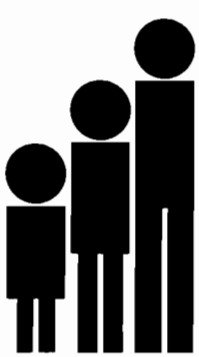
Examination Results in Selective and Nonselective Schools

Findings from the National Child Development Study

Jane Steedman

ISBN 90281 7221

Printed by Print Power (London) Ltd
01-739 9995/6



1983
National Children's Bureau

Acknowledgements

This study would not have been possible without all those in Great Britain born in the week of 3rd to 9th March, 1958, whose participation in the National Child Development Study is always appreciated.

I am grateful to everyone at the National Children's Bureau for contributing to this project's progress. The investigation reported here was funded by the Department of Education and Science, whose financial support has been gratefully received. For helping to supply the public examination results of members of the National Child Development Study, I should like to thank the teachers and headteachers of the schools involved, as well as Directors of Education and Chief Education Officers and many other people in Local Education Authorities throughout Great Britain. Within the Bureau, the collection and coding of results was ensured by Elizabeth Choice and her teams of tracers and coders. The data processing staff at the Bureau, Bob Wellburn (Senior Programmer), Puck de Raadt and Sheila Williams, toiled until Easter 1980 to put the data on to computer in useable form. Thereafter, Dougal Hutchison (Principal Statistician), who has provided the Statistical Appendix, ran the analyses (with help from Vanessa Simonite). Richard Ives was a cheerful and thoughtful assistant to the project from June 1979 to May 1980, and reported on relations between examination results and test scores.

Many people have typed various drafts; each one deserves particular thanks. Ken Fogelman (Assistant Director, Research) had oversight of the project. Any inadequacies in the work are, of course, my responsibility.

I should like in addition to thank Professor Halsey and the members listed below of the Advisory Group to the Secretary of State concerning this project; they have shown an active and encouraging interest in the work throughout.

The Advisory Group to the Secretary of State for Education and Science.

Chairman: Professor A.H. Halsey

Members: Miss M.E. Telford: formerly Head, Westwood High School, Leek.
Mr. A.R. Barnes: Head, Ruffwood School, Kirkby
Professor Harvey Goldstein: Institute of Education, London

Assessors: from the Department of Education and Science

CONTENTS

Tables

Diagrams

List of Appendices contained in Volume 2

All Appendices are in Volume 2, which is available from:

The Supplementary Publications Division, Reference SUP 81014, British Library (Lending Division), Walton, Boston Spa, Wetherby, Yorks. LS 23 7 BQ.

(The report by Richard Ives referred to in the text is also available from the same address, Reference SUP 81015)

A copy of Volume 2, Appendices, is also available for reference in the library of the National Children's Bureau.

Copies are available also in the following deposit libraries:

The British Museum, London; The Bodleian Library, Oxford; The University Library, Cambridge; The National Library of Wales, Aberystwyth; The National Library of Scotland, Edinburgh; and The Library of Trinity College, Dublin.

Summary

The progress of pupils in selective and nonselective secondary schools has been explored using data from the National Child Development Study (NCDS, 1958 cohort), in a DES-funded study since 1977. Results of the first stage were published in 1980*, and concerned test scores of sixteen-year-old pupils and other measures of progress, attitudes, behaviour, and plans for the future. This report presents results of the latest stage, chiefly concerning public examination results of the same pupils.

The sample was, throughout this project, NCDS pupils in English maintained schools, who stayed in one school throughout secondary years 1969-74. Pupils of grammar, secondary modern or comprehensive schools were compared. The last group was distinguished from pupils whose school went comprehensive while they were there (transitional pupils, 60 per cent of those in comprehensives at sixteen). Analyses also compared the mixture of grammar and secondary modern pupils (grammar+modern combination) with comprehensives as a whole.

NCDS has collected data on the same people throughout their lives, so it was possible to compare the school groups for characteristics of pupils at eleven, before the start of secondary school, as well as at sixteen years old. It was known from the earlier stage of this project* that pupils going into comprehensives at this time were more like secondary modern pupils than like grammar pupils who remained a large minority of the sample. At eleven, comprehensive pupils were lower scoring on tests of attainment or ability and more likely to be 'working class' than the average for the combination of grammars and secondary moderns. This reflected the high proportion of grammar pupils (26 per cent) in the sample. Indeed, the study could be regarded as comparing grammar pupils and half the secondary modern pupils with the other half of secondary modern pupils.

The attainment and examination results at sixteen of comprehensive pupils would therefore be expected to be lower, on average, at sixteen, but this would not necessarily have very much to do with secondary schooling. It was a premise of this study that performance could not be (even tentatively) attributed to secondary schooling without correcting results in some measure for preexisting and background differences. By basing the study on NCDS data and using standard statistical techniques (analysis of variance and of proportions) it was possible to arrive at figures for attainments and examination results which may reflect progress in the different types of school. These incorporate allowance for what was attributable to tested attainment and/or "general ability" at eleven, social class, parents' interest and sex ("intake-corrections").

Detailed results are presented and summarised in Chapter IV.

* Findings of this earlier stage are in 'Progress in Secondary Schools', Steedman, 1980 (National Children's Bureau).

The examination results broadly support the observations on attainment tests published earlier. Thirteen indicators of examination performance were used, including average number of '0' level 'pass' equivalents, proportions with five or more '0' level equivalents, and various measures of performance in two basic subjects, mathematics and English. The corrected figures suggest that progress measured by examination results was greatest in grammar schools and least in secondary moderns. Comprehensive pupils tended to be in between these two. The 'intake-corrected' averages for the grammar+modern combination and for comprehensive pupils were not statistically significantly different, on most measures of examination results. (The exceptions were grades of '0' level English language relative to eleven year reading, obtained by candidates for '0' level English). Transitional pupils tended to have somewhat lower examination results by some measures.

Additional detailed exploration indicated no clear evidence of departure from these overall patterns for those of different attainments, social class or parental interest. There were tentative suggestions of lack of support for two theories about comprehensives and grammar schools, in that it did not seem necessarily the case that high ability children in comprehensives did worse than equivalents in grammars, nor apparently demonstrable in these findings that 'bright working class' children did better in grammars. Whether or not these form counter arguments to theories of selection, it was clear that secondary modern pupils fared worse. This was true even of those of high ability before secondary school - not all high ability pupils went to either grammars or comprehensives.

Limitations on the evidence call for caution in discussing these results; the very highest-scoring 2 per cent or so at eleven still all went to grammars and could not be compared with other pupils, and examination results are not easy to interpret or to use as measures of aggregates of pupils, let alone as intake-corrected variables. This is part of the problem in constructing an artificial comparison of school groups as if types of school could be compared, when the types of school coexisted. There is no test here of what fully-developed, complete selective or comprehensive systems would be like, but rather a set of empirical observations on one system undergoing reorganisation. Some of the lack of evidence of differences is to do with limitations of examination measures, some to the artificiality of the attempt to compare 'selective' and 'nonselective' schools.

Simplification of presentation is aimed at by linking results to the question, 'Did selection make a difference?'. In so far as results have educational implications, they suggest a lack of difference attributable to being in the 'selective' combination as against comprehensives. There was no more examination taking or examination success in comprehensives. Nor, however, was there any less examination success in comprehensives than in the grammar+modern combination. In another sense, selection clearly made a difference, as results in grammar schools were 'better', relative to eleven year scores, than those of equivalent pupils who were not selected. The findings are open to alternative interpretations, and some suggestions are made for alternative approaches to analysis. Nevertheless, that oversimple question can be answered by saying that it did not appear, from the evidence there was, that comprehensive pupils were doing better or worse than pupils in the combination of grammars and secondary moderns which was, in theory, the alternative.

Chapter I

Introduction

The question addressed by this research is, at its most succinct, 'Did selection for secondary school make a difference?' The simplicity of the question belies the complexity of any attempt to answer it. It is of course worth asking, now that schools in the selective, tripartite system (grammar, technical and secondary modern schools) are, with increasing momentum since the 1960's, being replaced with comprehensive schools. Is there evidence that this reorganisation has helped or hindered the academic progress of children in secondary schools?

Such questions turned out to need very careful definition and to contain a number of different questions. Answers can be given to some of them but, as should become clear in this introductory chapter, only some of those questions are amenable to answers based on empirical evidence. Any exploration of an educational system has to choose when to look at a dynamic, indeed rapidly changing scene, and has to concentrate on one stage in its development. This study was of a particular period (1969 to 1974) in the piecemeal changeover to comprehensive schooling, at which time there was no fully comprehensive system. It is part of the purpose of this chapter to convey some characteristics of that era in detail. It must, however, be emphasised at the start that the findings of this study concerning schools called comprehensive may not generalise to present day comprehensive schools. They are observations of how pupils were faring in schools which were not true comprehensives but which coexisted with selective schools.

(i) Definitions of 'selection'

It was therefore necessary to clarify what comparisons of 'selective' and 'non-selective' secondary schools meant in this study. It will be recalled that under the old selective system of secondary school organisation in Britain, a minority (about a fifth) of the children in any given Local Education Authority were selected at eleven for a specially academic form of education, in grammar schools (or, occasionally, technical schools). The majority of children went to secondary moderns. Which of these schools were, for the purposes of comparison, to be defined as the 'selective' schools? If grammar schools were selecting the 'top' fifth, then the important comparison might be between grammar pupils, on the one hand, and secondary modern (or secondary modern and comprehensive) pupils, on the other. An alternative way of defining the purpose of this investigation was as asking whether the progress of pupils in schools called comprehensive at this period was as it would have been if they had gone to another type of school.

This would require comparisons between comprehensive pupils and those in the other two types of school attended by substantial numbers of children, grammars and secondary moderns. Accordingly, one major set of comparisons here was between three distinct groups of pupils according to whether they were in grammars, secondary moderns or comprehensives. Thus the study considered whether selection for grammar school made a difference, as well as whether 'rejection' by failing the eleven plus and going to a secondary modern made a difference. The point of those questions is an evaluation of the view that selection was required for the advancement of those most suited to academic work, in separate schools from the majority, whose educational needs were best suited to less academic, practical, broader, less examination-oriented work.

A further set of explorations was directed towards answering the question whether the pupils in comprehensives were faring as they would have in the old selective system. This required a definition of 'selective' schools as the combination of grammars and secondary moderns which in theory formed an alternative to comprehensives. Comparisons were thus made between two groups of pupils, the 'comprehensive' and 'grammar+modern' groups. Despite the attempts to compare the pupils in that way, this research was not a contrast between two systems. Still less could it be in any way a test of the ideals of proponents of either comprehensive or selective education. This introduces a fundamental observation to be made at the start, concerning the relation between policy and empirical data. Empirical evidence is not the sole basis for judging changes that take place in educational provision, but it can offer feedback on their implementation. Evaluation cannot be worthwhile, however, until it is clear whether the supposed changes have actually taken place. In this case, it is necessary to dwell on the sense in which selective and comprehensive schools did or did not exist at the time the study was carried out. Thus it will become clear that anyone whose desire to advocate a system led them to believe that there should be definitive evidence here for or against the desirability of a comprehensive system would be misguided.

In fact it was not possible to evaluate a change to a comprehensive system, since there remained, at the time of the investigation, grammar and other selective schools, and this makes the category of schools called 'comprehensive' less than fully comprehensive, in the comparisons here. Even in the 1980's it remains true of some comprehensive schools that they coexist with selective schools. Not only do some grammar schools remain, but also the independent schools take a proportion of pupils. Though this study was restricted to the maintained school population and pupils of independent and direct-grant schools were not part of the account, the fact that they existed influenced the intakes of the schools which are considered here.

(ii) Outline of the report

This report is the second to emerge from an evaluation of aspects of educational progress in selective and non-selective schools carried out at the National Children's Bureau with funds from the Department of Education and Science. The first report* described the progress of a sample of pupils between 1969 and 1974 in English comprehensive, secondary modern and grammar schools, as measured by tests of attainment and seventeen other indicators. This second report is a sequel, principally concerned with an attempt, using a similar research approach, to explore the public examination results of the same sample.

* Published in 1980 as 'Progress in Secondary Schools' (Steedman; NCB)

Having set the stage with those prefatory remarks it is part of the purpose of this introductory Chapter to describe the context of this investigation, with a little of the educational thinking and the background of debate and research on outcomes of comprehensive and selective schooling (section iii). This is followed by a clarification of the nature of this piece of research which may dispel some misconceptions of what the study aimed to do (iv). In sections v to xiii, details of the problems encountered and strategies adopted are given. The characteristics of the sample (v) at eleven years old, in the various categories of school, will be described (vi), and that will reveal the high proportion of children whose schooling was being reorganised while they were at secondary school, who cannot be regarded as receiving a comprehensive education, as well as the numbers who went to grammar schools. The consequences for the intakes of the schools involved (attainments and social class of their pupils) are then summarised (vii). The problems which this created for comparisons between pupils of the various sorts of school are explained, also in this first Chapter, section (viii), and the strategies for overcoming these difficulties are described in sections (ix) to (xiii). A fuller discussion of the historical context and the state of education between 1964 and 1974 in England, as suggested by the pupils in this sample in each type of school, is in the publication stemming from the first phase of the project. This first Chapter seeks to elaborate the warnings already given on the lack of a fully comprehensive system, and the selective nature of the secondary schooling system under study.

Chapter I presents some essential technical information to explain how intake differences between school groups could be taken into account by using longitudinal data. The techniques used to measure progress in secondary schools in isolation from earlier characteristics and home background will be described in as accessible a fashion as possible. (A more technical description for the researchers and statisticians is contained in the Statistical Appendix, and the most detailed material on designs for analyses is also appended).

Much of this first Chapter is a summary version of the background detailed in the 1980 publication. The latter can be referred to, in addition, for a full account of the findings on 'outcomes' of secondary schooling other than public examination results. In order to make this sequel somewhat self-sufficient, though, Chapter II contains an abbreviated version of those findings from the first phase which seem most closely linked to examination results, or to indicators of the academic 'output' of schools. These 'scholastic' outcomes include scores on tests of attainments at sixteen years old, pupils' ratings of their own academic progress and the age at which they intended to leave school. Chapter II also contains supplementary results and a discussion of these measures which bring up to date those findings which can be compared with examination results presented later.

The results of other, non-academic 'outcomes' of secondary schooling explored in the first phase are contained in Appendix F. Like all other Appendices referred to in this document, that is available on application to any of the libraries shown at the head of the list of Appendices, page (vii).

Chapter III is an introduction to the problems and assumptions inherent in measuring 'outcomes' of schooling. In particular, there is a discussion of the adaptations of the research approach which were necessary to cope with measuring examination results. Chapter III clarifies the necessarily incomplete nature of the study, which must be viewed as an exploration of a few of the pitfalls in using examinations as measures of a large sample. Some of the decisions to be made when using examination results as indicators of average characteristics of school populations, and an outline of the measures adopted for this study, are described.

There is some explanation of technical problems with general implications for the interpretation of results. Chapter IV gives the results of the use of such measures in investigations of the exam results of these pupils. That chapter contains the main detailed presentation of the hitherto unpublished findings on examinations. Chapter V sketches out what the data seemed to show and attempts to draw together the findings of the two stages of this study, particularly concerning "scholastic" outcomes. Brief reference is made in that concluding section to the bearing these results may have on the appropriate safeguards to be adopted in publication of examination results by schools and local authorities. Finally, mention is made of the scope and limits of possible educational implications of the study.

But, first, a brief description of the context of the study, of research and educational thinking, is offered.

(iii) The context for the investigation

Part of the rationale behind the selective schooling system of post-war England and Wales was that a 'tripartite system' of grammar, technical and secondary modern schools would best cater for the fact that children differed. There were two assumptions on which the scheme rested, that three groups of children could be distinguished and that the schooling which would best suit each child could be determined by testing at the age of eleven. The three types of school were to have parity of esteem; the schools for more academic children, fostering more academic skills, were not to be given higher status than the schools for technical or practical skills. Though selection was intended to be by ability, some research, by the 1960's, suggested social class effects were not avoided by the selection process (Douglas, 1964; Douglas et al., 1968); and the arbitrary nature of the classification of children was suggested by regional disparities in the proportions going to grammars which were not explained by attainments at eleven (Eggleston, 1977). So there was room for assessment of the efficiency of the selective organisation.

In view of the supposed differing aims of grammars and secondary moderns, it may seem very one-sided to assess schools and, particularly, to compare these different types of schools, on academic criteria. On the other hand, the selective system was regarded by some as capable of maximising opportunities for academic achievement given limited resources. Certain inefficiencies in the selection process were suggested by research. It was argued that a certain amount of academic 'talent', which in a period of educational expansion demanded

leave early from grammar schools, in addition, suggested that the selection process might not guarantee the maximum educational advantage even in academic qualifications and even to those selected (Floud ed., et al 1956; Hopkins, 1978; Eggleston; 1967).

Some advocates of comprehensives, too, held that they could extend what was valuable in grammar school education to a broader range of children. As part of this, a common school was seen as potentially more flexible and responsive to changes in pupils' capacities and interests after eleven. The aims of comprehensive schools and the combined aims of grammars and secondary moderns, therefore, were not necessarily different. Indeed, one temporary measure during comprehensive reorganisation involved comprehensives being organised to permit selection at thirteen, replacing an 11 plus by a 13 plus. This was sometimes called the 'Leicestershire system' and its possible contribution to the findings of this study is discussed in Chapter IV and in Appendix C. In that context, it is not wholly inappropriate to explore for possible 'effects' of comprehensive as against selective schooling, using the same indicators of outcomes of schooling. Nevertheless, that is a bias brought about by the demands of empirical research.

It may be argued that any such exploration is inappropriate, if the purposes of the two school systems differ. Such a study might be seen as unfair to comprehensives, if it is believed that more weight will be attached by readers to results on measures of academic attainment than to other indices of school progress. It could be thought that, by the nature of quantitive measurement of education, too much stress will be laid by the research design on narrow criteria of performance and attainment and even competitive success. Some may consider that these are not sufficient criteria for evaluating comprehensive or any other schooling.

The edge can be given to those criticisms by the limitations of some of the comparisons of comprehensive and other schools which have been published in the past. These can confuse the issue by producing simple averages for one type of exam result of comprehensive and grammar pupils, or even slightly less simple averages (Baldwin, 1979, 1981) which, because of the differences between the two samples, do not compare like with like. At least the study which was carried out here was able to do better than that, as the next section shows. That description itself will confirm that no comparison of crude averages is any basis for judging pupils' progress.

(iv) The nature of this investigation, in relation to what a study of selective/nonselective schooling might be

It is necessary to dispel possible misconceptions of what a comparison of selective and non-selective schooling might be. An empirical study, it was said, cannot compare two ideal, theoretical systems one of which has not existed. It might, even so, be believed that a research study could compare the 'old' selective system which existed in the 1950's and later, with a 'new' comprehensive system that had replaced it "nowadays". This is not what was done here. This is not only because there has not been a comprehensive system, since comprehensivisation has proceeded gradually and a full comprehensive system is

still being completed. It is also because to compare outcomes of schooling then with outcomes now would be a different study; all the attendant confounding factors would tend to make it other than a comparison of selective and nonselective schooling.

It might be thought. secondly, that comprehensive reorganisation took place in such a way that simple comparisons could be made, using data from 1974 for instance, when about half the sixteen year olds were in comprehensives, between those pupils in Local Authorities which happened to have gone comprehensive, and the other half as typical of 'the alternative', a selective system with about a fifth in grammars and the rest of the children in secondary moderns. There could not, however, be a simple comparison because Local Authorities continued to have a mixture of schools, and comprehensives did not exist in isolation from grammar schools.

(v) The National Child Development Study

The approach to overcoming those problems here was to design a longitudinal study. That also permitted great precision in the definition of samples in each type of school and description of the intakes of different school groups or average characteristics of the school populations at eleven years old. The approach was possible only with longitudinal data, measuring the same children at eleven and again at sixteen. Hence the value here of using data from the National Child Development Study (NCDS), which has collected information on all the children in Great Britain born in a particular week in 1958, at birth, and at seven, eleven, sixteen and twenty three years old. Because NCDS has detailed material on characteristics of those children both before the start of secondary school and at the end of their secondary school years, it is a valuable source for a study of secondary schooling.

The members of the study cohort went through secondary school between 1969 and 1976. So NCDS is a source uniquely suited to a comparison of those in different categories of school, as the stage of comprehensive reorganisation meant there were representatives in each type of school. Nevertheless, by the same token, the pupils studied were all of a particular year group, and at school during a period of rapid change in school organisation. So results from this study may not generalise to any other era of the history of educational provision, past, present or future. Moreover, 1974, the year in which these children were sixteen, was an unusual year, in that the school leaving age was raised. This sample were among the first pupils to be required by law to stay on at school until sixteen, and this undoubtedly created special conditions, if not special effects, in schools. What has to be remembered, too, is that this study is of children, 'child-based', not of schools, nor of systems nor of Local Authorities.

In being based on data from NCDS, this piece of research has advantages over a number of relatively unsatisfactory comparisons of selective and comprehensive pupils. It looks at national averages, for the whole of England, and compares sufficient numbers of pupils in the various secondary school categories for what might reasonably be attributed to their development during secondary school years, rather than earlier in their lives. Because it is a study of national averages, though, no one should look here for recommendations on the 'best' school for any particular child; there will always be more variation, and information, at local level than a national picture can reveal.

(vi) The pupils who went to each type of school

In order to be able to characterise pupils as belonging to any one of the categories of school, it was necessary not only to know what school they were in at sixteen, but also to know that they had been in the school throughout secondary school. So only those who had been an appropriate number of years in the same school were included in the analysis samples. In the case of most pupils, this meant since the age of eleven. Where their schools appeared to start at age 12 or 13, for instance in areas where comprehensives were organised as middle and upper schools, pupils who had been since 12 in '12+' schools or since 13 in '13+' schools were also included. (The possible role of this factor of the age of the youngest pupils in a school, which may also be relevant to the discussion of selection, is explored further in Appendix C).

Knowing what type of school pupils were in at sixteen was also a matter for careful definition. Teachers had characterised the schools whose pupils were studied as grammar, secondary modern or comprehensive. Although 53 per cent of NCDS sixteen year olds were in schools called comprehensive (and 56 per cent of those in one school throughout were in schools called comprehensive) over half of those pupils had been in a school which had been in the process of reorganisation while they were there ('transitional' pupils). The number of 'comprehensive' pupils, whose school had been called a comprehensive at least since 1969 when they arrived at it, turned out to be relatively small.

The pupils of the samples explored in analyses fell into their categories in the following numbers:

Table 1a

Number (n) in each school sample:

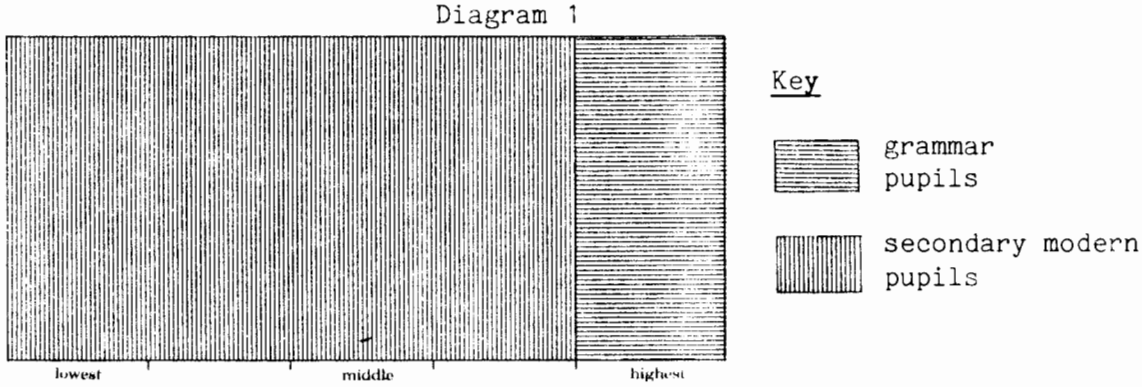
	as percentage of the three school types	total n	as percentage of all school pupils including transitional
grammar+modern	68%	1960	45%
grammar	26%	747	17%
secondary modern	42%	1213	28%
comprehensive	32%	936	21%
transitional		1479	34%

Even though, by 1974, the national figures (DES Statistics of Education) gave the proportion of grammar school pupils as about 12 per cent, it can be seen that the grammar school pupils formed 23 to 26 per cent of the sample for most comparisons in this study (15 to 17 per cent of the sample including transitional pupils). This was roughly the proportion of fifteen-year-olds in grammars under full selection (though that may have been lower, at about 20 per cent, according to Hopkins 1978). The significance of this in terms of intakes of schools will now be explained.

(vii) Characteristics of pupils going to each kind of school

It is possible to explain the stage of comprehensive reorganisation studied here by diagrams of the theoretical and actual allocation of pupils to secondary schools. If the population of eleven year olds were to be divided up according to their performance on tests at eleven into five groups, the theoretical allocation of pupils under a fully selective system, with (for example) 20 per cent in grammar schools, would look like this:

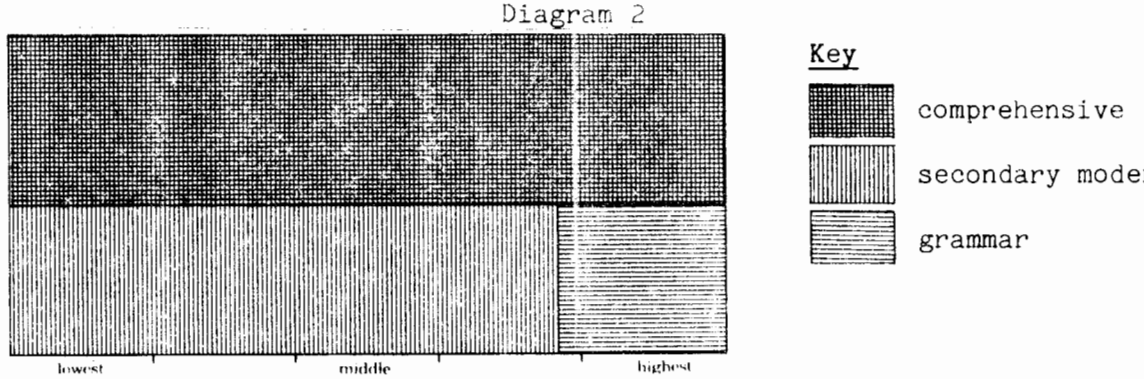
Diagram 1



If selection were perfectly reliable, and the eleven plus test reflected ability without any ambiguity or error, in a nationally uniform way, there would be no overlap between grammar pupils and secondary modern pupils. No one would be allocated to one kind of school, who could have gone to the other kind. So it would be impossible to compare the progress of children in the two types of schools. One could have no examples of exactly similar children in each of the two types of school, in order to be able to compare like with like.

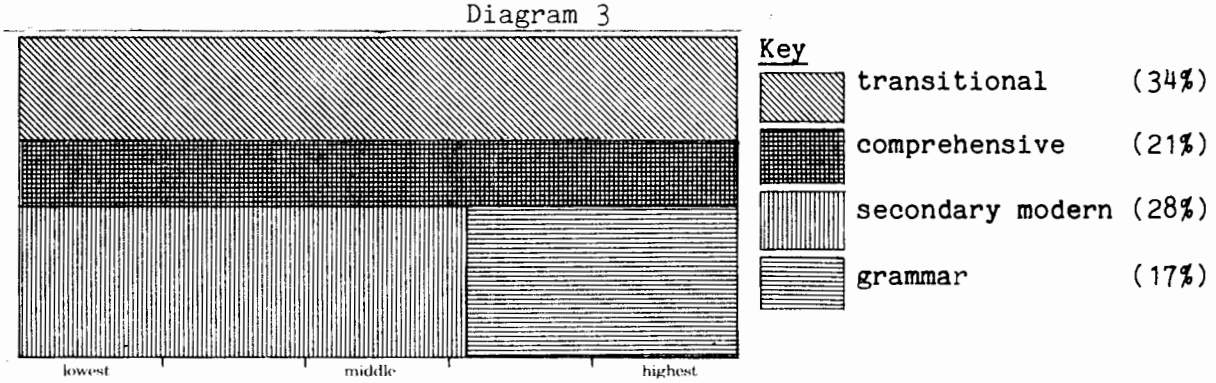
In the NCDS sample, however, there was not such a fully selective allocation, since 53 per cent of pupils were in comprehensives. It might mistakenly be believed that the pupils who went to comprehensives were the equivalent of those who went to grammars and secondary moderns, i.e. that the distribution was something like this:

Diagram 2



On the same analogy, given the actual proportions who were in grammar, secondary modern, comprehensive and transitional schools, for the appropriate number of years, the map would look like this:

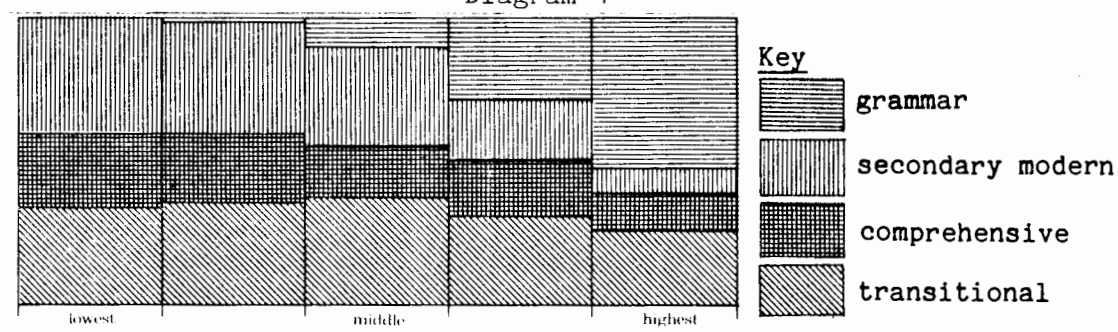
Diagram 3



fifths of range of tested attainment at 11

In fact however, allocation to comprehensive and selective schools was not equivalent in this way, and the actual proportions in each fifth of the range were as follows:

Diagram 4



fifths of range of maths scores at 11

It can be seen from these diagrams that, among comprehensive pupils, the actual chances of being in the lower attainment groups at eleven were higher than in theory. When the samples were fully described with all the necessary constraints, it emerged that the average test scores at eleven years old for comprehensive pupils were somewhat closer to those of secondary modern pupils than to those of grammar school pupils.

Table 2: Intakes of schools: tests at eleven

average test score at 11:	maths(0-40)	reading(0-35)	general ability(0-80)
grammar	29	22	60
secondary modern	13	14	39
comprehensive	14	15	40
transitional	16	16	42

The average test scores for transitional pupils at eleven, when they were starting secondary school in grammars and secondary moderns, were a little higher, but not much, than those of comprehensive pupils. Overall, the average tested attainments of the grammar+modern combination before the start of secondary school were higher than those of the comprehensive sample.

Table 3: Intakes of schools: test at eleven

	test score at 11	maths (0-40)	reading (0-35)	general ability (0-80)
grammar+modern combination	mean n = 2632	18	17	46
comprehensive	mean n = 1279	14	15	40
transitional	mean n = 2038	16	16	42

(The examples shown in the diagrams are those involved in the most recent analyses of public examination performance. They differ very slightly from the samples used in tables above and in the earlier publication of this project, though their implications are the same. The principal reason for the discrepancies in sample sizes is that the sample of everyone for whom the NCDS had information on their examination results is not identical to the sample of those with appropriate information for earlier analyses).

In addition to those differences in attainments, the samples had rather different social class characteristics. The proportion of comprehensive pupils whose fathers were in "nonmanual" occupations in 1969, for instance, was at least as low as that proportion among secondary modern pupils.

Table 4a Intakes of schools: social class at eleven

	percentage 'nonmanual' worker father	
grammar	57	per cent
secondary modern	27	per cent
comprehensive	26	per cent
transitional	30	per cent

The similarity in social class composition between comprehensives and secondary moderns is noteworthy, as is the high number of grammar schools tending to take pupils of higher social class and academic potential, who thus did not reach comprehensives. It is possible that areas which adopted comprehensive

organization fastest were those with higher proportions of manual working class pupils and relatively high numbers of secondary modern pupils. These phenomena would help to create the similarity between comprehensive and secondary modern intakes shown above. Perhaps they reflect the fact, if it is true, that the classic, early comprehensive authorities were inner city Labour-controlled areas. It is nevertheless, worthy of note that the areas which went comprehensive early, even if they were disproportionately'working class' or 'secondary modern', were not necessarily the Labour-controlled local authorities but were as likely to be rural authorities (often under Tory control) according to Eggleston (1977). Of relevance to a child-based study such as this, though, is the possibility that the number of pupils in inner city areas which went comprehensive might be higher than the numbers of pupils in rural authorities which went comprehensive. This would tend to produce the high proportion of manual working class pupils in comprehensives , but is a speculation which needs checking by research more thoroughly. In any case, it was impossible to evaluate possible outcomes of secondary schooling without taking account of the characteristics of the pupils before secondary school. If simple averages indicated that the grammar+modern combination was 'ahead' of the comprehensive pupils at sixteen, for example, it would not be possible to draw any conclusion about the influence of secondary schooling, since they had been 'ahead' by eleven. Therefore, some means of calculating the progress of pupils net of their initial attainments at eleven and of the advantages or disadvantages of their home backgrounds was called for.

Before the method of doing so is described, though, consider the high numbers of grammar school pupils relative to secondary moderns in the combination of grammar+modern pupils.

(viii) The weight of grammar and secondary modern pupils in the combination.

The figures have demonstrated that there was a relatively high proportion of grammar pupils in the selective group. It is odd to tackle the question of whether being in a comprehensive seems to make a difference when the comparison is not with a system which ever existed, but with a mixture of schools having an extraordinarily high ratio of grammars to secondary moderns. It is arguable that a comparison would require reweighting, to restore the grammar sample to its correct proportion as a fifth or so of the combination. Indeed, it has been argued (Goldstein, 1981) that, to make the comparison meaningful, the grammar+modern combination should be reweighted. The weight of grammar school pupils in the 'selective' group meant that the comparison did not correspond to anything like a 'real' selective versus nonselective comparison. If there were relatively low performance among secondary modern pupils, a more realistic, higher proportion of secondary modern pupils might lower the grammar+modern average. Had the comparison been between comprehensive pupils on the one hand, and a selective grammar+modern combination recalculated to make grammar school results a fifth of the total, for instance, on the other hand, the comprehensive group average might have appeared differently related to the 'selectives'. This would depend on whether the secondary modern results were lower or higher than others, which is one of the questions explored. If they were lower, any lack of an overall difference between pupils in comprehensives and the rest could imply superior results in comprehensives, as the composition of the sample would

'overestimate' the mixture of grammar+modern pupils. Simple comparisons of average exam results at sixteen showed that, if the mixture had been more typical of a selective system, with about a fifth in grammars, the grammar+modern average would have been lower. For instance, the numbers of 'O' level equivalents in each school group, on average, were:

Table 5a

Type of school	uncorrected 'raw' mean number of 'O' level equivalents among pupils with one or more 'O' levels	total number of pupils with one or more 'O' level
grammar	5.69	672
secondary modern	2.76	430
comprehensive	3.52	425

The averages for the entire samples, including those with no 'O' levels, were:

Table 5b

Type of school	uncorrected 'raw' mean number of 'O' level equivalents among all pupils	total number of all pupils
grammar	5.12	747
secondary modern	0.98	1213
comprehensive	1.60	936

and the average for all grammar+modern pupils was:

Table 5c

	uncorrected 'raw' mean	total number
grammar+modern	2.56	1960

Recalculated as if there were more secondary modern pupils, the average for the grammar+modern combination would be lower. That is, if the number of secondary modern pupils were four times the number of grammar pupils, as if grammars were 20% of the 'selective' grammar+modern group, the average would be:

Table 5d grammar+four times the number of grammar as secondary modern pupils

	uncorrected 'raw' reweighted average	total grammar + (4x grammar)
grammar+modern average:	1.81	3735

Including more secondary modern pupils, then, would lower the average examination results of the grammar+modern combination. But this study was able to pursue the question further and ascertain the extent to which that reflected differential progress, changes in level of performance after eleven years old, in grammar schools as against secondary moderns. The lower examination performance after eleven years old in secondary moderns might simply be a function of the fact that secondary modern pupils, on average, were behind grammar school pupils in attainments, and disadvantaged in home background, before the start of secondary school.

The difference between the average examination results for the grammar+modern combination and the average for comprehensive, moreover, could also be attributed to differences before the start of secondary school.

It would be possible to alter the ratio, and reanalyse the comparisons between comprehensive pupils and their 'selective' counterparts. On the other hand, as was pointed out in the 1980 publication, it could then be claimed that the comprehensive sample, too, needed reweighting. It would not be clear how to reweight them to represent a true system. The figures as they were at least showed some historical state of play. There is some justice, then, in looking first, as was done here, at the actual (albeit arbitrary) numbers which were in each category. It is a task for future research to explore the possibilities in and results of reweighting comparisons. An attempt to answer empirically a crude version of the 'selection' question is all that was tried here, and so it is necessary to make explicit how it differs from what the theoretical comparison would be. Hence the comparisons between the combined grammar+modern 'selective' group, with its 'unreal' ratio of grammar to modern pupils, and comprehensive pupils. Since the comparison, unsatisfactory though it was, had to be made, it seemed no more arbitrary to use the samples as they were than to reweight. Allowing for initial differences at eleven years old was anyway more reasonable than not doing so. Comprehensive pupils were being compared with a group which had a higher average test score at eleven. The only way to explore whether progress in these particular schools was similar, therfore, was to recalculate the crude averages above, to allow for the contribution of attainment at eleven years old. But the initial differences were not accidental, they were part of the effect of the continuing existence of some selection at eleven.

(ix) The use of longitudinal data and corrections for pre-existing attainments and background factors

The preceding discussion showed that the overall average examination results at sixteen were lower in this comprehensive sample than in this particular combination of grammar and secondary modern pupils. The co-existence of these

three types of school, however, the numbers of pupils in each type and their characteristics meant that averages at sixteen were in part a reflection of average attainments at eleven before the start of secondary school.

A lower average for comprehensive pupils at sixteen might reflect similar progress in secondary school to the other group's, since comprehensive pupils had been lower scoring at eleven, on intake. It would not be particularly valuable for a study such as this merely to provide the average examination result by sixteen. Similar figures suggestive of the differences shown, between comprehensive and other pupils, have been readily available for many years from other sources, and, in particular, from published DES Statistics, along with the findings of independent research, Local Authority statistics and bulletins from examination boards, NFER and elsewhere. These all provide equivalent, more up to date accounts, and there would be no value in adding to those figures from simple averages of examination results from the sixteen year olds in NCDS in 1974. The purpose of the longitudinal study here was to ascertain the extent to which differences in average examination performance between comprehensive and other groups were attributable to initial attainments, social class, sex and parents' interest in schooling.

