to the student's achievement and attitudes, and also to certain aspects of the educational environment.

This study concentrated upon the 1969 Year 7 population of Canberra. Data collected by Keeves (1972) was used throughout the study. The performance of students was assessed by achievement tests and attitude scales. Home background information was obtained from a questionnaire and the classroom environment in science and mathematics classes was measured by interviews with teachers as well as by direct classroom observation. This information covered the background and practices of the classroom. The data provided information on the types of students and teachers, and the achievements and attitudes of students as well as on the activities pursued in the classroom for a wide range of class sizes.

Twelve Propositions Arising from the Study

Owing to the large number of classroom measures derived from the data, it was decided to sift through these variables and to consider only those that changed substantially with class size. This sifting procedure provided a shorter, but more useful, list of measures which were related to class size. These measures, in conjunction with information on antecedent variables concerning both home background and prior achievement, on criterion measures, achievement and attitude outcomes, and on class size were then considered in a series of causal models for both science and mathematics classes at different levels of analysis. Using regression analysis, the strengths of the relationships between these measures were estimated. A list of the major propositions which have resulted from the study is now presented as a summary of the findings. Each proposition is accompanied by a brief commentary. A more detailed description of the results was provided in the preceding chapters.

1 Classroom practices did not vary greatly with class size. The effect of class size upon teaching practices was greater in mathematics than science classes, but for both

subjects, only a few of the variables associated with classroom practices were strongly related to class size. This confirmed the reports of Ryan and Greenfield (1975). Reducing class size did not lead to dramatic changes in classroom practices. Instead, the teacher's own individual style appeared to be the main factor determining classroom activities. It seems that the teachers had not been taught how to exploit the opportunities that smaller classes provided and that teacher training concentrated upon the development of skills appropriate to classes of 25 or more students.

2 More able students were placed in larger classes. This proposition is supported by two factors known to be used in the allocation of students to classes. Remedial classes have been traditionally small and year co-ordinators have probably been more prepared to tolerate class size creeping up if they knew that the class contained more able students. A majority of the classes in the study were from schools using streaming procedures.

3 The classroom practices that changed with class size had little influence upon achievement outcomes. After controlling for prior achievement, father's occupation and class size, only a handful of the variables selected as being related to class size had a recognizable influence upon achievement levels in either subject. This seems to imply that the differing activities which resulted from changes in class size did not lead to achievement gains. This observation does not relate to all classroom practices, but only those that changed with class size. Previous research (Smith and Glass, 1979) had suggested that students in small classes would be more likely to participate in such activities as individualized and small group work which would enhance achievement. The small variation in teaching style for differing class sizes was a probable explanation of the lack of influence of teaching practices upon achievement. Only when more teachers are using techniques appropriate to small classes could we expect to find greater changes in achievement levels.

4 Classroom processes that changed with class size influenced the students' attitudes. The activities that teachers were able to pursue in classes of differing sizes did change student attitudes towards the subjects of science and mathematics. The effects of classroom practices upon attitudes were substantial, although not strong, for a number of classroom practices. It appeared that teaching practices which changed with class size could influence student attitudes although they did not influence achievement. The varying use of reward and punishment, support for students and the level of individual contact in smaller and larger classes were among the important practices influencing student attitudes towards their subjects.

5 Meaningful results could not be obtained from the causal models unless prior achievement measures were present. Particularly for achievement outcomes, a prior

achievement measure was necessary to account for sufficient variation in final achievement to make the other path coefficients meaningful. In the absence of a prior achievement measure, father's occupational status accounted for little variation in final achievement and this led to the other path coefficients in the model acquiring inflated values. This forced the discussion of the analyses to be restricted to science classes since prior achievement measures were not available for mathematics classes. The same effect was true for attitudinal outcomes although it was not as pronounced since the relationship between prior achievement and attitudes was not as strong as between prior achievement and final achievement.

6 Larger classes had enhanced occupational and educational aspirations. Although the relationship was not always strong, there was a consistent positive relationship between class size and these two aspirational measures in analyses at both the class and student levels. It was suggested that the general atmosphere formed by the assemblage of able students heightened student aspirations since such students interacted with other students also with high expectations.

7 Strong and favourable attitudes enhanced achievement levels. This observation might seem obvious but it warrants mentioning since attitudes, when considered in several models, helped to explain variation in achievement levels in circumstances where class size was producing a substantial positive influence upon achievement. Although the inclusion of attitudes in the causal model did not eliminate the link between class size and achievement, attitudes were important predictors of achievement. Attitudes were related to achievement at all three levels of analysis: between classes, between students and between students within classes.

8 Larger classes showed enhanced achievement levels. This represented the most controversial finding of the study. It should again be noted that this observation was made after controlling for prior achievement, father's occupation and attitudes. The original intention of the study was to look more specifically at the activities occurring within small and large classes. As already noted, an examination of the activities in classes of differing sizes produced little of consequence because of the similarities in teaching styles and practices in classes of differing sizes. This proposition clearly contradicts the finding of Glass and Smith (1978) that large classes inhibit achievement levels.

In Chapter 8, an attempt was made to explain this observation in terms of the types of students who were allocated to large classes. Using a non-recursive model, some evidence was obtained to suggest that students with both higher ability and favourable attitudes were placed in larger classes. As has already been suggested, the congregation of able students in larger classes appeared to amplify achievement to higher levels. It is contended that in larger classes a superior classroom atmosphere was possible because a higher proportion of the class were able students and this permitted the class to progress at a greater rate. The original study was primarily concerned with low inference activities within the classroom and therefore it did not contain high inference measures designed to assess the learning environment of a classroom. More specific measures of the way that students approached their work would be needed to assess this claim more accurately.

9 Increased class size has a detrimental effect upon student attitudes towards science. Because this result was only noted in the presence of prior achievement measures, the proposition was only examined for science classes. It was supported at the class level in Chapter 5 and at the student level in the non-recursive model presented in Chapter 8. This result supported the findings of Smith and Glass (1979). However, the specific activities in larger classes that damaged students' attitudes to the subject could not be identified since the teaching practices that enhanced attitudes were often more prevalent in larger classes. This proposition refers only to science classes, yet mathematics classes were considered to be of more interest with regard to teaching practices.

10 Similar relationships between most variables in the causal models were obtained at all levels of analysis. Analyses in conjunction with class size were conducted at the class and student level. Additional analyses for the purposes of comparison were conducted at these two levels as well as at the student within class level. Although some differences were observed and noted in Chapter 7, the similarity of the results at different levels provided confirmation of the findings with respect to the effects of class size.

11 The choice of an appropriate unit of analysis should consider the conceptualization of the problem and the specification of the research questions. Notwithstanding the reported similarity of the results obtained at the different levels of analysis, some differences were found that provided insight into the subtle mechanisms acting between variables and the factors measured by those variables. Thus as an understanding develops of the action of factors at different levels of observation, and the collection of data from treatments and through sampling at different levels of operation, it is necessary to give greater attention to the conceptualization of the problems being investigated and to the clearer identification of the research questions to which answers are being sought. In the tighter specification of propositions and hypotheses for investigation, it would seem essential that the appropriate units and levels of analysis should be stated. Without the specification of levels of analysis prior to the examination of data, significant dangers arise in the post hoc interpretation of findings which are associated with the effects of bias in both the aggregation and disaggregation of data. The problems of disaggregation of data, where the factors associated with disaggregated

data do not apply equally to all members of a group have had to be carefully considered in this report. However, it has been argued that class size is a contextual factor that operates equally with respect to all members of a class group.

12 Smaller classes are associated with more equal achievement outcomes. A substantial positive relationship between class size and the regression coefficient between prior and final achievement in science would seem to suggest that larger classes give rise to a greater variability in achievement and hence smaller classes produce a lower variability in achievement which could be associated with the attainment of more equal educational outcomes. It should be remembered that this result refers to the variability, and not the level, of achievement and that larger classes still appear to enhance achievement levels.