By using statistical techniques, it is possible to compensate in some measure for the nonequivalence of the samples, in order to investigate whether similar progress was made by comparable children in the various categories of school. In order to do this, techniques were used that permit comparisons between nonequivalent groups. Certain factors, on which pupils have been seen to differ were used to discriminate them in a multivariate analysis. That is, since the social class composition, for instance, of the groups of grammar, secondary modern, comprehensive and transitional pupils differed, it was necessary to discount in any differences between them the component attributable to social class differences. (This is done by techniques analogous to 'analysis of covariance', in most cases here. The techniques are described in more technical terms in the Statistical Appendix). Similarly, where groups differed on attainments on a particular test before secondary school, the average amounts by which the results seemed to be explained by attainment on that test before secondary school had to be allowed for.

This is done by splitting groups into many, many subgroups; the two groups of grammar+modern and comprehensive pupils, for instance, were subdivided into five groups according to test score at eleven, two according to sex, three according to social class, two according to their parents' interest in their schooling and two for the age of youngest pupils in the school. This is 240 groups. When these many subdivisions have been made, comparisons between comprehensive pupils and other groups are carried out which calculate averages* for all the boys, for all those of a given score, of manual social class, and so on, relative to the overall average, so that the amount due to any one factor is known. This is the equivalent of 'matching' groups on all the factors which were taken into account, but matching could not possibly have been done here. For one thing, matching could be done on relatively few factors - sex and one test score, for instance - or one could not interpret results, because there would be too many separate results. Aditionally, the populations do not match, except in a very

* or actually the logarithmic function of the scores; see Statistical Appendix

few cases which could be compared. This, it will be seen much later, is more and more apparent, and the degree of overlap is seen to be less and less, as controls for pre-existing attainment become more refined and incorporate more test scores. Instead, an average over the range, the component which seems to be explained by social class (in this example) is used to reckon the amount by which any crude average difference between the two groups (which might mistakenly be expected to indicate something attributable to schools) should be corrected so that it does not reflect this factor of social class but may reflect a school difference.

The use of such techniques and the allowances for components attributable to extraneous factors is technically and properly called 'adjustment'. The term 'intake corrected' rather than 'adjusted' will be used here (even though statisticians may object that this is an incorrect use of the term 'correction').

It can be seen from this explanation of correction and adjustment that the extent of correction for social class, for instance, is simply the result of calculating how much (on average) each social class group departed from the overall average, in that sample in that analysis.

The uncorrected averages can be termed 'raw data', in the sense that they are 'before corrections'. They are not, of course, raw data in the normal sense of untreated individual data, because they are based on well-defined groups and averaged.

Raw data on which the results published in 1980 are based were deposited with the SSRC Data Archive and can be re-analysed by any bona fide researcher.

Where relevant, the average results for school groups will be presented not only as intake-corrected figures but also without any of the corrections for extraneous factors. It will be seen that some of these 'uncorrected' figures appear quite dramatic, and, if one fails to appreciate their dependence on prior attainments and background, they can be mistaken for results of schooling. This impression can be avoided by inspecting the results of the more defensible comparisons, with appropriate corrections, which it was the purpose of this study to carry out.

(x) <u>The corrections for extraneous factors made by analyses</u>.

In every statement of the findings of the investigation, it should be assumed, unless it is specifically stated otherwise, that the observations include allowance for one or more measures of attainment and/or 'ability' at eleven, and for the components of exam averages attributable to social class, sex and parents' interest, whenever results are ascribed to school category. In no way, though, could this study claim to have allowed completely for such extraneous factors. In various Appendices, analyses are described in more technical detail for those who care to know the exact nature of the corrections for pre-existing attainments and home background. Appendix R gives average exam results, on a limited set of measures, for school groups <u>without</u> allowance for these non-school factors.

Extraneous factors: pre-existing attainments

Attainments at eleven were the most crucial factor to allow for in measuring secondary school progress. Each set of results is presented as averages relative to at least one test score at eleven. This may be included in the analysis as a quasi-continuous variable, (that is, the analysis had separate averages for each score from 0 to 35, or from 0 to 80 depending on the possible scores, and so variation between any scores is allowed for). For some analyses, though, all the scores were grouped into three, or five groups. (This has the advantage of substituting for some correction for measurement error, but the different proportions in each school group for each third, or fifth, of the range make it an imperfect indicator, particularly when interactions are presented across the attainment range). Some analyses included more than one test score, in recognition of the fact that the pupils in different school groups might differ before the start of secondary school in more than one aspect of their attainment (exact details of the test scores used, with particular details of whether linear and quadratic components were allowed for, are in Appendix P).

Extraneous factors: sex

If the proportion of boys and girls differed between school groups, the influence of sex might be confused with the influence of school if the former were not allowed for. Most analyses included corrections for what was ascribable to the sex of pupils.

Extraneous factors: social class

It is well known that educational performance is associated with social class. However crudely social class is measured, pupils whose fathers are in higher socio-economic groups or have jobs of higher social status do better (see, e.g. Douglas, 1964; Douglas et al., 1968) and this effect appears to be independent of schooling and other factors. It was therefore important to include some allowance for this factor in every analysis.

It is customary at least to distinguish those of nonmanual social class from manual workers (e.g. Rutter et al., 1979). In addition, in this study, the children whose fathers' occupations were classed as professional and managerial (or, more accurately, social classes I and II according to the Registrar General's classification of occupations), were distinguished from those with fathers whose jobs were classed as III nonmanual (routine clerical workers). To include all those children in one social class category can be regarded as 'throwing information away', as Heath has argued, for 'there is plenty of evidence that children from professional and managerial backgrounds have considerably higher educational attainments on average than those from routine clerical origins', and this may be associated in its own right with differential progress (Heath and Clifford, 1980).

The third social class category was by far the largest and included all the children whose fathers were in manual occupations (and those without a male head of household). This large group might loosely be termed the 'working class' pupils.

These three social class categories (I and II; III nonmanual; manual and no male head) are a little unconventional, but, in comparing pupils in the three types of schools, this was a useful grouping, as the representatives of each group were as follows:

Table 4b Intakes of schools: social class at eleven*

social class at eleven	I and II	III nonmanual	manual, n.m.h.
grammar (n=747=100%)	42%	15%	43%
secondary modern (n=1213=100%)	17%	8%	76%
comprehensive (n=936=100%)	16%	8%	76%

Extraneous factors: parents' interest in their child's schooling

Over and above the influence of social class (father's occupation) or parents' education, the interest in school progress which parents show is highly influential on attainments and other aspects of pupils' well-being. In order to offset the known contribution to sixteen year old's performance of parental support and interest in their children's schooling, and the fact that this could be expected to be present to varying extents in different classifications of schooling, an indicator of parents' interest was incorporated in each analysis. The indicator chosen was based on assessments by teachers of the sixteen-year-olds, as their responses to the question 'with regard to the child's education, do the father (or father figure) and mother (or mother figure) appear ... overconcerned, ... expecting too high a standard, ... very interested, ... to show some interest, ... to show little or no interest'. Pupils who had one or two parents described as 'very interested' in their child's schooling were distinguished from those who had no such parent, each of whose parents was described as showing some interest, or having little or no interest.

There were thus two categories, or levels, of this 'parental interest' variable. Fuller discussion of relevant considerations in using this variable is in Appendix B. Although parents' education was also allowed for in a very few analyses, it was seen as more important to allow for parents' interest. For parental encouragement (as judged by a teacher's rating) seems to have the highest correlation of a series of variables (including parents' education, parents' book reading, income, family size and parents' attitudes to their child's education and future employment) with secondary school examinations and also with scores on a verbal intelligence test, according to Fraser (1959). As Douglas (1964) demonstrated, at age 8 and age 11, parents' interest has the closest relation to attainment of four variables (the others being housing, family size and their primary schools' academic record). This may be particularly crucial in a study of selective and nonselective schools, for, as Floud and her co-authors showed, those whose parents' attitudes formed a

* The proportions vary slightly from those in Table 4a, owing to the inclusion of class IV nonmanual and those with no male head of household in the 'manual' group in Table 4b.

favourable climate of schooling were more likely to be selected (Floud et al., 1956). Using teachers' ratings of parents' interest in schooling, gathered when children were sixteen, may involve a confusion between an outcome, and an influence on secondary school progress. Teachers' perceptions, too, might differ systematically between the types of school explored here, as a function of intakes. That could make teachers in a particular type of school with an intake biased towards the 'disadvantaged' underestimate or overestimate parents' interest. Judging by the class composition of samples here this might affect comprehensives more than the mixture of grammars and moderns, but it is not clear in what direction estimates might be distorted.

Other extraneous factors

Limitations of a technical kind meant that little more than the factors listed above could be allowed for in analyses. Certain obviously important factors such as the social mix of the school*, parents' education, and region could be included in only a few analyses and will be mentioned, where relevant, but it was not thought that they would add a great deal to the explanation, given the allowances for social class and parents' interest in schooling.

(xi) A further important feature of every analysis

Each of the factors above, and, most importantly, the school variable, was tested as 'last one in'**, so there was no need to take into account the order of their appearance in analyses as there would have been with a stepwise procedure. For those concerned to know the exact nature of the statistical techniques used, the Statistical Appendix contains a technical description.

(xii) 'Yardstick' (continuous) and 'criterion' (dichotomous) measures: results presented as differences and as odds.

The measures used as 'outcomes' in the work reported here are of two kinds, either dichotomous or continuous. The dichotomous ones are the proportions, for which results are expressed as odds, and they represent 'yes/no' answers to questions like "do they like school?" or "do they want to be well-paid?". When used with measures of examinations, it will be seen, these present a criterion measure of performance, such as "did they pass 'O' level mathematics?" or "did they take any exams at all?" The continuous, or quasi-continuous, measures are the ones which emerge as corrected averages, of a more familiar kind, and are yardstick measures like "how many points did they score, between 0 and 35?", or "at what age did they plan to leave school, at 16, 17 or 18?" The two sorts of measure are not the only possible ones. Some short-comings of such measures, particularly for examination results, are discussed in Chapters III and IV. Results are presented in Chapter IV in ways which may be unfamiliar and the following is intended to make them comprehensible.

* See Morrison and McIntyre, pp. 68-9. (1971); Coleman et al. (1966) (p.65). The drawbacks to the NCDS indicator of social class composition of the school are detailed in 'Progress in Secondary Schools', page 80.

** See Statistical Appendix

It is always the relations between averages for school groups that are interpretable, and there are no meaningful absolute 'average scores' to be presented. Comparisons between groups of pupils on continuous measures are therefore expressed as differences (the higher adjusted mean, minus the lower one). So differences between, say, comprehensive and secondary modern pupils are obtained by subtraction. If they were no difference, the result would be zero.

Measures of the second type were dichotomous, based on whether or not a pupil had or had not reached a certain grade, for instance, or on whether or not a pupil had reported staying away from school. Analyses using such dichotomous outcome measures calculated proportions fulfilling the criterion (corrected for initial eleven year old attainment, sex, social class and parental interest, as usual). Because dichotomies were involved, the results of these analyses of proportions can conveniently be presented in terms of the 'odds', for each group of pupils, of falling into one or other 'camp'; as, for example, the odds of obtaining a 'pass' at 'O' level standard in mathematics, or the odds of staying away from school. As with the continuous measures, the results which are interesting are always those of comparisons between the various school groups. In the case of dichotomous variables, though, such comparisons are not calculated in terms of additive (or subtractive) differences between groups. It is necessary, instead, to indicate relations between groups by presenting a ration of one set of odds to another. If there were no difference, the result would be 1. Thus, results in the case of criterion measures are presented as 'odds ratios'.

For fuller explanation of these two different ways of presenting results, readers may wish to refer to Chapter III of 'Progress in Secondary Schools' (section g pp. 86-9).

(xiii) The categories of school studied:

As in the earlier study, comparisons were drawn between pupils in various school groups categorised in four ways. Firstly the three groups, of those in grammars, pupils in secondary moderns and comprehensive pupils, were compared ('type of school' comparisons). The three types of school were very different, so separate comparisons were important. But the existence of grammar schools necessitated secondary moderns, and vice versa; the selective system meant both.

Partly for that reason, it was also necessary to compare the comprehensive pupils not with grammar school pupils alone, but with the combination of grammar and secondary modern pupils ("grammar+modern combination"). So the second main question implied by earlier discussion of whether comprehensives managed to cater for pupils in the way the 'old' organisation did was explored by comparing pupils of comprehensives with the average for pupils in what, in theory, was the alternative, namely, grammars and secondary moderns. In order to approximate this comparison, the combination of grammar and secondary modern pupils was looked at relative to the comprehensive group.

A third and fourth set of comparisons were designed in addition, partly because the variety among comprehensives can mean that some are oversubscribed, some undersubscribed, and this can create a form of selection. It was of interest to

explore this, and also to see whether there were exceptions to general observations on comprehensives. So, in addition, a few distinctions among comprehensives were drawn. For one thing, comprehensives formed from grammars and comprehensives formed from secondary moderns were differentiated, as well as purpose-built comprehensives and comprehensives resulting from the amalgamation of two or more schools. Any differences or similarities between those subgroups and the grammar+modern combination were studied (grammar+modern/origins comparisons). A difficulty for a study such as this is that such fine distinctions between comprehensives are particularly likely to be specific to the early 1970's. Discussion of those subgroups is confined to Appendix A, therefore.

A fourth set of comparisons was designed to ascertain whether differences in exam results were related to the date at which a school became comprehensive. That is, the analyses sought to establish whether comprehensives differed according to whether they had been set up by 1965 and therefore might be expected to be relatively settled (early comprehensive) or had been set up between 1966 and 1969, only a few years before the children in the study entered the school (recent comprehensive). In addition, the samples for the analyses included, as another separate group, pupils in 'transitional' schools, which were going comprehensive between 1969 and 1974 while NCDS members were at school. Again, these groups were also compared with the combination of pupils in grammars and secondary moderns ('grammar+modern/dates of comprehensive' comparisons).

The numbers of pupils in each subgroup, for the purposes of analysing exam results, were as follows:

Table 1b Numbers in each 'grammar+modern/dates' sample:

	n	percentage of total %
grammar+modern	1960	45
early comprehensive	344	8
recent comprehesive	592	14
transitional	1479	34
total	4375	101%

What questions did these investigations of early and recent comprehensives aim to answer? One question was whether pupils in comprehensives were lagging, and whether this seemed likely to be due to 'teething' problems. Newer comprehensives might have problems. On the other hand, older comprehensives, set up at an earlier stage in comprehensive re-organization when re-organization itself had teething problems, might have problems. More recent ones could fit into an existing framework for comprehensives. Early comprehensives were quite likely to be rural schools (Eggleston, 1977; Wright, 1977) or may have started

in highly selective areas where the full quota of high-attaining eleven year olds was still creamed off to go to grammar schools (Bellaby, 1977), and it might be anticipated that this would show distinctive patterns of progress.

(xiv) Overview of Chapter I

These introductory remarks have, it is hoped, shown the context for the study and clarified the peculiarities of any study of 'selective and nonselective' schooling. Chapter I has described the state of secondary schooling in England in 1974, the year in which this sample was sixteen and in which most measures of school 'outcomes' were taken. It has thus shown that there was no evidence on which to draw for a comparison of two systems, a selective and a comprehensive, because the various types of school coexisted, and reorganization was as yet incomplete. More than half the pupils in comprehensives at 16 had started in selective schools which then became comprehensives. Those in comprehensives throughout were at age eleven closer to secondary modern pupils than to grammar school pupils of the same age, and on average lower attaining, before secondary school, than the combination of grammar and secondary modern pupils. So to compare 'raw' averages later would not be to compare the relative merits of schools so much as the relative capabilities of their intakes. In what follows, therefore, the comparisons which are useful are those which attempt to offset the extent to which school 'outcomes' are distorted by factors other than secondary schooling.

No empirical research technique can allow for the fact that all these schools co-existed in the education system of the early 1970's. Testing the academic progress of comprehensive pupils with data from 1974, a particular year in the piecemeal transition to comprehensivization, is not testing the merits of a comprehensive system. Comprehensives were never intended in theory to operate while grammar schools continued to exist, and the grammar school pupils who would form part of a 'comprehensive' intake were absent. So under consideration were schools called comprehensive in a climate of selection. This context has to be understood before any findings are presented. There was no lack of selection, to yield data for comparison with what took place under selection.

So the question for this investigation was whether the pupils in comprehensives were doing as they would under selection. Selection continued to exist so, in that sense, of course comprehensive pupils were doing as they would under selection. The fact that they were not, on average, much more likely to be selected at eleven than secondary modern pupils suggests how they would have done under selection.

Thus some difficulties in addressing this question empirically will already be apparent. It is impossible to do without debatable assumptions, which it was part of the purpose of this introductory chapter to convey.

Chapter II

Progress in Secondary Schools: A summary and discussion of earlier investigations of this project

(i) Introduction

The earlier stage of the study of progress in selective and comprehensive schools, from which results were published in July 1980, took place before the examination results of these pupils were available to the research team. It covered a range of aspects of interest to those wanting to know how, during a particular period in the development of comprehensive schooling, those who went to comprehensives did, compared to those who went to selective schools.

Outcomes of schooling explored in the course of the study

The characteristics of the pupils and schools studied here encouraged the use of a varied set of indices of their educational attributes at the age of sixteen. A number of different measures of possible 'outcomes' of school were chosen. As usual in empirical work of this kind, it was not possible to infer a causal connection from associations between school and attributes of pupils. Nevertheless, because the allowance for extraneous factors was intended to rule out some correlated factors in the attributes of pupils, extraneous to secondary schooling, the purpose was to look at aspects of sixteen year olds which might be ascribed to the influence of schooling. These are termed 'outcome' variables here for convenience, but nobody should infer that influences of school have been demonstrated. A mixture of 'outcome' variables was explored, because any school pursues a great number of aims and influences different aspects of behaviour, academic attainments, basic skills, skills for life or socialization. So it would be inappropriate to measure a single aspect of sixteen-year-olds. The particular school categories explored, too, might be expected, by definition, to have different aims. Furthermore, the pattern of intakes would itself tend to create distinctive aims. Given the differences between categories of schools in intakes, certain outcomes might be more fitting criteria than others of the fulfillment of school aims, if not of the progress of pupils. A final incentive to measure a number of different characteristics of sixteen-year-olds in looking for outcomes of schooling was that some measures were better suited to the techniques of analysis than others, and the drawbacks to any one measure are offset by approaching the same questions with a number of different measures.

In the light of the considerations above, a mixture of aspects of sixteen year old pupils were explored for what they might reveal of relative progress in each school category. At the risk of helping to perpetuate divisions in educational aims, these 'outcome' measures could broadly be divided into scholastic aims and less academic indicators. On the scholastic side, the results pupils obtained on tests of mathematics and of reading, taken at 16 years old, were compared. Other indicators, possibly of school influences on 'scholastic' aspirations, were also of interest. For example, as well as the 'objective' measurement of attainments, the pupils' own ratings of their performance in mathematics and English were analyzed. Not only might self-confidence or success in a subject breed success, but this might be fostered differently, even among children of similar attainments at eleven years old, in the different categories of school. After all, one of the disadvantages of selection, it has been argued, is that divisive self-images may be created. Such self-perceptions are dependent, too, on the context of the abilities of other pupils.

Furthermore, the proportions wishing to follow courses of advanced study at college, polytechnic or university were measured. The age at which the sixteen year olds planned to leave school was also taken as an indicator of 'scholastic' plans. And, for those planning to leave at sixteen, it was thought appropriate to consider whether or not they wished to continue part-time study (day release or other) after starting work.

Some details of scholastic outcomes explored by the earlier stage of the study may be useful.

'Academic' outcomes discussed in this Chapter

- tested attainment at 16 in mathematics; a 40 item test, designed by NFER and first used by NCDS.

- tested attainment at 16 in reading comprehension; the same test as was used at eleven years.

- self-rated performance in mathematics
- self-rated performance in English

) at sixteen, pupils were asked to rate their performance in a particular subject, relative to other pupils (as 'below average', 'average' or 'above average')

- plans for advanced courses of higher education

- plans for school leaving age.

- plans for part-time study if leaving at 16

Other, less academic, 'outcomes' of schooling, though, such as signs of disaffection with school, may be just as telling, though the links with examination performance are less clear. The concern of the earlier stage of study was not only with 'academic' or 'scholastic' outcomes of secondary schooling, but also with indices of possible behavioural, affective or social influences of school. As well as reporting the 'scholastic' measures (the scores on tests, the self-ratings of academic performance, the intended leaving age and the plans for future study of these pupils), the earlier publication contained 'non-scholastic' results of ratings of truancy and disturbed behaviour, self-reported absence from school and liking for school, pupils' attitudes concerning future employment and parents' satisfaction with their schooling. In the interests of brevity, though, it has proved necessary to confine this Chapter to those 'outcomes' of schooling which are termed 'scholastic' above.

A summary of observations on those behavioural, affective or social outcomes is therefore offered in Appendix F, and, for more detail, readers are referred to the 1980 publication. To exclude from the text here the results on truancy, attitudes to school and views on future employment runs a risk. The danger is that this investigation could be seen as overconcerned with narrow, performance criteria of school success. The risk is increased by the attention paid in

Chapter IV, the first publication of the examination results, to GCE and CSE performance. As long, however, as readers remind themselves that this Chapter is consciously restricted to scholastic outcomes other than examinations, and that a few other aspects have been reported elsewhere, the advantages of brevity make the risk worthwhile.

In this second Chapter, a short account and gloss on the chief findings concerning scholastic outcomes in that 1980 publication will be given. It is hoped that a brief summary will help to give a context to the public examination measures of educational attainment. Such a Chapter also provides an opportunity to include new findings on maths and reading test attainments which were not available for the earlier publication. Scholastic attributes of sixteen year olds have perhaps most obvious relevance for comparisons with examination results; and correspond most closely to the crude assessments of academic outputs of secondary schooling to which this type of research is best suited. The summary may also help to dispel any impressions that examinations are the only, or best way to measure schools or the only thing schools are aiming at.

The purpose of this summary of earlier findings is to make this account self-sufficient. The school 'outcomes' form a mixed picture, and this Chapter will attempt to provide a path through those observations. It is first necessary to give a note of how to interpret the figures, some of which are given in this Chapter and many of which are to come in Chaper IV on the public examination results of the same sample. Secondly, this Chapter summarises the material on scholastic outcomes presented in the earlier publication, with a little supplementary evidence which helps to clarify and support the conclusions. There is also some consideration of issues that have been highlighted by reviewers or other critics of that work since its publication. The comparisons of the three separate school types (grammar, secondary modern and comprehensive pupils) are presented first, and followed by the comparisons between the comprehensive pupils and the grammar+modern combination. In both cases, the indicators of 'scholastic' outcomes discussed will be: test scores, pupils' self-ratings in maths and English, plans for higher or further education and the age at which pupils intended to leave school.

Comparisons of the largest school groups are followed by consideration of the extent to which overall averages could be elaborated by variations according to ther initial ability of the pupils concerned. Whether those of 'average' test scores, for instance, were the same as those of high test scores at eleven was explored in so far as this was shown in their performance by sixteen and how it varied between school groups. Grammar school pupils were higher scoring, on average, than secondary modern pupils, for example; was this the case over the whole range of attainments at eleven, or were 'average' pupils apparently better off in secondary moderns (as the theory of selection would imply)? Again, these patterns are looked at here only in the 'academic' measures. The observations on pupils in comprehensives of different dates and transitional pupils, including any interactions with eleven year old test scores of the sort described in the previous paragraph, are next to be discussed. The Chapter concludes with a discussion of the implications and of the interpretations which are possible. Alternative perspectives, which could suggest other conclusions, are given consideration.

A note on interpreting observations

As was emphasised in the previous Chapter, characteristics of the sixteen-year-olds, as averages of groups according to school, were only possibly indicators of 'outcomes' of secondary schooling in so far as they were corrected for levels of attainment before secondary school, and for home background.

It would be excessively repetitive to keep stating that, in what follows. Unless otherwise indicated, though, every observation is to be read as an indication of an average over an entire group of sixteen year olds, after correction for the components attributable to extraneous factors (typically, an average associated with school group). The absolute levels of such corrected averages are not readily interpretable; it is the similarities or differences, or odds ratios, between group averages which are meaningful. The size of any differences can be judged with reference to the standard deviation from the mean, or standard errors, and these are therefore mentioned where it seems relevant to do so.

(ii) 'Scholastic' outcomes other than examination results

Comparisons of types of school: grammars, secondary moderns, comprehensives.

test scores

The three main groups averaged out as statistically significantly different in average test score in reading and average test score in mathematics at sixteen, with the appropriate corrections for test scores at eleven and other factors. Those in grammar schools had made the most progress, on average, between the two occasions on which they were tested. Their average score (corrected) was almost a quarter of a standard deviation higher than that for comprehensive pupils, on the reading test at 16.

In mathematics, the grammar school pupils' average was over one-third of a standard deviation higher than the comprehensive pupils. There was only small advantages, on average after corrections, to the comprehensive pupils over and above the secondary modern pupils. They made statistically significantly more progress than secondary modern pupils, nevertheless*.

self-ratings

The study also compared pupils' own ratings of how good they reckoned they were at English and at mathematics, compared to other pupils of their own age (on a scale 'below average', 'average', 'above average'). The averages in such ratings, relative to test score at eleven years old, are difficult to interpret, but it is important to consider how self-ratings may be influenced by the characteristics of other pupils in the school or peer-group.

* Even with allowance for maths, reading and general ability, the very small advantage to comprehensive pupils was over 2.5 standard errors. It was therefore statistically significant, since the number of standard errors equal to a statistically significant difference can be thought of as the square root of the X^2 value; here $X^2 \simeq 6$, in a three level comparison (with two degrees of freedom) when $p < .05$.

In this case, comprehensive pupils, on average, again resembled secondary modern pupils, but both groups had rather higher self-ratings in English and in mathematics than grammar pupils. (Secondary modern pupils had, if anything, even higher self-ratings in mathematics than comprehensive pupils).

Because these ratings are relative to initial attainments at eleven, it can be seen that higher ratings for lower scorers and relatively low ratings by higher scorers - that is, people rating themselves more 'average' than they were - could give rise to these findings. It was not necessarily the case that the 'raw' average self rating of secondary modern pupils was higher than the 'raw' average for grammar pupils, for example.

plans for advanced study

The proportions wanting to continue their studies with an advanced course in a College, Polytechnic or University were highest in grammars and lowest in secondary moderns, even after corrections for one earlier test score and other factors. The odds for comprehensive pupils of their wanting to take such courses were somewhat lower than those for grammar pupils, and only a little higher, overall, on average, than among secondary modern pupils.

plans for part-time, continuing study after leaving at sixteen

Plans for future study were not so variable among those (presumably less academic) pupils who planned to leave school at sixteen. There was no statistically significant difference among these pupils according to their school type, in proportions aiming at getting a job which offered part-time study or day release. (One difficulty, however, in interpreting this is that more boys than girls are able to consider the possibility of day release.)

intended school-leaving age

The last of the seven indicators of 'academic' outcomes was the leaving age pupils intended. With correction for 'general ability' at eleven, and the usual other factors, there was a clear pattern of grammar pupils on average planning the latest leaving age, markedly later than the secondary modern pupils. Comprehensive pupils averaged in between, estimating their leaving age as about half a year (on average) earlier than grammar pupils and about quarter of a year later than secondary modern pupils.

It may be useful to give an overview of the findings for each school type.

Grammar pupils

Pupils in grammar schools had the highest average test score on reading and by far the highest on maths, even after allowing for their related test score at eleven years old. They had the lowest self-ratings, relative to eleven year old score. They planned later leaving ages, on average, than pupils in either comprehensives or secondary moderns. Few grammar pupils of sixteen years old were planning to leave school shortly ('early' leavers); there was no school type difference in propensity to hope for day release or part-time study with a job.

Secondary modern pupils

Pupils of secondary moderns scored relatively lowest on mathematics and reading at sixteen, lower than equivalent comprehensive or grammar pupils. They had the lowest average school leaving age, and the lowest proportions wanting courses of advanced study. The sixteen year olds' opinions of their own performance averaged out, relative to their attainment by eleven years old, as similar in secondary moderns and comprehensives. (They were, if anything, a little higher in secondary moderns in self rated mathematics). Neither group average was at the relatively low level of grammar pupils. Secondary modern pupils showed no difference from other groups in proportions of early leavers wanting part-time study. But they seemed to make less progress than children who had started at the same level on academic criteria and had apparently lower academic orientation, according to measures like planes for advanced study and expected leaving age.

Comprehensive pupils

The average, after corrections, for comprehensive pupils, on several of the 'scholastic' measures was in the middle of these three groups. Their corrected averages were thus a little higher than those for secondary modern pupils on maths and reading at sixteen, in average school leaving age expected, and in plans for advanced study. Comprehensive pupils were similar to secondary modern pupils in self-ratings, and the three groups resembled each other in proportions of early leavers wanting part-time study.

Summary of school type comparisons (including material on affective, social and behavioural outcomes detailed in Appendix F).

The comparisons of types of school, above, seem to yield rather predictable observations on grammars and secondary modern pupils. The polarisation of academic and nonacademic pupils by selection at 11 plus is accentuated by looking at their separate attributes. The comprehensive pupils, it could be suggested, fell somwhere in between the two, on average.
Certainly, the patterns above showed that the 'school type' categorisation was related to some differences, on reading, maths, self-ratings, truancy, staying away from school, school leaving age, and wanting advanced study or a profession. There were no such significant differences in desire for day release, nor in pupils' emphasizing good pay or promotion opportunities as valuable in future jobs. No great difference in liking for school or parental satisfaction was found, nor in aspects of future jobs. So perhaps, it is not too fanciful to suggest that there were school type differences on the more 'school related' outcomes (tests, self-ratings), and perhaps too on indicators of 'middle class' aspirations (advanced study, professions). It was less obvious, if there were school differences, in 'outcomes' on which it is easier to imagine home background being a most dominant influence, such as plans for jobs which required less academic background.

However, having seen those differences and similarities which did exist between the three types of school, readers will want to know whether the comprehensive averages differed from the averages for the 'selective' group which was, in theory, the 'alternative' that is, the grammar+modern combination.

Comparisons of grammar+modern combination with comprehensive pupils

The inequality in intakes between comprehensives and the grammar+modern combination, in characteristics of the eleven year olds who were to go to each type of school, highlighted the fact that the comparison was not one between two systems. No corrections for those inequalities could create such a comparison. At the very least, however, differences or similarities of sixteen-year-olds could reasonably be hypothesized as attributable to secondary schooling if they took into account differences which existed before the start of secondary school. There was no simple story, in the selective/nonselective comparisons, either of the comprehensive group as a whole being 'superior', or of the grammar+modern group being 'superior'. By many of the indices of school attainment the two groups were similar. Sometimes differences existed between subgroups of comprehensives and the grammar+modern mixture, after the usual corrections for attainment before secondary school, social class and so on.

'Scholastic' outcomes

By the test score measures of attainments at sixteen years old, the grammar+modern and comprehensive groups did not differ, among those of given earlier attainment and background influences. The corrected mean reading score at 16, relative to reading at 11, was similar for two groups, and there was no statistically significant difference. The average mathematics test score, when corrected for the contribution of mathematics and reading skills and 'general ability' at eleven, did not differ either*. There was no overall difference of a statistically significant kind between the grammar+modern and comprehensive pupils in self-ratings, nor in plans for advanced study. There was a slightly higher overall leaving age in comprehensives.

Summary of grammar+modern v. comprehensive comparisons on affective, social and behavioural outcomes detailed in Appendix F.

In view of the overall lack of difference on most outcome measures, it appeared that, broadly speaking, comprehensive pupils were doing no better and no worse on progress. This did not seem to be reflected in measures of truancy or parents' satisfaction, on which comprehensives appeared 'worse'. It must be borne in mind that 1974 was a special year, and that reorganization and raising of the school leaving age may have sensitized teachers and parents, while actual progress among those already reorganized in comprehensives was adequate.

Perhaps parents expected better of comprehensives, rather than that they would match the old system. But the explanation of parents' satisfaction with school might be less interesting, and simply a function of the measure used and its exclusion of those whose parents were 'satisfied in some ways but not in others'. It seemed possible that measures which were chiefly indicators of low attainers - truancy, disturbed behaviour - were showing 'worse' averages in comprehensives, while measures which were reflections of the whole range

* The lack of difference was as follows: schooling difference
grammar+modern - comprehensive .03
(χ^2 = 1.6, df = 1, ns.)
The figures in 'Progress in Secondary Schools' are without the correction for reading skill and general ability at eleven.

showed less difference. To explore this was the purpose of the later sections of this chapter.

Different groups of pupils.

The preceding observations were based on averages for the entire school groups - all comprehensive pupils, all grammar pupils, all secondary modern or all grammar+modern pupils. Were the findings equally true for low and high attaining pupils? Did they seem to hold regardless of pupils' attainment at eleven? A number of interactions were tested, chiefly in analyses where the 'outcomes' were proportions or test scores at 16. By no means all were tested and few were significant, so there are relatively few remarks that stand as qualifications to the previous observations on averages for the three types of school or the grammar+modern v. comprehensive comparison.

Those of various attainments at eleven years old:

Type of school

Interactions were tested to see whether any similarities or differences between school groups on scholastic outcome measures seemed to vary according to earlier attainment (test score) at eleven years old. These interactions were tested and found nonsignificant, in the school type analyses, when scholastic outcomes were measured by:

self ratings in mathematics
self ratings in English
proportions wanting part-time study after leaving at 16
proportions wanting advanced course of study.

So there was no reason to qualify the overall observations above, on the association between these aspects and school type. No such interaction was tested, moreover, in the analysis of intended school leaving age. So there was no opportunity there to inspect whether the leaving age differences were among high attainers or low attainers. There were significant interactions with test score at eleven years old in the findings for tested attainments at sixteen (both mathematics and reading test score at sixteen). Thus, there were only two interactions known to be significant, among these scholastic outcomes. The picture they offer is a partial one, therefore, and any conclusion about types of school must be put in the context of the overall averages for grammar, secondary modern and comprehensive pupils.

The interactions in test scores at sixteen showed very interesting patterns which, if they indicate anything of influences of schooling rather than statistical artefacts, have noteworthy implications. These apparant variations in relations between school types indicated that, over most of the below average range, comprehensive and secondary modern pupils resembled each other in test scores at sixteen as on the other measures. Among average pupils they still resembled each other in progress by sixteen, while the few pupils who had average test scores at 11 and got into grammar schools had made more progress than the majority of children.*

* It will be found by reference to Appendix F that, for average children, the pattern for truancy or parents' satisfaction did not resemble this. There was no evidence here suggestive of common underlying factors in school-related patterns of attainment, truancy and parents' satisfaction.

Above that range of eleven year old attainment, comparisons include pupils who, if those tests had been used as an eleven-plus examination, might have gone to grammar schools. Attainments on tests at sixteen suggested that the ones who went, instead, to secondary moderns had made less progress than they would have done in either comprehensives or grammar schools. Pupils in the next-to-top fifth on scores at eleven may have made slightly more progress still in grammars than in comprehensives, on these tests. But, among those in the top fifth on either maths or reading at eleven, those who went to grammar schools did not appear to make statistically significantly more progress in either mathematics or reading tests by sixteen than equivalent pupils who went to comprehensive schools. Secondary modern pupils of equivalent scores at eleven seemed to make less progress.

The conclusions above have provoked a good deal of attention, and the qualifications which surround them need to be carefully considered. The implication is the 'comprehensive' pupils, in this era of partial comprehensivisation, were doing no worse, and no better, than they would have done under selection. The discussion of this and of the observations of critics of the earlier report deserves a certain amount of space, but the discussion will be put in abeyance to consider whether the other grouping, with grammars and moderns grouped together and compared with comprehensives, also showed interactions.

Grammar+modern vs. comprehensive: no interactions with test score

Neither test at sixteen showed an interaction between test score at eleven years old and the grammar+modern vs. comprehensive grouping.* So the implication that comprehensive pupils were doing much as they would in the selective schools, in mathematics and reading by 16, did not seem to need qualifying. The lack of difference was consistent for low, middle and high attainment eleven year olds. There were no other significant interactions found with test score, for this grouping of the schools. The interaction of the grammar+modern vs. comprehensive variable with eleven year score was tested and found nonsignificant for the following scholastic 'outcomes':
self-ratings in mathematics and in English; likely school leaving age; proportions wanting part-time study after leaving school at sixteen; and proportions wanting advanced study.

Comparison of different dates of comprehensives

Before going on to discuss the conclusions of this stage of the study, a few observations can briefly be made on the results of comparing pupils on scholastic outcomes, in comprehensives of different dates, and transitional pupils. Those in early comprehensives will be described, in so far as they differed from those in recent comprehensives, and from the grammar+modern group,

* In the analyses of mathematics at sixteen with allowance for mathematics, reading and general ability at eleven, it was not possible to explore an interaction with any one of the eleven year test scores, as the picture for one test score would be relative to two other scores. Further research could establish the relation of maths test score at 16 to the combination of all three tests at eleven, in each of the two school groups.

each of which is given attention. The section ends with observations on the results for transitional pupils, whose schools changed over to being comprehensives during their secondary school years.

'Non-scholastic' outcomes will be commented upon in Appendix F.

Early comprehensives

The minority of comprehensive pupils whose comprehensives were set up by 1965 had started secondary school at a relative disadvantage, even by comparison with other comprehensive pupils, in the sense that, on average, their test scores at eleven years old were lower, and closely similar to those of pupils in secondary moderns.

Despite this, the early comprehensive pupils did not produce a different corrected average on the reading test by sixteen. They had, however, apparently made less progress in mathematics test attainments, corrected for maths at eleven; at least, the averages for those early comprehensives which were either purpose built or ex-grammar comprehensives were lower. This did not seem related to pupils' ratings of their own performance in mathematics, which did not differ, but they had the highest average self-ratings in English, relative to their reading at eleven.

The early comprehensive pupils planned the latest leaving age, on average, and were relatively likely to plan a course of advanced study later.

Recent comprehensives

Those in comprehensives set up between 1966 and 1969, a few years before they got there, were slightly higher attaining on intake than early comprehensive pupils, though still not very different from secondary modern pupils. They did not differ markedly on progress in test attainments, apparently, from those in the grammar+modern combination, and, like them, they had slightly lower self-ratings in English than early comprehensive pupils. They did not differ on reading progress or self-ratings in maths, from other groups. The proportion wanting advanced study was about average and they were in the middle, above grammar+moderns and transitionals, on school leaving age.