Implications for Future Research

Although these twelve propositions may represent the major findings of the study, many other findings of interest have been reported throughout the previous chapters. Instead of providing further comments upon these results, it would be preferable to address two of the major findings in the sense of suggesting further investigations necessary to clarify some of the propositions already stated. The lack of importance of the process dimension must be reconsidered. The general observation that teachers rarely adjust their teaching style for changes in class size may have been true in 1969. However, it is possible that the in-service and teacher training seen to be necessary has occurred, and it is important to ask whether teachers now modify their behaviour in response to changes in class size.

The procedure that Keeves (1972) used to assess the process dimension would still be relevant although several items could be added to make the instruments more sensitive for assessing individualized instruction. In particular, measures of individual and independent work by students must be included to assess the activities of individual students within the class as distinct from the activities of the entire class. This modification would seem desirable since it has been noted that not all students within a given class would share equally in the learning experiences provided. By investigating some of these issues, it should be possible to re-assess the effects of class size upon teacher and student behaviours.

A more central question which requires explanation is the positive relationship between class size and science achievement. Although the findings of this investigation are supported by other previous studies, they contradict the work of Glass and Smith (1978) which has suggested that class size damages achievement levels. Why were the findings for the group of students or classes considered in this study so different as to contradict the research of Glass and Smith? In particular, what activities or practices occurred in these classes to produce the conflicting results? The most plausible explanation would seem to involve the effects of grouping more able students in larger classes, but the findings also reported achievement gains beyond those expected solely from a consideration of differences in achievement levels.

Reference has been made to that largely undefined factor of classroom climate which was alleged to enhance both competition and the student desire to learn. Classroom climate is harder to measure than classroom practices because it clearly reaches into the affective domain of the classroom. Several issues associated with the climate of the classroom were considered in Chapter 8, but a more extensive list was prescribed by Smith and Glass (1979) during their summary of class size and affective outcome research. The measures that were considered in this present report sought to assess all the facets of student interest in their subject, their schooling, their family and their peers and it was also proposed that these interests should be related to the activities that they pursued at school and at home. The intention was to identify exactly why belonging to a class containing able students should make students perform better than previous achievement tests suggested that they could. By thoroughly investigating the affective links within a classroom it was hoped that the improved achievement levels observed in larger classes could be more fully understood. Clearly, further work is required.

In conclusion, this study has attempted to identify the factors associated with variation in class size and their influence upon educational achievement and attitudes. That the findings have not agreed with Glass et al. (1982) is proof that their summary work has not put the class size issue to rest. An explanation of their findings and the findings of this study is still very unclear. The reasons why a particular class size should produce achievement gains or losses still remains obscured by the diversity of activities, personalities and materials available in a single classroom. The findings of this study point towards a need for the continued examination of both the process, and particularly the attitudinal dimensions of the classroom. It is clear that an increased understanding of these features of the classroom is a necessary step towards teasing out the relationships between class size, teacher activities and student motivation, all of which appear as central themes of the class size question.

APPENDIX I

LIST AND DESCRIPTION OF AVAILABLE MEASURES

APPENDIX I

LIST AND DESCRIPTION OF AVAILABLE MEASURES

In Chapter 3, reference was made to a list of variables employed and reported by Keeves (1972). These included performance on achievement tests and attitude scales, teacher and classroom characteristics and a set of classroom observation measures. To indicate the nature of the assessment of the classroom environment, a complete listing of the variables is given below using the variable title used in this report. In addition, a brief description of the measure including its direction, type and scale is provided.

This Appendix is restricted to a description of the measures employed. Reliability coefficients, wherever appropriate, are quoted in Appendix II in connection with a discussion of the relationship between each measure and class size. Furthermore, this Appendix presents only a brief summary of each measure. A fuller description has been provided by Keeves (1972, 1974a, 1974b). To facilitate easier consideration of the variables, the measures are categorized in the same manner as used by Keeves (1972). The variable name as used throughout the study is given, followed by a description.

1 Home background, achievement and attitude measures

These measures were obtained from a general information questionnaire, achievement tests and attitude scales that were given to each student. Throughout Chapter 5 and for several analyses in Chapters 7 and 8 these measures were aggregated to form class averages. This was necessary where the class was used as the unit of analysis.

Sex of class. Coded on a two point scale with boys (1) and girls (2).

Ethnicity of home. The student's response to the question 'In which country were you born?' was coded into one of the following six categories. Over 80 per cent of responses were coded as (1).

1 Australia
2 Britain or New Zealand
3 English-speaking Commonwealth country
4 English-speaking non-Commonwealth country
5 Non-English-speaking European country
6 Non-English-speaking non-European country

Father's occupation. The students were asked their father's occupation. Occupations were coded using a 6-point scale developed by Broom, Jones and Zubrzycki (1965 and 1968) with a small residual group of unclassifiable responses. The categories and assigned scale values were as follows:

Table A.1 Results of Analysis of Variance of Father's Occupation Categories with Prior Achievement in Science

	df	Sum of squares	Mean square	F-value	Significance
Occupational categories	6	6544.6	485.9	32.4	$p < 0.001$
Error	1979	20808.6	15.0		
Total	1985	32623.7			

Occupational category	N	Group mean	Criterion scale value assigned
Unclassifiable	49	10.98	1
Unskilled workers	134	11.51	2
Semi-skilled workers	209	11.69	3
Skilled craftsmen	397	12.41	4
Clerical workers	450	13.37	5
Managerial workers	411	14.40	6
Professional, sub-professional workers	336	14.90	7

1 Unclassifiable

2 Unskilled workers

3 Semi-skilled and process workers

4 Skilled craftsmen and foreman

5 Clerical and non-commissioned servicemen

6 Managerial, farmer or shop-keeper

7 Professional, grazier or semi-professional

It should be noted that the direction of the scale has been reversed from the original study used by Keeves (1972). This was done to equate a greater score on the scale with a higher status occupation. Evidence for the assigning of scale values by criterion scaling procedures is given in Table A.1.

Student's occupational aspirations. Students were asked what occupation they expected to enter after finishing their schooling. The occupation given by the student was scaled using the same scale values as for father's occupation.

Student's educational aspirations. Students were asked the level of education they expected to attain. The categories and assigned scale values were as follows:

1 Unclassifiable

2 Four years secondary

3 Five years secondary

4 Matriculation

5 Tertiary diploma or studies

6 University degree

7 Higher degree and research

Table A.2 Results of Analysis of Variance of Educational Aspiration Categories with Prior Achievement in Science

	df	Sum of squares	Mean square	F-value	Significance
Educational categories	6	6544.6	1090.8	82.8	$p < 0.001$
Error	1979	26079.1	13.2		
Total	1985	32623.7			

Educational category	N	Group mean	Criterion scale value assigned
Unclassifiable	98	9.46	1
Four years secondary	344	10.96	2
Five years secondary	72	11.33	3
Matriculation	486	13.02	4
Tertiary diploma or studies	308	13.42	5
University degree	597	15.36	6
Higher degree and research	81	15.44	7

Evidence for the assigning of scale values by criterion scaling procedures is given in Table A.2.

Number of hours homework per week. Students were asked how many hours they spent on homework in all subjects in a week.

Student's participation in pop culture activities. This scale was designed to assess the effects of participation with friends in the 'pop culture'. It was suggested that participation in these activities would be contrary to educational goals. Four items were used in an 'Activities with Friends' questionnaire which assessed the frequency of participation in pop culture activities on a three-point scale. In general terms, scaling involved the assignment of the values indicated - Rarely (0), Occasionally (1) and Frequently (2).

Student's participation in mathematics and science activities. This scale was designed to measure the influence of students participating with their friends in activities which fostered an interest in science or mathematics. Twelve items were used in an 'Activities with Friends' questionnaire which assessed the frequency of participation in the activities using the same coding procedures as for pop culture activities.

Like mathematics. This attitude was assessed with a 10-item Likert scale in which students were required to respond to statements in the categories, 'agree', 'disagree' or 'uncertain'. The scale included a range of statements intended to identify the degree of interest and enjoyment in learning mathematics.

Academic motivation. This measure was assessed by a 20-item Likert scale in which students were required to respond in three categories to a range of statements concerning their motivation to learn at school.

Like school. This attitude was assessed with a 17-item Likert scale in the same manner as attitudes to mathematics.

Like science. This variable was assessed in the same manner as the other attitude measures with a 20-item Likert scale.

Self regard. This variable assessed the students' respect and confidence in themselves. A 17-item Likert scale with three response categories was used with a range of statements involving self-esteem with respect to the peer group.

Science achievement pretest. The initial science test was constructed from 25 items chosen from a pool of items collected for the IEA Science Project. The content of the test was as follows: twelve items tested knowledge; six items required the application of information; and seven items required analysis or evaluation. The items were field tested prior to use in the study.

Science achievement. The final science test contained 55 items, also chosen from the pool of items for the IEA Science Project. The items tested knowledge, comprehension, understanding, application, analysis and evaluation. These items were also field tested before use in the study.

Mathematics achievement. This was assessed by a 55-item test. The items covered the content areas of computation, knowledge, definitions, translation or interpretation, analysis and application. The items were field tested before being used in the study.

2 Structural dimension measures

The second category of variables represent measures of teacher, classroom and school characteristics. Many of the variables from the structural dimension are relevant to both science and mathematics classes and distinctions are only made between the subjects where necessary. If the variable is appropriate to both subjects, the variable name for science classes will be given first, followed by the equivalent name for mathematics classes in brackets. The structural dimension information was obtained from an interview and questionniare completed by the class teacher. Additional information was obtained from discussions with the principal, vice-principal or subject co-ordinator. A list of all these items is presented below.

Sex of teacher. Coded as male (1) and female (2).

Teacher years at school. The number of years of teaching at the present school was recorded.

Teacher marital status. Coded as single (1) and married (2).

Teacher science (mathematics) specialist. The teachers were asked if they regarded themselves as a science or mathematics specialist. Their responses were coded as Yes (1) and No (2).

Teacher years of education. The number of years of full-time education received by the teacher was coded in years.

Teacher years tertiary education. The number of years of full-time education beyond secondary school was recorded in years. In addition, the amount of training in specific subject areas that the teacher received was recorded in terms (1/3 year). For science teachers, the specialist areas were physics, chemistry, biology or geology, while for mathematics teachers, the specialist areas were general mathematics, pure mathematics, applied mathematics or statistics.

Total teacher training in science (mathematics). This variable measured the total number of terms of training in all of the above specialist areas. As such, it represented an aggregate of the previous measures.

Teacher training institution. Four categories were used to classify the training institution attended by the teacher. They were university (1), teachers college (2), university and teachers college (3) and other (4). These categories did not form a monotonic scale.

Teaching load. Three categories were used to assess the teaching load. They were part-time (1), full-time (2) and extra full-time (3). The extra full-time category was rarely used.

Teacher lesson preparation. The number of hours per week that the teacher spent preparing lessons was recorded.

Teacher laboratory preparation. The number of hours per week that the teacher spent preparing apparatus or materials for laboratory classes was recorded. This variable was only relevant to science classes.

Teacher hours marking. The number of hours per week that the teacher spent marking work or other teaching-related activities was recorded.

Sex of class. This item was a class average obtained by coding boys (1) and girls (2).

Proportion of class from foreign language homes. This item was calculated from the responses to the home background item which asked whether or not English was the language normally spoken in the home. The ratio of the number of students from foreign language homes to the size of the class was used to form the proportion.

Size of form cohort. This item recorded the total number of Year 7 students at the school.

Proportion of male teachers on staff. The item used the number of full-time (or equivalent) male teachers divided by the number of full-time (or equivalent) teachers in the school.

Years school open. The age of the school was recorded in years.

Degree of streaming. Streaming practices were divided into three categories. They were coded as non-streamed (1), setted (2) and fully streamed (3).

Time on science (mathematics). The time spent each week by students in class on the subject was recorded in fifteen minute intervals.

Time on science (mathematics) homework. The time spent each week by students doing homework was recorded in fifteen minute intervals.

Total time on science (mathematics). This item was the sum of the two previous items.

Time on all homework. This item was calculated as a class average using the responses received from students to the question concerning the number of hours spent each week on homework in all subjects.

Periods contact with male teachers. This item used only the basic subjects. The number of periods each week in which students were taught by a male teacher was recorded.

Number of teachers in year. The number of teachers who, during the year, had taught the class group in the subject for more than 10 lessons was recorded.

Number of regular teachers. The number of teachers who regularly shared the teaching of the subject to the class group was recorded.

Size of school. This item recorded the number of students attending the school.

Age of teacher. This item was coded on a 5-point interval scale.

Teacher years experience. This item was coded on a 5-point interval scale.

Teacher in-service training. The number of weeks that the teacher attended full-time in-service training courses during the past five years was recorded.

Teacher attends lectures on science (mathematics) teaching. The number of lecture sessions attended by the teacher during the current year concerned with the teaching of the particular subject was recorded.

Teacher attends lectures on science (mathematics). The number of lecture sessions attended by the teacher during the current year concerned with the particular subject was divided by two to form this measure.

Member of Science (Mathematics) Teachers Association. Membership by the teacher of a subject association was coded as No (1) and Yes (2).

Periods in laboratory. The time spent per week by students in the laboratory was recorded in fifteen minute intervals. This item referred only to science classes.

3 Process dimension measures

Another interview and questionnaire session was used to obtain information on items associated with teaching practices and the process dimension. The items assessed in this way differed from those obtained by direct classroom observation since they covered the use of assessment, homework and educational aids. Interviews with teachers were considered necessary since the use of all of these features would not have been detectable in only four lessons of direct classroom observation. The practices examined in this way included the nature and amount of assessment used by teachers, the presentaton and monitoring of homework, the types of educational aids employed by teachers and the frequency with which they were used. Science and mathematics teachers were asked the same questions such that the item descriptions applied to both subjects.

Assessment practices. The teacher was asked how frequently they used various forms of assessment when marks were given and recorded. Their responses were coded as rarely or never (1), sometimes or occasionally (2) and always or regularly (3). The various forms of assessment were as follows:

1 Short answer tests,
2 Extended answer, either problem or essay, tests,
3 Multiple choice objective tests,
4 Performance on homework or weekly assignments,
5 Performance in workbooks or notebooks, and
6 Performance in assessed projects or major assignments.

The sum of the scores on the six assessment items was used to indicate the range of assessment procedures employed by the teacher. Additional assessment procedures as well as homework measures are listed below.

Assessments in Term 2. The number of assessments recorded in the teacher's markbook during Term 2 was used as a measure of the frequency of assessment.

Assessments in year involving student choice. The number of assessments during the year in which the student had some choice in either the topic or the nature of the assignment was used as an index of the freedom of choice that the student was allowed.

Frequency of revision homework. The number of times each week in which the class was expected to do non-specified or revision homework was recorded.

Frequency of homework set. The number of times each week in which the class was expected to do any homework, including written, reading, learning, revision or non-specified reading, was recorded. The teachers were also asked if they checked that the homework that they had set had been satisfactorily carried out. This question was posed in regard to a wide range of aspects concerned with both the setting and monitoring of homework. Teacher responses were coded on a 4-point scale as never (0), occasionally (1), usually (2) and always (3). The questions asked of the teachers were divided into three main categories dealing with homework: work habits, guidance or instruction and satisfactory completion of homework.

In regard to work habits associated with homework, teachers were asked if they supervised the following points:

1 Making a written record of homework,

2 Making a record of set homework in a special notebook, and

3 Doing homework in a prescribed book.

For guidance and instruction associated with homework, teachers were asked if their supervision related to written homework, specified reading or learning homework or non-specified reading and revision work. The points requested of the teachers are given below.

For written homework:

(a) Written homework discussed in class,

(b) Written homework marked outside of class,

(c) Written homework examined in class.

For specified reading or learning homework:

(a) Learning of homework checked by short test,

(b) Learning of homework checked by questions in class.

For non-specified reading and revision homework:

(a) Homework checked from pupil record.

Teacher supervision of the completion of homework was checked by two items:

(a) Reprimands given for unsatisfactory homework,

(b) Homework must be completed satisfactorily.