The grammar+modern group relative to others

The combination of grammars and secondary moderns had higher scores and a higher proportion from 'nonmanual' social classes at eleven than comprehensives. Although similar overall to comprehensive pupils in progress by 16 on tests, their average tended to resemble that of the higher scoring comprehensive sub-groups. On mathematics, relative to maths at eleven, they appeared to fare better than the pupils of early comprehensives (though not better than those in early ex-secondary moderns.) The grammar+modern pupils rated themselves along with pupils of recent comprehensives, slightly lower in English than did early comprehensive pupils, but there was no difference in maths self-ratings. As a group, they anticipated leaving school somewhat earlier, on average, than either comprehensive group. They tended to have slightly lower proportions intending an advanced course of study, on average, than any subgroup of comprehensives, though not as low as the transitional pupils.

Transitional pupils

The schools going comprehensive during this time tended to have slightly higher attaining eleven year olds, slightly more of whom were of 'nonmanual' social class, than comprehensives did. In terms of progress on tests by sixteen, they were not different, apparently, from recent comprehensives at least. That is, they had a similar average, corrected for earlier mathematics attainment, on the maths test at sixteen, to those in recent comprehensives, slightly above that for early comprehensives, slightly below that for the grammar+modern combination.* On reading progress, the school groups did not appear to differ, nor did self-ratings in either subject show differences. Transitional pupils, on average relative to their 'ability' at eleven; they resembled grammar+modern pupils in this, but they also had the lowest proportions intending to take a course of advanced study.

Interactions between school and parental interest or social class

The possible contributions of parents' interest and class to truancy rates are discussed in Appendix F.

A statistically significant interaction in plans for advanced study suggested that grammar+modern pupils who lacked an 'interested' parent were particularly less likely than any comprehensive subgroup (especially than the ex-grammar comprehensive pupils)to wish for advanced courses of study. Among pupils with a very interested parent, who were, as might be guessed, relatively likely to plan an advanced course of study, school differences were less striking (though it did appear that amalgamation and ex-grammar comprehensive pupils might be a little likelier to plan courses of advanced study than the grammar+modern combination). Other outcomes of schooling showing interactions between some school variables and social class were mentioned in the previous publication. The overall slightly later leaving age predicted by comprehensive pupils, relative to general ability before secondary school, was a function of the relatively late leaving planned by comprehensive pupils of III nonmanual homes, relative to grammar+modern pupils of equivalent social class (and also, partly, of a less marked difference in the same direction among manual workers' children). Among the pupils of social classes I and II, the leaving age planned in the grammar+modern group was fairly like that in comprehensives, though a little later. This was outweighed in the overall comparison by the results for pupils of less high social class.

Social class differences were also found in the relation of job aspirations and parents' satisfaction to schooling. These are discussed in Appendix F.

Summary and discussion of findings of earlier stage

Progress was measured by comparing, after correction for preexisting differences between the groups on intake, several attributes of the sixteen year old pupils relevant to the findings to be presented in Chapter IV on public examination results. Perhaps most relevant were the 'scholastic' measures, and most of the

* The maths progress in purpose-built transitional comprehensives (n = 134) was relatively good; transitional ex-secondary modern pupils (many of whom were at eleven, presumably, sent to secondary moderns) made less progress in maths tests.

discussion focusses on these: attainment in mathematics and reading comprehension, as assessed by tests at sixteen, self-rated performance in maths and English, plans for leaving school, and plans for further, part-time or higher, advanced courses of education. The other, nonacademic measures are detailed and discussed in Appendix F, but are also mentioned here when relevant. This summary will first pick out some principal findings from the material in Chapter IV; then go on to discuss some characteristics of the measuring instruments, and some criticisms which have been made of the interpretations of that stage of the study. Third and lastly, the account will draw out some of the main implications of the results.

Grammar school pupils and secondary modern pupils differed in progress, beyond the marked differences already seen at eleven years old in the average attainments and social class of those two groups. Grammar school pupils were relatively much more 'successful' academically, in mathematics and reading test progress, in plans to stay on at school, and plans for advanced study.

Secondary modern school pupils, even when attempts were made to compensate, statistically, for their earlier less successful academic standing, made less progress on tests, planned to leave school earlier and were less hopeful of advanced courses of study.

Comprehensive pupils were neither worse off, as a whole, than secondary modern pupils nor better off than grammar school pupils. They were very slightly more 'successful', academically, than secondary modern pupils, but both groups were less academically successful, on average, than the grammar pupils.

There was a departure from this in the 'subjective' measure of academic success, the pupils' own ratings of their performance in maths and English, in which secondary modern and comprehensive pupils both had relatively high ratings, given their eleven year old scores.

It is not adequate, though, as an appraisal of school progress, to compare the three groups of pupils. Account has to be taken of the numbers of pupils who were in each type of school. The majority of pupils, under the old selective system and in the combination of grammar and secondary modern schools here, were in secondary moderns. It has already been shown (Chapter I) that the comprehensive pupils were close to resembling secondary modern pupils on average, at intake. This was not simply because those groups were similar in social class composition, but also because of relatively high proportion (23 per cent of those in the three school types) still going to grammars, and of the effect that had on the average attainments of pupils who went to comprehensives and to secondary moderns. With that proportion remaining in grammars, it would have been as interesting (if less satisfactory to those with policy concerns) to compare progress in comprehensives with progress in secondary moderns. That would have shown marginally better test attainment, slightly higher proportions wanting advanced study and a later leaving age in comprehensives, but virtually identically self-ratings, as well as similar plans for part-time study among early leavers in the two groups. (In addition the two groups showed similar truancy levels and plans for working life, though comprehensive pupils had less liking for school and lower levels of 'parents' satisfaction' with school).

More relevant to the concerns of those who would find fuel for policy here is the fact that overall averages were sometimes found to need qualifying because of interactions. The position of the school type groups varied with the level of attainment before secondary school of pupils. A great deal of discussion of these results has focussed on the more informative pictures of differences between pupils in test score progress in the three types of school (grammars, secondary moderns and comprehensives) based on such interactions. These show that the size of these differences was dependent on the sort of children being compared, that is, on whether they had been of below average attainments at eleven years old, in the middle of five groups on attainment, among the next to top group of eleven year olds, or scoring in the top 20 per cent by eleven. Those who were below average, in comprehensives and secondary moderns, performed similarly on tests of attainment by sixteen. The ones in the top 20 per cent who went to comprehensives performed better than the ones of that high attainment at 11 who went to secondary moderns. Among these 'bright' children, in fact, the ones who went to comprehensives were as high scoring on attainment tests at 16 as equivalent grammar school pupils, after allowing for social class differences and so on. Grammar school pupils scoring lower than the top 20 per cent at eleven were a minority of grammar pupils; relatively few grammar pupils had scored in the middle group at eleven and none to speak of had scored below average. Those few who did made more progress in tested attainments by 16 than their equivalents among either secondary modern or comprehensive pupils. Comprehensive pupils who had above average scores at 11 but were not in the top fifth appeared to have progressed to a level between that of grammar and that of secondary modern pupils in these test attainments by the age of sixteen.

For an understanding of the implications it is also necessary, taking account of the fact that a majority of 'selective' pupils were in secondary moderns, to combine the grammar and secondary modern data and compare their average with the comprehensive pupils.

Summing up a number of observations, it was possible to state that pupils of comprehensive schools and the combination of grammar and secondary modern pupils did not differ as a whole in a clearcut fashion on test score measures of attainment. The mathematics test results shown in the 1980 publication indicated that, with allowance for preexisting mathematical ability, but not for reading skill or general ability before secondary school, the average maths test score for comprehensive pupils was just statistically significantly lower than the grammar+modern average. When the comprehensive average was looked at more closely but still without such added control for reading skills and 'general ability' at eleven, it was the pupils of comprehensives set up by 1965 (early comprehensives), rather than those in recent comprehensives set up between 1966 and 1969, who seemed to make less progress in mathematics than the combined average for grammars and secondary moderns. It was possible in presenting those findings to suggest that this might reflect other aspects of pupils' capabilities before secondary schools. This was because analyses carried out before the publication had showed that, once all three measures of 11 year skills (that is, mathematics, reading and 'general ability') were taken into account, there was no overall difference at all in maths test attainment at sixteen between grammar+modern and comprehensive sixteen-year-olds. These confirmatory averages were available to the research team but unpublished at the time of the first publication and this was part of the background to the conclusions.

On most other measures, the grammar+modern and comprehensive averages were similar. The 'scholastic' measures showed no overall difference, apart from a slightly later leaving age planned in comprehensives. The overall conclusions to be drawn are based on these impressions, but they also depend very much on the subgroups chosen - where the comparisons are drawn - and on the measure used. It was noted, for instance, that the maths progress in early comprehensives was, relative to maths at 11, less than in recent comprehensives. Pupils in early comprehensives, however, might be thought of as more 'academic' (if that is the correct interpretation) in the light of their intended school leaving age. Pupils of recent comprehensives and those whose schools were in transition to becoming comprehensive were similar to the grammar+modern group on both maths and reading test measures by sixteen years old. Transitional pupils had been on average somewhat higher scoring on intake than comprehensive pupils but not up to the level of the grammar+modern combination. They were similar to those whose schools carried on uninterruptedly as grammars or secondary moderns by 16, as regards their wish to leave school on average a little earlier than pupils in early comprehensives (and in being somewhat less likely to truant than comprehensive pupils). It is possible to surmise that their slightly earlier leaving age was an effect of the disruption of school organisation, but this must remain speculative at this stage, in the absence of other findings of that sort. However, transitional pupils also had the lowest proportion wanting courses of advanced study.

(iii) Speculations, criticisms and reservations concerning the measures used.

It could be valuable at this stage to draw attention to a few issues that surround the interpretation of findings presented in this Chapter. Some of the issues have been raised or given a new emphasis by criticisms, since the original publication, from a variety of sources. Where it may be helpful to do so, the criticisms will be analysed and explanations offered. In some cases the need for clarification of the original presentation has been highlighted by these criticisms, whether they are based on misunderstandings or justified. In an attempt also to provide a better context for discussion of the implications of these findings, and for the presentation in Chapter IV of examination results, a little of the relevant work of other researchers will be mentioned in this section.

Test scores

The major advantage of test scores is that they are interpretable. The two tests used at sixteen have been discussed elsewhere,* and copies are in Appendix X. It is worth noting here, nevertheless, certain aspects of the reading and mathematics test results.

The tests were administered to everyone in the study so they apply across the board. Particularly when findings are matters of relative performance, it is important that the different groups of pupils related to each other are all pupils to whom the measure applies. In addition, one of the assumptions of the statistical techniques used to relate performance at sixteen to performance at eleven, as was pointed out in Chapter I, is that sixteen year olds' test scores are well distributed, approximating a normal distribution. Moreover,

* Steedman, Fogelman & Hutchison (1980); Goldstein (1981).

'transformations' of the original scores were used, as is conventional in such work, to make the distributions of scores acceptable for analysis (instead of scores running from 0 to 35 or 0 to 40, they ran from about - 3 to about + 3) with a mean of 0 and a standard deviation of 1, creating a valid measure for the statistical technique used. The tests were therefore reasonable indicators, in a statistical sense, of what they appeared to indicate.

The tests here were thus more telling and more acceptable statistically than some of the measures of examination performance to be described in Chapter III.

Attention was drawn in the 1980 publication to a potential source of distortion in results on the reading test at sixteen. If a reading test does not contain any items too difficult for the very best readers at sixteen, those very best readers are not differentiated, as they are unable to get higher marks than (about) 100 per cent correct. If this had happened the very best readers would be at the optimal, 'ceiling' level of performance as indicated by that test and it would not be possible to show differential progress between, say, the very best in one school and the even better in another. If that had been the case among the top ten per cent of pupils in this study, there would have been no way of telling whether the similarity between grammar and comprehensive pupils was a true similarity or a function of the measuring instrument. However, though this comment was taken by a few critics as casting doubt on the conclusion concerning reading progress, there are several reasons why it is unlikely to have applied here. For one thing, the slope of the progress curves at the top end of sixteen (and eleven) year reading score was quite marked, so scores were still increasing among the high scoring pupils. The picture* for those who had been in the top fifth at eleven, looked at separately, moreover, showed increasing reading score among sixteen-year-olds whose scores at eleven were higher than any comprehensive or secondary modern pupils - they were all in grammar schools - and these had been the very best readers at eleven. There did not seem to be a crude or obvious ceiling effect stopping them altogether from 'improving'. It seems, therefore, that, in order to affect the comparison between grammars and comprehensives, any ceiling would have had to be lower. In any case, the crude distributions of sixteen year reading scores of pupils who had been in the top twenty per cent of reading comprehension score at eleven have a fair spread; if there were a marked ceiling effect, these scores would be expected to be crammed up against the ceiling.**

It is worth mentioning, as a mater of interest, that the overall level of scores might be affected by practice, in that the reading comprehension test used at sixteen was the same test as the pupils had taken five years earlier, at eleven. Such a practice effect, though, would apply to everyone in the sample, and no one can say whether it might enhance the results of the average pupils, or those of the poorest readers, for instance, more than others.

Because the possibility of a ceiling effect received much attention and little careful analysis by critics of what such a ceiling effect would do if it

* These were reliability-corrected regressions; see page 104 of the 1980 publication.

** See Steedman, Fogelman and Hutchison, 1980.

existed, several readers of the earlier publication have concluded that more weight should be attached to results on the mathematics test. The mathematics test at sixteen was different from that used at eleven, and it is clear that at least some of the mathematical skills to be assessed at those two ages will differ.*

Cox and Marks (1980) stated that the mathematics test used for the sixteen year olds was bound to be inappropriate because its original purpose was to be used in a survey of fifteen year old pupils. Those authors, of course, were not as familiar as the research team with the distributions of scores on the test. Their criticism will be seen to be misplaced by a look at the relevant distributions (some of which were set out in Steedman et al., 1980). If greater weight is attached to mathematics test results, though, it may now be clearer that their overall implications were broadly similar to those of results in reading, as they showed no significant difference in progress between the grammar+modern and comprehensive groups once allowance was made for the contribution to mathematics performance at sixteen of reading and general ability at eleven.

This raises a general point, which is that the findings depend critically on what 'extraneous' factors are incorporated in the account. For example, the mean maths test score for pupils in comprehensives was seen to be higher than that of equivalent secondary modern pupils, among those whose scores at eleven were below the top fifth but above average. The observation was qualified by a supplementary (hitherto unpublished) analysis, which allowed (in addition to normal controls) for the interest in schooling shown by parents as reported by the primary school teacher. When this was included, even though parental interest at sixteen was also allowed for, there was a significant contribution to explaining results. Allowing for the level of interest perceived before secondary school appeared to raise to the level of equivalent comprehensive pupils the performance of secondary modern pupils from the next-to-top group at 11 years old. (See Chapter II here, and page 76 of 'Progress in Secondary Schools', for the reasons for this choice of variable.) There undoubtedly will be many influences, extraneous to schools, on 'outcomes' such as those explored here, which are not allowed for in the analyses. It seems likely that such influences might favour the more favoured intakes, and be a source of disadvantages to disadvantaged intakes.

Despite the preceding remarks, exploring some actual or potential drawbacks to the tests, it must be repeated that these tests were superior as measures of progress in attainments to examination measures. Test scores obtained by sixteen-year-olds could be corrected for comparable test scores at eleven; examinations could be corrected for test attainment, something quite different, at eleven. Above all, tests applied to everyone, whatever their ability, while there is no examination (or even collection of examinations) that is taken by all people in this (or any) sample.

* One possible minor drawback is that to use a different mathematics test at sixteen from that used at eleven as a covariate may have statistical disadvantages, compared to the reading test analyses.

None of these comments materially affects the main point, which is that some allowance could be made for pre-existing attainment in a way that has not been possible except with longitudinal data. That applies to all the 'outcomes' studied in the course of the investigation.

Self ratings

The exploration of self ratings at academic subjects was inspired by the question of whether evidence could be found on the question of how student self images might be affected by 11-plus-failure and rejection. The overall findings of no difference between grammar+modern and comprehensive pupils fits with attainment tests results, overall. The relation to type of school may be more challenging to assumptions, though, and it has to be remembered that findings are all relative to 11 year old attainment, after allowance for social class.

The ratings pupils used were deliberately related to their eleven year old performance, as the intention was to see whether self-ratings showed the same patterns as 'objective' tests at sixteen, relative to the same attainment tests at eleven. This design has been criticised (Gray, 1981b; Cox and Marks, 1980) for not relating self-ratings to existing sixteen-year-old levels of performance. To do that would be to see whether, for example, one type of school was associated with lower or higher self-images as if those were independent of current attainments. While that would be of interest (if rather hard to interpret) and could be tackled by another researcher, it was not the question approached here. This analysis was designed to shed light on how self-images might develop during secondary years, among children who had been in some sense objectively similar at eleven (by test scores). It is, of course, open to any reader to object that such delicate psychological traits are hardly revealed by the crude measure used here.

Though it is superficially surprising that secondary modern and comprehensive pupils had high self-ratings and grammar pupils lower self-ratings, it supports an observation of Soares and Soares (1969), who found that 'disadvantaged' students in their neighbourhood schools had, if anything, more favourable self concepts than advantaged students from schools in more prosperous middle class areas.

It seems very likely that pupils asked to rate themselves relative to 'people of their own age' will base their self perceptions on the peer group they know in their own school. So those in schools with lower attainers will rate themselves 'high' and with higher attainers, 'low'. Comprehensives had about the same 'peers' for this purpose, presumably, as secondary moderns. But there may also be a human tendency for people to rate themselves more average than they 'really' are.

School leaving age

Findings concerning school leaving age have occupied a central place in the history of debate and research about education and particularly about selective schooling. Their implications have altered somewhat in light of the raising of the school leaving age, by which this year group was the first to be affected. 1974 was a peculiar year in which to ask about school leaving age, therefore,

and comparisons with other studies are slightly hampered. In looking at whether pupils planned to leave at 16, or stay on, it could not be judged who, in earlier years, would have left at fifteen. It was not obvious, for example, whether any of the sample were comparable with the pupils who in earlier years were 'staying on' till sixteen.

Moreover, the measures here were of pupils' intentions. A few sixteen-year-olds who were expecting to leave at the end of the year might stay on or return to school as a consequence of exam results or unemployment; some who planned to stay on might decide to, or have to leave. There were not, then, actual leaving ages.

It has emerged since then that the proportions intending to leave at sixteen were a little lower than the actual proportions who left school at sixteen. The figures for NCDS members aged 23* indicate that the proportions who stayed on at school, while similar to the proportions predicted by the sixteen year olds, were a little lower.

Table 6 School leaving age of NCDS members

	Intended	Actual
16 years old	62%	71%
17 years old	7%	9%
18 or older	23%	18%

In the data here on intended leaving ages of sixteen-year-olds, the highest average leaving age was found among grammar school pupils, but the average for comprehensive pupils exceeded that for equivalents in the grammar+modern combination. While one could predict the later staying on at school in grammars from the higher progress in attainments in grammars, the overall slightly lower average for the selective combination was less easily explicable. Future research will be able to determine whether actual leaving ages from different groups of schools were constant for various levels of attainment or examinations at sixteen years old.

Some studies on the relation between leaving age and school have been unable to allow for class differences.** It was possible in this study to give results after allowance for social class, removing this confusion. There was in fact a suggestion in these data that the comprehensive pupils who exceeded the grammar+modern combination in plans to stay on were those of social classes III, IV and V, rather than those of higher social class background.

* The figures for actual leaving age are based on unedited data. Eight per cent of the sixteen year olds were uncertain of their intentions and 2 per cent of the 23 year olds gave uncertain answers to this.

** The correlation, for instance, between the percentage in each area allocated to grammar schools and the proportion still at school at 17, indicated by the Robbins Report (1963) to be as high in 1961 as 0.82, was described by that report as attributable to social class differences.

In the NCDS data, moreover, it seemed to be the pupils of early comprehensives, particularly, whose intended average leaving age was a little later. This is not in conflict with the observation of the Oxford Social Mobility Project on pupils of comprehensives at an even earlier stage, that 'the comprehensive schools were somewhat better than the overall average in holding children beyond 16' (Halsey et al., 1980). In that study, though, comprehensive pupils tended to have fewer 'O' levels. A slightly higher proportion staying on later at school could be an indication of higher motivation to study or gain qualifications, but it could also be a result of a greater tendency to defer taking 'O' levels until the sixth form. The question of whether comprehensive pupils in the NCDS reached sixteen with fewer such qualifications is explored in Chapter IV.

One feature of comprehensives during the early 1970's would have seemed, contrary to the implications of these findings, likely to lead to earlier rather than later leaving. The necessity to make transfer arrangements, to schools with sixth forms, which affected some comprehensives, might have been thought a deterrent to sixth form work. The hypothesis, though, that the availability of a sixth form in the same school might make a difference was, if anything, supported by the finding for early comprehensives.

Plans for advanced study

Grammar pupils, as might be expected from their academic progress generally, had the highest proportions planning advanced courses of study in colleges, Polytechnics or Universities.

One finding which was less predictable from attainments was that transitional pupils seemed somewhat less likely to plan advanced study (and pupils in comprehensives formed from a grammar school were somewhat more likely to do so).

Plans for part-time study and occupations

No differences were revealed in plans for part-time study among early leavers which were interpretable as school influences. Perhaps the determinants of early leaving and the availability of day-release opportunities, as well as attitudes to part-time study and occupational choice, were affected more by economic factors and home background.

'Non academic' outcomes'

The observations on truancy and staying away from school indicate worse ratings of absence from comprehensives than from other schools. A difficulty in interpretation arises on study of the interactions, as the differences seem to lie in the ratings given to high ability or high social class pupils. Among those who might be expected to truant more often, there were no such noticeable school differences.

The particular circumstances of the raising of the school leaving age have to be borne in mind, as well.

The preceding section has tackled a few of the issues and problems, and some of the criticisms of this study since its earlier publication. To repeat what was underlined in Chapter I, it is only possible to look at these results as evidence of how equivalent children might fare under each type of schooling with a number of cautionary reservations and provisos in mind. The assumptions of the study and the limits to what can be measured have to be understood, as does the necessarily relative nature of the comparisons.

(iv) Comment

The remainder of this chapter is a brief comment on the possible implications of this stage of the study. The findings for each type of school could be regarded as predictable functions of the existence of three types of school. The academic outcomes, at least, tended to suggest good progress in grammars and worse progress in secondary moderns. Subjective measures, like self ratings and liking for school, did not reveal the depressed 'failure' symptoms that can be ascribed to secondary moderns, but progress in performance terms did point to 'worse' secondary moderns. Comprehensive pupils tended to be a little better off than equivalent secondary modern pupils. On balance, though, it was not possible to say that progress in comprehensives was worse or better than would have been predicted from the findings on the members of the sample who went to grammars and secondary moderns. This observation is the nearest that one can get to a comparison of progress under a selective system and progress in comprehensives. It is not such a comparison, but is closer to a reasonable comparison than some others that have been made elsewhere.

The conclusion is based on the mixed collection of findings summarised earlier. When, in addition to compensating statistically for some of the misleading components in comparing schools with such different intakes, the study also took the necessary step of comparing comprehensive pupils with the combination of secondary modern and grammar pupils to represent a selective system, findings balanced out. On most measures, the two groups were equal, on a few comprehensives were worse and, on one, comprehensives were, if anything, 'better'. There are some qualifications to those tallies, in light of interactions and the relation of the averages for subgroups of comprehensives to the overall average. For instance, in early comprehensives the leaving age was later, and this was balanced by the possibility that the early comprehensive pupils made less maths progress. Another example of a finding needing careful interpretation was the variation of average test score at sixteen for each school type according to test score at eleven. This interaction does have salience for any discussion of selection, in that the findings for those in the 'top fifth' at eleven appeared to challenge assumptions not only of the merits of the selective system but also of how the continued existence of grammar schools might affect the performance of those in other schools. They also highlight how the continued existence of grammar schools affected the intake of other schools. The study was in part exploring whether the potential for wider availability of high academic performance at sixteen did exist under the 'mixed economy' of selective and 'comprehensive' schools explored. If anything, test results for reading and maths did imply a larger proportion reaching 'grammar school' levels of performance than would have passed the eleven plus examination. There was, too, a suggestion of 'wasted talent' in the lower performance of secondary modern pupils who had been of high attainment at eleven.

The concentration on the attainments of this top-scoring minority to the exclusion of the majority of the sample exemplifies the debate over selection, but the objections to comprehensives have often focussed on the supposed disadvantages to children who would have been selected under the old system.

There are, however, some limits on the application of the findings to the debate on selection. The major limitation, which cannot be over-stressed, is that the pupils in 'comprehensives' were not representing a fully comprehensive system, and that comprehensives coexisted with grammar schools which selected the majority of the most academic pupils in the sample. That gave rise to a second limitation on the study's relevance to the discussion of selection. Comparisons between school types could not be made among the very top scorers at eleven, as the 'top one or two per cent' of this sample all went to grammars. For this reason, test score results for the highest scoring pupils who were compared have always been presented as 'for the range within which comparisons can be made'. It is within that range that it was impossible to find evidence of pupils in comprehensives in the top twenty per cent of scores at eleven 'doing worse' than equivalent grammar school pupils.

It is partly because the top one or two per cent of this sample were all in grammar schools that the finding was not precisely relevant to discussions concerning the Assisted Places Scheme. The study is less informative for any evaluation of that scheme, more importantly, because these pupils were in secondary schools at a time when direct grant schools continued to exist, and to take a proportion of pupils at least as large as that for the Assisted Places Scheme. (The argument over selection, however, is seldom so precise as to whether two or ten or even twenty per cent are to be selected, and, as these findings were seen as informative about selection, they were of some relevance to many issues involved in the Assisted Places Scheme).

Another reservation about the implications or about misinterpretations of the implications of the study, should be mentioned. The study has been taken to mean that bright children will do well wherever they are. This is not one of its implications, but it would fit with a model of the rationale for selection. For, if there is no change after age eleven, separate schools for 'non-academic' and 'academic' might be appropriate. It would also fit with the idea that secondary schools cannot make a difference and that all progress is attributable to father's occupation, or genes, or primary school. The results here suggested, if anything, that 'bright children' did not do well regardless; those from the top-most group who went to secondary moderns did worse in test scores than those in comprehensives or grammars.

Chapter II has provided, it is hoped, a summary and discussion of the earlier stage of the study, which may aid interpretation of the rather detailed and technical material on the same pupils' examination results which is to come in Chapter IV. Before that material can be understood, however, it is necessary to describe the particular problems which arise in measuring the relative progress of different school groups by means of examinations, and the strategies adopted to overcome those problems.

Chapter III

Measuring school progress with examination results: problems and strategies

(i) Introduction

The previous chapter discussed the findings published in 1980 concerning the progress of pupils in the various kinds of secondary school studied here, progress being measured by indices other than public examination results. In that discussion, many assumptions about the properties of the measures used, and the extent to which they might or might not be indicators of what they appear to show, went unchallenged. A few were tackled in the discussion at the end of the Chapter. It is now necessary to consider particular problems associated with the design of measures. Some are problems common to exam results and to outcome variables used in the earlier stage. In particular, before the public examination results for those in different kinds of school are presented, in Chapter IV, the special difficulties in choosing what to measure of public examinations must be pointed out.

This Chapter sets out some of the relevant considerations in designing measures of secondary school progress, particularly when those are to be used with allowance for prior attainments. A further aim is to delineate certain particular drawbacks of public examination results as measures and demonstrate some pitfalls in their use as national (or even local) statistics. It becomes necessary to go into quite considerable technical detail (further elaborated in the various Appendices signalled throughout) concerning the exact nature of the measures used to study examinations as outcomes and the ways in which results are corrected for attainments before secondary school. It has to be said, however, that no study of public examination results of this kind could ever be complete, even in its own terms. This study can only offer a series of explorations, and perhaps examples, of how research in this area might be done and make clear the limits to the investigation. These limitations compound the difficulties created by the history of selection and developments in public examinations which are not independent. Whether or not the study can be revealing is a question best left until after Chapter IV, in which the findings on examination results will emerge.

Examinations as measures of groups rather than individuals: a few problems.

Some of the measures designed for the study of examinations were less adequate measures of progress or performance in secondary school than the test scores obtained for each member of the sample at sixteen years old. Test score measures had not raised the same problems and need for definitions as the examination measures did. Certain criticisms below of examination measures could equally be made of some of the other school 'outcomes' explored in the earlier stage of the study.

Examination results may have a greater superficial plausibility than test scores. Because they are widely-recognised, public criteria, rather than an unknown test, they appear to be more important indicators, or to signal attributes of pupils that are more important than their scores on tests. They may even seem to have more 'objective' validity. We think we know what having four CSE grade ones means, for instance. However, examinations are designed to be revealing about individuals and are not as easily averaged as tests. As

soon as it becomes necessary to use aggregates or averages of examination results to measure groups, problems arise. Some of these difficulties will be sketched out to remind readers that it is not a simple matter to measure examinations; and then the strategies adopted will be detailed. One set of problems has to do with the nature of examinations. A further set arises from the nature of measures of examination performance. Last to be mentioned are problems arising from the fact that schools differ.

(ii) The examination system

In the English educational system there are different levels of examinations ('O' level/CSE, 'A' level) and it is a matter of reasoned choice which or how many to use as indicators. Within that system, too, there are different sorts of examination, even at equivalent levels. The two exams taken about age sixteen and studied here are General Certificate of Education (GCE) and Certificate of Secondary Education (CSE). There are formal rules for comparability of standards which allow some aggregation of the data, but how this is done will always be open to discussion. The details of how that difficulty was tackled in this work are given later in this Chapter.

In addition, problems arise as soon as one wants to compare individuals' results. Is one 'A' level with four 'O' levels 'better' than eight 'O' levels, for instance, and which would look 'better' on a measure of the proportions with five or more 'O' levels? The problems increase as soon as one wishes to compare large numbers of pupils, as here, and tries to devise measures that will be revealing about a wide range of children.

The first difficulty is that examinations measure only certain ranges of children. Those of various abilities take different types and levels of examinations, and some take none at all. The children, for instance, at the 'lower' end of the scale of academic performance who take no examinations or gain no graded result in GCE or CSE will not directly be measured by results. In theory, by examination design, these pupils would be the lowest 40 per cent, in any given subject, presumably. In the sample analysed here only 9 per cent took no exams at all, however. At the other extreme, there are great deficiencies in judging solely by 'A' level results (attempted, in the sample for this study, by only 14 per cent). The same reservations apply to all measures of examination performance; exams measure those who take them.

(iii) Measures of examination results.

The examination measures used here, moreover, were in some cases designed to apply to some people. In particular, performance in a single, specified subject revealed nothing about performance in any other subject. In investigating the grade obtained in English CSE, for example, it is necessary to exclude from analysis those without a grade at CSE. This may seem circular, but there could not be a measure which applied to everyone, because of the difficulty of assigning a value to the performance of someone without a graded CSE in English. Such a person might have done the course and failed to gain a grade in that CSE; it seemed unreasonable to put such a candidate on the same point on a continuum as someone who had no grade at CSE English because they did English 'O' level, or someone who got 8 CSE grade 1's but took no exam called English. It is necessary to recognise that such measures ignore the performance of those who

did not take the examinations concerned. Also, because measures (including measures of those who took no exams) recognisably apply or are likely to apply to particular ranges of ability there is an additional problem in interpreting results. Differences in overall figures, which reach statistical significance, may have applied to areas of the range of children, at least as defined by eleven year old scores, in which there should be few children (see section III xviii).

Interpretation of examination results is risky when each grouping of results or choice of measure obscures differences between those who gained different results within that group or by other indicators of exam performance. When this study considers the proportions with an 'O' level pass, for example, the question remains unanswered whether, for example, schooling types differed, as a whole, in the grades obtained within the range regarded as a 'pass' (in 1974) or its equivalent at CSE. It will be seen later that the measures used included the proportions obtaining five or more 'O' levels, since it was of interest to study people obtaining a high standard of a kind. But the same question was left unexplored, that is whether, within the proportion getting that level of qualifications, there were differences in the numbers producing higher numbers of very high grades between the various school groups. Were comprehensive pupils reaching this standard with great numbers of grade C's while grammar+modern pupils who were going in for, or who got any 'O' level equivalents were in fact performing better and getting grade A? Were some schools putting in pupils for eight 'O' levels and getting 'pass' grades, others putting in pupils for five 'O' levels and getting grades A? A pupil who did not reach the standard of five 'O' level equivalents, but nevertheless obtained grade A in four subjects at 'O' level, might be considered to have demonstrated academically superior performance to a pupil with five grade C's. The possibility of such detailed variations had to remain unexplored by this study and would better be explored with different measures of examinations, but it was possible to look at different attainment groups in case that threw light on the question above.

(iv) Differences within and between schools

It is a problem for research, though not for education, that there are many different aspects of schools, and that schools try to do a variety of things, both within the same school and as between schools. For example, pupils study different areas of curriculum and one result (and one of the causes) is that exams are taken in different subjects. The average grades obtained in one particular subject are to be understood as telling us nothing about performance in any other subject.

Then, too, schools aim both to equip pupils with qualifications and to educate them, keep them interested, and so on. So, in so far as examinations are relevant, it is necessary to look at measures of how many pupils reach what levels of qualification or certification. But it would be important, as well, to look at, for instance, whether schools tended to aim at a wide variety of subjects, or at providing examination opportunities for 'less able' pupils. The qualifications indicator is more appropriate to this sort of study but it can be borne in mind that variety in the curriculum, say, is a relevant matter outside the scope of such a study.

Because the range of application of exam results was limited, the different types of school would vary in the extent to which exams measured their pupils. Thus, when the sample was reduced to pupils obtaining a grade in maths 'O' level, for instance, 72 per cent of the grammar sample were included, while the sample contained only 17 per cent of comprehensive and 16 per cent of secondary modern pupils. This is a reflection of the continuing existence of selection, but it also meant a problem for analysis. The problem became more marked when pupils were categorised according to fifths of eleven year score. In any analysis involving the whole sample of those in the top fifth at eleven on maths, 68 per cent were in grammars, 20 per cent were in comprehensives and 12 per cent were in secondary moderns. (The equivalent figures for the top fifth on reading were 64, 22 and 14 per cent). Not only do types of school vary, but also examination entries differ from one school to another. The frequency with which people are put in for exams can depend in small part on school policy and resources, as can the number of subjects, the level at which they are studied and the type of exam.

So, while an individual's results may be adequate as a criterion of performance of that child, examinations are not intended as absolute measures of the comparative merits of types of school averaged across the nation, divorced from the essential contextual information of what the school aimed for and whether it put pupils in for those exams in those subjects.

(v) General strategies and cautions.

Measuring examination performance in a variety of ways did not lessen the difficulty that no single measure applied to all children in the sample. But there was strength in variety. For one thing, some investigation was thus possible of everyone in the sample, from all levels of ability. For another, different examinations represent different facets of pupils, testing different skills, performance in a variety of subject areas, various levels of performance, and quite different perspectives on the 'attainment' of any one. Pupils might vary in having, one, two or nine 'O' level equivalents with no 'A' levels, or one, two or three 'A' levels; they might still resemble each other in obtaining a grade B in Biology 'O' level. This is the problem in taking national measures of examination performance, and, to compensate for it, it is necessary to take more than one viewpoint. Previous research work comparing examination results has not always succeeded in taking a broad range of examination measures. Yet it is unjustified (and can be misleading) to take one index and generalise about 'examination performance' or even academic performance or attainment in general. No single measure is very meaningful for a broadranging survey of all sorts of children, let alone for comparisons between different types of school which may have different traditions, aims, or policies and which may cater for very different subsections of the range of children. It would not be right to make a single comparison as if matching like with like, even after allowing for some preexisting differences, if different yardsticks are needed for different schools. The need for a multiplicity of measures and vantage points from which to judge examination results should incline the reader to caution concerning the authority of examination results for comparing schools. Examination measures are bound to be only partially informative, and some will be rather insensitive, crude or narrow indicators. As a result, no measure of examination results can tell as much about the performance of all the children in the sample as a test, like those of reading

and mathematics used by the NCDS, which is applied to everyone in the sample.

These reasons for caution in interpreting results apply not only to the findings here, but also to those which, under current legislation, are to be published in future by Local Education Authorities. In reading the results reported here it will perhaps become clearer that, even if examination results were quite good as characterisations (for headteachers, advisers or inspectors) of individual schools, they do not lend themselves to averaging across schools. In this national study, averages of results in CSE and GCE of different pupils are less revealing than the averages of their scores on universally applied tests.

Nevertheless, examination results are widely and publicly recognised and available, and are used to compare schools. They seem meaningful perhaps because, at an individual level, they do signify achievements and qualifications (see III vi). But, apart from the question of which children, at what age or range of ability, are to be measured, there are other aspects to be chosen from all the possible measures of examination performance. The important thing to bear in mind is that no single measure tells the whole story. Few measures based on the English examination system are satisfactory indicators of the performance of the majority of the age group to which they apply (and none is an indicator of the performance of other age groups, even within secondary school). The defence here is that the same measures were applied to different school groups - they may not have applied in the same way to individual members of the sample.

Before the details of the measures designed for the investigation, two further general comments on the assumptions and strategies of the study can be made.

(vi) Examinations as qualifications

Though many research studies have relied on test scores as valid indicators of school performance (for instance, Coleman, 1966) there is a particular justification in looking at examination results, arising from their ability to indicate qualifications. Measuring qualifications is not equivalent to measuring educational progress, though, and examinations are not equivalent measures to test scores. The tests, of which results were reported earlier, are better measures of relative academic success, because they are taken by all pupils. The numbers gaining a qualification will be affected by whether or not pupils enter for an exam. As Rutter et al. (1979) have pointed out, though, measuring the extent to which pupils possess qualifications is measuring one important aspect of what schools aim to provide. Some readers may even set greater store by examination results than tests, as indicators of 'tangible' results of schooling which pupils actually take away with them into later life. It will be clear, however, that it would be a mistake to value them more highly than tests as measures of school progress on a national scale.

(vii) Conventional not actual indicators

Most measures were designed as indicators of conventional qualifications or as scales with conventional or theoretical values. That is, if in theory twenty per cent got grade A to C, the value assigned to grades A to C was in some cases designed to reflect this (see Appendix W). There are many alternative possibilities which were not tried here, involving grouping, weighting or

scaling measures of examination results to fit the actual characteristics of the sample. It would be valuable to extend the study of examination results to use such measures, to reflect the actual distributions of exam results and the related attainment test performance, rather than to base the measures on examination definitions and official characteristics. The preliminary, exploratory nature of this work left room for many such developments. In the course of reporting results it will nevertheless be possible to describe some of the actual characteristics of the sample's distribution on various examination measures, and these will have implications for interpretation of results.