The scores on each item in each of the three categories were summed to form three general measures. A homework work habits score, a homework guidance and instruction score and a homework completion score. These three items were intended to assess the importance that teachers attributed to homework.

Several items were concerned with the type of reporting on a student's progress at school that was sent to parents.

Achievement items in reports. The number of achievement items on the reports throughout the year was recorded.

Work habit items. The number of items associated with work habits throughout the year was reported.

Total items in reports. This measure was obtained by summing the two previous measures.

Use of educational materials. Finally, teachers were asked to report upon the use that they made of various educational materials. Their responses were coded as no use (0), intermittent use (1) and regular use (2). The list of educational materials that was considered is given below:

(a) A main textbook,
(b) A subsidiary textbook,
(c) A printed workbook,
(d) Duplicated work sheets,
(e) Programmed learning materials,
(f) Television programs,
(g) Films,
(h) Slides and similar visual aids,
(i) Field trips and visits,
(j) Commercial achievement tests,
(k) Pupil note book,
(l) A spelling book or list, or a table book,
(m) An Individualized Mathematics Programme or the use of small group practical work involving scientific equipment, and
(n) Structured aids or demonstration experiments involving scientific equipment.

In items (m) and (n), the alternatives for mathematics and science classes, respectively, were presented. The scores on all fourteen items were used to form a measure of the range and number of educational materials provided for the student to work with and learn from.

4 Classroom observation measures

The final method for assessing the classroom environment was by direct classroom observation. The activities occurring in the classroom were recorded in two ways. The time spent on different activities was recorded. Each minute, on the minute, an observer would 'sweep' the class and record the behaviour of the teacher and the class, providing time measures of various activities. In addition, each instance of a specific behaviour was recorded throughout a lesson, providing a frequency measure of the activities in the class. A discussion of the reliablity of the observation schedule is given in Appendix II. For further details of the development of the schedule and the reliability trials, Keeves (1971) should be consulted.

A class was observed for an equivalent of four forty-minute lessons. The teachers were asked to conduct their lessons as normally as possible under the unusual circumstances of an observer being present in the classroom. The observers were instructed to position themselves in an unobtrusive position while still being able to view the students' faces. A list and description of the items on the observation schedule is given below. It should be remembered that the given item label is the same as that used in subsequent appendices.

Time used by teacher. This item recorded the total time over the four lessons that the teacher was active in the teaching process. This represented the majority of the time available in most lessons for it included the time spent on all activities where the student could be learning. Examples of activities where the teacher was inactive in the teaching process included absence from the classroom, conferring with another person (not a student) and carrying out an administrative or organisational task without supervising the work prescribed for the class.

Students talk. The time that students either talked, read, discussed or explained work was recorded. The teacher either listened or supervised.

Question and answer session. The time spent by students interacting with the teacher in a question and answer session was recorded.

Students write. This item recorded the time that students wrote, drew, an[illegible]ed a written test, or wrote while the teacher dictated.

Students read. The time spent by students reading silently to themselves on an assigned task while the teacher supervised the class was recorded.

Students investigate. This item recorded the time spent by students undertaking an investigatory activity or a practical task while the teacher supervised.

Students mark work. The time spent by students marking their work under the direction or supervision of the teacher was recorded.

Unclassified. The time for which the activities in the classroom could not be classified into any of the previous seven categories was recorded. Most commonly, this would refer to times when students were engaged in activities not prescribed by the teacher or no activity had been prescribed by the teacher and the students were doing nothing.

Number of activity changes. The number of distinct changes in student and teacher activity in the interval of time between successive sweeps was recorded.

Teacher reviews work. The time spent by the teacher for the purpose of reviewing work previously taught was recorded.

Teacher contacts student. The number of instances where the teacher spoke to a student or a student spoke to the teacher in connection with academic work was recorded.

Teacher asks question. The number of instances where the teacher asked a student a question was recorded.

Student asks question. The number of instances where a student asked the teacher a question was recorded.

Invitation to participate. The number of instances where the teacher invited the student to ask a question or to participate in academic work was recorded. This included the student making suggestions, explaining work to the teacher or class, performing a demonstration or checking the work of another student.

Use of language. The number of discussions on the use of language was recorded. This included discussing the meaning, the spelling or the pronunciation of a word.

Invitation to inquire. The frequency with which students took part or were invited to take part in an activity that involved investigation and inquiry was recorded. This may have involved the student using equipment or reference books, estimating an answer, finding reasons for results, solving problems with non-unique solutions or discovering new methods for solving problems.

Consider work habits. The frequency with which students were encouraged to consider their work habits was recorded. The teacher usually emphasized the value of work and recommended more work.

Raise aspirations. This item recorded each occasion that the teacher encourged students to lift their level of aspiration. Typically, teachers mentioned the value of work and education.

Praise and Rebuke. The instances where teachers rewarded or praised a student as well as rebuked or punished a student were also recorded. Both praise and rebuke were classified into more specific categories. These categories were:

(a) Casual praise or rebuke,

(b) Deliberate praise or rebuke,

(c) Another type of reward or punishment, and

(d) Total praise or total rebuke.

The two measures described in (d) were obtained by summing the appropriate measures in (a), (b) and (c). The general distinction between casual and deliberate praise or rebuke depended upon whether the comment was made privately or publicly. Both casual and deliberate rewards and rebukes were given verbally while other praise or punishment involved more definite behaviour from the teacher.

Positive support. The number of instances where the teacher provided positive emotional support for a student was recorded.

Negative support. The number of instances where the teacher provided negative emotional support for the student was recorded.

Laughter with. This item recorded the number of instances where laughter occurred with the teacher or a student, not at the expense of the teacher or student.

Laughter at. This item recorded the number of instances where laughter was at the expense of a student or the teacher.

Autonomy. The number of instances where students were required or encouraged to make decisions for themselves or to act with autonomy was recorded. This involved the student selecting an activity, project, topic or experiment for themselves.

This is the complete list of variables that is used throughout the study. It should again be mentioned that this appendix only provides a summary of the variables employed by Keeves (1972). An extensive description of the classroom observation schedule has also been provided by Keeves (1974a).

As a concession to space, several tables presented in the following appendixes use abbreviated variable names. Although the intention of the abbreviated title was to suggest the nature of the original measure, this may not always by obvious. Hence, the variable labels used throughout this appendix were intended to be as close as possible to those titles used in the following appendixes. As such, this appendix provides a point of reference for the measures used throughout the study.

APPENDIX II

IDENTIFICATION OF MEASURES RELATED TO CLASS SIZE

APPENDIX II

IDENTIFICATION OF MEASURES RELATED TO CLASS SIZE

Appendix I provided a brief description of all the variables reported by Keeves (1972) that were used in this present investigation. To indicate the suitability of these measures for inclusion in the study, reliability coefficients for these measures, where appropriate and available, are given in this appendix. In particular, reliability coefficients relevant to variables obtained for tests, attitude scales or by use of the classroom observation schedule have been recorded. The other variables used in this study are measures of teacher, class or school characteristics obtained from teacher questionnaires and hence measures of reliability were inappropriate.

In addition to reliability coefficients, since we are primarily concerned with class size, the product moment correlation coefficients between class size and each variable using the class as the unit of analysis are also given. These coefficients have been used in Chapter 4 to determine which variables were significantly related to class size, as a first indicator of the differences between the teachers, the students and the processes present in small and large classes. Since nearly all Year 7 classes in the Australian Capital Territory were considered, the data provided population measures. However, the 10 per cent level of significance of the correlation coefficient, assuming the use of a sample random sample, was employed as the main criterion for the selection of a variable as being related to class size. As the selection of variables was being conducted at the between classes level, with class as the unit of analysis, the application of these procedures for the calculation of the sampling error to population measures was defensible. However, this was not the sole criterion for selection of variables. Analysis of variance was conducted between each variable and a class size measure formed by creating six class size categories. A variable was selected for inclusion in subsequent regression analyses if it was significantly correlated with class size, and the analysis of variance suggested that there was not evidence of a curvilinear or highly skewed relationship. Several other variables were also included in the discussion in Chapter 4 although their correlations with class size were not found to be significant at the 10 per cent level. They were included for their suspected importance, as suggested strongly by previous research. These few variables included some attitude and individualized instruction measures, and these have been commented on at a later stage in this appendix.