(viii) General strategies; range of possible measures of exams and measures used.

The information available included details of Certificate of Secondary Education (CSE) examinations and General Certificate of Education (GCE) examinations at Ordinary ('O') level and Advanced ('A') level. A number of different measures of examination attainment had to be used; no one measure would be satisfactory, since, as was discussed above, different pupils take different subjects and types of exams, and each examination measure applies to only a part of the range of children in the sample. It was necessary to choose a reasonable collection of measures, sensitive to a variety of aspects of examination performance. But one could choose to study performance in any single subject, or a given combination of subjects. If it were decided to study pupils' performance in a combination of subjects, the combination could be in a particular field, like 'sciences', or a measure of four or more passes in a good sprinkling of subjects, or a representative collection of 'core subjects'. Or one can measure general performance, by the total number of examination 'passes', for instance, in any subject, or by combining the numbers of exams and their grades, either as a total figure or as a scale on which each person is indicated by their 'best' achievement.

There are countless possible measures of examination performance, and doubtless readers will invent their own. It would have been valuable, for instance, to make findings comparable with the work of National Survey (Douglas et al., 1968) and have a criterion of four 'passes' including main academic subjects. It might be important too, to look at a broader range of subjects, or excellence in one subject, or numbers of people with a qualification in some subject. As the range of children among whom distinctions according to numbers of 'A' levels could be drawn was narrow, the work here was chiefly concerned with examinations which tend to be taken during compulsory school years ('O' level GCE and CSE) as these apply to more people. It would, however, also be valuable to relate performance at 'A' level to earlier performance at 'O' level/CSE. None of this has been tackled here, and other researchers could pursue these topics. The measures which were explored are listed below with some explanation, and then necessary details on the definitions of those measures are given.

The pattern of outcomes for different school groups may be influenced by the way in which a measure is constructed. A measure could be a criterion or a threshold measure - the possession of 5 or more 'O' levels, for example - or a yardstick type of measure on which it is possible to score from zero to 40, or from 1 to 5, or from grade A to grade E, for example. With a yardstick type of measure, results can quite justifiably be grouped in a variety of ways which may lead to different results. Different weights can be assigned to given levels of

the measure, and aggregates and averages alter as a result. And choices have to be made between various possible samples; should one confine the comparison between school groups of the average grades obtained at 'O' level English to pupils who took 'O' level English? Or see whether groups differed as a whole in the average 'O' level English grade they produced, collectively, whether or not members of the group took the exam? Many such choices had to be made, and the resources of the project were finite. As full as possible a description of the measures used will therefore be provided. Any selection of measures is bound to be arbitrary, but at least the combination of measures can be regarded as appropriately complex, since, by looking at a variety of aspects of examination performance, it was possible to inform interpretation of any particular results. Moreover, some of the factors that limited the generalisability of particular indices of examination performance are offset by the fact that analyses make adjustments for attainment and ability test scores, at least before the start of secondary school. So the fact that measures apply to subsamples, and perhaps to pupils falling in a restricted range or attainment, is partly offset.

(ix) The indicators of examination performance used

Performance in examinations was measured in a number of different ways, in order to tap different aspects of performance and find out who had various sorts of qualification. These fell into three main groups; general measures at sixteen, not specific to one academic area; subject-specific measures at sixteen years old; 'A' level measures beyond sixteen.

General measures:

- proportions in the group who did not take any public exams.
- average numbers of 'O' level equivalents ('pass', grades A to C or CSE grade one).
- proportions with five or more 'O' level equivalents.

Measures of achievement in examinations in particular subjects: English and mathematics:

English (language)

- proportions reaching 'O' level 'pass' standard ('pass', grades A to C or CSE grade one)
- average grade on combined scale of GCE 'O' level English Language grade or CSE English grade.
- average grade on scale of CSE English grades (1 to 5).
- average grade on scale of GCE 'O'level English Language grades (A to E)

mathematics

- proportions reaching 'O' level 'pass' standard ('pass', grades A to C or CSE grade one).
- average grade on combined scale of GCE 'O' level or CSE mathematics grade.
- average grade on scale of CSE mathematics grades (1 to 5).
- average grade on scale of GCE 'O' level mathematics grade (A to E).

Measures of GCE Advanced ('A') level performance

- average product of grades and numbers of 'A' levels.
- average number of A level passes

(x) Discussion of measures

To give a context to what was observed with these measures it is worth briefly explaining what prompted their use in this study. Why was it of interest to compare school categories on these measures?

General measures at sixteen-plus

No exams

The investigations had to be of the whole sample and to include, for instance, pupils not measured by any examination, because they did not go in for any exams. The proportion of school leavers each year with no exam qualifications at all has been diminishing. In 1964, according to DES Statistics of Education, the estimated proportion of the age group with no 'O' level passes, when there was still no CSE, was about 60 per cent. Even after the introduction of CSE's, it remains a commonplace to speak of the 'bottom 40 per cent' as the ones who do not take examinations, but this reflects the theory of examinations rather than 'actual' proportions of age groups. The growth of CSE examinations was part of greater attention to the provision of certificated study for the vast majority of pupils who had done no certificated work at school under selection. Ordinary level GCE examinations are officially* said to be 'designed for the most able 20% in each subject', and ''O' level and CSE together are aimed at the top 60% in each subject'. That would suggest 40 per cent as the proportion with no exam entry in a given subject. It is an indication of the range of pupils regarded as capable in some subject that the proportion who take no exam is much lower. (The proportion who take no exam in geology, of course, is much higher.)

* Schools Council (1980)

By 1974, for example, it was not 40 per cent but 20 per cent, according to DES figures, who took no examinations; 'four out of five attempted CSE or 'O' level'. The Tenth Report of the Expenditure Committee expressed the change that had taken place by 1976 by pointing out that the balance of certificated to uncertificated pupils had shifted from 20:80 to 80:20. But this was not a sudden change; the National Survey had drawn attention by 1968 to the growth in exams in secondary moderns, as one symptom of wastefulness of selection. It might be expected, nevertheless, that those who went to the kind of school that had existed in large numbers before the advent of CSE, a grammar or a secondary modern, would be more likely to be channelled by their school in the more traditional pattern, of exams for the few rather than the many. Secondary moderns, for example, might be expected to be less ready to put pupils in for CSE if they were used to pupils taking no exams at all. Was the fact that growth in numbers of exams taken and developments in exam policy (CSE) went hand in hand with the changeover to comprehensivisation reflected in any higher rate of exam entry in the comprehensives than in the old organisation of grammar+secondary modern?

To answer such questions, proportions with no exam entries were explored. A notable feature of NCDS data is the high proportion overall who had some exam entry. Among the children in NCDS, 14 per cent were reported as having taken no exams at all*. In the smaller sample studied here, of those in various school categories, the proportion without exam entries was lower, at 9 per cent. This group of 'nonexam' pupils was unevenly distributed over the four categories of school studied.

Table 7 Proportions with no exam entries

	Pupils taking no exams	total
grammar	1.1%	747
secondary modern	13.5%	1213
comprehensive	10.4%	936
transitional	9.7%	1479

* That is, no GCE or CSE; a few might have taken RSA exams alone. However, when looking at school performance, it may be reasonable to assume that people who took RSA, City and Guilds, and so on, would have taken GCE or CSE exams as well.

Given what is known of these children, particularly the lower attainment of those going into comprehensives relative to the combination of eleven year olds entering grammars and secondary moderns, it might be expected that the proportion with no exam entries later would be higher in comprehensives (as long as no correction were made to allow for earlier attainment).

But, when the 'selective' group was combined, the proportion of those grammar+modern pupils with no exam entries was 9 per cent, only slightly smaller than the proportion of 'nonexaminees' in the comprehensives group. Were the comprehensives more likely to put children of lower ability in for some exam? Were the 'nonexam' pupils similar in the grammar+modern group and in comprehensives?

'O' level pass or equivalent

The demand for certificated courses, already increasing when Douglas and his co-workers wrote in 1968 that 'the number of successful GCE candidates..... has risen more rapidly than anyone could have anticipated in 1944', had grown by the time of this study. At an earlier stage in comprehensive reorganisation, according to Benn and Simon (1972), comprehensives may have placed particular emphasis on 'O' levels, feeling thay had to emulate grammar schools to gain recognition and respectability. Was there any evidence from NCDS data that comprehensives were still aiming for 'O' levels more frequently than would be expected?

In 1970, after the introduction of the CSE exam, the proportion leaving school with no 'O' level pass or CSE grade 1 was down to 54 per cent, and in 1974 was 49 per cent.

The effect was indicated in the introduction to Statistics of Education Volume 2 (1970); 'Since the introduction of the CSE examination in 1965, the percentage of school leavers with a 'paper qualification' has risen to over 50 per cent'. Using the term 'paper qualification' to mean an 'O' level pass or CSE grade 1, the proportion in the NCDS sample explored here was similar. Many more pupils obtained other grades at CSE.

The fact that more people were offered exam courses for CSE's than in earlier times was mentioned above. But the innovations during this time did more than cater for the 'broad band' neglected under the old system; they made it necessary to think rather differently about who did what exam over the range of ability. It was not simply that the top 25 per cent were 'O' level pupils while the title 'CSE pupils' was applied to the next 40 per cent. The number of pupils in the NCDS who went in for a mixture of CSE 's and 'O' levels shows that, to many teachers at least, the same child could be a 'CSE type' in one subject and an ''O' level type' in another. The definition of grade 1 at CSE as indicative of 'O' level potential makes it clear that 'CSE pupils' and ''O' level pupils' are not necessarily different people. The very high proportion of pupils regarded as suitable to be put in for at least one 'O' level, almost a half of the school population, is another indication of increasing flexibility in definitions of children as suited for 'academic' work. Was this any more apparent in comprehensives, where, in theory at least, there was more chance within one school for varying categorisation of children after eleven years old?

Number of '0' level equivalents

One measure of examination performance at sixteen which it seemed useful to apply was the number of exam qualifications obtained at a level equivalent to the old 'O' level pass. Half the sample did get the equivalent of at least 'O' level grade A, B or C or a grade 1 at CSE. One set of comparisons among school groups carried out for this study was confined to pupils with at least one result at this level. Conclusions were thus about the pupils who had shown themselves capable of 'O' level standard in at least one subject, and not about all the pupils.

Another set of analyses also compared the school groups on numbers of 'O' level equivalents, but averages included all the pupils with no results at this level, as well as those with one or more.

The 'O' level standard, generally regarded as differentiating the top twenty or twenty five per cent of pupils, was in the past regarded as a 'grammar school' sort of examination. That is, the proportion of the year group selected at 11+ and the proportion who 'passed' in a given subject at 'O' level were similar. However, the results here suggest a far higher number of pupils capable of performance at this standard in some subject than the selection at 11+ of a limited number of 'O' level candidates - and perhaps even a superficial interpretation of the standard - would suggest.

The proportion taking exams had been increasing since 1944, but, if the figures here represented any greater increase, was it associated with comprehensives? The number who gained one or more 'O' level pass or equivalent in the samples analysed here were:

Table 8 Proportions of pupils gaining one or more 'O' level equivalent:

grammar	90%	(n=747)
secondary modern	35%	(n=1213)
grammar+modern	56%	(n=1960)
comprehensive	45%	(n=936)
transitional	46%	(n=1479)
all schools (exc. transitionals)	53%	(n=2896)
all schools (inc. transitionals)	50%	(n=4375)

The number of 'O' level equivalents obtained might measure how well schools cater for the top performing quarter or so of pupils. By means of statistical adjustments it was possible to attempt comparisons of school even though they differed in the extent to which they had taken pupils in the top performing quarter. But though the 'O' level equivalent standard is an indicator of relatively good performance in the top quarter or so in a given subject, the number of pupils reaching that standard in some subject in this sample was more like a half.

The use of numbers of 'O' equivalents as a dependent variable in the analyses can be understood as one of the less unsatisfactory approaches to measuring exam performance in this study. In particular, the use of this measure restricted to the sample with one or more 'O' levels was tolerable, in that the distribution of pupils on the measure approximated a reasonable distribution for the techniques used. The figures below show this. It will be noted that restricting this examination measure to the 'passes' gained by 1974 (rather than including all 'passes') depressed the apparent performance of comprehensive pupils slightly more than that of the combination of grammar and secondary modern pupils (before corrections for prior attainment and background). This restriction to exams taken by sixteen also underestimated the 'transitional' pupils, by a rather more marked amount.

Table 9

Number of 'O' level equivalents obtained at specified school: the upper rows show exams passed by 1974; the lower rows all exam passes known about by 1978 (note: these are slightly smaller samples than the previous table shows, as there is the added constraint that the pupils below all had test scores at sixteen as well)

		0	1	2	3	4	5	6	7	8	9+		total = 100%
grammar	by 1974	9	5	8	7	10	10	12	12	14	13	)	730 =
	anytime	6	4	8	6	8	8	12	12	14	24	)	100%
secondary modern	by 1974	64	14	7	6	3	3	2	1	0.2	0.7	)	1155 =
	anytime	61	14	7	6	3	3	2	2	0.3	1	)	100%
comprehensive	by 1974	53	14	7	7	5	3	4	3	3	2	)	875 =
	anytime	50	13	7	7	4	4	4	4	4	3	)	100%
grammar+ modern	by 1974	42	10	7	6	6	6	6	5	6	5	)	1885 =
	anytime	40	10	7	6	5	5	6	6	5	10	)	100%
transitional	by 1974	53	13	8	6	5	4	4	3	3	2	)	1393 =
	anytime	49	13	8	5	4	5	4	4	3	5	)	100%
Total all schools													
	by 1974	48	12	7	6	5	4	5	4	4	4	)	4153 =
	anytime	45	12	7	6	5	5	5	5	4	7	)	100%

Table 10

Proportions with five or more 'O' levels:

	by 1974	at any time
grammar	62%	68% (10% increase)
secondary modern	7%	9% (28% increase)
comprehensive	15%	19% (27% increase)
grammar+modern	28%	32% (14% increase)
transitional	16%	21% (31% increase)

Several measures used in this work did not have appropriate distributional properties for analysis. In particular, distributions of grades were odd. The restriction of analyses to the samples of those who obtained graded results to some extent alleviates this problem. Some details of distributions of grades by school type are in Appendix R.

The counts of 'O' level equivalents for this study differentiated the number of passes gained up to 'nine or more', so that a pupil with seven passes would not contribute so many 'points' to the average as a pupil with nine passes. This is worth noting as possibly artificially counting against those schools which did not value taking nine 'O' levels more highly than taking seven, in line with a tendency in higher education to count seven, eight or nine 'O' levels as comparable.

Five or more 'O' levels equivalents

The highest scoring 20 per cent by this measure of 'O' level equivalents by 1974 were those who gained five or more 'O' levels. It might commonly be accepted that the criterion of five or more 'O' levels differentiates the sort of performance that used to be expected from the majority of grammar school pupils. It therefore seemed a fairly direct way to look at how many pupils in each schooling group would be among the top fifth or quarter in exam attainments, to explore the proportions achieving five or more 'O' level eqivalents. The proportions obtaining five or more 'O' levels in each type of school were as shown. (These figures are reduced by the exclusion of the pupils without scores on tests at sixteen from the samples analysed and described later).

Table 10 Proportions with five or more 'O' level equivalents

	proportion of each group gaining five or more 'O' level equivalents		Number in each group
	by 1974	anytime	
grammar	62%	68%	730
secondary modern	7%	9%	1155
comprehensive	15%	19%	875
transitional	16%	21%	1393
all schools	21%	25%	4153

The nearest equivalent in national figures is to be found in DES Statistics of Education for 1974:

Table 11 School leavers, 1973-4: proportion with five or more GCE 'O' level passes* in each type of school

DES figures	proportion gaining five or more 'O' levels	Number in each group
grammar	65%	70520
secondary modern	2%	160400
comprehensive	14%	380150

But that is without including any CSE grade ones. Quite high proportions did get CSE grade ones (DES Stats. of Education 1974 Table 4), so this would be significant.

If no account is taken of the abilities of children before secondary school and of their backgrounds, fewer comprehensive pupils obtained this set of qualifications than might be predicted, if comprehensives were expected to equal the combination of grammars and secondary moderns.

Table 12 Five or more 'O' levels, grammar+modern and comprehensive; NCDS and DES figures compared

NCDS figures:	proportion gaining five or more 'O' level equivalents by 1974	anytime	Number in category
grammar+modern	28%	32%	1885
comprehensive	15%	19%	875

DES figures (school leavers 1974*)	proportion gaining five or more 'O' levels	Number in category
grammar+modern	25%	230920
comprehensive	14%	380150

As has been explained, the purpose of the investigation here was to look at whether that difference could be explained by initial attainments, social class, sex and parents' interest in schooling. There are limits to the number

* From DES Statistics of Education Volume 2 1974 Table 6. These figures do not include CSE grade one. They are for England and Wales. The 'comprehensives' include schools which were not comprehensive when pupils started secondary school (transitional).
The DES figures are not restricted to pupils who were in the same school during their secondary years.

of measures any one study can use, but further work along the lines of this investigation could be done to explore yet other criteria of examination performance. One reason for giving priority to the proportions gaining five or more 'O' level equivalents is that the measure is used currently by other researchers. Results can be compared, for instance, with those obtained by Rutter and his colleagues (Rutter et al., 1979) (though they worked with samples with a lower likelihood of gaining five or more 'O' levels than average). It has been argued that four or more 'O' levels are a better criterion, and the data could be explored further to study what discriminates best among the sample. The use of five or more 'O' level equivalents, though, facilitates comparisons with figures published by DES, and by some local authorities e.g. ILEA.

Single subject measures: It was recognised that there were advantages in having some of the measures of exam performance restricted to results in a single subject. Studies of performance in the specific subjects of mathematics and English were obviously necessary supplements to the general measures, some of which masked the nature of individuals' exam performance and proved difficult to interpret.

To have single subjects as measures of exam performance was particularly appropriate, too, because measures like that reflect the kinds of variety between pupils that are seen as relevant to examiners. In that respect they complement the universal or wider application of the general measures. The number of exams equivalent to an 'O' level pass, or overall measures of CSE's, are not on their own satisfactory indices of school attainment. At one extreme, a traditional regard for basic skills could prompt the question of whether two grade one CSE's in Greek and fibreglass modelling are signs of adequate schooling. The overall measures of exam performance explored above left rather vague the question of how well academically pupils were doing because they contained no information on subjects. The information which was available for this study made it possible to investigate more closely how people were doing in two particular subjects, however. These were English and mathematics, which could be regarded as key, if not core, subjects. The two subjects of mathematics and English were chosen as being specially relevant to certain sorts of comparison with mathematical attainment and reading comprehension respectively, which had been measured by tests at sixteen years old at the third follow-up of the National Child Development Study. The choices were partly also determined by practical considerations. The two subjects most commonly taken are English (language in 'O' level) and maths. These were seen, therefore, as good subjects on which to base measures, as they would apply to many people in the sample. Another advantage to studying maths and English, apart from the fact that these are seen as skills applicable to many other areas of life, was that adjustments for scores on mathematics and reading comprehension at 11 were particularly appropriate. Such adjustment was, as usual, necessary in order to judge sixteen year olds' attainments attributable to experiences in the secondary years, in the light of existing relevant attainments before the start of secondary school.

Results studied were those of exams called 'mathematics', and exams called English, or English language in the case of GCE. Anyone who took English literature or additional mathematics at GCE 'O' level could be expected also to have a grade on English language or mathematics examinations and thereby to contribute to these results. However, those candidates who, for example, did

not take an examination called English but obtained CSE (even a grade one) in Creative Writing, or those who did not take an examination called mathematics, but obtained even a CSE grade one in 'mathematics in society' would not contribute to the results presented here. This could be expected to exclude more CSE candidates, therefore, than GCE candidates. Despite the broad range of ability represented by the full scope of grades 1 to 5 CSE, the number of people involved in analyses confined to those with some graded result at CSE mathematics was rather small. This may be because of the variety of different subjects at CSE. Restriction of the sample in the fashion of these analyses, though, should not make results uninterpretable, since allowance was made for maths attainment scores at eleven years old.

Even with the single-subject approach, offering a closer study of an aspect of academic performance, there were two sorts of questions to anwer for which two types of measurement were needed. On the one hand, it was important to answer questions about the proportions in each type of school who gained a recognisable qualification. A similar need to look at a criterion had prompted inspection of the proportions gaining five or more 'O' levels. So analyses were designed to look at the proportion who obtained a 'pass' or the equivalent of a 'pass' at 'O' level. Thus, one measure of performance in a single subject, chosen to combine both CSE and 'O' level achievement, was whether pupils had the equivalent of an 'O' level 'pass' (A to C grades or CSE grade 1). Whether or not an 'O' level is passed, however, is a question which applies to the highest attainers in a particular subject. The measure did not discriminate among other people. It also did not discriminate among grades above the 'O' level 'pass' level, as current GCE grades do.

So, on the other hand, it was also important to have a broad-ranging measure, a yardstick which discriminated performance over the whole range of graded achievement recognised by GCE 'O' level and CSE examinations. Broader ranging measures have disadvantages in that many levels of examination performance, at any of which there might be differences between school children, are masked in the composite average. Nevertheless, combined measures, over all grades of GCE 'O' level and CSE, might reveal differences in performance at lower attainment levels. Accordingly, combined scales of 'O' level and CSE grades were compiled. These ordered pupils' exam achievements on a six-point scale from CSE 4 or 5 up to 'O' level grade A.

The ordering of results for one set of analyses of the combined scales took into account the status of 'O' level results below grades A to C in 1974. Officially, the equivalent of grades D and E, 'failure', could well be regarded as less of a qualification than CSE grades 2 or 3. So, one set of analyses treated CSE grade 2 or 3 as worth 3 points and 'O' level 'fail' as worth 2 points. However, it can be thought, on the basis of test results at the same age, that CSE grade 2 or 3 was a lower result. It therefore was important to vary values assigned. It was also important to try the effect of giving the various assigned grades alternative weights for different analyses, since there was no reason to expect the distribution of results to be at equal intervals along a continuum. Another analysis, therefore, reweighted grades. The ordering and grouping of the grades was in most respects the same, as follows:-

GCE 'O' level grade A; grade B, or 'pass' or CSE grade 1; 'O' level grade C; CSE grade 2 or 3 or 'O' level D or E; CSE grade 4 or 5; no graded result. With

weighting, though, CSE 2 or 3 and GCE 'O' level D or E (or 'fail') were treated as equivalent. The actual weights assigned are explained in Appendix W, and in Section III. xv below.

Thus, there were criterion and yardstick measures, designed to reflect both CSE and GCE 'O' level performance, whichever pupils happened to take. It was likely to help in interpreting a combined measure of CSE and 'O' level, however, to be able to look at pupils who gained a graded 'O' level in the subject separately from pupils who gained a graded CSE in the subject. (Those who gained both appeared only once, and had their 'better' result included in analyses). This might help to explain underlying characteristics of the pupils' performance averages. Accordingly, additional analyses were carried out on results in terms of the grade obtained in CSE (English or maths) by those who took the exam, and, separately, on the grade obtained in English language or in maths at 'O' level, by the pupils who took a GCE 'O' level examination in the subject.

Another reason for exploring average grades in GCE 'O' level mathematics within the whole graded range A to E was the need for more detailed investigation within the higher grade range of achievement represented in 1974 by a 'pass' and, since 1975, by grades A to C. Leaving out of account CSE results for the moment, then, a set of analyses was carried out on the average grades obtained by different groups, within the sample of those with some GCE 'O' level grade.

'A' level achievements beyond compulsory school age

In addition to all the measures of 16-plus attainment described so far, two measures of 'A' level performance were used. Both varied only within the range of 'pass' grades at 'A' level, i.e. grades A to E. One measure took into account each grade, assigning five points for grade A, four for a B, and so on, and totting up an overall figure according to numbers of passes at each grade. This created a scale from 0 to 20+. The scale may be recognised by those familiar with the Universities Central Council on Admissions (UCCA), and is here termed a UCCA scale. UCCA and DES frequently regard 9 points on that scale as an indication of good 'A' level standard.

The second measure of 'A' level performance was less discriminating about grades within the pass range, but simply looked at numbers of 'A' level passes. It was hoped by using this measure in conjunction with the UCCA scale to distinguish whether any differences in 'A' level performance reflected differences in numbers of subjects taken or differences in the standards attained in those subjects. Were some types of school more likely to put people in for three 'A' levels at the risk of their performing at a lower standard in each one? Were other schools more likely to restrict the numbers of 'A' levels they put pupils in for, in order to better the grades obtained? Given that it was known from the overall NCDS figures that most 'A' level entrants took three or more 'A' levels, did different (or similar) levels on the UCCA scale reflect different (or similar) standards of achievement, amongst those passing similar numbers of 'A' levels? Or were results on the UCCA scale sufficiently informative as a measure of aggregates of pupils?

Results studied were confined to those 'A' level passes obtained at the same school as was attended in 1974, in order to ensure that the nature of the institution in which 'A' level study took place was the same as the 'sort of

schooling' about which comparisons were being made. The sample was slightly larger than that of the 'UCCA-scale' analyses, since it was not restricted to pupils who obtained at least one 'A' level grade E but included all who went in for an 'A' level, at school. Those who went to another school or to a sixth form college to do 'A' levels were not included. So depending on local provision, the 'A' level results of comprehensive and perhaps secondary modern pupils could be excluded because of the organisation of sixth form work in their area.

It seemed important in addition to look at 'A' level performance as a reflection of the whole sample. For one thing, the numbers of 'A' level passes per entrant are a function of schools' policies on who should be put in for 'A' level examinations as well as of the attainment of pupils of given potential. It remained possible that, even after correcting in the normal way for intake and background differences, the schooling groups differed in 'A' level attainments in a way which had not been revealed. In case grammar+modern and comprehensive schools differed as a whole in this respect, it seemed essential to compare within each group of all pupils in the school groups studied, not just the 'A' level entrants. It was clearly very important to take into account the level of attainment of the pupils who were put in for the exam. Numbers of 'A' level passes obtained were therefore studied by comparing those of given attainments at sixteen, from the two schooling groups - or, rather, in terms of the techniques used by this exploration, after allowing for differences in sixteen year old test scores. In addition, the number of 'A' level passes at 11, and the number of passes obtained by those in the top fifth at 11 on all three such scores, were compared between school types. And the three types of school were also compared, on average numbers of 'A' levels obtained by the entire set of pupils, including the vast majority who did not take any 'A' levels at all.

(xi) Further definition: exams taken by conventional time, in school.

One crucial detail of the definition of exam performance is that, for this study, 'O' level or CSE examinations were limited to those taken in 1974 or before. This was another 'traditional' bias of the study, which would be expected to underestimate the schools offering opportunities to take exams later. There is a certain amount of evidence, in Scottish schools at least, of an increase in examinations taken after sixteen, in the year after compulsory age (Ryrie, 1981). If the same is true in England, constraining 'O' level and CSE results to those taken by 1974 might explain some differences between the results here and DES figures for schoolleavers.

That limitation on examinations counted by the study might also be expected to bias results in favour of more traditional schools - perhaps some grammar schools - and against newer, less conventional schools - perhaps some comprehensives? Comprehensive pupils especially may stay on to take '16+' exams in the sixth form. (See Table 9 page 56, this Chapter). It hardly needs emphasising that this restriction on examinations, to those taken in the school by sixteen (or 'A' levels taken in the school), confined the study to a concept of educational qualifications as entirely school-based. No account was taken of opportunities for later school examination work or for continuing education elsewhere, in these measurements. The main point, it will be recalled, was to study categories of schools and those examinations which could safely be attributed to conventional patterns of examination entry in those schools.

The 'A' level measures, as was mentioned, were also restricted; only those examinations taken in school were counted (though they could be taken at any time). (There was an exception which is noted elsewhere.)

(xii) Samples to whom measures applied:

Since only some people go in for given exams, it is difficult to generalise about a particular sample. As crucial to the results as grouping and weighting of the yardsticks by which exam outcomes are measured, is the sample among which comparisons are made. Sometimes the sample had to be restricted to the candidates for that exam.

Each of the analyses which looked at numbers of 'O' level equivalents was carried out on two different samples. Firstly, the comparisons were drawn within that half of pupils who had one or more 'O' level equivalent. Thus, it was possible to see whether, among pupils theoretically 'in the running' for 'O' level equivalents, there was variation according to schooling. But it was clearly also valuable, particularly given the high proportion who did achieve at least one 'O' level equivalent, to see whether the averages for different school groups differed overall, taking all the pupils who did not get any 'O' level 'passes' or equivalents into consideration as well. So the second sample chosen for each analysis was the entire study sample of secondary school pupils, including those who never went in for any exams, as well as those who took a CSE or GCE but did not obtain an 'O' level equivalent.

All the analyses of proportions, looking at whether or not a certain criterion level was reached, explored proportions of the entire study sample, whether they took any exams or not. The entire study sample consisted of those who, by the definitions used here, went to grammar, secondary modern or comprehensive schools, and, for some analyses, those whose schooling was transitional. So the analyses which measured the proportions who did not have any exam entries, the analyses which measured whether pupils had five or more 'O' level equivalents, and the analyses of proportions obtaining an 'O' level pass (or CSE grade one) in mathematics or English (Language) were all comparing, between the schooling groups, the number among all the pupils in each group who fitted that criterion.

The other, yardstick, measures which used scales of total, general achievement in terms of numbers or grades of exams, were most commonly used in analyses confined to those who obtained a 'score' of at least one on the measure in question. That is, analysis of grade in maths CSE was confined to those with some graded CSE maths result. This seemed essential to make sense of the comparison. There were so many uncontrolled influences on a pupil's not having a graded CSE result - whether they took 'O' level maths, whether they took other CSE mathematics exams, whether they took CSE's in other subjects and no maths exam, whether they took no exams at all - that to include in average CSE grades for school groups the component attributable to the proportion who had no CSE maths grade at all would distort the relative positions of the school groups. Similarly, the numbers taking 'A' level varied so much according to school, and 'A' level performance was regarded as specialised, so that it made sense to restrict most observations to those who took 'A' level or those with a graded 'A' level result (grade A to E at 'A' level). But confining analyses in this way also had disadvantages. It meant that, in discussing 'O' level or CSE

maths grade, for example, it had to be borne in mind that those without a graded result, or with no entry for either exam in maths, are not included. To interpret grades, results are also needed on whether or not people obtained an 'O' level pass or equivalent in maths. This demonstrates the value of the variety of measures by which this study was able to compare exam results; it was possible to check conclusions from restricted but meaningful samples against the results of proportions from the entire sample. Ideally, it would have been advisable to try all analyses on the entire sample as well as on the groups scoring positively on a measure, but it made sense to give priority to some restricted sample analyses, because of the impossibility of knowing what factors went into not scoring on a scale or graded measure.

(xiii) 'O' Level GCE; a pass/fail exam or a graded exam?

The fact that NCDS members were 16 in 1974, and that most of their 'O' level examinations were taken in that year, created a special difficulty in analysing their results. In 1974, GCE 'O' level exams were officially awarded a 'pass' (or fail) and any grades were unofficial. From 1975, the pass/fail distinction was officially abolished and replaced by a grading system running from A to E (with an ungraded category for exceptions). Grades A to C correspond to the standard of the former 'pass', grades D to E (and ungraded) to that of the 'fail'. Deciding on appropriate equivalences between 1974 and 1975 examinations was necessary, both because some NCDS members took some 'O' levels after 1974 and because current interpretation of results may require comparisons with modern versions of 'O' level grades. There was little difficulty over equivalences in the 'pass' range: unofficial grades, which were supplied for these data in most cases of exams officially awarded a 'pass' in 1974, translate straightforwardly into current A's, B's, and C's. In the few cases where teachers reported exams simply as 'pass', those were translated to grade B (as the 'average' of the present three grades equivalent to the old 'pass'). It will be clear from this how it was possible in this document to use the term ''O' level grade A to C' and the term ''O' level pass' interchangeably. The 'fail' mark and unofficial grades below a pass do not translate directly as the present D and E. Rough equivalences were set up, however; though unlikely to make a great deal of difference, their roughness was one factor in the decision to group grades D and E, or 'fail' at 'O' level, together for some analyses (chiefly, those using grades in single subjects as measures) (see Ives, ibid., pp. 11-12 for more detail).

One implication was that the measure of 'O' level graded performance (A to E) was in two respects the least satisfactory measure employed by this project. For one thing, it awarded points for exam results still in 1974 regarded as failures to obtain 'O' level standard, which means that interpretation is questionable. Secondly, the 1974 grades (or 'fail' awards) do not have exact equivalents in the current grades used for analysis here, so that it was not clear what averages of results translated as grades A to E would represent. There was no reason to expect that this would affect one type of school more than another. However, the translation to grades following the current pattern from the 1974 grading system may not be quite irrelevant* to the problems of error/reliability in measurement.

* Mitchelmore, (1981)

(xiv) CSE grade one.

It is widely recognised that a grade one in a CSE exam is the equivalent of an 'O' level pass. More officially, a CSE grade one is regarded as an indication that the candidate would be capable of obtaining a grade A,B or C in a GCE 'O' level examination in that subject, given that he or she had followed an appropriate course of study. Relatively small numbers obtained a grade one CSE. Because of the theoretical equivalence of a CSE grade one to a 'pass' at 'O' level, and because some of our 'O' level sample were being assigned a grade B for an award of 'pass' without grade information, it seemed appropriate to assign an equivalence of grade B to CSE grade 1 when details of grades within the 'O' level 'pass' range were being explored. To assign a grade C to a CSE grade one seemed likely in our data to lead to underestimation of the performance of high scoring CSE candidates. (See also Appendix Y for discussion of 'double entry' for CSE and GCE exams in the same subject at the same time).

This is relevant only to the average grades in particular subjects on combined scales of CSE and GCE 'O' level attainment (English (language) and mathematics). The term ''O' level equivalent' is used throughout this study. A CSE grade one and an 'O' level 'pass' (i.e. grades A to C at GCE 'O' level) are both counted as an 'O' level equivalent; this follows official definitions.

The measures used here are sometimes based on separation of the 'pass' from the 'fail' at 'O' level standard, because, at that time, the examination system itself retained the concept of a 'passed' examination discriminating 20-25 per cent of pupils from the rest. Even if in fact the change of awards made little difference, it would be unsurprising if, in 1974, some schools concentrated on getting pupils over the 'pass'/'fail' hurdle. The more recent establishment of an examination grading system may mean that more attention is given to the distinction between D and E in examination grades. It remains true that grades C and above are accorded 'pass' status by employers, pupils and schools. But the NCDS data are to be treated cautiously when they include results in the range called 'fails' in 1974, because the grading changes since then mean that results in that range are not easy to translate into current D or E grades.

(xv) Distribution of examination performance.

Even with the definitions of examination performance above, a variety of different measures would be possible. The values assigned to the levels of examination performance, on continuous measures, will affect averages and may even alter the interpretation of results. It is therefore, essential to give further details of groupings, weightings and orderings used and of the actual and theoretical distributions of examination results.

Of the two sorts of examination measure used, the analyses of proportions avoided violating some of the statistical assumptions which were upset by the continuous variables, limited though the proportions analyses were. Three possible treatments existed for continuous or quasicontinuous measures.

Each level could be given an arbitrary value, (from one to six, for instance). This was done in the case of measures like numbers of 'O' level equivalents, which were treated as 0,1,2,3 ..... 9 or more, or as 1,2,3,4 ..... 9 or more. It was also the strategy in the case of all analyses of grades in mathematics

or in English and for analyses of numbers of 'A' levels and the 'A' level 'UCCA-scale'.

A second possibility is to assign values to levels of the measure based on conventional or official views of what the exam results correspond to in the distribution of ability.

That is, if the 'top' 20 per cent of the population is expected to obtain a grade A, then a grade A can be assigned a value of 90 (midpoint of 80-100 per cent) on a scale of 0 to 100. With any such system of expected equivalences, there are areas of decision. The problem is encountered, for instance, of assigning a value to those with no such examination result. Such people may have 'lower' exam performance - such as no exams - than the pupils scoring higher on the scale in question. But it may be that such pupils have higher results in other subjects, if for instance, 'Physics' results are the only ones under consideration. The derivation of values from hypothesised characteristics of the population is difficult, when the characteristics of those taking or not taking one particular subject may differ from the supposed characteristics of candidates for exams in general. Nevertheless, the use of such hypothetical values was explored for just two measures of examination performance. These were combined scales of English and of mathematics performance, as was explained in section III(x). Importantly too, any of the combined scale results included examination grades not universally recognised as qualifications. This was necessary in order to include substantial numbers of pupils; grade 4 at CSE is an indicator of average performance, for instance. So the scale had application to a broader sample of the sixteen-year-olds, even if not the total range, and differentiated at different points over a wider range than the single dichotomous 'pass/fail' mark. There was little scope within this project to attempt alternative groupings of the exam measures. Alternatives would have been possible in theory, since there is no a priori reason to suppose that, say, scores run evenly from 0 to 20. It might be that a quarter of the population score 0, 1, 2, 3, 4 or 5; a quarter scores 6, 7 or 8; a quarter scores 9 or 10; and the top quarter scores 11 to 20. In that case, one sensitive grouping might be into four, as 0 to 5; 6 to 8; 9 or 10; and 11 and above. As it was, only one grouping for each measure was possible. What every scale did do, though, was to exclude from analysis any ungraded results; these were, presumably, attempts at exams which were officially regarded as equivalent to not having attempted the examination, and this study treated them accordingly.

There is a third way of assigning values to levels of measures, which is much to be preferred if it is possible. In order to use measures as if they were continuous measures in a regression, they should be scaled by discerning which pupils, of what attainments, got which results. That is, values could be based on actual distribution of results in the study sample and ordered, at least, according to the mean test score of those obtaining each type of result.

Scaling of measures might have made a difference to the implications of results. It will be apparent throughout the material later in this report on examination results that all the observations are subject to the proviso that they might vary if one altered the weights, groupings or orderings of measures. For example, the decision to accord 'O' level grades D or E lower 'status' in the ranks than CSE 2 or 3, on the combined scales, was questionable in view of the mean reading scores at 16 of those achieving each of these two standards.

Table 13 English grades and reading scores in each type of school.
Mean reading scores at 16

English result pupils	GCE 'O' level grade D or E	CSE 2 or 3	No graded result	No exams.
grammar	.346 (58)	.410 (4)	.799 (41)	.216 (5)
secondary modern	.465 (29)	.027 (348)	-.303 (258)	-1.099 (146)
comprehensive	.292 (46)	-.037 (199)	-.393 (175)	-1.072 (72)
transitional	.291 (77)	-.012 (297)	-.274 (251)	-1.055 (123)

The table above shows that the average scores in reading at 16 of those obtaining grades D or E in English language 'O' level were higher than those of pupils with a grade 2 or 3 at CSE, among pupils in transitional, comprehensive or secondary modern schools. In all but grammar schools, relatively few got grades D or E at 'O' level, compared to the numbers getting CSE grades 2 or 3. In grammar schools, the relatively high number getting 'O' levels D or E (failures) were lower scoring on the reading test, on average, than the few with CSE 2 or 3. These patterns would tend to reflect the sort of pupils put in for each examination in the different types of schools. There was no reason, in the definitions of the exams, to think of grade 2 at CSE as a lower achievement than 'O' level grade D. In one set of analyses of combined scales in maths and English, grades D and E at 'O' level were treated as 'lower' achievements than CSE 2 or 3, because, in 1974, the former were not officially regarded as achievements, nor as anything other than 'failure'. But in the absence of evidence that grades 2 or 3 at CSE were higher achievements than 'O' levels D or E, the second set of analyses of English grades and of maths grades on combined scales equated grades D and E at 'O' level and CSE grades 2 and 3, by 'weighting'. In the event, the two sets of analyses with different weightings produced very similar results.

However, the use of the first two approaches above, the 'arbitrary' values and the 'official' weighting, can be defended in the light of results obtained in that, as will be shown, the results on combined scales of mathematics and English did not seem different, as far as was explored, according to the values assigned (which are described in Appendix W). Nevertheless, it is worth mentioning a fourth possible strategy for analysing examination results, which could have obviated the problem raised by lack of scaling of measures. This would be to use a technique for analysing exams as categorical variables, rather than either dichotomous or quasi-continuous ones. Some more recent research at NCB, exploring other questions in NCDS data, is using McCullagh's approach as a method. This has the advantage of being designed for categorical variables.

Thus, the assumption that, say, grade A is as much above grade B as grade B is

above grade C, can be avoided. The assumption that, if two 'O' levels are twice as good as one, four are twice as good as two, which is perpetuated here by the use of unscaled numbers of 'O' level equivalents, is more defensible. The problem arises as soon as one includes in the analysis the pupils with no 'O' level equivalent; this is 50 per cent of the sample, and giving them a value of zero on the scale is (statistically or educationally) questionable.

These are relatively subtle questions of scaling, however, compared to the crude discontinuity that is one feature of measures of examination performance in this study, a function of lack of any measure linking performance at 16+ to 'A' levels. This is a necessity, however, since relatively few people take 'A' levels and since, in relation to attainment test scores at sixteen, there is no clear continuity between 'O' and 'A' level measures. That is, the average reading and mathematics test scores at sixteen of those who got 'O' level grade A in English or maths and did not go on with that subject, are higher than the average test scores at sixteen of those who later got grade E in 'A' level English or mathematics. So the 'highest' exam grade in a subject may not be a sign of 'highest' achievement in other subjects.

Table 14

Maths and English grades and test scores at sixteen

"highest" grade in maths	mean maths test score at 16
'O' level grade A	1.19
'A' level grade E	1.01
"highest" grade in English	**mean reading test score at 16**
'O' level grade A	1.36
'A' level grade E	0.84

(xvi) Allowance for intake.

It will be clear that an initial rationale for this study and for an investigation based on the National Child Development Study was that schools could not be compared on examination results without taking into account the fact that their pupils differed on intake.
Schools also differed in the extent to which they gave priority to exam results. They presumably would vary, too, in the degree to which that factor is correlated with their intake.

That is, some schools with a great deal of 'creaming' of the top attainers would emphasise the inappropriateness of concentrating on examination work, others would concentrate on examination results as a way of attracting pupils. Either of those schools might actually put more children in for exams. But all these factors are left unexplored by such a study as this. What this study did was to attempt to make comparisons meaningful by allowing for differences before secondary school. This makes it necessary to point out that results are not meant to correspond to the superficial version of what average results look like, but are a sign of what that school was doing, relative to another school, with a child of given attainment. Results, then, are not 'good' or 'bad' but

'better' or 'worse', relative to what the average was for children who were equivalent at eleven. This might go without repeating but for the erroneous idea that examination results are public, familiar measures. The implications for particular graded results are that a relatively good grade of D might 'surpass' a relatively bad grade of B.

Another difficulty for the interpretation of results in single subjects is that one score at eleven could have been a poorer predictor of examination performance than another. The relation between English exam results and reading test scores at 11, for instance, might be expected to be not as close as that of mathematics results with mathematics test scores at 11. It would be worth looking at the relation of English results to general (verbal and non-verbal ability at 11), and of mathematics results to reading and general ability at 11, before drawing firm conclusions about English or mathematics exams in secondary schools. To allow for reading skills alone, in this example, may not be as adequate a control for intake for English exam results as controlling for mathematics score at eleven is for maths exam results.

In single subject exam results, adjustments for attainment were for one test score only (or, at times, for scores at different ages but on mathematics in both cases or on reading in both cases). Ways in which the tests might be differently related to exam results can be conjectured from the preliminary analyses of relations between test scores and exam results, reported in Ives (1980). It was known from those analyses that CSE entry differed from 'O' level entry in its relation to test scores. The relation between mathematics examinations and tests differed from the relation between English examinations and reading test score. Entry for CSE declined over the 'above average' range of pupils in both subjects, as top attainers were more likely to be put in for 'O' level. But entry for English language 'O' level was more common at average levels of sixteen year old reading test attainment (Ives, p. 120; 1980) than entry for maths 'O' level. Liberal entry policies for English language 'O' level might explain the lower average grade in comprehensives.

There is a further complication, illustrative of the caution necessary in inspecting examination results in relation to test scores. It was possible that allowing for eleven year test scores in fifths had different influences on results depending on the subject in question. Perhaps treating eleven year scores as continuous and correcting for a quadratic component would have made a difference to results. From figures 13.1 to 13.4, pp. 117 to 119 and p. 124 of Ives (1980), it looked as if allowing for a non-linear component in mathematics results might correct in a different fashion from similar allowance in English results for attainment-related factors, indicating differing entry strategies.

It was emphasised earlier that the range of application of examination result outcomes was limited, and that they applied to differing extents depending on the school group explored. A further difficulty is that the range of eleven year attainment to which exam measures apply is limited, and that will vary according to school. Moreover, when samples were restricted by the examination measure (to candidates or those with a graded result, for example) the school groups were less comparable. And when more than one eleven year score was included, or all three eleven year scores are used in a composite measure, the range in common, for application of exam results measures to schools, was diminished further.

In addition, the nature of examinations (especially those which are a criterion distinguishing the 'top' or 'bottom' scores at eleven) means that one would not predict a linear relation between exam results and test score at eleven. That is, though it can in general be said that, the higher the earlier test score, the higher the exam result, it is possible to be more sophisticated. The statistical techniques used can take account of the fact that exams may vary little over the lower half of test scores, but increase rapidly with higher test scores, for instance. This would be a non-linear, perhaps 'quadratic' pattern. Examinations are not unique in this respect, and similar considerations may apply to measures of truancy, for example. There was more of a case for expecting a linear relation between test scores at different ages. In the analyses carried out, some attempts were made in the case of continuous outcome exam measures, to correct for quadratic components of the relation of exam results to eleven year score in addition to the general allowance for linear components. The analyses which incorporated eleven year score in fifths did not have this feature. Grouping scores in fifths has some compensations compared to linear corrections.

The complete list of which test scores at eleven years old were corrected for by incorporating them in analyses, of how the eleven year scores were grouped, or of whether linear or quadratic components were allowed for, and of whether a combination score was allowed for, is in Appendix P.

(xvii) Reliability and measurement error.

Given all the arguments above for caution in interpreting results, because of the inadequacy of examination results as measures and the imperfect allowance for initial attainments, it may surprise readers that considerable care was nevertheless taken to deal with a finer point of allowance for initial attainments. Each test score at eleven years old has its own degree of reliability (as does any measurement) and it is possible to approximate an allowance for unreliability in eleven year scores, to compensate for the fact that tests are not perfect indicators of children's performance, i.e., 'measurement error'. The estimated 'reliability' of each test score, as divined by one of the possible methods for calculating this, is given in the Statistical Appendix (see also Douglas, 1964). These estimates were used to correct the picture for the 'outcomes' maths and reading test scores at sixteen, but not for most other measures in the examination results or in the earlier stage of the study. One analysis of examination results, though, incorporated allowance for measurement error in a comparable way. This was the analysis comparing average numbers of 'O' level equivalents (0 to 9 or more) in the three types of school, taking as the baseline the range of combined scores at eleven years old. There was otherwise no allowance for the statistical lack of 'reliability' in the eleven year old scores. Readers familiar with statistical techniques and the arguments used in the presentation of the reliability-corrected regressions of numbers of 'O' levels and in publication of the mathematics and reading test results, will appreciate that this could make a difference to the results, perhaps particularly for the top and bottom fifths. Grouping eleven year scores in fifths of the range and graphing the points at means for each group separately were the alternative to correcting for measurement error.

The above are arguments for caution in interpreting overall results. There are also reasons for caution over results for pupils of different levels of initial attainment.

(xviii) Subgroups of pupils (attainments and background): difficulties for interpretation.

Distinctions according to exam performance apply to varying degrees over different ranges of ability. Differences in proportions passing 'A' level maths among those who had been on the lowest fifth for maths at eleven, for instance, would be less likely to be meaningful than differences among those in the top attainers at eleven. The patterns of pupils with no exams could be more important among the low attainers at eleven, for whom no exams were likely.

A further difficulty in interpretation is caused by the fact that the pattern of interactions might be otherwise with a different grouping of the underlying eleven year scores. It may be the case that examination measures which involve proportions reaching a given criterion (pass or fail, five 'O' levels or not) need to be treated in a different way from continuous measures of grades, 'yardstick' measures. Perhaps criterion measures need different groupings of eleven year old scores (and other independent variables) from those appropriate for continuous measures, if only because the latter may apply over a wider range. Perhaps overall findings as well as interactions would be modified by such alteration in groupings of the underlying eleven year scores. Moreover, in the case of continuous variables, adjustment was possible for any quadratic components in results, whereas the same was not true when the examination measure was a dichotomous one (See Appendix P).

The interactions with test scores grouped into fifths when the dependent variable is dichotomous create a problem. In the top fifth of grammar school pupils, there were many with higher scores at 11 than any comprehensive pupils. This does not present a problem in the continuous variable regressions where like can still be compared with like over a common range.

(xix) Selection/nonselection and exams.

The above considerations are deterrents to any overconfidence in examination results as measures. Three final, general reminders of reservations expressed earlier can be given here. One is that such an exploration may be inherently biased against comprehensives. There are finer points of examination measures which make them of 'traditional' bias; these include: having a 'high' criterion of five or more 'O' levels; the exclusion of 'O' level or CSE exam results taken after the age of sixteen; and the restriction of 'English' to examinations called 'English' or 'English language' (rather than creative writing, say) and of 'mathematics' to examinations called 'mathematics'.

The second point draws together several detailed observations of examination measures made in this Chapter. It is that the measures with the most appropriate properties as far as statistical analysis is concerned are not necessarily the most telling educationally.

The third general remark is that the difficulties for measurement caused by the continuing existence of selection, and the focus of the study on selection/nonselection, were not independent of the characteristics of the examination system. This was touched on both in the introduction to this report as a reflection of educational history and in this chapter as a reason for regarding the study as violating statistical assumptions.

There was a link in theory, originally, between 'O' level and grammar schools. The 'O' level GCE was originally a grammar school sort of exam, in the sense that it was meant to discriminate among the most academic quarter of the population. At the time of this study, a 'pass' in mathematics at 'O' level or CSE went to about twenty per cent of the NCDS sample. It is still regarded as assessing the top fifth or quarter.

Part of the bipartite or tripartite tradition was that only in grammar (and technical) schools were academic exams likely to be given attention. Formerly, the 20 to 25 per cent in grammars were generally the ones who were certificated. There had been a rapid increase in the examinations taken and in certification with the development of the Certificate of Secondary Education and the growth of examinations in secondary moderns and comprehensives. So, with the advent of CSE, the relation of exam entry to type of school changed, as, too, the types of school changed with comprehensive reorganisation. Some of the changes can be seen in the proportions with particular exam results in NCDS (Appendix R). Of course, even with such developments, there were necessarily limits to the number who would obtain the higher levels of examinations, by definition. But a major consideration for readers interpreting the examination results presented in the next Chapter is that, while the educational system continued to contain selective schools, there was no way of gauging how pupils would perform at examinations in the absence of selection. This was true even of pupils in schools called comprehensives.

Chapter IV.

Exams and schools: Findings

(i) Introduction

Comparisons between pupils categorised according to their schools were carried out using their examination results. It has been emphasised that there could be no exploration of 'whether selection made a difference', consisting of a comparison of exam results between pupils in a selective system and pupils in a system without selection, because there was no nonselective system. On the other hand, it was clear that 'selection made a difference' to the conditions of pupils, in that the continued existence of selection obviously made a great difference to the intakes of the schools (see Chapter I). There were several questions to be answered, then, all addressed to whether selection made a difference. To explore, for example, how being in a grammar school, as against in a secondary modern, may contribute to examination performance is also to study the possible effect of selection. Was being in a comprehensive, and, in a sense, 'not selected', associated with differences in performance from that of equivalent children in grammars? Were the two 'nonselective' groups, secondary modern and comprehensive pupils, different?

It is already clear that the intakes of the school groups differed. The eleven year olds differed in average test score and social class in the categories of school studied. Since the grammar, secondary modern, comprehensive and transitional groups differed at eleven, and the grammar+modern combination differed from the comprehensive average at eleven, the groups could be expected to differ at sixteen in examination results. The following table shows 'raw' examination results for each school group, and how those related to preexisting attainment (fifths of combined score) and to social class. Averages within each attainment band at eleven are shown. Results within each fifth of attainment, also, independently related to social class.

It will be clear that, in the lowest attainment groups, all pupils were in either comprehensives or secondary moderns and did not differ.

It can be seen, for example, that within each fifth and social class, grammars tended to be 'best' and secondary moderns 'worst'. The comprehensives pupils' raw average number of 'O' level equivalents matched or exceeded that of the grammar+modern combination, with the exception of the averages for 'nonmanual' pupils who had been in the top fifth at eleven. For 'manual' pupils in this group, the averages for comprehensives and grammar+moderns were very similar, though the proportion obtaining five or more 'O' levels in comprehensives exceeded that proportion in grammar+moderns. For 'nonmanual' pupils that relation was reversed.

The average number of 'A' levels, too, was similar in grammar+modern and comprehensive groups, while the grammar average was highest and the secondary modern average was lowest. The proportions getting one 'A' level or more were the same or, if anything, higher in the comprehensive sample among the middle group. For the top two groups the overall average proportion was lower in comprehensives (though more detailed data in Appendix R reveal that this was only in recent comprehensives; the average for comprehensives set up by 1965

Table15. Raw exam results in each school group, by fifths of combined score at eleven and by social class

	n	% no exam	mean no of 'O' level equivalents by 1974	percentage with five or more 'O' level equivalents by 1974	mean no. of A level passes at school	% with one or more A level passes at school
top fifth						
nonmanual						
grammar	295	0%	6.5	81%	1.5	61%
secondary mod	22	5%	5.5	64%	0.7	27%
grammar+mod	317	0.3%	6.4	80%	1.5	58%
comprehensive	63	0%	5.2	59%	1.2	48%
transitional	124	0%	5.3	69%	1.1	47%
manual						
grammar	177	2%	5.0	59%	1.1	42%
secondary mod	40	0%	3.4	40%	0.2	10%
grammar+mod	217	1%	4.7	56%	0.9	36%
comprehensive	64	3%	4.8	64%	0.6	28%
transitional	108	0%	4.3	47%	0.9	38%
next to top fifth						
nonmanual						
grammar	111	0%	4.6	53%	0.7	35%
secondary mod	80	1%	2.4	19%	0.1	3%
grammar+mod	191	0.5%	3.7	39%	0.5	21%
comprehensive	44	0%	3.3	29%	0.5	25%
transitional	114	2%	3.1	28%	0.4	18%
manual						
grammar	116	3%	3.6	34%	0.4	23%
secondary mod	130	2%	1.9	15%	0.1	3%
grammar+mod	246	3%	2.7	24%	0.2	13%
comprehensive	126	2%	2.7	19%	0.2	7%
transitional	213	3%	2.1	17%	0.2	8%
middle fifth						
nonmanual						
grammar	18	0%	3.4	39%	0.2	11%
secondary mod	90	1%	1.4	7%	0.1	4%
grammar+mod	108	0.9%	1.7	12%	0.1	6%
comprehensive	54	4%	2.1	17%	0.2	11%
transitional	96	0%	1.7	13%	0.1	9%

cont'd

Table 15 cont'd

	n	% no exam	mean no of 'O' level equivalents by 1974	percentage with five or more 'O' level equivalents by 1974	mean no. of A level passes at school	% with one or more A level passes at school
middle fifth						
manual						
grammar	25	0%	1.7	12%	0.4	16%
secondary mod	241	7%	0.9	5%	0	1%
grammar+mod	266	7%	1.0	6%	0.1	3%
comprehensive	143	6%	0.9	3%	0.1	3%
transitional	210	5%	1.0	5%	0.1	3%
next to low fifth						
nonmanual						
grammar	-	-	-	-	-	-
secondary mod	66	9%	0.6	0%	0	2%
grammar+mod	66	9%	0.6	0%	0	1%
comprehensive	38	8%	0.7	0%	0	0%
transitional	62	6%	0.5	2%	0	3%
manual						
grammar	5	20%	0	0%	0	0%
secondary mod	229	10%	0.3	0%	0	0%
grammar+mod	234	10%	0.3	0%	0	0%
comprehensive	169	9%	0.4	2%	0	0.6%
transitional	250	12%	0.4	0.8%	0	0.4%
lowest fifth						
nonmanual						
grammar	-	-	-	-	-	-
secondary mod	35	20%	0.1	0%	0	0%
grammar+mod	35	20%	0.1	0%	0	0%
comprehensive	27	15%	0.4	0%	0	0%
transitional	27	11%	0.4	0%	0	0%
manual						
grammar	-	-	-	-	-	-
secondary mod	280	37%	0.1	0%	0	0%
grammar+mod	280	37%	0.1	0%	0	0%
comprehensive	208	28%	0.2	0.5%	0	0.5%
transitional	275	32%	0.1	0.4%	0	0%

equalled that of the 'selective' combination). The average number of 'A' levels was similar in the comprehensive sample and the grammar+moderns, except in the top fifth. There it was lower in comprehensives. It could not be seen from this table to what extent this might relate to differences between school groups in initial attainments within the fifths of ability at eleven.

It will be seen later that this was a problem, even with sophisticated statistical techniques, for the interpretation of analyses which grouped eleven year scores into fifths or thirds.

The observations above are detailed and so do not provide an answer to a simple question like 'which group is better?' One aim in this study was to obtain an overall average for each school group which took into account such variation between attainment groups and social classes but allowed an overall generalisation about the group. The advantage of the techniques used in this study was that they permitted correction for a number of other extraneous factors (sex, parents' interest, age range of school) which would have been shown only in a very much more complicated table.

This simple table may help interpretation of the corrected results presented in this chapter. The account of overall corrected averages for school types and for the comparison of 'grammar+modern' and comprehensives is intended to provide the nearest equivalent to a statistical 'summary' of the various school groups. This will be followed by discussion of the various subgroups of comprehensives and schools of different age ranges. Then a description will be provided of the incomplete data which were yielded on the different subgroups of pupils, grouped by eleven year attainment and social class. Table 15 above may illuminate that discussion, particularly.

Corrected averages were explored using a number of different measures of examination performance. Firstly, it was important to analyse the contribution that being in a grammar school as against a secondary modern, relative to being in a comprehensive, might make. This meant establishing the extent to which differences in examinaton results were attributable to that threefold distinction between types of secondary school, rather than to background and pre-existing initial attainments of individual pupils. It emerged that, on most measures, there were statistically significant differences between some school categories, differences not accounted for by pre-existing attainments, social class, parents' interest in schooling or sex. In most cases, these indicated that the grammar pupils had, on average, got further ahead in exams by 16 and in 'A' levels, after allowance for initial attainments, than had pupils elsewhere, on average. Beyond that, however, the 'selection' question was further explored by comparing the 'selective' grammar+modern combination's average exam results with the average for comprehensive pupils. The results of those comparisons are given after the school type comparisons, below.

In what follows it should be assumed, unless it is specifically stated otherwise, that the observations include allowance for one or more measures of attainment and/or 'ability' at eleven, and for the components of exam averages attributable to social class, sex and parents' interest, whenever exam results are ascribed to school category.

In Appendices S and P, analyses are described in more technical detail, for those who care to know the exact nature of the corrections for pre-existing attainments and home background. The average exam results on a limited set of measures for school groups without allowance for these non-school factors are included in this Chapter where thay may illuminate the intake-corrected figures. (Analyses are numbered as in Appendix S.)

(ii) Types of school: grammar, secondary modern, comprehensive

'School type' comparisons, using each exam measure in turn, were carried out (except in the case of proportions with no exams).

General measures of examination performance.

No exam results:

Did the propensity to go in for some exam differ between schools?

Uncorrected proportions entering no exam were:

grammar (100% = n = 747) 1%

comprehensive (n = 936) 10%

secondary modern (n = 1213) 14%

As would be expected, the proportions with no exam entries in each type of school differed. More importantly, there were virtually no grammar pupils with no exam entries. As a result, any more sophisticated analysis which compared the three types of school would have been virtually impossible, and certainly unreliable. A comparison was made, however, between comprehensives and secondary moderns. Given allowances for intake, the comprehensives did not differ significantly from secondary moderns.

1. Odds of taking no exams (statistically, no significant difference) *

	type of school (without grammars) odds ratio
secondary modern : comprehensive	1.09

* from now on, a lack of statistically significant difference will be signalled here as (n.s.). The statistical significance of any overall differences will be signalled as follows:

'5% level' - odds are one in twenty that differences occurred by chance-p<.05
'1% level' - odds are one in a hundred that differences occurred by chance-p<.01
'.1% level' - odds are one in a thousand that differences occurred by chance-p<.001

Exam results

On most other measures, there was a statistically significant relation of average examination achievement to type of school, with grammar schools achieving the highest average, and the comprehensive pupils either at a somewhat higher level than secondary moderns or at a similar level (after allowance for pre-existing attainment and so on). The exceptions, on which there were no statistically significant differences, were the average grade in CSE English, the average grade in 'O' level mathematics and 'A' level measures. On all three measures, the comparisons of school type groups were among candidates for the examination, rather than all pupils.

Numbers of 'O' level equivalents

Without allowances for attainment and social class at eleven years old, before the start of secondary school, the average number of 'O' levels/CSE grade 1's for grammar pupils was much the highest, and the mean for comprehensive pupils, while higher than that in secondary moderns, suggested that, on average, they were getting three and a half fewer 'O' levels than those in grammar schools. The figures were as follows:

grammar	5.12 'O' level equivalents	(n = 747)
secondary modern	0.98 ' '	(n = 1213)
comprehensive	1.60 ' '	(n = 936)

After allowance for initial attainment, social class and other factors, the intake-corrected mean for comprehensive pupils remained below that for grammar pupils. The comprehensive pupils gained, on average, slightly more 'O' levels than their secondary modern 'equivalents'.

2.

Numbers of 'O' level equivalents (whole sample) (statistically significant differences; $p < .001$)	type of school differences
grammar - comprehensive	.92
grammar - secondary modern	1.20
comprehensive - secondary modern	.28

The average numbers of 'O' level equivalents above are averages for the entire samples, including pupils with no 'O' level equivalents. Thus they measure 'O' level standard performance among the whole sample, rather than averages for those who had any 'O' level equivalents. The differences above therefore cannot be expressed in terms of 'the number of extra 'O' level passes' to be expected in one rather than another group, in the way that the next set of differences can. Further comparisons between the three school type groups were carried out to look at the average numbers of 'O' level equivalents gained by the pupils who had obtained at least one 'O' level equivalent. These were about half the pupils in the study. The results again showed that grammar school pupils, even after the customary corrections for initial scores and other background

factors, had the highest number of '0' level equivalents. Among secondary modern pupils with at least one such exam result, the lowest average number of '0' level equivalents was found, even with allowance for their average attainment at 11 and social class. The intake-corrected differences were:

3. Numbers of '0' level equivalents (those with one or more) ($p < .001$)	type of school differences
grammar - comprehensive	.61
grammar - secondary modern	.98
comprehensive - secondary modern	.37

To give an indicaton of the nature of those intake-corrections, the uncorrected differences were, in numbers of '0' level equivalents:

grammar - comprehensive	2.17
grammar - secondary modern	2.93
comprehensive - secondary modern	.76

From these crude comparisons it looked as if going to a grammar rather than a comprehensive would contribute a difference of just over two '0' level equivalents to the performance of the more academic half of the school population. Once the extent to which that was dependent on extraneous factors was lessened, the difference was reduced. The figures above can be regarded as suggesting that factors other than initial test attainment and background, which were connected with being in a grammar as against a comprehensive, contributed almost two thirds of an '0' level to progress by 16. Factors connected with being in a grammar rather than a secondary modern contributed roughly one '0' level pass, in these terms. Something to do with being in a comprehensive, rather than a secondary modern, similarly, contributed the equivalent of over a third of an '0' level to pupils' progress, provided those pupils already had one or more '0' level pass.

Differences between grammar pupils and the two other groups in the latter set of intake-corrected differences were smaller than the earlier set in which all pupils were included, simply because the many non '0' level pupils were in a majority in comprehensives and secondary moderns, but only a fraction of the grammar school sample. The differences above do not generalise to the other half of pupils, those with no '0' level equivalent. Confining the sample to the pupils with '0' level equivalents was regarded as a more satisfactory use of the statistical techniques employed here, since the numbers of pupils with each number of '0' levels were closer to those assumed by the techniques. Nevertheless, both sets of results for numbers of '0' level equivalents had similar implications, in that grammar pupils had the highest average and secondary modern pupils the lowest.

Grammar school pupils were also the likeliest to pass five or more 'O' level equivalents. Presenting results as intake-corrected odds ratios between the groups indicates that grammar pupils' odds of passing five or more 'O' level equivalents were 3.5 times those of equivalent secondary modern pupils but about twice those of equivalent comprehensive pupils. The odds for comprehensive pupils were thus one and a half times the odds for secondary modern pupils.

4. Odds of getting five or more 'O' level equivalents*

($p < .001$)	school type odds ratios
grammar : comprehensive	2.39
grammar : secondary modern	3.45
comprehensive : secondary modern	1.45

Without the corrections for differing attainments and social class at eleven, the proportions reaching this level in each type of school were as the left hand column below indicates. The nearest equivalents in corrected figures to the 'raw' proportions passing five or more 'O' levels in each type of school are shown in the right hand column below.

proportions with five or more 'O' level equivalents:

	'raw' proportions	'corrected' proportions
grammar	62%	45%
secondary modern	7%	19%
comprehensive	15%	26%

The figures suggest the advantage to those in a comprehensive as against in a secondary modern and the advantage to those in a grammar as aginst in other schools, in terms of chances of performance of this kind. There was, apparently, considerable variation in the likelihood of getting five or more 'O' level equivalents attributable to factors other than initial attainment or background, associated with school type.

Performance in single subjects: mathematics, English.

Introductory overview of findings

How did the three types of school compare on the individual subjects chosen for special attention by this study? Were there any 'school type' differences in mathematics or English examinations?

* The relative positions of the three school groups were more or less unchanged when general ability scores at 11, themselves contributing a highly significant amount, were also allowed for, in a supplementary analysis; Appendix E for detailed results of analysis No. 5.

Again, the chief observation was that, even after allowing for some background differences between the three groups of pupils and for the attainment differences that existed before the start of secondary schools, there was every sign that grammar school pupils had outstripped the rest, secondary modern pupils were at an overall disadvantage in exam results by sixteen, and comprehensive pupils had usually, on average, higher exam results than equivalent secondary modern pupils. On most measures of 'O' level standard performance, when comparisons were relative to all other pupils, grammar school pupils were ahead. The main exception was a lack of school type difference in the grades achieved by candidates in 'O' level mathematics. Comprehensive pupils tended to have somewhat better results than equivalent secondary modern pupils in analyses of proportions, which involved entire samples, and thus the full range of pupils. The intake-corrected proportions reaching 'O' level pass standard, for example, were appreciably higher in both maths and English in comprehensive than in secondary moderns, whereas, on the combined scales, which measured only pupils with graded results, this advantage to comprehensive pupils over secondary modern equivalents was not very marked. On no measure involving the full range of pupils was the difference between comprehensive and secondary modern averages as great as the differences between each of them and the average for grammar pupils. The figures on which these observations are based now follow.

Performance in mathematics 'O' level GCE & CSE by 16 years old

Proportions obtaining an 'O' level pass or equivalent in mathematics

Consider first the minority (overall proportion of 23%, or 26% of the sample which included transitional pupils) who obtained the equivalent of a pass at 'O' level in mathematics. The proportions explored were of the whole sample, including pupils who did not enter for public exams.

The highest proportions obtaining an 'O' level pass standard in mathematics were found in grammar schools, and the lowest in secondary moderns. The odds of grammar school pupils reaching this level were over twice the odds for comprehensive pupils, and over three and a half times the odds for secondary modern pupils.

6. 'O' level 'pass' in mathematics
(whole sample) ($p < .001$)

types of school	odds ratios
grammar : comprehensive	2.30
grammar : secondary modern	3.53
comprehensive : secondary modern	1.53

The equivalent 'raw' and 'corrected' proportions were:

	'raw' proportions	'corrected' proportions
grammar	64%	43%
secondary modern	9%	18%
comprehensive	18%	25%

Grade at CSE mathematics or 'O' level mathematics

The combined CSE/GCE scale

The overall comparison between those grammar, secondary modern and comprehensive pupils who got a grade in CSE or GCE 'O' level mathematics indicated statistically significant differences, with grammar pupils achieving most highly on this measure, and comprehensive pupils considerably lower, though the latter were not scoring quite as low, on average, as secondary modern pupils.*

8. Grade of maths at CSE or 'O' level (weighted 6,5,4,3,2,1 among those with a grade in either exam)

($p < .001$) type of school differences	
grammar - comprehensive	.53
grammar - secondary modern	.68
comprehensive - secondary modern	.15

Mathematics grade at CSE separated from mathematics 'O' level

In order to see whether components of the findings on the combined measure were attributable specifically to the performance of CSE candidates, it was necessary to separate and analyse the grades obtained in CSE exams. This also permitted analyses to be more sensitive to variation among those achieving lower grades at CSE (though averages included pupils with any grade in CSE mathematics).

The sample consisted of 950 pupils (1554 when transitional pupils were included). Of these, relatively few (under 10 per cent) were grammar school pupils (more than would be expected to go in for 'O' level). Nevertheless, those who were in grammar schools produced an overall average higher than either secondary modern or comprehensive pupils. The latter two groups were closely similar to each other in average grade in mathematics at CSE.

* Two alternative weightings were tried. The results had similar implications: see Appendix W, and Appendix S for other details of analysis No. 7.

9. Grade in CSE mathematics (among those with a grade in CSE maths) ($p < .001$)

type of school differences	
grammar - secondary modern	.62
grammar - comprehensive	.66
secondary modern - comprehensive	.05

Grade at mathematics 'O' level, separated from CSE

There was no statistically significant difference between the groups of pupils who got a graded result at 'O' level in the average grade obtained, whether they were in grammars, secondary moderns or comprehensives.*

10. Grade in 'O' level mathematics (among those with a grade in 'O' level maths) (n.s.)

type of school differences	
secondary modern - grammar	.07
secondary modern - comprehensive	.11

It is worth recalling that the majority of pupils with a grade at 'O' level mathematics (pass or fail) were in grammar schools. Those with a grade in other schools did not seem to have worse or better grades. The numbers of pupils in each school type were as follows:

Number with graded 'O' level mathematics (for purposes of analysis)

	grammar	secondary modern	comprehensive	total in one type of school
	539	74	159	772
row per cent	70%	10%	21%	100%
per cent of school type group	72%	6%	17%	27%

Performance in English 'O' level GCE and CSE by 16 years old

The English results were for the most part parallel to results in mathematics. Grammar pupils had the highest odds of passing at 'O' level standard and highest average grades on the combined scale. A small advantage to comprehensive pupils over their equivalents in secondary moderns was observable. In separate graded CSE performance and graded 'O' level, the obverse of the mathematics results, in a sense, was found, in that the significant school type difference occurred in 'O' level grades rather than CSE.

* Appendix E contains fuller results of additional controls for other test scores, which do not affect this observation.

11. '0' level 'pass' in English language (or CSE grade 1 English)

($p < .001$)	school type odds ratios*
grammar : secondary modern	5.39
grammar : comprehensive	3.28
comprehensive : secondary modern	1.64

13. Grade at CSE English or '0' level English language (among those with grade) (weighted 6,5,4,3,2,1)**($p < .001$)

	type of school differences
grammar - comprehensive	.61
grammar - secondary modern	.79
comprehensive - secondary modern	.18

English grade at CSE separated from English language '0' level

There was no significant relation of average CSE grade in English to type of school. (Only 18 grammar school pupils obtained any grade at CSE English and could be included in the total sample for this analysis of 1136, or 1796 including the transitional pupils). Though grammar pupils obtained the highest average score, this was statistically non-significant. Comprehensive and secondary modern pupils were virtually identical in their overall average grade.

14. Grade in CSE English (among those with grade in CSE English) (n.s.)

	type of school differences
grammar - comprehensive	.19
grammar - secondary modern	.18
secondary modern - comprehensive	.01

Grade at English language '0' level separated from CSE

The investigation was confined to those with some grade at '0' level English language (a sample of 1076, or 1513 including the transitional pupils). This included a far higher proportion of grammar school pupils than among the CSE English candidates; indeed they formed an overall majority of those in one type of school with a grade A to E in English language.

*The corresponding 'raw' and 'corrected' proportions were:

	'raw'	'corrected' proportions
grammar	79%	61%
secondary modern	16%	23%
comprehensive	27%	32%

**Weighting: See Appendix W.

Numbers with graded 'O' level English language (for purposes of analysis)

	grammar	secondary modern	comprehensive	total in one type of school
	656	152	268	1076
row per cent	61%	14%	25%	100%
per cent of school group	88%	13%	29%	37%

The grammar school pupils gained a significantly higher average grade at 'O' level English language. The average for comprehensive pupils was between that and the average for secondary moderns. Grammar pupils were getting significantly higher grades, while comprehensive pupils were close to the secondary modern pupils' level.

15. Grade in 'O' level English language (among those with grade) ($p < .001$) type of school differences

grammar - secondary modern	.37
grammar - comprehensive	.28
comprehensive - secondary modern	.10

'A' level results

UCCA Scale

One analysis compared groups of those pupils who passed one (or more) 'A' level, on the numbers of passes and their grades. The relevent 'raw' averages, among those with an 'A' level pass, were:

	'raw' UCCA score mean	n
grammar	6.90	326
secondary modern	5.25	24
comprehensive	5.90	80

18. Number and grades of 'A' level (among those with one or more pass. Score ranges 1 to 20) (n.s.) type of school differences

grammar - secondary modern	.71
grammar - comprehensive	.38
comprehensive - secondary modern	.33

There was no statistically significant overall difference related to the type of schooling received, once allowance had been made for background factors and levels of attainment before the start of secondary schooling.

Number of 'A' level passes obtained by all 'A' level candidates

The lack of differences associated with school category, observed when numbers and grades of 'A' level were combined, was also found when the criterion was the number of 'A' level passes.

16. Number of passes at 'A' level (all candidates)

	(n.s.) type of school differences
grammar - secondary modern	.18
grammar - comprehensive	.08
comprehensive - secondary modern	.10

Within that limited sample of high attaining, 'A' level standard pupils, then, school type did not seem to be associated with overall differences. This was despite the fact that the overall average number of 'A' levels among candidates differed for the three school types, when there was no correction for pre-existing attainment and background.

'raw' average number of 'A' level passes among candidates (n = 510)

	mean	n
grammar	2.13	368
secondary modern	1.32	34
comprehensive	1.69	108

Another comparison of pupils in the three types of school explored the average numbrs of 'A' level obtained, taking the average across all pupils. The crude averages for the whole sample, too, including all those who did not attempt 'A' levels, differed between schooltypes.

'raw' average number of 'A' level passes among all pupils (n = 2896)

	mean	n
grammar	1.05	747
secondary modern	0.04	1213
comprehensive	0.19	936

Even after corrections for pre-existing attainment, there was a significant difference associated with school type, with that sample; the average for grammar pupils exceeded both that for comprehensive pupils and that for secondary modern pupils.

17. Number of passes at 'A' level (all pupils) ($p < .001$)

type of school differences	
grammar - secondary modern	.39
grammar - comprehensive	.31
comprehensive - secondary modern	.08

(iii) Grammar+modern combination vs comprehensive

The following table summarises the comparisons between 'selective' and 'nonselective' pupils, that is, between grammar+modern and comprehensive pupils. It will be seen that, on most measures, there was no difference between the two groups, on average, after correction for attainment and background at eleven. Details and a similar table of all analysis results are in Appendix S, Volume 2.

Key: prop. = proportion
dep. var. = dependent variable, outcome measure of examinations

Table 16 corrected examination results in selective/nonselective schools - grammar+modern vs comprehensive comparisons:

examination measure	number in analysis sample	result	'variance explained'	comment
25. proportions with no exam (overall prop = 9%)	n=2896	grammar+modern = comprehensive	53%	no difference, school X^2 is 0.5% of total X^2
26. number of 'O' level equivalents (0-9+) (dep. var. mean = 2.25)	n=2896 (sd=2.87)	grammar+modern = comprehensive	59%	no difference
27. number of 'O' level equivalents (1-9+) (dep. var. mean = 4.26)	n=1527 (sd=2.66)	grammar+modern = comprehensive	46%	no difference
28. proportions with five or more 'O' level equivalents (overall prop = 23%)	n=2896	grammar+modern = comprehensive	80%	no difference school X^2 is 0.01% of total X^2
29. proportions with five or more 'O' level equivalents (overall prop = 23%)	n=2896	grammar+modern = comprehensive	78%	no difference
30. proportions passing maths at 'O' level standard (overall prop = 26%)	n=2897	grammar+modern = comprehensive	88%	no difference school X^2 is 0.05% of total X^2
31. grade in mathematics on combined 'O' level/CSE scale of grades (weighted 6 to 1)(dep. var. mean = 3.17)	n=1719 (sd=1.71)	grammar+modern = comprehensive	44%	no difference reaching statistical significance
32. grade in mathematics CSE (dep. var. mean = 2.66)	n=950 (sd=1.30)	grammar+modern = comprehensive	25%	no difference
33. grade in mathematics 'O' level (dep. var. mean = 3.42)	n=772 (sd=1.18)	grammar+modern = comprehensive	19%	no difference
34. proportions passing English at 'O' level standard (overall prop = 34%)	n=2897	grammar+modern = comprehensive	87%	no difference school X^2 is 0.2% of total X^2
35. grade in English on combined 'O' level/CSE scale of grades (weighted 6 to 1) (dep. var. mean = 3.42)	n=2130 (sd=1.5)	grammar+modern slightly higher	44%	small difference, comp slightly lower
36. grade in English CSE (dep. var. mean = 3.04)	n=1136 (sd=1.18)	grammar+modern = comprehensive	24%	no difference
37. grade in English language 'O' level (dep. var. mean = 3.43)	n=1076 (sd=1.13)	grammar+modern slightly higher	27%	comp about one-sixth of a standard deviation lower
38. number of 'A' level passes (dep. var. mean = 1.98)	n=510 (candidates sd=1.18)	grammar+modern = comprehensive	10%	no difference
39. UCCA scale (dep. var. mean = 6.63)	n=430 (sd=4.14)	grammar+modern = comprehensive	16%	no difference
40. Number of 'A' level passes (incl. sixth form college) (dep. var. mean = 0.38)	n=2919 (all pupils)	grammar+modern = comprehensive	36%	no difference

Various measures of examination performance, considered indicative of school attainments, were explored for overall relative attainment of grammar+modern as against comprehensive pupils.