Home Background, Achievement and Attitude Measures

The set of classroom background, achievement and attitudinal variables are provided in Table A.3 for both the mathematics and science sets of data. It should be noted that the

Table A.3 Properties of Achievement and Attitude Measures

N = 72 Variable/reliability	Reliability[b]	Correlation with class size		Level of significance[a]		
		Science	Maths	Science	Maths	Inclusion
Sex of class	NA	0.17	0.17	NS	NS	No
Ethnicity of home	NA	-0.20	-0.22	<0.10	<0.10	Yes
Father's occupation	NA	0.35	0.35	<0.01	<0.01	Yes
Student's occupational aspirations	NA	0.57	0.57	<0.01	<0.01	Yes
Number of hours of homework per week	NA	0.39	0.40	<0.01	<0.01	Yes
Student's educational aspirations	NA	0.60	0.60	<0.01	<0.01	Yes
Student's participation in p/c activities	0.69	-0.31	-0.31	0.05	<0.05	Yes
Student's participation in m/s activities	0.69	0.18	0.17	NS	NS	Yes*
Like mathematics	0.83	-	0.17	-	NS	Yes*
Academic motivation	0.81	0.25	0.25	<0.05	0.05	Yes
Like school	0.89	0.29	0.30	<0.05	<0.05	Yes
Like science	0.90	0.17	-	NS	-	Yes*
Self-regard	0.79	0.15	0.14	NS	NS	No
Science achievement pretest	0.69	0.57	-	<0.01	-	Yes
Science achievement	0.84	0.61	-	<0.01	-	Yes
Mathematics achievement	0.89	-	0.59	-	<0.01	Yes

a Correlation coefficients which were not significant at the 10 per cent level are indicated by NS.
b NA = Not appropriate.

results are very similar for both subjects. This is, in part, because the classes in these subjects were the same for all except three of the 72 cases. It should be remembered that some measures were aggregate measures for the class formed by averaging the students' scores on each test or questionnaire.

In the tables, each variable is named, its reliability coefficient and its correlation coefficient with class size are given and then it is stated if it is to be included in subsequent analyses. If inclusion or exclusion differs from that suggested by the correlation coefficient, a brief explanation is provided by the following code:

- * This means that although the correlation was not significant enough for automatic inclusion it was large enough to warrant further consideration since previous research had indicated that variables of this type were often associated with differences between large and small classes.
- ** This means that the correlation coefficient was significant but results obtained from an analysis of variance had shown that the relationship was curvilinear or skewed and hence not suitable for subsequent analysis.

Reliability coefficients were found by using KR-20. Only correlation coefficients which are significant at the 0.10 level are reported.

Of the 16 variables concerned with general class background, achievement and attitudes, 11 variables related significantly to class size. Three additional variables were considered sufficiently important to warrant further analysis.

Structural Dimension Measures

The set of variables from the structural dimension are presented in Tables A.4 and A.5 for science and mathematics classes respectively. These tables present the measures of both teacher, classroom and school characteristics and as such they were considered to influence the behaviours that occurred in the classroom although not as directly as process variables. These variables were collected by teacher interviews and questionnaires as well as from discussions with school principals and subject co-ordinators. For this reason reliability coefficients are not applicable. Since the characteristics were expected to differ greatly between science and mathematics classes, the results for each subject are presented separately. As noted in Chapter 4 and elsewhere (Gage, 1963; Rosenshine, 1971), teacher characteristics rarely relate to educational outcome measures. The same conventions as used in Table A.3 apply to these two tables and a fuller account of all variables was given in Appendix I.

The data available provided 74 variables from the structural dimension: 38 for science and 36 for mathematics. For science, seven variables correlated significantly with class size but three of these were considered unsuitable. For mathematics, nine variables were found to have significant correlations with class size, but again three

Table A.4 Properties of Structural Variables for Science Classes

N = 72 Variable	Correlation coefficient[a]	Level of significance[b]	Inclusion
Sex of teacher	0.02	NS	No
Teacher years at school	0.07	NS	No
Teacher marital status	-0.11	NS	No
Teacher science specialist	0.06	NS	No
Teacher years of education	0.03	NS	No
Teacher years tertiary education	0.05	NS	No
Teacher training - Physics	-0.04	NS	No
Teacher training - Chemistry	-0.07	NS	No
Teacher training - Biology	0.01	NS	No
Teacher training - Geology	0.16	NS	No
Teacher training - Other Science	0.00	NS	No
Total teacher training in Science	-0.01	NS	No
Teacher training institute	0.02	NS	No
Teaching load	0.13	NS	No
Teacher lesson preparation	-0.05	NS	No
Teacher laboratory preparation	0.04	NS	No
Teacher hours marking	0.23	<0.10	No**
Sex of class	-0.19	NS	No
Prop'n class from foreign language homes	-0.23	<0.10	Yes
Size of form cohort	-0.06	NS	No
Propn of male teachers on staff	-0.10	NS	No
Years school open	-0.15	NS	No
Degree of streaming	-0.19	<0.10	No**
Time on science	0.10	NS	No
Time on science homework	0.44	<0.01	Yes
Total time on science	0.37	<0.01	Yes
Time on all homework	0.37	<0.01	Yes
Periods contact with male teachers	-0.02	NS	No
Number of teachers in year	0.09	NS	No
Number of regular teachers	-0.04	NS	No
Size of school	-0.13	NS	No
Age of teacher	0.12	NS	No
Teacher years experience	0.01	NS	No
Teacher inservice training	-0.23	<0.10	No**
Attends lectures science teaching	0.05	NS	No
Attends lectures on science	0.00	NS	No
Member of Science Teachers Association	+0.02	NS	No
Periods in laboratory	-0.08	NS	No

a Correlations with class size recorded.
b Correlation coefficients which were not significant at the 10 per cent level are indicated by NS.

Table A.5 Properties of Structural Variables for Mathematics Classes

N = 72 Variable	Correlation coefficient[a]	Level of significance[b]	Inclusion
Sex of teacher	0.02	NS	No
Teacher years at school	-0.01	NS	No
Teacher marital status	-0.13	NS	No
Teacher maths specialist	0.03	NS	No
Teacher years of education	0.19	NS	No
Teacher years of tertiary education	0.18	NS	No
Teacher training - General Maths	0.11	NS	No
Teacher training - Pure Maths	0.02	NS	No
Teacher training - Applied Maths	-0.06	NS	No
Teacher training - Statistics	0.18	NS	No
Teacher training - other Maths	0.17	NS	No
Total teacher training in Maths	0.10	NS	No
Teacher training institution	0.05	NS	No
Teaching load	0.07	NS	No
Teacher hours preparation	0.13	NS	No
Teacher hours marking	0.21	<0.10	No**
Sex of class	-0.19	NS	No
Prop'n class from foreign language homes	-0.23	<0.10	Yes
Size of form cohort	-0.06	NS	No
Propn of male teachers on staff	-0.11	NS	No
Years school open	-0.15	NS	No
Degree of streaming	-0.24	<0.05	No**
Time on mathematics	0.04	NS	No
Time on mathematics homework	0.43	<0.01	Yes
Total time on mathematics	0.36	<0.01	Yes
Time on all homework	0.36	<0.01	Yes
Periods contacts with male teachers	-0.02	NS	No
Number of teachers in year	-0.31	<0.05	Yes
Number of regular teachers	-0.34	<0.01	Yes
Size of school	-0.12	NS	No
Age of teacher	0.10	NS	No
Teacher years experience	0.09	NS	No
Teacher inservice training	0.16	NS	No
Attends lectures on maths teaching	0.15	NS	No
Attends lectures on mathematics	0.18	NS	No
Member of Mathematical Association	+0.24	<0.05	No**

a Correlations with class size recorded.

b Correlation coefficients which were not significant at the 10 per cent level are indicated by NS.

Table A.6 Properties of Process Variables for Science Classes

N = 72 Variable	Correlation coefficient[a]	Level of significance[b]	Inclusion
Use of short answer tests	0.29	<0.05	Yes
Use of extended answer tests	0.13	NS	No
Use of multiple choice tests	0.25	<0.05	Yes
Assess homework or assignments	-0.18	NS	No
Assess workbooks	-0.16	NS	No
Assess projects or major assignments	0.17	NS	No
Range of assessment	0.16	NS	No
Assessments in Term 2	0.19	<0.10	No**
Assessments involving student choice	0.09	NS	Yes*
Frequency of revision homework	-0.06	NS	No
Frequency homework set	0.20	<0.10	No**
Written record of homework	0.02	NS	No
Record homework in notebook	-0.11	NS	No
Homework in special book	0.08	NS	No
Homework discussed	-0.03	NS	No
Homework marked	0.12	NS	No
Homework examined	-0.03	NS	No
Homework tested by short test	0.23	<0.10	No**
Homework questioned in class	0.04	NS	No
Homework checked from record	-0.03	NS	No
Homework reprimand given	0.24	<0.05	Yes
Homework must be completed	0.14	NS	No
Homework work habits score	0.00	NS	No
Homework completed score	0.22	<0.10	No**
Homework guidance score	0.12	NS	No
Frequency reports sent home	0.26	<0.05	Yes
Achievement items in report	-0.10	NS	No
Work habit items in report	-0.12	NS	No
Total items in report	-0.11	NS	No
Use of textbook A	0.03	NS	No
Use of textbook B	0.31	<0.05	Yes
Use of printed workbook	0.01	NS	No
Use of duplicated worksheets	-0.03	NS	No
Use of programmed instruction		No instance recorded	
Use of TV	0.00	NS	No
Use of films	0.01	NS	No
Use of slides and visual aids	-0.06	NS	No
Use of field trips or visits	-0.15	NS	No
Use of commercial tests		No instance recorded	
Use of pupil notebook	0.32	<0.05	Yes
Use of spelling book	-0.09	NS	No
Use of group practical work	-0.07	NS	No
Demonstration experiments	-0.08	NS	No
Total teaching aids	-0.07	NS	No

a Correlations with class size recorded.

b Correlation coefficients which were not significant at the 10 per cent level are indicated by NS.

were unsuitable. The small proportion of variables from the structural dimension which were related to class size was not surprising considering the disappointing history of previous research findings in this area.

Process Dimension Measures

Another interview and questionnaire session was used to obtain information on variables associated with the process dimension. The practices assessed in this way were distinct from those obtained by direct observation of the classroom. These measures were used to assess other facets of the classroom environment which would have gone unnoticed during the four periods of direct observation. These measures included the amount and methods of assessment used by teachers, the emphasis upon homework and how it was recorded and checked, the diversity of educational materials employed in the classroom as well as the amount and type of reporting to parents. All these variables were intended to assess the type of educational environment that teachers sought to generate in the classroom, in particular, concentrating upon the stimulation and monitoring of a student's academic progress. Reliability coefficients were not appropriate. The correlations between class size and these processes are reported for both science and mathematics classes in Tables A.6 and A.7, in the same manner as in previous tables.

For the process variables not measured by classroom observation, 44 measures were obtained in each subject. For science, 10 were found to correlate significantly but four were deleted from this list for the reasons mentioned earlier. The measure of the extent of student choice in assessment was included to detect student autonomy and the presence of individualization in the classroom. The inclusion of this variable was supported by a moderately strong correlation and a significant difference between groups using analysis of variance. For mathematics, 16 variables were found to correlate significantly, but two were not considered for further analysis. Student choice in assessment was not included for mathematics as the correlation was not strong and analysis of variance was not supportive. The greater incidence of significant results in mathematics would seem to suggest that assessment and homework procedures, as well as the use of educational materials varied more with class size in mathematics than in science classes.

Classroom Observation Measures

The final method for measuring variables was the classroom observation schedule. The schedule assessed classroom activities in two ways. An observer noted the behaviour of the teacher and the students every minute, on the minute, during a class period. The time spent on different activities was thus obtained. Also, the observer noted each instance of certain behaviours as they occurred throughout the lesson. Hence, the frequency of specific activities was recorded. A major problem was whether an observer

Table A.7 Properties of Process Variables for Mathematics Classes

N = 72 Variable	Correlation coefficient[a]	Level of significance[b]	Inclusion
Use of short answer tests	0.13	NS	No
Use of extended answer tests	0.26	<0.05	Yes
Use of multiple choice tests	0.20	<0.10	No**
Assess homework or assignments	-0.07	NS	No
Assess workbooks	0.10	NS	No
Assess projects or major assignments	0.00	NS	No
Range of assessment	0.24	<0.10	Yes
Assessments in Term 2	-0.08	NS	No
Assessments involving student choice	0.06	NS	No
Frequency of revision homework	-0.21	<0.10	Yes
Frequency of homework set	0.41	<0.01	Yes
Written record of homework	0.25	<0.05	Yes
Record homework in notebook	0.31	<0.05	Yes
Homework in special book	0.16	NS	No
Homework discussed	0.42	<0.01	Yes
Homework marked	0.01	NS	No
Homework examined	0.09	NS	No
Homework tested with short test	-0.19	NS	No
Homework questioned in class	-0.06	NS	No
Homework checked from record	-0.20	<0.10	No**
Homework reprimand given	0.10	NS	No
Homework must be completed	0.33	<0.05	Yes
Homework work habits score	0.35	<0.01	Yes
Homework completed score	0.31	<0.05	Yes
Homework guidance score	0.13	NS	No
Frequency reports sent home	0.26	<0.05	Yes
Achievement items in reports	-0.10	NS	No
Work habit items in reports	-0.11	NS	No
Total items in reports	-0.10	NS	No
Use of textbook A	0.56	<0.01	Yes
Use of textbook B	-0.07	NS	No
Use of printed workbook	-0.32	<0.05	Yes
Use of duplicated work sheets	-0.14	NS	No
Use of programmed instruction	No instance recorded		
Use of TV	0.42	<0.01	Yes
Use of films	No instance recorded		
Use of slides and visual aids	0.03	NS	No
Use of visits and trips	No instance recorded		
Use of commercial tests	-0.08	NS	No
Use of pupil notebook	0.08	NS	No
Use of table and spelling books	-0.19	NS	No
Use of individual program kits	-0.03	NS	No
Use of structured aids	-0.05	NS	No
Total teaching aids	-0.03	NS	No

a Correlations with class size recorded.
b Correlation coefficients which were not significant at the 10 per cent level are indicated by NS.

would record the same activities in each class in the same way each time the schedule was used, and would different observers agree upon their interpretations of various classroom activities. These problems generated some uncertainty about the reliability of each item included on the observation schedule. Reliability coefficients for the observation of specific behaviours on the schedule were obtained using intra-class correlation coefficients. Wherever possible, these reliability coefficients are given to indicate the confidence available in any particular variable from the schedule. The relationships between class size and the classroom processes recorded in the schedule are given for science and mathematics in Tables A.8 and A.9 respectively. A fuller description of each item on the observation schedule is given in Appendix I.

For the process variables obtained from the classroom observation schedule, 32 different measures were available for each subject. For science, eight variables were found to correlate significantly with class size, but one was deleted from this list. A measure of the number of positive support statements made by teachers was added as smaller classes were believed to have higher levels of interpersonal relationships (Vincent, 1967; 1968) and it was desirable to test this finding. The inclusion was justified by a moderately strong negative correlation and a significant F-ratio. For mathematics, 17 variables correlated significantly with class size, but one was deleted since it did not satisfy all criteria required. Two variables were added to the list. The number of activity changes, as a measure of instructional diversity, and the number of teacher invitations for students to participate in academic work were both added as they had moderate correlations and significant F-ratios. Their inclusion was further justified by previous research which claimed that both of these measures were more prevalent in small classes. As noted earlier, class size appeared to have a greater impact upon the behaviours and activities that occurred in mathematics classes than in science classes. Also, reliability levels were satisfactory for the majority of the classroom observation measures.