General measures:

No exams

It was of interest firstly to look at proportions not entering for any exams at all. Without corrections for pre-existing attainments and social class the two groups did not differ; ten per cent of comprehensive pupils and nine per cent of the grammar+modern group took no exams. The negligible proportions with no exams among grammar pupils and the higher proportion among secondary modern pupils were such that, if grammar pupils had been one fifth of the grammar+modern combination, the proportion with no exams in the grammar+modern combination would have been higher:

'raw' proportions with no exams

grammar+modern	9%
grammar+modern reweighted	11%
comprehensive	10%

Without such reweighting but after 'intake' corrections, the relative odds of taking some examination did not differ, it was found, between the 'selective' and 'comprehensive' groups.

25. Odds of taking no exams (whole sample) — school odds ratio

comprehensive : grammar+modern (n.s.)	1.19

Exam measures

As measured by 'O' level 'passes', or equivalent standard (grade 1) at CSE, uncorrected for initial attainment before secondary school, the average for grammar+modern pupils at sixteen years old exceeded that for comprehensive pupils. Moreover, even when the grammar+modern average was reweighted to base it on something like the proportions of grammar and secondary modern pupils found under full selection, their average at sixteen, as was shown in Chapter I, remained at a level very slightly above that of the comprehensive pupils.

	'raw' average number of 'O' level equivalents (among pupils with one or more)
grammar	2.56
reweighted grammar+modern	1.81
comprehensive	1.60

However, without any reweighting but with the usual allowance for extraneous factors, it was found that there was no difference between the comprehensive pupils as a whole, on average, and the grammar+modern combination.

26. Numbers of 'O' level equivalents (whole sample)

	school difference
grammar+modern - comprehensive (n.s)	.04

It was also the case among those with at least one 'O' level equivalent that grammar+modern and comprehensive pupils had made the same sort of progress. The 'raw' average number of 'O' level equivalents among those with at least one 'pass' of this standard averaged out higher, without corrections for attainment, for the grammar+modern mixture.

	'raw' average number of 'O' level equivalents
grammar+modern (n = 1102)	4.55
comprehensive (n = 425)	3.52

The 'raw' schooling difference was about one 'O' level. The extent to which this was a function of prior attainments and background was then reduced by corrections, and the result was a lack of any statistically significant difference.

27. Number of 'O' level equivalents (among those with one or more 'O' level equivalent)

	school difference
grammar+modern - comprehensive (n.s.)	.10

Though a statistically acceptable measure, the numbers of 'O' level equivalents were based for this analysis on 90 per cent of grammar pupils, 35 per cent of secondary modern pupils, 56 per cent of the grammar+modern combination and 45 per cent of comprehensive pupils. Nevertheless, to the extent that this result can be interpreted, it would suggest that, within a group demonstrating their ability to do some subject to 'O' level standard, 'equivalent' children were getting comparable numbers of 'passes' in comprehensive and grammar+modern schools. Nor did the proportions obtaining five or more passes at 'O' level (grades A to C or CSE grade 1) differ between those two groups, after corrections for pre-existing attainment:

28. Odds of getting five or more 'O' level equivalents (whole sample)

	school odds ratios
grammar+modern : comprehensive (n.s)	1.04

This was despite the fact that, as was seen in Chapter III, the proportion with five or more 'O' levels at sixteen was far higher in the grammar+modern group, before 'intake correction'. Even when reweighted to make grammar pupils one fifth of the 'selective' combination, the grammar+modern average, without corrections, was very slightly higher than the comprehensive average.

	'raw' proportions with five or more 'O' level equivalents
grammar+modern	28%
reweighted grammar+modern	18%
comprehensive	15%

Scrutiny of general examination attainment at 16-plus, therefore, revealed no overall difference attributable to being part of a 'selective' combination.

Performance in single subjects: mathematics, English

Without taking into account whether results were attributable to attainment before secondary school, the crude numbers of sixteen year olds' 'O' level standard 'passes' in mathematics and English (language) were higher in the grammar+modern combination than in the comprehensive sample.

	'raw' proportion with 'O' level 'pass'	
	in mathematics	in English
grammar+modern	31%	41%
comprehensive	18%	27%

When the grammar+modern group was reweighted to make grammar pupils one fifth and secondary modern pupils four-fifths, however, the raw, reweighted proportion passing maths was less markedly higher than the comprehensive group's and the proportion passing English was even closer to that of the comprehensive pupils.

	'raw' proportions with 'O' level 'pass'	
	in mathematics	in English
reweighted grammar+modern	20%	29%
comprehensive	18%	27%

Without such reweighting, but with corrections for background factors and for either mathematics or reading test performance at eleven, there were no consistent differences between the two groups in average performance in examinations in the two subjects.

In mathematics, no overall schooling differences emerged between the grammar+ modern and the comprehensive pupils. That is, the corrected proportions of children achieving an 'O' level pass (grades A to C or CSE grade 1) in mathematics were comparable in the two groups:

30. 'O' level "pass" in mathematics (whole sample)

	school odds ratios
grammar+modern : comprehensive (n.s.)	1.10

The average grades, moreover, for the groups also showed no statistically significant difference attributable to secondary schooling. This was true for three measures; firstly, the scale which combined the CSE grades obtained with grades for 'O' level candidates:

31. grade of mathematics CSE or 'O' level (among those with grade in either exam

	schooling difference
grammar+modern - comprehensive (n.s.)	0.13

It was also true of averages on the separate scale of grades obtained by CSE candidates:

32. grade in CSE mathematics (among those with grade in CSE maths)

	schooling difference
grammar+modern - comprehensive (n.s.)	0.01

and in terms of average 'O' level grades looked at alone:

33. grade in 'O' level mathematics (among those with grade in 'O' level maths)

	schooling difference
grammar+modern - comprehensive (n.s.)	0.05

Two measures, of the four used to explore English exam results, also showed no difference. The proportions 'passing' at 'O' level standard were comparable; and the average grades of CSE candidates were similar.

34. 'O' level "pass" in English (whole sample)

	schooling odds ratio
grammar+modern : comprehensive (n.s.)	1.19

36. grade in CSE English (among those with grade in CSE English)

	schooling difference
grammar+modern - comprehensive (n.s.)	0.01

By the two other measures, though, the intake-corrected averages differed slightly. Comprehensives produced a statistically significantly lower average grade in English 'O' level, when all grades (A to E) were included. The difference was not large - perhaps representing a fifth of a grade at most.

37. grade in 'O' level English (among those with grade in 'O' level English)

	school difference
grammar+modern - comprehensive ($p < .01$)	.19

The results of comprehensive pupils were statistically significantly lower, too, on average, on the combined scale of 'O' level and CSE grades in English.

35. grade of English CSE or 'O' level (among those with grades in either exam)

	schooling difference
grammar+modern - comprehensive ($p < .05$)	0.12

The lack of difference on the two former measures of English performance suggest that the schooling difference found on the combined measure might reflect the grades obtained by those pupils who did get a grade at 'O' level English language.

'A' level results

Using two measures of 'A' level performance - the numbers of 'A' levels each pupil obtained and the combination of grades obtained for all 'A' levels - the results for the two groups were compared. The mean numbers of 'A' level passes without corrections for extraneous factors were rather higher in the grammar+modern combination than in comprehensives, both among the candidates and among the whole sample.

	'raw' average number of 'A' level passes: among candidates: (N=510)		among whole sample: (N=2918)	
grammar+modern	2.98	(n=402)	0.42	(n=2002)
comprehensive	1.69	(n=108)	0.19	(n=916)

In order to reweight the average for grammar and secondary modern candidates, it would be necessary to make some dubious assumptions about the proportion who 'ought' to have taken 'A' levels in secondary moderns and in grammars, under selection. It is possible, however, to reweight the raw average for the entire samples, as if four-fifths went to secondary moderns. This brings the grammar+modern average down nearer the level of comprehensive pupils, but a little above.

	'raw' average number of 'A' level passes: among whole sample
reweighted grammar+modern	0.24
comprehensive	0.19

On the UCCA scale of numbers of 'A' levels by their grades, the 'raw' average for the grammar+modern pupils with A level passes was a little higher than that for comprehensive 'A' level pupils.

	'raw' average on UCCA scale among those with one or more 'A' level pass.	
grammar+modern	6.79	(n=350)
comprehensive	5.90	(n= 80)

Again, reweighting was not attempted for averages among such a restricted sample. Without reweighting, however, but after correction for 11 year old attainment and other background factors, neither measure of 'A' level performance appeared to show a difference among candidates for 'A' levels, attributable to schooling.

38. Number of passses at 'A' level obtained by 'A' level candidates

	schooling difference
grammar+modern - comprehensive (n.s.)	0.06

39. Combination of grades of all 'A' levels: (of all those with one or more 'A' level pass)

	schooling difference
grammar+modern - comprehensive (n.s.)	0.29

So equivalent pupils who went in for 'A' levels obtained equivalent results in either group. The comparison of the entire groups (with or without 'A' levels) of 'selective' and comprehensive pupils was not carried out in the course of the analyses. It was, however, possible to approximate a likely figure for that comparison by subtracting the comprehensive pupils' average from a weighted* average of the means for grammars and for secondary moderns, derived from the school type analysis reported earlier (17):

Number of passses at 'A' level (all pupils)

	schooling difference
grammar+modern - comprehensive	0.07

(The use of this recalculated grammar+modern average assumes a lack of interaction of school type and eleven year score, which was not tested for the original analysis.) The results show a lack of any difference attributable to the grammar+modern/comprehensive distinction.

In addition, the entire groups were compared on numbers of 'A' levels after correction for test scores at sixteen years old. This gave corrected averages for all pupils including those who were not 'A' level candidates.

The pupils in the two groups looked at as a whole seemed to create equivalent averages for 'A' level results, even though most of them were not 'A' level candidates, once allowance was made for reading and maths test attainments at sixteen:

40. Number of passes at 'A' level (out of all pupils)

	schooling difference
grammar+modern - comprehensive (n.s.)	0.06

So there was no evidence that, given pupils of equivalent attainment at 16, the school groups would produce any different 'A' level exam results two years later.

(iv) Different sorts of comprehensive relative to grammar+modern pupils.

It would be wrong to draw conclusions about comprehensives in general, without pointing out where exceptions existed or conclusions applied only to an identifiable subgroup. So attempts were made to compare different subgroups with the grammar+modern combination, and it will be recalled that in addition to overall averages for comprehensive pupils a number of differences between comprehensives, including the ways in which they were set up ('origins') and the

*Weighted according to actual numbers of grammar and modern pupils

length of time they had existed ('dates') were investigated by this study. A further distinction drawn between subgroups of comprehensive pupils separated those in schools starting at eleven years old ('11+' comprehensives) from the pupils in schools which appeared to be part of a middle school system, in that their youngest pupils were 12 or 13 years old ('12+13+' comprehensives). Most results of the comparison are reserved for Appendix C, where they are given in detail. It is possible in this section though, to give a brief summary of the role of this factor of 'age of youngest pupil' or 'age range of school'. The results for pupils in comprehensives of different origins are presented in Appendix A. Discussed here, and in more detail in Appendix D, are the results for pupils in comprehensives of different dates, as compared to grammar+modern pupils and to the pupils whose schools were reorganised during their secondary school years, the 'transitional' pupils.

One overall, if somewhat contrived, rationale for this, in a study exploring different facets of the question of whether 'selection' made a difference, is that comprehensives differ in the degree to which they themselves are selected from or select. At least, certain types of comprehensives in particular localities may be oversubscribed, some undersubscribed. Then, too, there seemed to be differences in intake between comprehensives. Were there also differences reflected in outcomes after allowing for these intake differences? The comparisons which were made depended on whether distinctions between comprehensives could be explored among the information collected by NCDS. The question was whether any distinct type of comprehensive was consistently associated with more or less progress, better or worse behaviour, or higher or lower examination results. And were the most striking differences to be found among groups of comprehensives, or between certain types of comprehensives and the combination of grammar and secondary modern pupils?

It will be remembered that findings here are specific to that particular period of comprehensive reorganisation in the early 1970s. Especially when detailed claims may be made concerning pupils in comprehensives of a particular era, it is important to remember how different sorts of schools can change within a matter of years.

Subgroups of comprehensives: grammar+modern/dates of comprehensive

It seemed worthwhile, nevertheless, to explore whether differences in results seemed to be associated with whether comprehensives had been set up recently (1966 to 1969) or at least four years before NCDS members started secondary school. It could be conjectured, for example, that comprehensives set up immediately before the study child entered might have teething problems. Most comprehensives in the study were fairly recently established, at most two or three years before the arrival of the children whose progress is studied here. Some, though, had been set up in 1965 or earlier. Assuming the latter group would have had more chance to settle, perhaps these 'early' comprehensives would be better examples of 'comprehensives' than the, as yet unsettled, recent comprehensives. As well as distinguishing recent from early comprehensives, this study also distinguished the schools which were still in the process of reorganisation while study members were at them, the transitional schools. Their results, as it happened, were of particular interest.

The comparisons between early and recent comprehensives will be summarised first, then transitional pupils will be compared to the rest. The relevant results for grammar+modern pupils, relative to the previous three groups, are then mentioned. The figures on which these summaries are based, with graphs of interactions not included in this section, are in Appendix D.

Some small differences emerged between early and recently-established (1966 to 1969) comprehensives, but their direction depended on the measure of examination results used, and there were many similarities. The areas of difference and similarity will be summarised next, and then an impression of results for each group will be offered. Details are in Appendix D.

Summary of measures on which differences and similarities were found:

First, the similarities; on six of the fourteen exam measures, the early and recent comprehensives did not differ from each other or from the grammar+modern combination. The corrected proportions passing 'O' level maths and 'O' level English resembled each other, and there was no significant overall difference in average grades at CSE in either maths or English, nor in 'A' level results (by the two measures used).

On two other indices of exam performance, the combined 'O' level/CSE scale of grades of English exam and the average number of 'O' levels obtained by those who got one or more 'O' level, early and recent comprehensive groups were similar to, but slightly lower scoring - perhaps not importantly so - than the grammar+modern combination. This was not reflected in the recent comprehensive pupils' 'O' level equivalents averaged over the <u>whole range</u> of pupils, including the half who got no exam result at this level.

This last was one of the six measures of exam performance on which there did seem to be differences between early and recent comprehensives. Recent comprehensives seemed to produce <u>slightly</u> higher numbers of 'O' level equivalents, averaged over all the pupils, and in this respect were like the grammar+modern group. So, among all early comprehensive pupils, there were <u>slightly</u> fewer 'O' levels (relative to background factors and attainments) by this measure. Yet the proportions obtaining <u>five or more</u> 'O' level equivalents in early comprehensives were relatively high, perhaps slightly higher than among pupils in recent comprehensives and than the grammar+modern combination, but notably higher than transitionals. The odds of having <u>no</u> examination entries, on the other hand, were also somewhat higher in early comprehensives - probably significantly higher than in the grammar+modern combination and somewhat higher than in recent comprehensives. Perhaps this was behind the varying implications of the three measures of 'O' level equivalent performance.

There were only two other measures on which it was worth distinguishing an apparent difference between early and recent comprehensives. These were two measures on which performance was confined to candidates getting a graded result, regardless of the numbers with no exam entries. The early comprehensives were slightly higher scoring than more recent comprehensives, and similar to the grammar+modern combination. Pupils of recent comprehensives were, like transitional pupils, averaging out slightly lower than the grammar+modern combination. These measures were the combined scale of maths

grades at 'O' level or CSE, and the scale of graded 'O' level performance in English. (A similar pattern, though not at an acceptable level of statistical significance, was found in grades in 'O' level maths among those with grades in 'O' level maths.) So, on those last five measures, though the pupils of early comprehensives were not exactly similar to those in recent comprehensives, such small variation as there was did not bear a consistent implication that either group was relatively higher scoring. Nor did they suggest any clear interpretation. Most of the significance of the results, on which statistically significant variation was found according to the 'grammar+modern/dates' categorisation, lay in the lower performance of pupils of transitional schools.

Of the fourteen analyses using this grammar+modern/dates grouping, in only four were interactions between test score and that school grouping tested, and in each analysis, for some, at least of the test scores, interactions were indicated. That is, in proportions with no exams there were statistically highly significant interactions between school categories and each of the three test scores (corrected for both other scores). In odds of five or more 'O' levels, the interaction with maths score reached a lower level of statistical significance, after allowing for reading and general ability; there were no interactions with other scores. The proportions passing at 'O' level standard in maths and in English also showed interactions with the one test score involved.

Early comprehensives

Pupils of schools which went comprehensive by 1965 had, overall, relatively high proportions with no exams and slightly lower average numbers of 'O' levels, than grammar+modern or recent comprehensive pupils, after correcting for all test scores at 11. Nevertheless, they had the highest corrected odds overall, of any group, of getting five or more 'passes' at 'O' level standard. Apart from that, they did not differ, overall, from the grammar+modern group, on any other measures, either in the individual subjects, maths and English, or on 'A' level measures.

The findings for non-exam pupils were somewhat modified by interactions, in so far as these could be interpreted (see Appendix D). Among those in the lowest third on eleven year tests, early comprehensives did not appear to differ from the grammar+modern combination as they had overall. The overall higher odds of no exam appeared to be among those of a higher eleven year old score in early comprehensives, and it is not clear what the implications of these findings are. The children amongst whom one might expect meaningful average differences or similarities in whether or not they took any exam would be the lower attainers.

The interactions in proportions passing five or more 'O' levels are a little clearer. The overall higher odds of those in early comprehensives seemed chiefly to derive from the performance of those of middle, average scores or even low scores in maths at eleven, and perhaps the next to top fifth, at least after allowance for reading test score (and an echo of this pattern was found, at a level of statistical significance too low for acceptance, in interaction with reading test score allowing for mathematics; see Appendix D). The odds of five or more 'O' levels among pupils whose scores at eleven put them in the top

scoring third (and who were therefore, the more likely to pass five or more 'O' levels) were closely similar in both groups of comprehensive pupils to those of grammar+modern pupils (figure 1)*. In maths and English passes, the early comprehensive pupils in the next-to-top fifth seemed to have a slightly higher proportion 'passing', but, for those in the top fifth, being in an early comprehensive was associated with somewhat lower odds of obtaining an 'O' level standard pass in these subjects than being in a recent comprehensive or the grammar+modern mixture, or even a 'transitional' school.

Recent comprehensives

The pupils of comprehensives set up between 1966 and 1969, shortly before they started secondary school, were similar overall, on average, in general measures of exam results corrected for attainments, to the grammar+modern combination. They were closely similar, too, in odds of passing 'O' level maths or English to grammar+modern and early comprehensive pupils. Those in recent comprehensives who obtained the relevant grades had slightly lower averages in English 'O' level and on the combined scale in maths, but they were otherwise like those in early comprehensives overall. Candidates for 'A' level were not different, apparently, in recent comprehensives.

The few interactions available did not qualify this picture. Recent comprehensive pupils who had average and above average scores at eleven, even the top fifth, were fairly similar to grammar+modern pupils. The possible exception was in odds of passing English at 'O' level standard, where recent comprehensive pupils were similar to transitional pupils (as those with an 'O' grade had been overall).

Transitional pupils

One of the most consistent and striking findings was of a tendency for pupils in schools undergoing reorganisation while they were there to have lower exam attainment. Though similar to the grammar+modern groups in proportions with no exams (and this was so when distinctions over the range of eleven year scores were made), they had lower average numbers of 'O' level equivalents and lower odds of passing five or more, of an 'O' level pass in mathematics and perhaps even of passing English at 'O' level standard, than any other group. (The grades obtained by those with grades in English or mathematics did not differ systematically; in general, transitional pupils were similar to, or sometimes higher scoring than pupils in recent comprehensives. Nor did 'A' level candidates differ significantly).

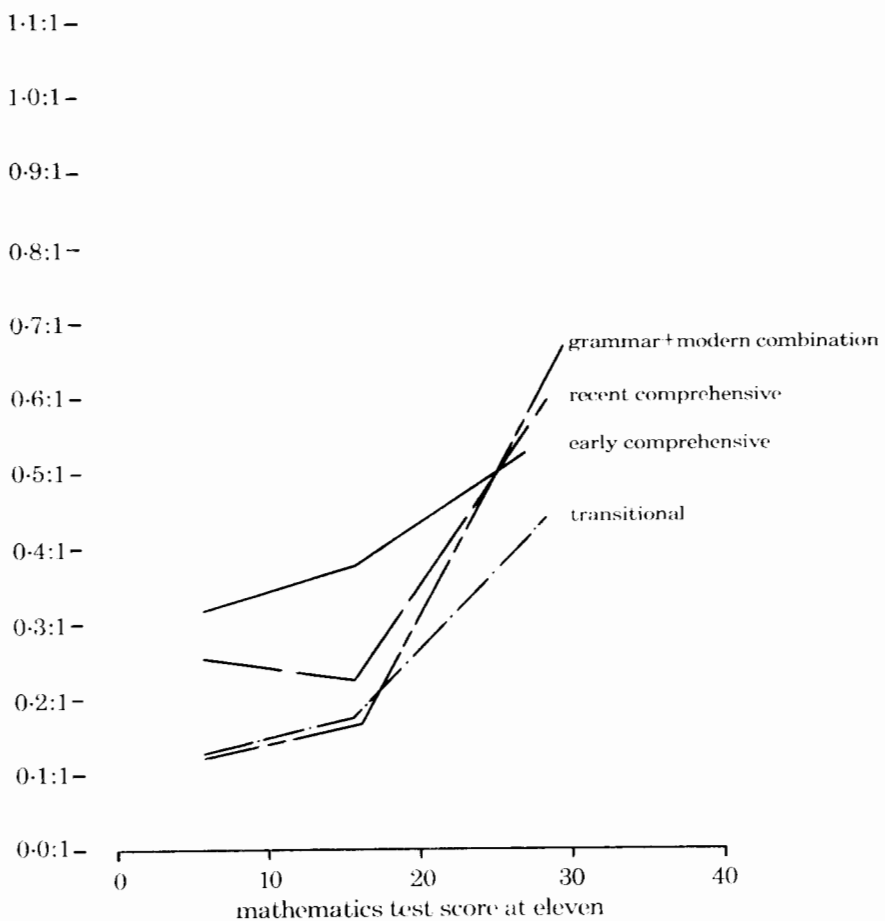
* It will be recognised that patterns in some of the graphs could be deceptive. The presentation of relative odds as corrected constants rather than ratios tends to exaggerate findings, since a ratio of 6:4 looks twice as large as a ratio of 3:2, even though it is an exactly similar difference in terms of the odds ratios which matter. This exaggeration is more marked the higher the examination performance, so is most deceptive for the top scorers at 11.

figure 1

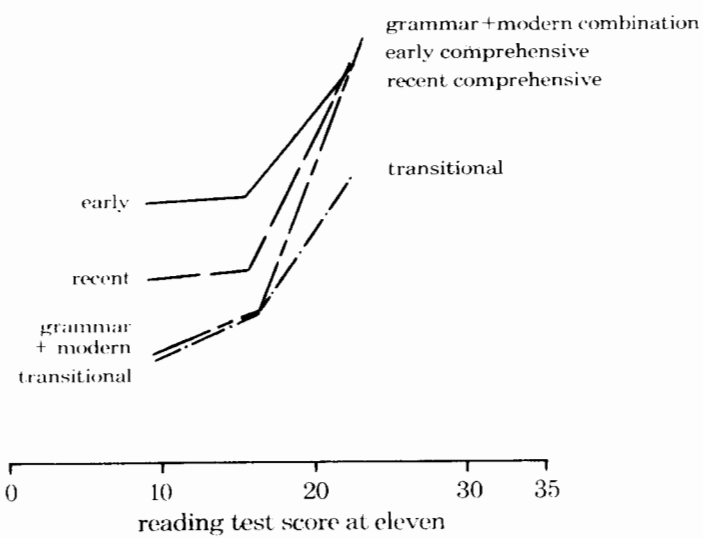
Odds of obtaining five or more 'O' level equivalents : interactions of school variable with test score at 11

grammar+modern/dates × mathematics at 11 (thirds) with corrections for reading and 'general ability' at 11 (thirds), parents' interest and social class

($p<0.01$)



grammar+modern/dates × reading score at 11 (thirds) with corrections for maths and 'general ability' at 11 (thirds) parents' interest and social class. Statistically non-significant (p approximately 0·06)



Pupils whose schools were undergoing reorganisation along comprehensive lines scored lower on all measures of 'O' level and CSE grade 1 performance. Transitional pupils produced the lowest average, after adjustment for intakes, on numbers of 'O' level equivalents. This was so, though neither pattern reached a high level of statistical significance, among all pupils, where they scored even lower than pupils in early comprehensives, and among those with at least one 'O' level equivalent. They were also particularly unlikely to reach the qualification of five or more 'O' level equivalents, if anything less likely than the pupils of the grammar+modern combination, and certainly apparently less likely than either subgroup of comprehensive pupils.

It thus appeared that, according to several of the single subject measures, where analyses were restricted to those with a graded pass in the subject (or in 'A' level measures, to the minority who were entrants), the transitionals were not necessarily distinguished from others. On some measures with a fairly wide range, such as the combined maths scale and the 'O' level grades in English, they were equal lowest with pupils of recent comprehensive. But on the five measures which compared among wider ranges of attainment, the transitionals were lowest achieving. That is, on average numbers of 'O' level equivalents, proportions 'passing' five or more 'O' level equivalents, proportions reaching 'O' level standard in English, proportions reaching 'O' level standard in mathematics and average grade on the combined scale of 'O' level and CSE in English, transitional pupils were lowest. A slightly different pattern in the comparison of different groups on proportions with no entries for examinations at all was found. In that respect, transitional schools did not differ from the grammar+modern combination, who were the least likely to have non examinees.

Since the different school groups varied in proportions obtaining five or more 'O' levels over the range of eleven year score, it could be seen that, while transitional pupils were very like the grammar+modern mixture among those of average and below average scores at eleven, with similarly low odds of five or more 'O' levels, those in the top third (for maths at least) had relatively low odds (after allowing for other scores). They had relatively low proportions in the next-to-top fifth, it seemed, getting five or more 'O' levels, compared to those in 'early', relatively well-established comprehensives. But those two groups were similar when pupils in the top fifth on maths were compared. (There was no statistically significant interaction distinguishing the 'top fifth' on reading at 11, in this way). A similar pattern was found in proportions passing 'O' level maths, with the next to top fifth of transitional pupils particularly low scoring, but on that measure of exam results those in early comprehensives with highest maths scores at 11 had even lower odds of passing maths than the transitional group. While pupils of next-to-top scores in the transitional group had rather low odds of English 'O' level passes, their pupils who had been in the top fifth on reading at 11 had relatively high odds of English 'O' level. So the overall pattern in English 'passes', where a slightly but not definitely lower average was seen for transitional pupils, was elaborated.

These few interactions, then, slightly qualify the overall picture for transitionals, but it could still be concluded that they had relatively low performance in odds of five or more 'O' levels.

To summarise the implications of findings for the subgroups of early, recent and transitional pupils, it is worth commenting on any departure of their results from those of grammar+modern pupils. (Differences from comprehensives of particular origins will also be mentioned, where relevant).

The grammar+modern group

Since there was no overall difference of statistical significance, apart from in the English grades, between the gramamr+modern group and comprehensives as a whole, the overall average for the grammar+modern group might be expected to fall somewhere among averages for different sorts of comprehensive. The grammar+modern group sometimes was somewhere in between the various comprehensive groups.

Despite the lack of overall difference, the grammar+modern group had marginally higher proportions going in for at least some examination than were found in most subgroups of comprehensives.

The grammar+modern group was also less likely to gain five or more 'O' level equivalents than pupils in early comprehensives. The only group of pupils with lower odds of five or more 'O' level equivalents than the grammar+modern pupils was the transitional group. But the grammar+modern pupils were more or less the equivalent of recent comprehensive pupils in numbers of 'O' level among all pupils; both groups had a somewhat higher average than early comprehensive pupils. Compared with other pupils with one or more 'O' level, though, the grammar+modern group had a marginally, though probably not significantly, higher average than early and recent comprehensives. The grammar+moderns, then, were least likely of all sub-groups bar the transitional pupils to reach the qualification of five or more 'O' levels, yet emerged with a rather average number of 'O' level equivalents. The simplest findings from which to draw conclusions, in that they are consistent, at least, are the overall grammar+modern and comprehensive averages, on which groups were closely similar.

The only differences found in English exam results attributable to the grammar+modern/dates distinctions were in average 'O' level grade. Smallish but statistically significant patterns indicated marginally superior performance by pupils with a graded English 'O' level in the grammar+modern group than by pupils in recent comprehensives who had a grade in English 'O' level. This was not very revealing of English attainments, as neither the odds of 'passing' 'O' level or equivalent, nor the average CSE grade, nor the combined scale showed statistically significant differences.

Overall, the grammar+modern pupils' average was in the middle range, close to or in between subgroups of comprehensive pupils, on many measures of examination performance. In the grammar+modern mixture, those of low eleven year attainment had relatively low exam attainment. In odds of 'O' level maths or English passed by 1974, those of high eleven year scores had particularly high examination attainment. Among higher scorers at eleven, grammar+modern pupils tended to have similar odds of five or more 'O' level passes, though, to those obtained by various comprehensive subgroups (though corrections for other eleven year scores left this unclear).

The considerable range of grammar pupils at the top end of the scale for whom no equivalents existed in comprehensives helped to form the proportions with particular numbers of 'O' level equivalents, for example, at the top end of the scale.

The age range of schools

A further distinction between types of school had to be explored in these data. This was whether the pupils started school at eleven years old or at 12 or 13. The distinction was an important one at this time, as it still is, because many comprehensive systems are organised with middle schools up to 12 or 13 years old and transfer to comprehensives at that later age. For whatever reasons, the comprehensives whose pupils started at 12 or 13 (whether NCDS pupils had been in those schools since 11 or only since 12 or 13) had more 'advantaged' intakes. That is, they seemed to take in pupils who, on average, were higher scoring on tests at 11, and contained a higher proportion of non-manual workers' children, than the pupils who went into schools starting at 11*. Almost a third of the comprehensive sample went to 12+ 13+ schools. Because the difference in intakes was noticed from the start, it was explored in every analysis of exam results.

Grammar and secondary modern pupils were also distinguished for the purposes of analysis according to this feature of the age range of schools, for it appeared from these data that, at the particular stage of piecemeal comprehensive reorganisation under study, there were some grammar and secondary modern schools whose pupils started secondary school at 12 or 13. Most grammars and secondary moderns, in the days when selection was universal, started at eleven years old. Perhaps '12+' or '13+' selective schools were part of an interim arrangement during comprehensivisation. Although it was not clear in NCDS data what these grammars and secondary moderns were, the distinction between schools that started at eleven years old, and schools whose youngest pupils were 12 or 13, was included in corrections for extraneous factors, as an independent variable. It is reported in some detail here, with more detailed evidence and graphs in Appendix C, because it was apparently the source of differences in examination performance.

To the extent that there was a finding of exam performance related to age of youngest pupil, and this was to only a limited extent, the evidence tended to suggest slightly better exam results in comprehensives if they started at 12 or 13 years old rather than 11 years old. There was also a suggestion of some aspects of exam performance in 12+ 13+ secondary moderns being better than in 11+ secondary moderns. This did not hold for grammar school pupils. Where

* One minor point detracts from the argument put forward by critics of this work who thought the selective intakes had artificially high scores at eleven because of practice for the '11 plus'. The comprehensive pupils who were least likely to have 11 plus practice, because they went into schools which did not start at 11, had higher scores than the comprehensive pupils who transferred at eleven, some of whom may well have taken '11 plus' examinations.

there was a difference, as in proportions with five or more 'O' levels, better performance by comprehensive pupils was found in 12+ or 13+ schools. This was particularly clear in the case of those in the top fifth at eleven. Whether the school ran from age 11 up, or from 12 or 13, did not seem associated with such large differences in the grammar+modern combination. The differences between grammar+modern and comprehensive groups among those whose secondary school started at eleven were smaller, in proportions with five or more 'O' levels or maths 'O' level passes. Comprehensive pupils were, if anything, slightly lower attaining. However, grammar, comprehensive and secondary modern pupils were more similar to each other in 12+ 13+ schools in these results, while those in 11+ schools were more clearly differentiated by type of school (see Appendix C).

In sum, the age of youngest pupils did not seem related to such substantial differences among the grammar+modern group as among comprehensives. Those in comprehensive 'upper' schools appeared to stand a better chance of obtaining five or more 'O' level 'pass' equivalents or maths 'O' level by the end of their compulsory school years.

One possible interpretation of the comprehensive results is that they might have reflected the distinctions between schools made by a 'Leicestershire' system (a system of comprehensivisation with selection at 13+, in which some comprehensives go up to age 18 and some have no sixth form).

The relations of exam results to the age of youngest pupil in the school cannot be fully explored in this report, as they are complex and a distraction from generalisations about the schooling categories. The fact, however, that the age range of the schools is itself often connected with selection, and that it may have as close an association with exam results as the schooling classifications which form the major focus of the study, means some attention to the 'age of youngest pupil' factor may repay study in future.

Other outcomes of schooling which showed interactions with the age range of schools

Test scores at 16 did not show the same patterns as the examination results, of school differences varying according to whether schools started at 11 or later. Scores were higher in 12+ 13+ schools, on average, but the grammar+modern group and comprehensives did not appear to differ in this respect. The only aspects of educational progress other than examinations to show school differences in their relation to the age range of schools (apart from inconclusive features of intentions concerning future employment) were the two measures of truancy or staying away from school (see Appendix F). This was one instance where examination results and other outcomes seemed to support each other, in that pupils in 12+ 13+ comprehensives appeared 'better' and resembled grammar pupils in 12+ 13+ schools.

(v) Different groups of pupils
Relations of exam results to school category: variation according to attainments before secondary school, social class and parents' interest

So far the results have been presented for each school group or subgroup as an average for all the pupils in that sample (or all the ones to whom the measure applied). Thus, the differences and similarities that have been discussed have been those which were found after the components due to attainments before secondary school and to home background have, as far as possible, been corrected out of the picture. In that sense, any 'school type' differences, and the lack of selective/nonselective differences, reveal what can be thought of as relations between average exam performance and school. It is possible, in addition, to have a more elaborate explanation of relations of exam results to schooling, where these were not constant but varied according to some other factors, that is, where there were interactions. Allowing for social class differences and attainment differences before secondary school is necessary if one is to generalise about possible school influences. But it would be premature to discuss those overall averages for school groups without indicating whether observations might be qualified by interactions.

To be tackled, then, was whether these overall patterns could be illuminated by exploring whether they applied equally to all groups of children. For instance, was the lack of overall difference between the grammar+modern combination and comprehensives, in proportions passing five or more 'O' levels, a lack of significant difference over all pupils, whatever their test scores at eleven? Were the patterns of results the same for pupils of social classes I and II as for those of equivalent eleven year attainments with manual worker fathers? It is essential to bear in mind in following the account which follows, of interactions with test scores, parents' interest and social class, that a minority of such interactions were tested and that an inordinate amount of discussion to given to those which were statistically significant because of the possibility that they complicated the account already given (and that a proportion of those, though it cannot be said which ones, can be expected to have been statistically significant by chance, and therefore not meaningful).

Then, too, some of the interactions incorporated adjustment for so many related factors that they were difficult to interpret in terms that had any reference to real life, let alone educational or policy implications. The inconsistencies between patterns, and the relativity of those patterns which were found, were deterrents to relying on most interactions. In addition, the continuous outcome measures of exam performance may be more sensitive measures of exam performance, particularly in interactions with a continuous measure of earlier attainments, than a dichotomous, criterion measure. So, although two criterion measures, odds of obtaining five or more 'O' level equivalents and odds of passing mathematics at 'O' level standard, did show interactions, these have to be described very exactly and treated with caution. (These are given in Appendix T, where there is fuller discussion of them). Interactions involving continuous 'outcome' variables may be more illuminating. The results on combined scales of English and maths grades, for each of the reading or maths test ranges, and the relation of numbers of 'O' level equivalents to a combination of scores, were of interest. Interactions, nevertheless, may give an unrepresentative idea of the findings, unless the overall view already provided is retained. The results to be presented, on differences between subgroups of children in particular

categories of schools, should, therefore, be regarded as a jigsaw with most of the pieces missing.

A summary will be given here, containing salient findings on school differences which varied according to attainments at eleven. This will be followed by brief notes on patterns of school differences related to parents' interest in school and to social class. Because of the reservations about the generalisability of findings on different subgroups of children, many details are necessary, most of which are offered in Appendices T (interactions with test scores), B (parents' interest) or G (social class).

Attainment group at 11; interactions with test scores at 11: overview of a very incomplete investigation

There were some variations in relations between exam results and test scores for different school categories. Most patterns were not explored, however, and in the majority of cases, patterns which were tested revealed no statistically significant interactions. Many of the interactions which were significant themselves incorporated corrections for other test scores at 11, so they were highly relative. In some cases, it seemed that the allowance for test scores as grouped variables (in thirds or fifths of the range) might be inadequate. Lastly, some analyses had significant interactions between school category and one, but not another, test score.

The complete list of interactions tested is in Appendix T, from which it can be seen that the majority of those tested showed no statistical significance. So there were few robust interactions, and any observations must be treated with circumspection.