Of the 242 variables on which data were obtained, 76 variables or 31 per cent of the variables were found to be significantly related to class size. After adjusting the list, 68 variables remained for further consideration using regression analysis. As only the structural and process variables referred specifically to each subject, it is interesting to note that 19 science variables and 38 mathematics variables from this group were to be considered for use in regression analyses. The greater incidence of mathematics variables has been noted several times, but it would seem that science teachers developed much of their teaching style independently of class size while mathematics teachers, intentionally or not, changed their style as class size changed. However, it must be noted that there was a strong positive relationship between class size and achievement and it would be possible that teaching style was dependent more on the ability of the students in the class than on the size of the class.

Table A.8 Properties of Process Variables from the Observation Schedule for Science Classes

N = 72 Variable	Reliability[c]	Correlation coefficient[a]	Level of significance[b]	Inclusion
Time used by teacher	0.99	0.10	NS	No
Students listen	0.96	0.23	<0.10	No**
Students talk	NA	-0.07	NS	No
Question and answer session	1.00	0.11	NS	No
Students write	0.98	-0.16	NS	No
Students read	NA	-0.02	NS	No
Students investigate	0.99	-0.07	NS	No
Students mark work	NA	-0.09	NS	No
Unclassified	0.97	-0.07	NS	No
Number of changes in activity	0.37	0.13	NS	No
Teacher reviews work	0.23	0.26	<0.05	Yes
Teacher contacts student	0.90	-0.41	<0.01	Yes
Teacher asks question	0.99	0.04	NS	No
Student asks question	0.95	0.06	NS	No
Invitation to participate	0.99	0.08	NS	No
Use of language	0.91	0.18	NS	No
Invitation to inquire	0.91	-0.03	NS	No
Consider work habits	0.89	0.00	NS	No
Raise aspirations	0.25	-0.05	NS	No
Casual praise	0.99	-0.05	NS	No
Deliberate praise	0.99	0.06	NS	No
Other reward	0.99	-0.21	<0.10	Yes
Total praise	0.99	-0.05	NS	No
Casual rebuke	0.98	-0 32	<0.05	Yes
Deliberate rebuke	0.98	-0.31	<0.05	Yes
Other punishment	0.98	-0.30	<0.05	Yes
Total rebuke	0.98	-0.39	<0.01	Yes
Positive support	0.79	-0.17	NS	Yes*
Negative support	0.64	0.00	NS	No
Laughter with	0.92	0.19	NS	No
Laughter at	0.53	0.06	NS	No
Autonomy	NA	0.11	NS	No

a Correlations with class size recorded.

b Correlation coefficients which were not significant at the 10 per cent level are indicated by NS.

c Reliability coefficients which could not be calculated are indicated by NA.

116

Table A.9 Properties of Process Variables from the Observation Schedule for Mathematics Classes

N = 72 Variable	Reliability[c]	Correlation coefficient[a]	Level of significance[b]	Inclusion
Time used by teacher	0.99	0.15	NS	No
Students listen	0.88	0.11	NS	No
Students talk	NA	0.14	NS	No
Question and answer session	0.98	0.40	<0.01	Yes
Students write	0.97	-0.35	<0.01	Yes
Students read	NA	-0.06	NS	No
Students investigate	NA	-0.21	<0.10	No**
Students mark work	0.98	0.10	NS	No
Unclassified	0.89	-0.36	<0.01	Yes
Number of changes in activity	0.76	0.19	NS	Yes*
Teacher reviews work	0.00	0.04	NS	No
Teacher contacts student	0.94	-0.45	<0.01	Yes
Teacher asks question	0 92	0.35	<0.01	Yes
Student asks question	0.64	0.26	<0.05	Yes
Invitation to participate	0 95	0.19	NS	Yes*
Use of language	0.95	0.14	NS	No
Invitation to inquire	NA	0.21	<0.10	Yes
Consider work habits	0.95	0.20	<0.10	Yes
Raise aspirations	0.64	0.12	NS	No
Casual praise	0.99	0.30	<0.05	Yes
Deliberate praise	0.99	0.18	NS	No
Other reward	0.99	0.16	NS	No
Total praise	0.99	0.30	<0.05	Yes
Casual rebuke	0.94	-0.21	<0.10	Yes
Deliberate rebuke	0.94	-0.28	<0.05	Yes
Other punishment	0.94	-0.16	NS	No
Total rebuke	0.94	-0.28	<0.05	Yes
Positive support	0.83	0.29	<0.05	Yes
Negative support	0.87	-0.01	NS	No
Laughter with	0.77	0.32	<0.05	Yes
Laughter at	0.89	0.32	<0.05	Yes
Autonomy	NA	0.13	NS	No

a Correlations with class size recorded.

b Correlation coefficients which were not significant the 10 per cent level are indicated by NS.

c Reliability coefficients which could not be calculated are indicated by NA.

APPENDIX III

FACTOR ANALYSIS OF ATTITUDE MEASURES

The importance of five attitude and practice measures was noted in the first set of regression analyses reported in Chapter 5. The five attitudes or practices were a student's occupational and educational aspirations, academic motivation, attitude to school and level of participation in mathematics and science activities. These five attitudes were chosen for special consideration since they were related to both class size and academic achievement. As such, they were considered to be suitable for improving the predictive power of the causal model, particularly in the absence of prior achievement measures. It was envisaged that factor analysis could be used to construct a general attitude measure which would act as a single measure of attitude, instead of five individual scales.

The factor loadings for each attitude measure were calculated by using the SPSS Factor program (Nie et al., 1970) with the class averages on the attitude measures. With only five variables, at most five factors were necessary to explain all the variance, but only the first factor was significant using the criterion that the associated eigenvalue must exceed one. Nevertheless, the second factor was examined because it provided evidence of important differences between the five attitude measures. The factor loadings recorded in Table A.10 represent the initial unrotated factors using principal components analysis.

The alignment of all attitude measures along the first principal component is consistent with the first factor being a general attitude factor. The second factor, although not significant, has factor loadings which distinguish between a student's occupational and educational aspirations and the other three measures because the like school, mathematics and science activities and academic motivation measures are all aligned positively on the second factor while the aspirational measures are assigned negative factor loadings. Although it is often hard to identify the nature of the artifacts

Table A.10 Attitude Measure Principal Factor Pattern (Between Classes)

N = 72 Attitude measure	Factor loading 1	Factor loading 2
Occupational aspirations	0.87	-0.40
Educational aspirations	0.80	-0.48
Like school	0.84	0.41
Maths/science activities	0.51	0.15
Academic motivation	0.73	0.42
Eigenvalue	3.08	0.94

produced by factor analysis, it does appear that the second factor distinguished between aspirations for the future and current attitudes and experiences. This distinction was also apparent when the attitude measures were included in the regression analyses. The two aspirational measures were more influential in the regression model than the other three attitude measures. These findings are reported in Chapter 5.

REFERENCES

Balow, I.H.
A longitudinal evaluation of reading achievement in small classes. Elementary English. 1969, 46, 184-187.

Barr, R. and Dreeben, R.
Instruction in classrooms. In L.S. Shulman (Ed.), Review of Research in Education (vol. 5). Itasca, Illinois: Peacock, 1977.

Blalock, H.M.
Casual Inferences in Nonexperimental Research. Chapel Hill: University of North Carolina Press, 1964.

Broom, L., Jones, F.L. and Zubrzycki, J.
An occupational classification of the Australian workforce. Supplement to Australian and New Zealand Journal of Sociology, 1965, 1(2), 1-16.

Broom, L., Jones, F.L. and Zubrzycki, J.
Social Stratification in Australia. In J.A. Jackson (Ed.), Social Stratification: Sociological Studies I. Cambridge: Cambridge University Press, 1968, 212-33.

Brophy, J.E. and Good, T.
Teacher-Student Relationship: Causes and Consequences. New York: Holt, Rinehart and Winston, 1974.

Burstein, L.
The analysis of multilevel data in educational research and evaluation. In D.C. Berliner (Ed.), Review of Research in Education (vol. 8), 1980, 158-233.