Interactions between test score and type of school (comprehensive, grammar, secondary modern)

For several cases of school type effects, the relation with eleven year test score just reached a not very convincing level of statistical significance. Where there were interactions with school type, and these were not consistently found, they did not for the most part modify the observation that pupils differed on average according to school type, since grammar school pupils were usually highest attaining, with comprehensive pupils in the middle and secondary modern pupils getting the lowest exam results, on average, among children treated as equivalent here. Over the 'below average' range of eleven year attainment, where there were no grammar pupils because no one of low attainment was selected, there were only small differences between secondary modern and comprehensive averages. More detail is given in Appendix T, but there are few qualifications to the overall school type findings that can be given.

Figure 2
The odds of mathematics 'O' level passses for those of various mathematics test score ranges at eleven were not corrected for other test scores at eleven. The odds of mathematics 'O' level passes suggested that, for those of middle and higher scores on maths at eleven, being in a comprehensive was associated with slightly higher odds of passing maths at 'O' level standard than being in a secondary modern. The odds were slightly higher in grammar schools, for pupils over the middle range, but substantially higher - hence the significance of the interaction - when comparisons were confined to those in the top fifth of maths at eleven.

Figure 2

Odds of mathematics 'O' level (A to C) pass or CSE grade 1

- interaction of school type × mathematics test score at eleven years old ($X^2_8 = 18{\cdot}9$, $p<0{\cdot}025$)
- mathematics grouped in fifths of range
- incorporates correction for parents interest in school at 16, sex, social class and age of youngest pupil in the school
- graphed on mean of maths test scores for each call for each school type group

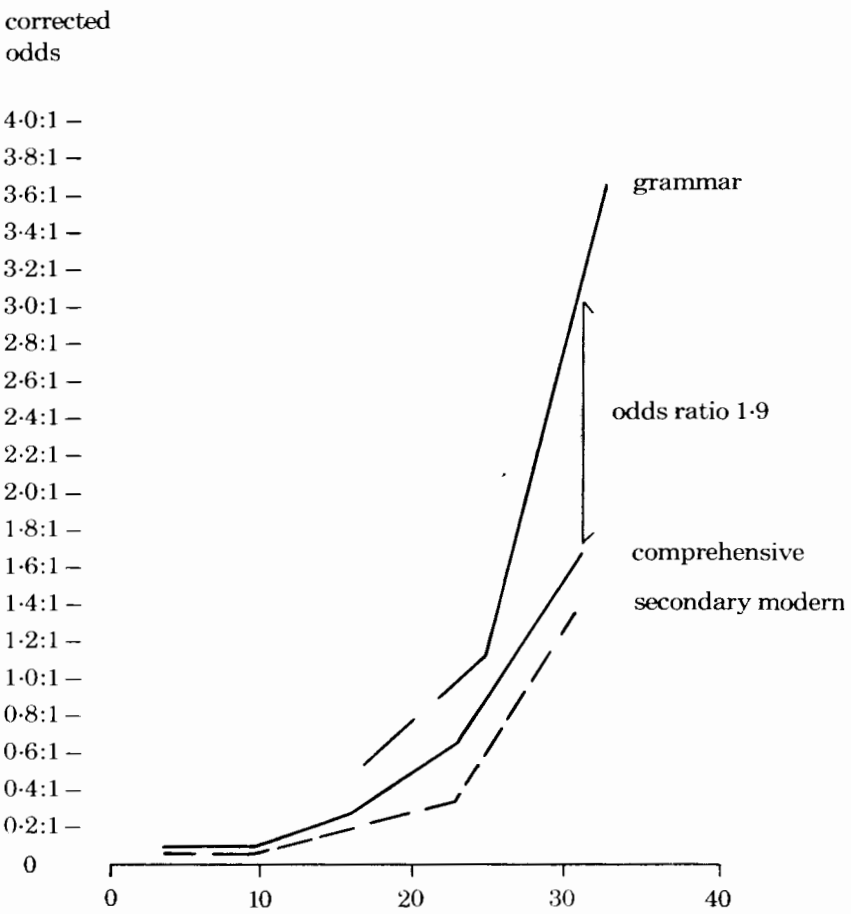


Figure 3
How did these observations on proportions passing maths relate to the continuous measure of maths performance, the combined scale? It is not possible to compare results closely as the samples differ. There were slightly different implications for the pupils who had grades in maths in the three schooltypes. While comprehensive and secondary modern pupils appeared closely similar among below average scorers at 11, comprehensive pupils of average and above average mathematics scores at eleven seemed to have substantially higher mathematics exam grades. An exception to this, though was apparent for those with the very highest scores on maths at 11. Comprehensive and secondary modern pupils with those scores and grades were not significantly different. For above average and average scores at eleven, the average maths grade for grammar pupils was just significantly higher than the comprehensive average. Such as it was, the implication for the top fifth was similar to that of the proportions with 'O' level maths, that the average for the grammar top fifth was the highest one. It is necessary, however, to treat these findings with caution, as there was no correction for the reliability of eleven year olds' scores.

Figure 4
The same reservation applies to the interaction on the equivalent scale for English results. A further uncertainty arises from the fact that these are the figures with the combined scales weighted so that grade A was equal to 0.98, and so on. The interaction was not tested with the other weighting and the shape of any graph is therefore unknown. Secondary modern pupils who had grades in CSE (or GCE) English resembled equivalent comprehensive pupils though the latter had slightly higher grades on the combined scale, among those who had been the lowest fifth of readers at 11. For those of average reading score, there was not a reliable difference - the averages were just two standard errors apart - between comprehensive and secondary modern pupils. (The mathematics grades had tended to be different among those pupils). Grammar pupils of average and above average attainment in reading at 11 were getting significantly higher grades in English on average, than comprehensive pupils in the same fifth of reading attainment. Comprehensive pupils in the top fifth for reading at eleven were obtaining significantly higher grades in English than those in secondary moderns in the same fifth for reading at 11. There was no interaction in proportions passing 'O' level English language or CSE grade 1 English to qualify the overall observation that grammar pupils had the highest odds and secondary modern pupils the lowest, with comprehensive pupils in between, another respect in which mathematics and English results were different. These observations however, suffer from the additional drawback that they may apply only to the subjects English and mathematics, that, even for those two subjects, results differ and that the interaction in English results is limited to the pupils who obtained grades in those subjects.

Far more informative and substantial results were available to suggest whether relations between schooltype results differed over the range of ability, from an exploration of numbers of 'O' level equivalents (a continuous measure including those pupils with no 'O' levels). Pupils were differentiated according to their combination of scores on maths, reading and general ability tests at eleven. This had the advantage of some correction at least for each of the three available indices of attainment at eleven, without making interactions less meaningful by adjusting each interaction for other highly correlated, test scores. It also had the advantage of permitting relatively sophisticated corrections for the reliability of scores.

Figure 3

Average grade on combined scale of mathematics at CSE or 'O' level

type of school : interaction with mathematics test score at 11

(weighted : grade A at 'O' level = 0·97; grade B or CSE 1 = 0·90; grade C at 'O' level = 0·83; grade D or E or CSE 2 or 3 = 0·67; CSE 4 or 5 = 0·48)
($X^2_8 = 15{\cdot}54$, $p<0{\cdot}05$)
n.b. there is no correction for reliability of eleven year scores

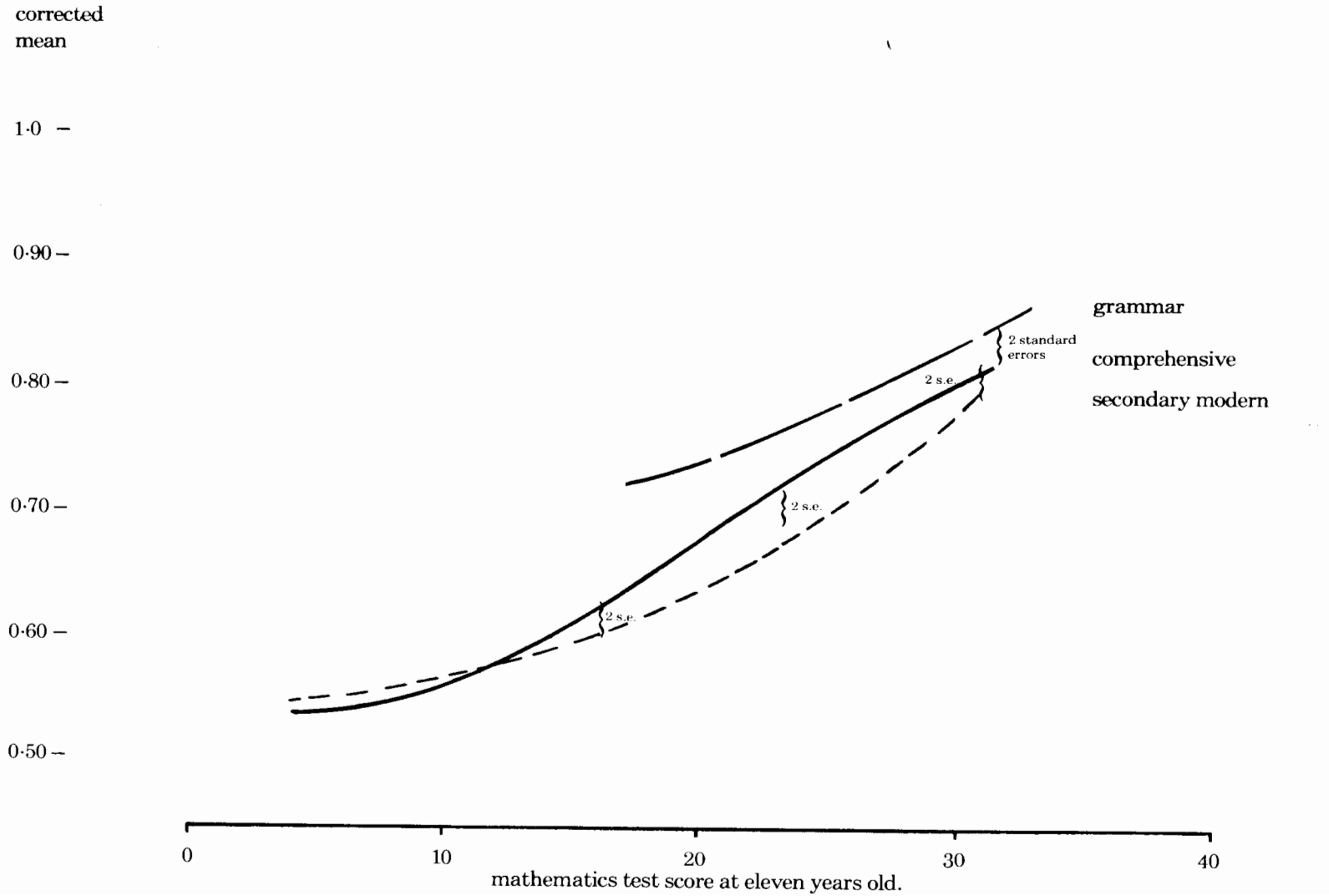


Figure 4

Average grade on combined scale of English Language at CSE or 'O' level

(sample = 2130, = 668 grammar pupils (89 per cent of grammar sample) + 801 secondary modern pupils (66 per cent of secondary modern pupils) + 661 comprehensive pupils (71 per cent of comprehensive sample)
weights, grade A at 'O' level = 0·98, grade B = 0·92, etc. $X_8^2 = 23{\cdot}7$, $p < 0{\cdot}01$

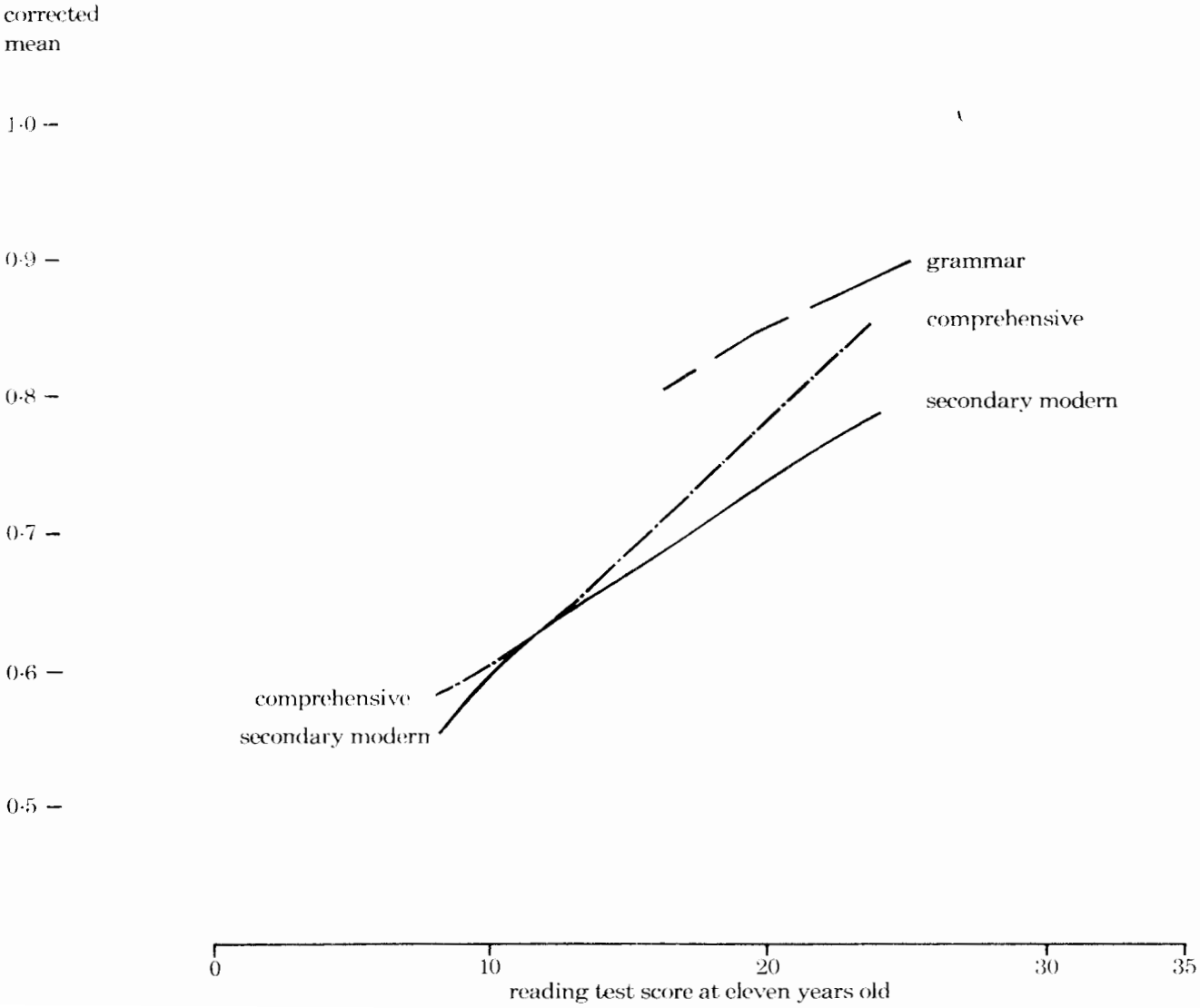


Figure 5

Regressions on composite of test scores at 11 years old (mathematics, reading, "general ability")

Reliability corrected regressions; number of 'O' level equivalents ('O' level A-C or CSE 1)

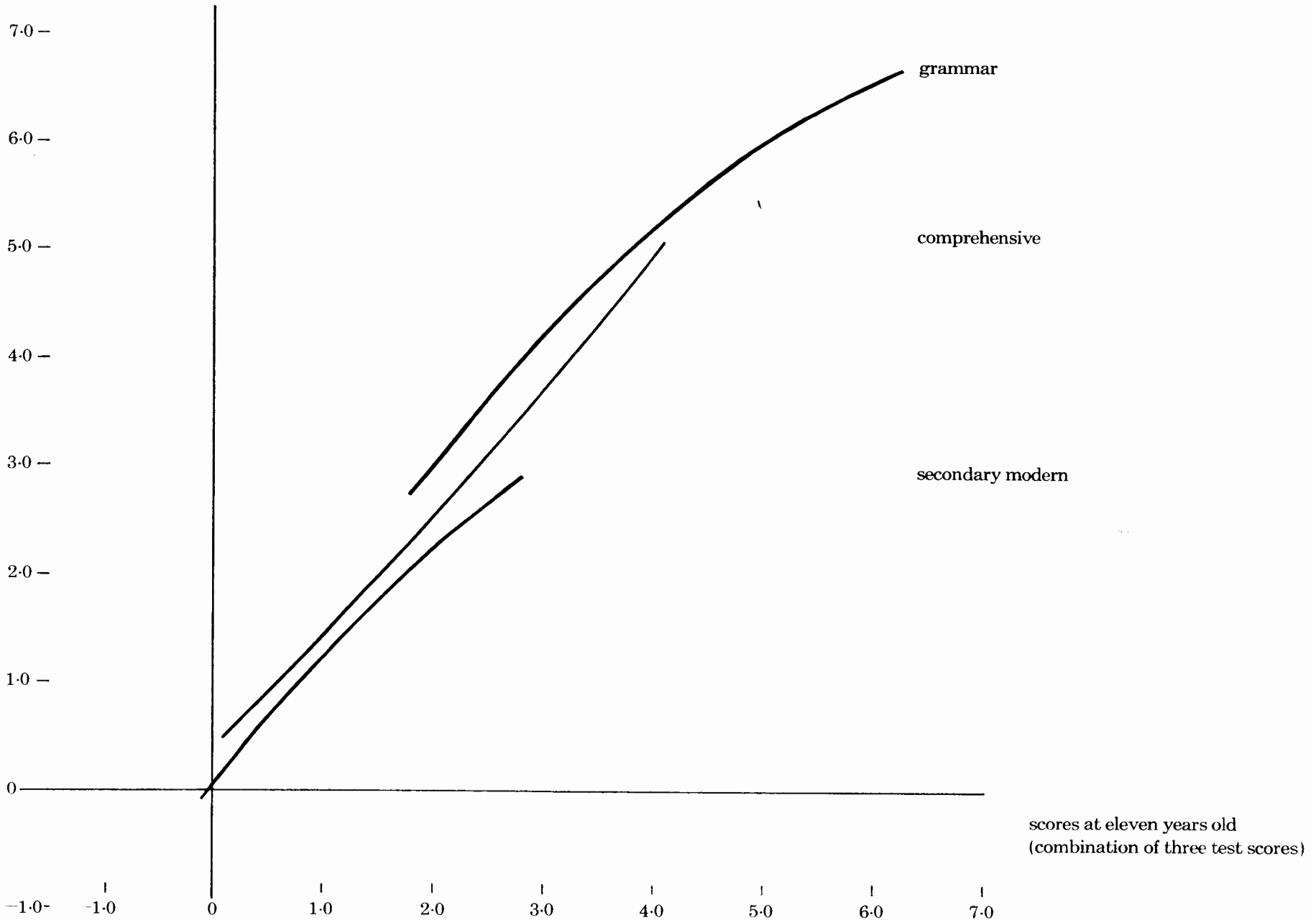


Figure 5
The implications were that, over the lower half of the range, the comprehensive pupils' average tended to be close to, if anything a little higher than that of the secondary modern pupils. Around the average of the combination score, which was the narrow band where there were three school types to compare (that is where there were pupils in grammars, secondary moderns and comprehensives), the figures resembled the overall average. Those in grammars had the highest average numbers of 'O' levels, those in secondary moderns the lowest, and the pupils of comprehensives were in between. Above that narrow band there was a range of scores shared by both comprehensive and grammar pupils. Over that range, as the eleven year attainment rose so did both groups' averages, but the comprehensive pupils' average rose slightly faster. At the very top of the range over which comprehensive pupils existed, their average number of 'O' levels was within two standard errors of the average for grammar pupils. Thus, at that top end, there did not appear to be a statistically significant difference between those in grammars and those in comprehensives.

There were plenty of grammar pupils whose combined scores were higher at eleven than pupils of any other school group. Their progress could not, therefore, validly be compared with that of pupils in other school groups. But, over the range for which there could be comparative evidence, it did seem that, with the most appropriate correction for attainment and ability at eleven years old, results were obtained which somewhat modify the overall observation of significant schooltype differences.

The grammar+modern vs. comprehensive distinction

The second question to be tackled is whether the lack of overall difference on most measures of examination attainment, between the grammar+modern group and comprehensive pupils, needed to be qualified.

There are few relevant patterns, as few schooling (grammar+modern v. comprehensive) interactions were tested. Most details are Appendix T. Interactions of the grammar+modern v. comprehensive schooling variable with each of the three test scores in the analysis were significant in the case of proportions with no exams. Since interactions were statistically significant with each of the three test scores involved, all showing increasing distinction between the grammar+modern and comprehensive groups in proportions with no exam as pupils of higher scores were involved, there might be grounds for inferring a genuine effect. But it is hard to see what this would be. The average for those in each third of the eleven year olds is relative to two other scores each of which is grouped in thirds. Interpretation would be folly.

There were few other interactions to suggest that the lack of grammar+modern v. comprehensive distinction in results failed to apply over some part of the range. Even when looked at over the range of eleven year attainments, the average number of 'O' level equivalents did not seem to differ significantly between the grammar+modern group and the comprehensive pupils. When the criterion of examination performance was the acquisition of five or more 'passes' ('O' level A to C CSE1) on which no overall difference had been found, only two of six possible interactions with test score at eleven reached statistical significance at the five per cent level. The other four were non-significant. There was no interaction with reading test score or general

ability score, so, although there was a mathematics test interaction, its implications are obscure (see Appendix T).

Figure 6
The results in single subjects had some similarities (though relative to only one score at eleven). The similarity between the grammar+modern and comprehensive pupils in odds of passing maths at 'O' level standard varied slightly over the range of maths attainment, as odds of passing five or more 'O' levels had. The lowest attaining and highest attaining groups seemed to differ, if anything. Grammar+modern pupils in the top fifth had the higher odds of a maths pass. That group's result was predominantly a function of grammar pupils' performance. No such variation in odds of passing English, relative to reading at 11, was found, just as, again, there had been no such variation in odds of five or more 'O' levels relative to reading at 11. (It also has to be remembered that these observations were without corrections for measurement error, which might be most likely to affect the low and high scorers).

Figure 7
The combined scales of mathematics and English grades were the only other measures compared, and they showed no interactions to qualify the overall findings. These analyses, though confined to the pupils with grades in CSE or GCE 'O' level, had the advantage of using continuous exam measures, though not continuous eleven year score. On the scale of mathematics exam grades, the grammar+modern and comprehensive averages were closely similar (within two standard errors of each other) at all points where there were pupils to be compared. If there was a difference, it seemed to appear among top maths score pupils, but, again, there were no reliability corrections. So there were no grounds for modifying conclusions based on the overall lack of difference.

Figure 8
Comprehensive pupils had averaged out doing less well, after allowing for reading at 11, on graded performance at 'O' level English language and on the combined 'O' level/CSE grades in English. If the comprehensive pupils who got the equivalent of graded results at 'O' level had a greater tendency to get the 1974 version of grade E when grammar+moderns got grade D, or grammar+modern pupils had a greater tendency to get A's rather than C's this might reconcile these findings with the fact that there was no difference between the grammar+modern and comprehensive groups in their relative odds of passing English at 'O' level. But no inference of that kind could be drawn, since the overall difference between grammar+modern and comprehensive pupils on combined English grades was not qualified by any interaction with eleven year score. If there was any difference, it was among the highest attainers, the top 40 per cent or so at eleven in reading, but, even then, no statistically reliable difference existed.

The generalisability of any observations on the combined scale must be limited in view of contrasting implications of the criterion measure of proportions passing 'O' level English. It is consistent with the overall findings and the pattern over the eleven year range to conclude that the difference lay in those comprehensive pupils who went in for 'O' level not getting as high grades, on average, as grammar+modern pupils, after suitable corrections. Nevertheless, in that sense, average readers at 11 in comprehensives matched their equivalents in grammar+modern schools in grades at sixteen.

Figure 6

Corrected odds of mathematics 'O' level equivalent (A-C, CSE 1) grammar + modern v. comprehensive with allowance for means within fifths

$X_4^2 = 11{\cdot}49$ ($p < 0{\cdot}025$)

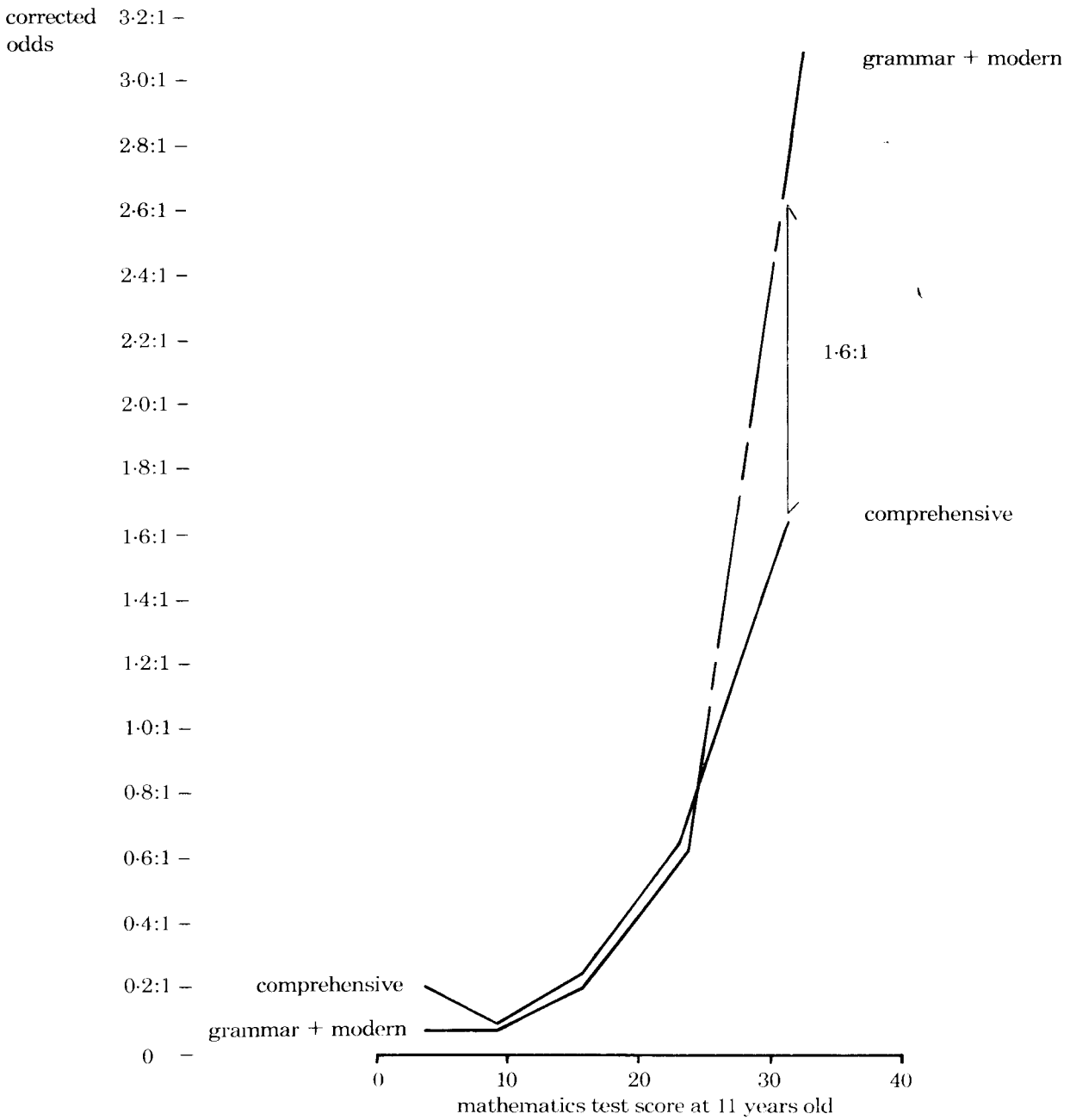


Figure 7

Mathematics grade at CSE or 'O' level, on combined scale grammar + modern recalculated from school type

overall n = 1719 (grammar 630, 84 per cent; secondary modern = 588, 48 per cent; comprehensive = 501, 54 per cent of comprehensive pupils)
no statistically significant differences and no interaction. The two groups are within two standard errors of each other, over the entire range

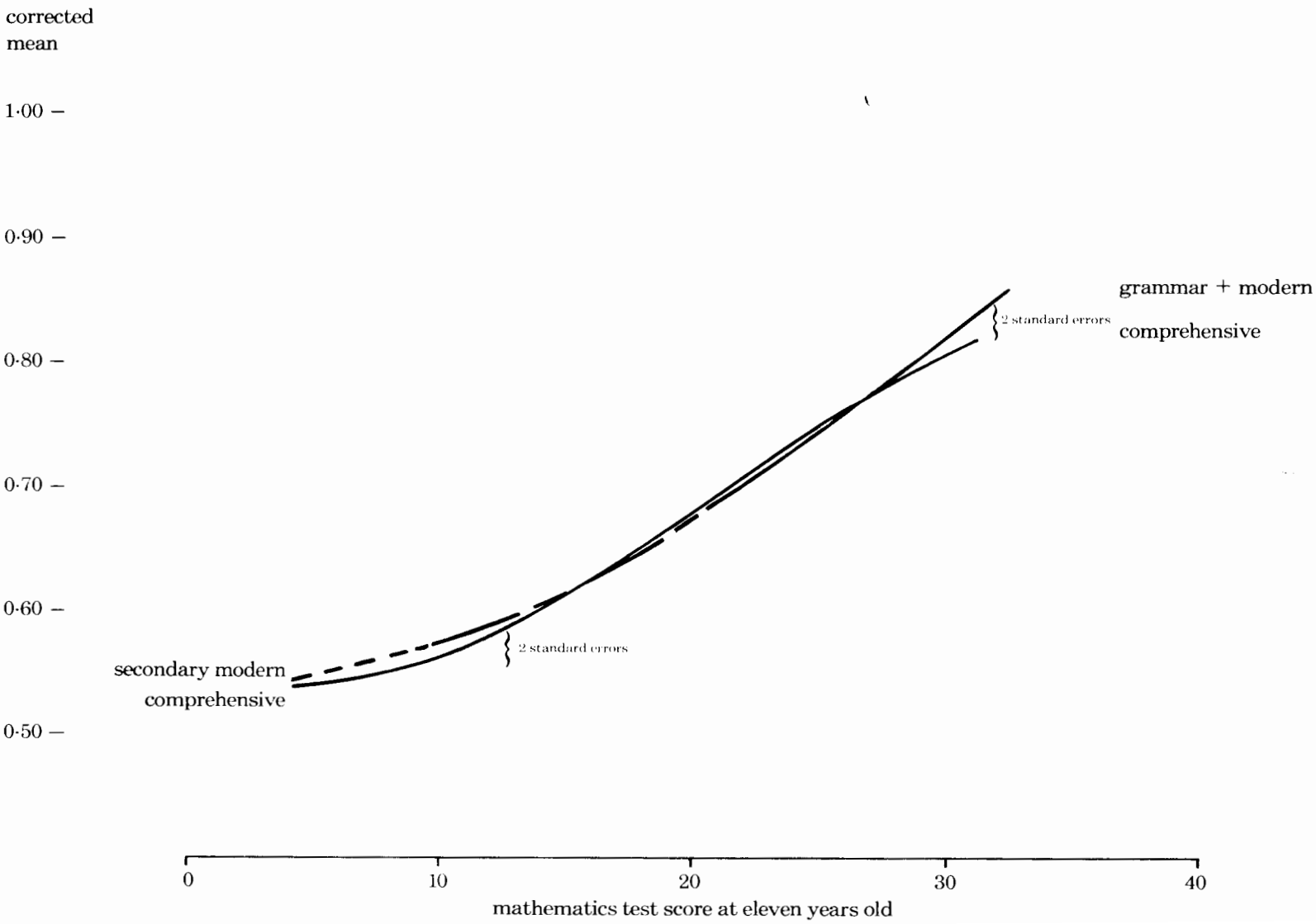
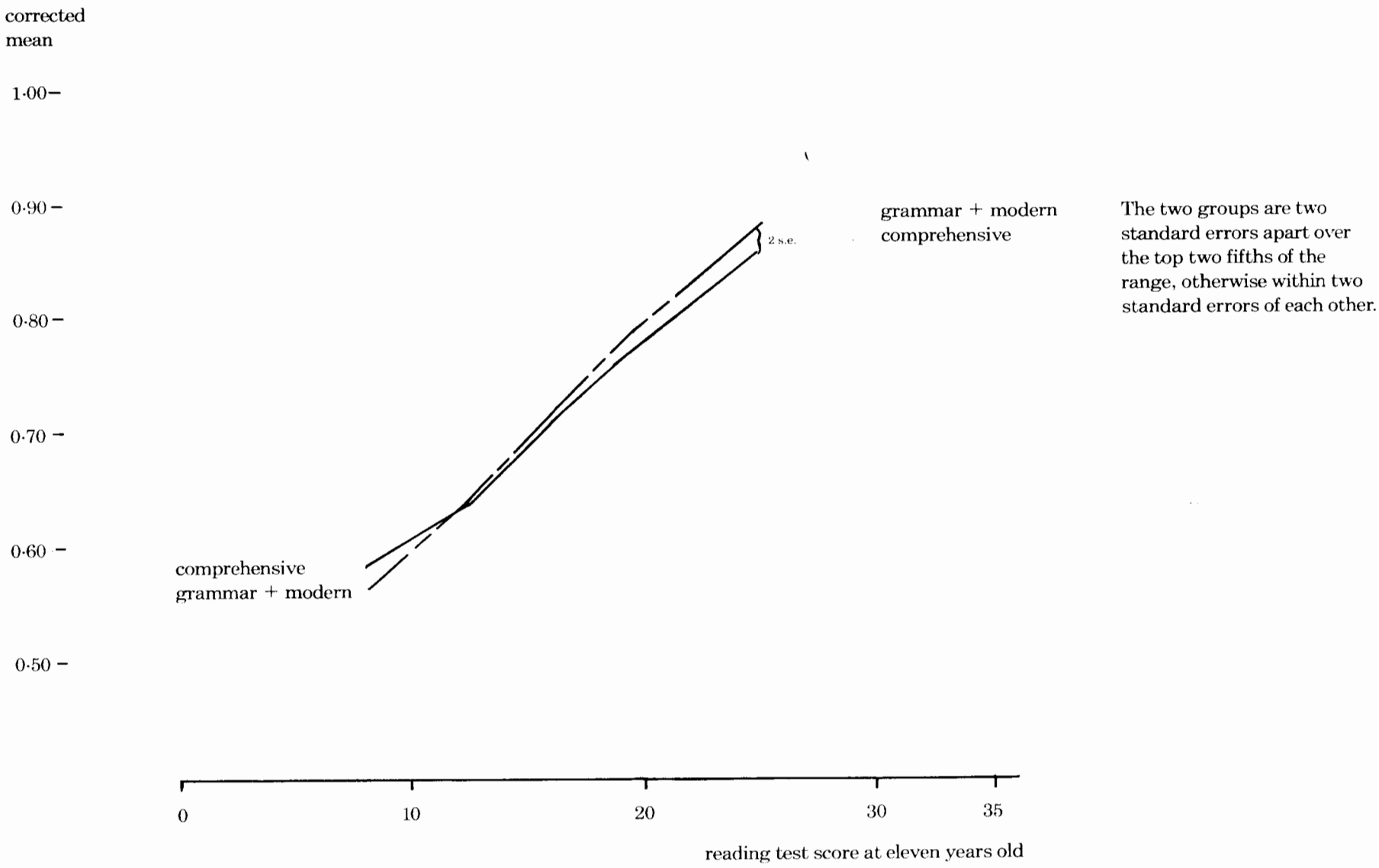


Figure 8

Combined GCE 'O' level grades in English language and CSE grades in English



The two groups are two standard errors apart over the top two fifths of the range, otherwise within two standard errors of each other.

Parallel mathematics results did not show differences attributable to being in the grammar+modern or comprehensive group. Because results appeared to apply in this limited way only to a single subject, and relative to only one aspect of ability before secondary school, they are to be treated with caution.

The overall implication was that there was not sufficient evidence of any sufficiently substantial differences between the grammar+modern and comprehensive pupils, after suitable corrections for pre-existing and background influences, to warrant modification of overall conclusions. So the predominant finding, of a lack of statistically significant differences (apart from the grades in English language, probably at 'O' level) between grammar+modern and comprehensive pupils, were unaffected.

Parents' interest: interactions

The indicator of parents' interest in school and its influence on examination results is defined in Chapter I. The factor of parents' interest is discussed and findings of relevance are detailed in Appendix B. Although illuminating in its own right, the role of parents' interest is particularly relevant to the study to the extent that it interacted with school variables. A brief note on its implications is given here.

The overall lower relative proportions in the grammar+modern and transitional groups with no examination entries were only really clear among those with a 'very interested' parent. For those without such a parent, being in the grammar+modern combination was not asssociated with lower odds of taking no exams than for the majority of comprehensive pupils.

While the odds of getting five or more 'O' level equivalents had not been seen to differ between grammar+modern and comprehensive groups, there was a suggestion in the data that there were marginally higher odds in comprehensives than grammar+moderns among those without an interested parent and marginally lower odds than in grammar+moderns among those relatively few comprehensive pupils who had an interested parent. There was no interaction of this kind in odds of five or more 'O' levels with three types of school distinguished, but a very similar pattern was found in odds of mathematics 'O' level. If anything, the patterns for subgroups of comprehensive pupils varied, so the fractionally lower odds among those with a 'very interested' parent found for 'comprehensive pupils' seemed to be a reflection of only some subgroups of comprehensive pupils. Those in early comprehensives, for instance, matched grammar+modern pupils among the children of 'interested parents' in odds of five or more 'O' levels. Equivalent pupils in exgrammar comprehensives exceeded the grammar+modern pupils.

In general, the overall school results, corrected to offset the influence of parents' interest, did not differ greatly from the patterns for those with or those without interested parents, looked at separately. So the inclusion of parents' interest as a variable may not have altered conclusions.

It can be seen from Appendix B, too, that when 'parental interest' was not allowed for, then odds of five or more 'O' levels for some subgroups of comprehensive pupils in the top fifth at eleven may have been slightly higher than was otherwise the case. Possibly, then, results for grammar+modern pupils were slightly overestimated, if anything, by inclusion of the 'parents' interest' variable, but this is by no means clear.

Social class: interactions

The question arose of whether different sorts of schooling were associated with particular levels of performance for those of given social class background (fathers' occupations). It was clear that the social class differences in the intakes of the different school samples here could, if not controlled for, be confused with the influence of school. What was necessary for this study was to allow for differences among eleven year olds going to the different schools in their social class characteristics.