Burstein, L.
Units and levels of analysis. In T. Husen and T.N. Postlethwaite (Eds), International Encyclopedia of Education: Research and Studies. Oxford: Pergamon (in press).

Burstein, L., Linn, R.L. and Capell, F.
Analysing multilevel data in the presence of heterogeneous within-class regressions. Journal of Educational Statistics, 1978, 3, 347-83.

Coleman, J.S. et al.
Equality of Educational Opportunity. Washington: US Government Printing Office (2 Vols), 1966.

Comber, L.C. and Keeves, J.P.
Science Education in Nineteen Countries. International Studies in Evaluation, I. Stockholm: Almqvist and Wiksell, 1973.

Cooley, W.W., Bond, L. and Mao, B-J.
Analyzing multi-level data. In R. Berk (Ed.) Educational Evaluation Methodology: The State of the Art. Baltimore, Md: The Johns Hopkins University Press, 1981.

Cronbach, L.J. and Webb, N.
Between class and within-class effects in a reported aptitude X treatment interaction: reanalysis of a study by G.L. Anderson. Journal of Educational Psychology, 1975, 67, 717-24.

Davie, R., Butler, N. and Goldstein, H.
From Birth to Seven: A Report of the National Child Development Study. London: Longman, 1972.

Educational Research Service.
Class Size Research: A Critique of Recent Meta-analysis. Arlington: Educational Research Service, 1980.

Flinker, I.
Optimum class size: what is the magic number? Clearing House, 1972, 471-73.

Furno, O.F. and Collins, G.J.
Class Size and Pupil Learning. ED025003, October 1967.

Gage, N.L.
Paradigms for research on teaching. In N.L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rank McNally, 1963.

Glass, G.V., Cahen, L.S., Smith, M.L. and Filby, N.N.
School Class Size: Research and Policy. Beverly Hills: Sage Publications, 1982.

Glass, G.V. and Smith, M.L.
Meta-Analysis of Research on the Relationship of Class Size and Achievement. San Francisco: Far West Laboratory for Educational Research and Development, 1978.

Haskell, S.
Some observations on the effects of class size upon pupil achievement in geometric drawing. Journal of Educational Research, 1964, 58(1), 27-30.

Hauser, R.M.
Socioeconomic Background and Educational Performance (Rose Monograph Series). Washington: American Sociological Association, 1973.

Husén, T. (Ed.)
International Study of Achievement in Mathematics (vol. 2). Stockholm: Almqvist and Wiksell, 1967.

Karmel, P.
Education in the Eighties: Some More Economic Aspects of Education. Australian Educational Review, No. 15. Hawthorn, Victoria: Australian Council for Educational Research, 1981.

Keeves, J.P.
The Home, the School and Educational Achievement, unpublished PhD Thesis, Australian National University, 1971.

Keeves, J.P.
Educational Environment and Student Achievement. Hawthorn, Victoria: Australian Council for Educational Research, 1972.

Keeves, J.P.
Can Teachers Make a Difference? Hawthorn, Victoria: Australian Council for Educational Research, September 1974(a).

Keeves, J.P.
Some Attitude Scales for Educational Research Purposes. Hawthorn, Victoria: Australian Council for Educational Research, September 1974(b).

Keeves, J.P. and Lewis, R.
Issues in the analysis of data from natural classroom settings. Australian Journal of Education. 1983, 27(3) 274-287.

Lafleur, C.D., Sumner, R.J. and Witton, E.
Class Size Survey. Canberra: Australian Government Publishing Service, 1974.

Lindbloom, D.H.
Class Size As It Affects Instructional Procedures and Educational Outcomes. ED 059532, Minneapolis, Minn., June 1970, 48pp.

Linn, R.L. and Werts, C.E.
Assumptions in making causal inferences from part correlations, partial correlations, and partial regression coefficients. Research Bulletin RB-69-6. Princeton, NJ: Educational Testing Service, 1969.

Little, A., Mabey, C. and Russel, J.
Do small classes help a pupil? New Society. 1971, 473, 769-71.

McCluskey, L.
Class Size or Mode of Instruction: What Makes the Difference in Classroom Processes. ED158409, June 1978.

Mason, W.M., Wong, G.Y. and Entwisle, B.
Contextual analysis through the multilevel linear model. In S. Leinhardt (Ed.) Sociological Methodology 1983-84. San Francisco: Jossey Bass, 1983.

Moody, W.B., Bausell, R.B. and Jenkins, J.R.
The Effect of Class Size Upon Learning in Mathematics: A Parametric Study. ED062138, April 1972.

National Education Association.
Class size. Research Bulletin. 1968, 35-36.

Nie, N.H., Bent, D.H. and Hull, C.H.
Statistical Package for the Social Sciences. New York: McGraw-Hill, 1970.

Olsen, M.N.
Classroom variables that predict school system quality. Research Bulletin, IAR, Columbia University, 1970, 2(1).

Olsen, M.N.
Ways to achieve quality in school classrooms: some definitive answers. Phi Delta Kappan, 1971, 52(1), 63-65.

Peaker, G.F.
The regression analyses of the national survey. In United Kingdom, Department of Education and Science, Children and their Primary Schools, (Lady Plowden, Chairman). London: Her Majesty's Stationery Office, vol. 2, appendix 4, 1967.

Peaker, G.F.
An Empirical Study of Education in Twenty-One Countries. International Series in Evaluation, vol. VIII. Stockholm: Almqvist and Wiskell, 1975.

Pedhazur, E.J.
Multiple Regression in Behavioral Research: Explanation and Prediction. (Second Edition). New York Holt, Rinehart and Winston, 1982.

Porwell, J.P.
Class Size: A Summary of Research. Arlington: Educational Research Service, 1978.

Queensland Teachers Union.
Report on Study to Determine the Merits of a High School Allowance. Brisbane: June 1973.

Rogosa. D.
Politics, process and pyramids. Journal of Educational Statistics, 1978, 3(1), 79-86.

Rosenshine, B.
Teaching Behaviours and Student Achievement. Slough: National Foundation for Educational Research in England and Wales, 1971.

Rosier, M.J.
Home background is major factor in school results. IEA (Australia) Report 1973:2. Hawthorn, Victoria: ACER 1973.

Ryan, D.W. amd Greenfield, T.B.
The Class Size Question. Toronto, Ontario: Ministry of Education, 1975.

Shapson, S.M., Wright, E.N., Eason, G. and Fitzgerald, J.
Results of an Experimental Study of the Effects of Class Size. ED151985, March 1978.

Smith, M.L. and Glass G.V.
Relationship of Class Size to Classroom Processes, Teacher Satisfaction and Pupil Affect: A Meta-Analysis. San Francisco: Far West Laboratory for Educational Research and Development, 1979.

Thorndike, R.L.
Reading Comprehension in Fifteen Countries. International Studies in Education, III. Stockholm: Almqvist and Wiksell, 1973.

Vincent, W.S.
Indicators of quality. Research Bulletin. IAR, Columbia University, 1967, 7(1).

Vincent, W.S.
Further clarification of the class size question. Research Bulletin. IAR, Columbia University, 1968, 9(1).

Walberg, H.J.
Class size and the social environment of learning. Human Relations. 1969, 22(5), 465-475.

Welch, W.W.
Evaluating the Small Group on a Component of Modular Schedules. ED064216, February 1971.

Woodson, M.S.
Effect of class size as measured by an achievement test criterion. Research Bulletin. IAR, 1968, 8(2), 1-6.

The purpose of this investigation was to examine the ways in which class size affected other facets of the educational environment of the classroom.

The central question of the study which required explanation was the commonly found positive relationship between class size and achievement. The most plausible explanation of the evidence would seem to involve the effects of grouping more able students in larger classes, but the findings also indicated achievement gains beyond those expected solely from a consideration of differences in achievement levels. It is clear that an increased understanding of these features of the classroom is a necessary step towards teasing out the relationships between class size, teacher activities and student motivation, all of which appear as central themes in the class size question.

AUSTRALIAN COUNCIL FOR EDUCATIONAL RESEARCH

ISBN 0 85563 390 5