Table 4c

Social class differences in the intakes of different school types

fathers' occupation at eleven years old	I+II	III non manual	manual	Total
grammar (n=747=100%)	42%	15%	43%	100%
secondary modern (n=1213=100%)	17%	8%	76%	100%
comprehensive (n=936=100%)	16%	8%	76%	100%
transitional (n=1479=100%)	19%	10%	71%	100%
total (N=4375=100%)	22%	10%	69%	100%
grammar+modern (n=1960=100%)	26%	10%	63%	100%

There were also social class differences in the relation of exam results at sixteen to test scores at sixteen (Ives, 1980). NCDS and other research has shown differences between the social classes in other aspects of progress and exam results in schools.

Social class differences have not been mentioned in the results so far. It was necessary to obscure, but not ignore them, in comparing school groups. Social class was, nevertheless, associated with a statistically significant set of differences overall, after allowing for other factors, on many measures of examination performance explored here.

The social class differences seen in Table 15 within different ability groups can be summarised as follows.

Table 17. 'Raw' examination results for different social class groups

% no exams/n — fifths of combined test scores at eleven:

father's occupation:	lowest	next	mid	next	top	total
nonmanual	16% / 89	8% / 166	1% / 258	0.9% / 349	0.2% / 504	2.5% / 1366
manual	33% / 763	11% / 653	6% / 619	3% / 585	1% / 389	12.6% / 3009

mean no. of 'O' levels / % with five or more 'O' level equivalents — fifths of combined test scores at eleven:

father's occupation:	lowest	next	mid	next	top	total
nonmanual	0.3 / 0	0.6 / 0.6%	1.8 / 13%	3.4 / 34%	6.0 / 74%	3.5 / 39%
manual	0.1 / 0.3%	0.4 / 0.8%	1.0 / 5%	2.5 / 21%	4.6 / 55%	1.4 / 12%

mean no./ of 'A' levels/ / % with one 'A' level pass or more — fifths of combined test scores at eleven:

father's occupation:	lowest	next	mid	next	top	total
nonmanual	0 / 0	0.2 / 2%	0.1 / 8%	0.5 / 21%	1.3 / 54%	0.6 / 27%
manual	0 / 0.1%	0 / 0.3%	0.1 / 3%	0.2 / 10%	0.8 / 35%	0.2% / 7%

These social class differences might be attributable to different schools or different average attainments within fifths of ability. Was there any evidence to link social class effects with the school attended? Would the data modify observations on school influences by indicating that overall averages applied to only one particular social class group?

The only indicator of examination performance to show statistically significant interactions with social class (fathers' occupation I or II; III nonmanual; manual) when the whole range of pupils was looked at was the proportion with no examination entries. (The interactions, however, were not tested in any cases

where the exam outcome was measured on a continuous dimension, e.g. numbers of 'O' levels or grades). It is necessary to refer to Appendix G for the details, if only to indicate what a peculiar pattern the relative proportions with no examination entries revealed.

The pattern seemed to suggest that, when proportions with no examination entries were explored for whether they differed more than would be expected from differing attainments, in ways that could be to do with school category or social class, the 'non-manual' pupils in comprehensives had unexpectedly high proportions of non-exam takers, compared to other schools. The school groups did not differ for 'working class' manual workers' children, who might be expected to be the pupils concerning whom the largest differences might appear. The differences among nonmanual pupils probably represented small numbers of pupils. Schools could vary in the chances they offered of taking one or two CSE's, say, as against no exams, to fairly low attaining pupils. If so, they did not do so in a way which was shaped by pupils' social class, at least not when pupils were grouped according to the categories of school here.

There was little evidence from any other measures of examination performance - to the extent that this was explored - to suggest that social class patterns differed according to school category. At least there was no sign, then, that any one organisational feature of schooling necessarily altered the overall class patterns. On the whole, examination performance (as distinct from the proportions with no exam entries) showed a pattern of he highest social classes (I and II) on average achieving more, even relative to their higher average achievements before secondary school, and those of 'manual' social classes achieving less, with the 'III' nonmanual' group in the middle. (See Appendix G).

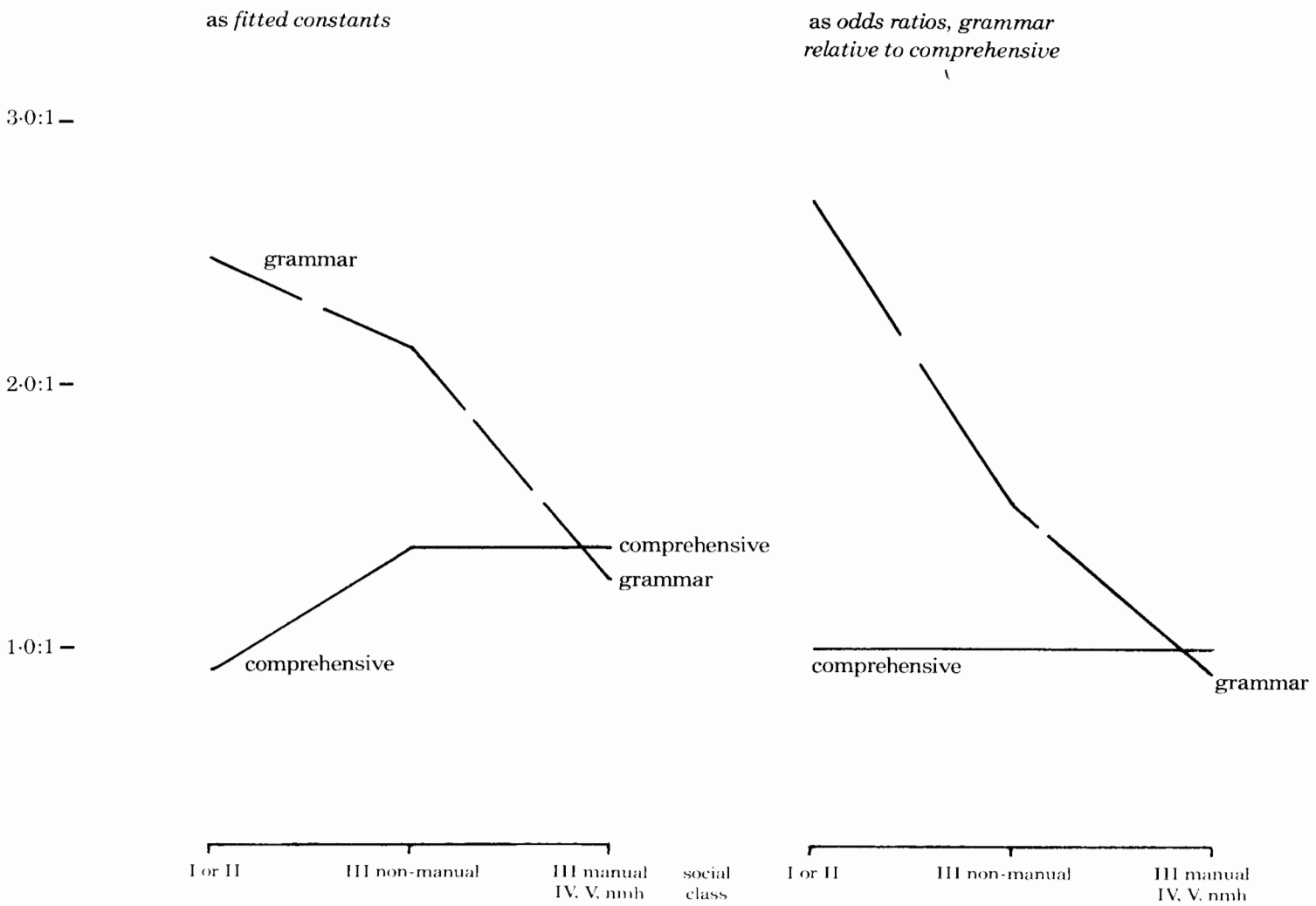
Pupils of social class III nonmanual were achieving results at or above the level of those of social classes I or II, though, on average, once allowance was made for background factors, school and eleven year attainments, in almost all the analyses of proportions. But it was consistently the case, whether in analyses of proportions or on continuous exam measures, that 'manual' workers' children did relatively poorly, given their attainments before secondary school. So there was a widening differentiation, by examination measures of performance, according to social class. It could be conjectured that such increasing differentiation of working-class from other pupils during the secondary years was likely to be shown in examination results where it would not be in other 'outcomes', as examinations tend to be for higher attainers. However, the same phenomenon was observed in test scores applied to the whole sample. So the social class gap widened in attainments (exams and tests), and not just in certification or qualifications obtained, during the secondary school years. (Of course, the existence of examinations designed for a minority may have determined the pattern of attainments, but the pattern extended to test attainments as well).

But the purpose of the inquiry reported here was to evaluate whether anything which could be attributed to school category appeared to alter this. Was there any evidence, for example, that the 'selective' grammar+modern combination tended to do more to offset such under-achievement among bright working class pupils? One defence of the selective system, against charges that it is socially divisive or perpetuating class boundaries, can be that grammar schools

Figure 9

Social classes within the top fifth

Odds of five 'O' level equivalents or more, for grammar and comprehensive pupils who were in the top fifth on a combination of three test scores at eleven years old.



offered a ladder whereby bright working class youngsters could climb out of their social class origins. Was there any evidence here of greater progress among 'working class' pupils in selective schools than in comprehensives? On the whole, the analyses showed no interaction of social class with schooling, so, once secondary moderns were included in the calculation, results suggested no sign of better or worse progress among working class pupils according to whether they were in the 'selective' or non-selective group. That observation however, is not enough to rule out the possibility that the different social classes may have fared differently in some aspects of exam performance in particular categories of school, since the interaction was not usually tested. Nor was one of the interactions which could have answered the question of how bright working class pupils got on by sixteen (schooling by social class by test score at eleven).

The only way to explore the possible fate of bright working class pupils, given the difficulty of exploring three-way interactions, was to look at the 'able' children separately, for whether school and social class showed an interaction. The data from the top fifth of children, as measured by a combination of reading, maths and general ability tests, at eleven, was distinguished for this purpose from the rest. They could be regarded for these purposes as 'able' at eleven. Among those children, the pattern of social class differences in proportions passing five or more 'O' levels in grammar schools appeared to be slightly different from the pattern in comprehensives*.

In this one limited respect, the study was able to pronounce on the fate of 'bright working class children' going to comprehensives when they might have gone to grammars. It could do this by exploring whether, among boys and girls in the top fifth, the ones of lower social class were any more or less likely to pass five or more 'O' levels in comprehensives than in grammars. Among those of social classes I or II, almost all (84 per cent) of whom were in grammars, odds of passing five or more 'O' levels were considerably higher in grammars. Among those of 'manual' social class background, on the other hand, odds were similar in grammars and comprehensives. The effect was small, just reaching an acceptable level of statistical significance, and not significant without correction for general ability at eleven years old; and the odds of five or more 'O' levels was the only exam measure on which this social class interaction was tested, among this 'able' top fifth, between grammars and comprehensives. But it meant that there did not seem to be an advantage in this respect, to 'bright working class children', from being in a grammar school as against a comprehensive.

(vi) Summary of examination results

The preceding chapter has described a broad range of measures investigated for evidence of any relation of examination results to the category of school attended. Certainly, various patterns of performance differing between groups of pupils suggested that several types of exam result could be linked in part to the school category. This was after allowance for some factors that differentiated the averages for pupils of different schools (such as what their parents were like) which might, in the absence of 'intake-correction', have

* See Figure 9

been mistaken for school influences. The interpretation of detailed findings is not easy, as results depended on how schools were grouped. They also depended on the subjects explored, the standard or range of grades inspected, and the grouping or weighting of measures of examination achievement. As in other outcomes of schooling studied for the same sample, grammar school pupils were of highest attainment overall, even after allowance for their higher pre-existing attainment. As before, secondary modern pupils were of relatively low attainment, lower than equivalent comprehensive pupils. Comprehensives as a whole were between grammars and secondary moderns in the numbers of 'O' level equivalents people obtained and in results of examinations both in mathematics and in English (language) at 'O' level and CSE. In most cases, this meant that the comprehensive pupils' average did not differ from that for the combination of grammar and secondary modern pupils.

There was no difference in numbers of 'O' level passes (or equivalents) attributable to being in a comprehensive as against the theoretical alternative of the grammar+modern combination. (This was so whether the children being compared were the half of the sample who had demonstrated their capacity to reach 'O' level standard in some subject, or the entire sample of children being studied.) Proportions with five or more 'O' levels were similar. The numbers of pupils passing at 'O' level standard (in GCE or CSE) did not differ in either English or maths results compared between the two schooling groups. Furthermore, there was no significant difference between grammar+modern and comprehensive pupils in maths results in terms of the average grade obtained on a scale combining grades at CSE and 'O' level. And the two schooling groups did not differ in the average grades obtained in mathematics, either in 'O' level by those with grades A to E, or in CSE by those who got grades 1 to 5.

By the criterion of possessing the equivalent of an 'O' level pass in English (language) or according to the yardstick of CSE grades (1 to 5) in English, there was no difference between pupils of comprehensives and those in the grammar+modern combination. It seemed therefore that the two groups were similar, over the spectrum of English exam attainment (that is, among both 'O' level standard and CSE pupils). That inference was tempered by the finding that, on a combined measure of 'O' level performance in English language and CSE English grades, comprehensive pupils were slightly, but statistically significantly, lower attaining than the grammar+modern combination. A lower average grade at English 'O' level - which could have been in the 'fail' grades or 'pass' grades or both - was found among the comprehensive pupils who obtained a grade, and seemed to underlie the combined scale findings. There was no echo in mathematics results, so it is not possible to draw consistent conclusions that 'O' level candidates differed.

The only overall generalisation possible was that there was no consistent difference between grammar+modern and comprehensive pupils. There were certainly no consistent grounds for concluding that comprehensives as a whole differed from the grammar+modern, selective category of schools, by these indications of examination performance. It seemed that, in so far as they were catering for comparable candidates at 16-plus, the two groups of schools, by most measures of exam performance here, produced comparable results, on average. Did the same go for results at 'A' level? Comparing among 'A' level candidates over the whole range of eleven year score, there was no statistically significant difference, and certainly not between grammar and comprehensive pupils. Nor was

there a difference when the comparison was among 'successful' candidates - those with one or more 'A' level pass. Comprehensive pupils were between equivalent grammar and secondary modern pupils in the number of 'A' levels they obtained, but there was no statistical significance to this. The grades obtained by those who did get 'A' level passes did not differ either, at least as far as could be inferred from the combined UCCA-scale measure of grades with numbers of 'A' level. Relative to sixteen year old attainment, the types of school did not differ; that is, children who were equivalent by sixteen did not stand different chances of 'A' level passes. When the entire samples of children were looked at, the school types did differ in numbers of 'A' levels for given eleven year score. Grammar pupils had the highest intake-corrected mean and secondary modern pupils the lowest. The grammar+modern and comprehensive groups did not appear to differ.

Overview of comprehensive subgroups (including those subdivided by origins; see Appendix A). A few overall observations on subgroups of comprehensives are possible. After the customary corrections for attainments and social class before secondary school, parents' interest and so on, there were few consistent marked differences. By most indices of examination performance, the grammer+modern combination was about average and the various comprehensive subgroups were sometimes higher, sometimes lower scoring but frequently closely similar. Former grammars and former secondary moderns showed a slight, and not entirely consistent, tendency to exceed the grammar+modern combination in exam results, and in proportions taking no examinations. Otherwise there were more similarities between averages than there were differences.

On only one measure were grammar+modern pupils lower scoring than any comprehensive subgroup (whether divisions were according to origins or dates of comprehensive). This was the proportions getting five or more 'O' level equivalents. There was no overall difference between them and the comprehensive pupils looked at as a whole, on this measure. The grammar+modern combination had only marginally lower odds of this 'O' level pass standard in five or more subjects than those in 'recent' comprehensives, or than pupils of purpose-built or amalgam comprehensives. However, there were more marked advantages in this respect to the early comprehensive subgroup and, in particular, to both ex-secondary modern and ex-grammar comprehensive pupils. The last subgroup was, on average, particularly likely to have pupils who passed five or more 'O' levels, for given social class and allowing for initial attainments.

If one is looking for characteristics of the various sorts of comprehensive that seem associated with general measures of examinations, it is possible to find them. That is, among the comprehensives it seemed that, if anything, purpose-built and amalgamation comprehensives had the lower averages and ex-selectives the higher averages, on general measures of examination results. But this was also the pattern for proportions of non-examinees. These inconsistent variations will not in any case be the central concern of those who wish to know whether comprehensive pupils in general were better or worse off at 16+ in exams. There was no consistent difference between the grammar+modern group and any particular subgroup of comprehensives that would allow for coherent conclusions, other than the implication of the overall results, that comprehensive pupils did not differ, by these lights, from the combinaion of grammar and secondary modern pupils.

Transitional pupils

A very common characteristic of pupils investigated by this study - true of the majority of those in comprehensives at 16 - was that their schools were still going comprehensive during the time they were at school. These pupils whose schools went comprehensive during their secondary years did not 'pass' exams as often or get as high grades, in English or mathematics, on average, as the grammar+modern group or early comprehensive pupils, though no significant difference appeared between them and other groups on CSE grades.

Pupils whose schools were undergoing reorganisation scored lower on all measures of 'O' level and CSE grade 1 performance. A few interactions may slightly qualify the overall picture for transitionals, but it could still be concluded that they had relatively low performance in a number of respects. The implication was that examination performance was probably hampered by reorganisation and changeover during the secondary years. The transitional group appeared to be the only group whose results were consistent enough for generalisation. The generalisation was that they were doing relatively poorly, after allowance for eleven year old attainment and given home backgrounds, at least on those measures which showed any statistically significant differences.

Interactions with eleven-year scores

An incomplete picture, and one which is difficult to interpret, arises from inspection of how the relation of exams to schools varied over the range of abilities. The findings have to be put in the context that most type of school interactions were non-significant, so, mostly, the overall observations apply. There were overall schooltype differences without interactions, in proportions with five or more 'O' levels, and this, together with separate analysis of the top fifth, can be taken to mean that, under the conditions of the analyses here, the odds of five or more 'O' levels for those who were in the top fifth in grammars were higher than for those in the top fifth in comprehensives, who were, in turn, relatively more likely to get five or more 'O' levels than those in the top fifth in secondary moderns.

That finding has to be interpreted in the context of the reliability-corrected regressions (Figure 5) of numbers of 'O' level equivalents. The former finding is, moreover, open to an objection which applies to averages over the entire top fifth of eleven year scores. The comparisons there are insensitive to the fact that a large number of those grammar pupils had higher scores at eleven than any comprehensive pupils. When comparisons were confined to equivalent grammar and comprehensive pupils, comparing like with like, the very highest scoring comprehensive pupils appeared to approach the level of equivalent grammar pupils on average numbers of 'O' level equivalents.

Interactions: subgroups of comprehensives relative to the grammar+modern group.

There were some interactions indicating differences over the various attainment subgroups in the relations between different sorts of comprehensive and grammar+modern pupils, over most of the range. While results in the individual subjects (maths and English) sometimes gave cause to doubt details of overall

conclusions, this was not usually true of more general measures with more adequate adjustment for attainment and 'general ability' before secondary school.

Among those in the top groups at 11, proportions passing five or more 'O' levels were high in recent comprehensives or comprehensives formed from a secondary modern or grammar, and in the grammar+modern group. Those in early comprehensives who were in the top fifth had scarcely lower odds of five or more 'O' levels. Pupils from the top fifth at eleven in purpose-built and amalgam comprehensives were not as high scoring, relative to other test scores, on five or more 'O' levels.

Early comprehensive pupils in the top fifth for maths at eleven were less likely to pass maths, it appeared, whereas top scorers at 11 were almost as likely to 'pass' at maths 'O' level standard in recent comprehensives as in the grammar+modern (mostly grammar) schools, despite the relatively high grammar+modern average. The patterns among top scorers in the individual subjects, maths and English, and in five or more 'O' levels were similar for ex-secondary modern pupils and those in early and recent or purpose-built comprehensives. For those in amalgam comprehensives, relative odds of maths or of English 'O' level were quite high (when odds of five or more, as was said, had not been). For ex-grammar comprehensives it was the other way round.

Among those who had been in the next-to-top group at eleven, the pupils of early comprehensives had higher proportions passing five 'O' levels. But some of these groups had relatively higher odds of maths or English 'O' level (amalgam, ex-secondary modern comprehensives) which the former grammar comprehensives did not.

Perhaps, underlying the overall lack of difference between grammar and modern and comprehensive pupils in proportions reaching the high standard of five or more 'O' levels at 16, the pattern for pupils who had been in the top group at eleven was similar among equivalent high-attaining grammar+modern pupils and 'recent', 'ex-grammar' and 'ex-secondary modern' comprehensive pupils.

Chapter V

Overview, reservations and interpretations

(i) Introduction: The context for the study reviewed

The research project reported here has explored the relation between performance in school and whether schools are selective or comprehensive. There are as many theories about how those in different types of school are faring as there are views on education. Some theories which might generate hypotheses for a study like this derive in part from the claims of advocates of the comprehensive or selctive systems. Some supporters of the selective system, for instance, may put forward the argument that selection 'caters better for able students and therefore for the preservation of standards at all levels of ability', as Husen suggests. Advocates of comprehensives, on the other hand, may argue that such a system would aim 'to enhance educational opportunities for students from lower social strata' (Husen; 1979). Advocates have a mixture of aims and theories, however. Some contributions to educational debate - the Black Papers on education, for example - defend selection as a means of social mobility, while pointing to alleged failures of the comprehensive system (Szamueli, 1969).

So the possible research hypotheses generated by such theories are numerous and sometimes inconsistent. The attempt here to provide empirical data on a question which provokes such speculation has at least helped to clarify which questions can and which cannot be answered.

First, a little of the history of this project will be described with an outline of the contents of earlier chapters of this report (Section ii). That is to be followed by an explanation of some of the assumptions on which this study was bases - Section (iii). In Section (iv) the findings will be discussed and interpreted. In Section (v) readers are advised to treat results here with caution, in view of reservations about public examinations as measures of groups and about the limited allowance for intake possible. Nevertheless, to the extent that exam results echo earlier observations on test scores, they suggest similar overall conclusions to those of the first publication from this project. The final section (vi) presents a summary and discussion of implications of the findings, and concludes with further reservations about the extent to which the debate on selection can be informed by empirical data, even when those data are the best available and are analysed in appropriate ways.

Other authors have attempted in the past to compare pupils of selective and comprehensive schools, but research in this area has been bedevilled by problems of identifying what is attributable to secondary schools and isolating progress in secondary schools from extraneous factors. The intention has usually been to show that something to do with the type of secondary schooling is associated with differential performance of pupils, sometimes regarded for this purpose as 'output' of schools. Rarely, however, can investigators adequately compare pupils of comprehensives with pupils of other schools. First, and crucially, it is not usually possible to distinguish 'outputs' from 'inputs', to disentangle school pupils' performance from what they were like before they went to secondary school. Different kinds of pupils tended to go to different kinds of school, and it is not a simple matter to compare like with like. Secondly, it is seldom possible to gather which of the school leavers went to the same kind of school throughout. Lastly, it has to be made very clear in what sense schools can be called 'comprehensive'. These three points will be elaborated in the next section.

(ii) Background to this project

This investigation was able to offer something of an improvement in all three respects, because it was based on longitudinal data from the National Child Development Study. Earlier researchers, even when able to make longitudinal comparisons, sometimes had to conclude that there were as yet too few comprehensive pupils to form a reliable sample (Douglas et al., 1968). The data for the investigation reported here are twelve years on from Douglas' National Survey data, and, although they date from the early 1970's and are thus specific to a particular, historical period, were collected at a stage at which comparisons between comprehensive and other schools were feasible. They also come from an era which is particularly practical for research comparing school pupils, as reasonably large numbers of pupils still went to grammar and secondary modern schools.

With adequate samples of that kind, based on national, longitudinal data, the comparison of pupils at sixteen was therefore practicable. The investigation was confined to members of the NCDS 'cohort' in England in maintained grammar, secondary modern, comprehensive or 'transitional' schools. The project started in 1977, with a grant from the Department of Education and Science. The findings of the first stage, published in 1980, were summarised in Chapter II. The main focus of the current report has been a repetition of the design used in that earlier work (outlined in Chapter I) which extends the comparisons between pupils in different sorts of school to provide an account of their public examination results. The examinations explored were detailed and discussed in Chapter III; all measures were confined to GCE ('O' and 'A' level) and CSE results. The findings on examination results were presented in detail in Chapter IV.

(iii) Definitions and assumptions

Who went to what type of school?

To isolate characteristics of sixteen year old pupils that might relate to their type of school, it was essential to know that they really had spent their secondary year in that type of school. With data from the NCDS, it was possible to specify which type of school was attended, which pupils spent all their secondary years in that school, which schools were comprehensive before these pupils started at them and which pupils went to secondary schools in the process of being reorganised while they were there ('transitional' pupils). This created 'pure' samples of genuine grammar, secondary modern, comprehensive and transitional pupils; but such purity had the disadvantage that it was artificial. Something like one fifth or a quarter of the sample in each type of school was removed, because they had been to other schools as well, and, with the exclusion of pupils from direct grant and independent schools, the sample of pupils in one maintained school throughout secondary years was about two thirds of the original total. Such constraints were the price of knowing the actual type of maintained school attended. They may seem elementary requirements of such research but they are not often fulfilled.

Defining 'comprehensive' and selective

The peculiar condition of English secondary schooling at the time of this study means that there are several terms to be defined, not least 'selective' and 'comprehensive'. As with any such piece of research, the limitations to the techniques of this study, its context in the history of school reorganisation and the artificiality of the attempt to compare comprehensive and other pupils all have to be understood.

The members of this study were elven years old in 1969, and most of them went through secondary school between 1969 and 1974. By 1974, over half the sixteen year olds were in comprehensives, a sufficient proportion to study their performance. It emerged, however, that over half the sixteen-year-olds in so-called 'comprehensives' at that time had spent their secondary years in schools in the process of reorganisation into comprehensives. These 'transitional' pupils had, therefore, started secondary schooling at a grammar or secondary modern.

There were also sufficient numbers who stayed throughout at a grammar school to mean that 'comprehensive' pupils, in their national average characteristics, were nothing like the 'all-ability' comprehensive collection of pupils intended for comprehensive schools. The average test scores and social class characteristics of eleven year olds in the sample here (see Chapter I) revealed that comprehensive pupils, at intake, were more like secondary modern pupils than grammar school pupils. The comprehensive pupils were, on average, lower scoring at eleven than their contemporaries in the combination of grammar and secondary modern pupils.

'Intake-correction'

The point of the research study reported here was to ascertain to what extent 'raw' averages were to do with what happened during secondary school and to what extent they were merely a function of the fact that, at eleven years old, before the start of secondary school, the average attainment of the grammar+modern combination was already ahead of the average attainment of the group of comprehensive pupils. For this reason, the level of performance which pupils reached by sixteen has always to be judged relative to the level they had reached by eleven, in order to allow any comment on secondary schools.

The major strength of this study is its use of longitudinal data, including information on the same children gathered when they were eleven, before secondary school, and when they were sixteen, towards the end of compulsory schooling. With these data it was possible to fulfil one of the chief assumptions on which study was based, the essential need to make some allowance for the level of attainment pupils reached by eleven years old, and to discount the contribution it made to the level of attainment reached by sixteen.

The earlier publication from this project did not present simple averages, for comparisons between sixteen year olds without allowance for attainments at the end of primary school (Goldstein, 1981; Gray, 1981). Throughout this report,

however, examples of 'raw' averages have been presented. Some averages were offered in the introductory Chapter I. There is a certain amount of material from which raw averages were offered in the introductory Chapter I. There is a certain amount of material from which raw averages can be derived, both in Chapter III, a discussion of the problems and the drawbacks of examination measures, and in Appendix R, which presents information on distributions of exams in NCDS data. Chapter IV, on the results of comparing schools on examinations, presents averages without 'correction' and averages 'corrected' to allow for components attributable to social class (father's occupation), sex, parents' interest in schooling at 16 and attainments and/or ability at eleven years old, before secondary school. Table 15 in Chapter IV elaborates the raw averages to show how the school groups differed in the relation of examination results to social class and attainment at eleven years old.

Overall averages, without corrections for home factors and for prior attainments, contain all the irrelevant material which could be confounded with, and mistaken for school influences. The 'raw' averages are relevant, however, because they did help to define the schools and indicate what 'comprehensive' meant. The 'raw' averages for each type of school showed that, in test scores and measures of exam results at sixteen, the grammar pupils were far higher scoring than the comprehensive pupils, who in turn were higher attaining than the secondary modern pupils. The combination of results from grammars and secondary moderns tended to be higher on average than that from comprehensives. These simple, 'raw' averages reflected, in large part at least, the intakes of the types of school.

Reasonable comparisons could be made between different groups of pupils only if adequate account was taken of the ways in which they differed independently of anything which could be ascribed to their secondary schooling. The solution was to employ a standard statistical technique for computing an average for each group of pupils 'net' of factors specified as likely to be confused with school influences, such as prior attainment, social class and other background factors extraneous to school. The comparisons of pupils at sixteen are therefore presented as averages for each school group, but should be assumed (unless described otherwise) to include corrections for earlier attainments, as well as for social class, sex and parents' interest.

It was one of the assumptions of this study, then, that an investigation of secondary schooling in relation to pupil progress has to include such control.

The interdependence of types of school

It is another assumption of this work that one of the important comparisons between comprehensive and other pupils should be with the combination of secondary modern and grammar pupils. It is not a sufficient test of progress in comprehensive schools to match their pupils against grammar pupils. A further premise of the work here, therefore, was that at the very least, progress in the 'selective' group had to be judged not solely from the progress of the selected, in grammars, but also from the progress of those who went to secondary modern schools. This is particularly necessary, since the secondary modern pupils were, it will be remembered, the majority of children, something like four fifths of the total under a fully selective system.

It is only in that context, and with the addition of careful provisos hedged about with reservations, that the questions addressed by this investigation can be crudely summarised as 'Did selection make a difference?' Readers familiar with the more valuable research in this field may find some of the basic explanation here superfluous, but some interpretations of the previous publication from this project have suggested the need to spell out the aims, assumptions and definitions again, and even to summarise them once more in this final Chapter.

Thus, it had to be re-emphasised that this study is no comparison of educational systems. There can be no empirical evidence, for England as a whole, on the effects of a comprehensive system, as comprehensivisation has, even today, not been completed. That kind of test of the aims or goals of non-selective schooling is impossible, and, certainly, there are no conclusions offered here concerning the merits of either ideal system, the selective or the nonselective. The NCDS data are from one particular stage in the reorganisation of secondary schooling in the early seventies. This helps to define what 'comprehensive' and selective or 'grammar+modern' mean in this report. Otherwise, the terms might have suggested, wrongly, the types of schools to which, in theory, they refer.

Although it is a central reflection that different kinds of pupils tended to go to different kinds of schools, it is not a purpose of this study to reflect the fact that different sorts of schools may have different aims. Some selective grammar schools may have concentrated on academic pursuits for academic children; some comprehensives may have aimed at widening academic opportunities, or broadening education; some secondary moderns may have tried to promote more concrete skills among less academic pupils. It is not the place of this study to attempt to define aims. Rather, the chief pursuit here was the measurement of attributes assumed for the purposes of the study to be part of the common aims of all schools.

(iv) Findings

Some assumptions and definitions have been set out. It is now necessary to swallow some of the reservations above and draw out whatever implications the findings had. It has been shown that it is not possible to tackle the real educational questions, but exploring 'whether selection made a difference' was initially tackled in two ways.

Type of School

First, it was necessary to make the nearest approximation to studying what difference being in a grammar school - and, in that sense, being selected - made, by comparing pupils in grammars with the 'nonselected' in secondary moderns and comprehensives. The purpose of comparing the three types of school was also to study any influence of being in secondary moderns, the major part of the selective system. Then too, the position of comprehensive pupils relative to those two groups could be examined.

On most measures, the 'intake-corrected' examination performance of grammar pupils exceeded that of comprehensive pupils, while that of secondary modern

pupils was relatively poorest. On general 'O' level/CSE grade 1 examination performance, such as 'passing five or more 'O' levels', grammar pupils were 'ahead' and secondary moderns pupils were performing relatively fairly badly, both compared to comprehensive pupils and, especially, in comparison to grammar pupils. Grammar pupils had the highest corrected average number of 'O' levels, secondary moderns pupils the lowest. There was no sign of secondary moderns being less likely than comprehensives to put equivalent pupils in for some sort of examination, however.

The two specific subjects explored, maths and English, can be regarded as two 'basics' and are certainly the most widespread subjects. A 'common-sense' question to interest an employer, for instance, might be whether a sixteen year old had the equivalent of an 'O' level pass in maths or English. The grammar pupils more commonly attained these recognised qualifications; even after intake correction they had highest 'relative odds', and secondary modern pupils had lowest. Comprehensive pupils had somewhat lower relative odds than grammar pupils, somewhat higher than those of secondary modern pupils. There was no sign that the grades obtained by candidates for 'O' levels in mathematics differed, while those for English did differ. Such candidates were, of course a minority of the population (and, in particular, minorities of secondary modern and comprehensive samples). Even on broader measures of 'O' level and CSE achievement, though, pupils in selective grammar schools appeared to make greater progress, while the nonselected majority in comprehensives and secondary moderns tended not to do as well.

In that respect, then, the exam results echoed the findings of the earlier phase of the project. They suggested, once again, that selection made a difference. Going to a grammar school was associated with increased chances of academic qualifications as well as relatively high attainments, while going to a secondary modern was associated with lowered chances of these qualifications and attainments. The position of comprehensive pupils, while it appeared to be in between the selective groups on most measures, was not as close to the grammar pupils as to the secondary modern pupils. Those in comprehensives, too, then, were relatively worse off in exam results, given their attainment at eleven, than pupils in grammars.

In that sense, the efficiency of the system was clear: selective grammar schools continued to be associated with greater progress, progress was related to selection. Those not selected, the majority, did not do as well as the grammar minority, even for given initial ability. Whether this was a result of treatment in secondary school or a function of the efficiency of selection cannot be known from this study. But, given the assumptions of the tripartite system, its allocation of children and resources to grammar and secondary moderns, and the differing extents to which they concentrated on academic work and examinations, the observations on secondary moderns and grammars are not unexpected. But what happened to the rest of the nonselected majority? Comprehensive pupils were, on average, 'worse off' overall, than grammar pupils; but a little 'better off' than the rather larger group, the secondary modern pupils. Could it be said that they were worse or better off than they would have been under selection?

Grammar+modern combination v. comprehensives

The question was approached first by exploring whether there was any evidence that they were doing worse or better than the average for grammars and secondary moderns, after correction for initial attainments and background.

There was no sign from these results that selection, in the sense of having a mixture of grammars and secondary moderns, could be said to have made a difference to examination performance. The comprehensive pupils did not differ from the grammar+modern combination in average numbers of 'O' level equivalents. With one or two exceptions, the measures at sixteen plus showed no statistically significant differences, and there was nothing to imply that comprehensives were worse or better.

In English exam results, there were two measures on which comprehensive pupils were lower attaining, but, on two other measures, no statistically significant difference was found, so there were no consistent implications. For mathematics, no overall difference appeared between the two groups. In terms of whether or not pupils 'passed' at 'O' level standard, there was no overall difference attributable to schooling, so a difference found in English 'O' level grades could not readily be interpreted. Nor was there a suggestion from CSE candidates'grades that any difference lay in their performance. It is not clear quite what it would mean for schools to produce candidates with lower average grades (including 'fail' grades) in 'O' level English while apparently remaining of equivalent standing in general. The results for English examinations present a good example of the lack of simple answers among examination measures when used as aggregate measures. Nevertheless, there was no evidence that the comprehensives were either failing to match or surpassing the performance of the grammar+modern schools, in terms of whether or not pupils obtained 'O' level passes or equivalent at CSE.

One striking feature of NCDS data was the large number of exams taken, by a very wide range of pupils. (Only 9 per cent of the sample here, it was seen, took no examination at all.) Any hypothesis that this might be associated with comprehensivisation, however, has no support from NCDS data. There was no evidence here of greater overall prevalence in comprehensives of taking some rather than no exam, given levels of attainment before secondary school. So, even in the odds of taking no exams, there was nothing to suggest the need to qualify the conclusion that grammar+modern and comprehensive groups did not differ.

Overall patterns; similarities between exam results and earlier findings

In the main, the findings here, to the extent that they suggest anything about schools, are broadly in line with those published in 1980. It seemed to make a difference which type of school pupils went to. The separate existence of grammar schools and secondary modern schools did appear to make a difference to their pupils' progress, in so far as that was suggested by a clear difference in intake-corrected test scores in the earlier work and in intake-corrected exam results here. Some of the expected results of rejection by the eleven plus were not apparent in the earlier work, since self-ratings and liking for school appeared higher, on average, among secondary modern pupils, once corrected for

eleven year scores. What could be suggested as an effect of non-selection on pupils of secondary moderns, though, were their lower results on more 'objective' measures. In test scores and examination results, secondary modern pupils were making less progress on average than their equivalents who went to grammars. Pupils who had been selected for grammar school were ahead of those who had not.

So secondary modern pupils were at an apparent disadvantage relative to grammar pupils. There, if nowhere else, is an old problem, highlighted by this project's results. Are educational interests best served by selecting rather few pupils for separate schooling, when the majority, the secondary modern pupils, are apparently, on average, underachieving? According to these data, there was no overall advantage in that selective 'system', in that comprehensive pupils did no worse than the grammar+modern combination.

That comparison is not to be mistaken for a genuine test of the relative merits of comprehensive and selective schooling. Grammar pupils were 26 per cent of the sample, including comprehensive pupils. The comparison between comprehensive and grammar+modern samples, as was made clear in Chapter I, was really, therefore, a comparison between a combination of the grammar and half the secondary modern pupils, on the one hand, and the remainder of the secondary modern pupils (in comprehensives), on the other hand. So another comparison of progress can be made, between secondary modern and comprehensive pupils. There may have been an advantage in going to a comprehensive rather than a secondary modern.

The main finding of the earlier stage of the study was borne out, that there was no clear evidence that progress could be regarded as better or worse in comprehensives than in the grammar+modern combination. In the sense that there seemed no overall advantage or disadvantage to comprehensives as against the selective, grammar+modern group, selection did not appear to have made a difference.

Subgroups of pupils (ability, social class)

Ranges of ability before secondary school

It is necessary to comment on how different groups fared in various school arrangements, particularly those of different abilities at 11, as the results of the previous stage were interpreted as suggesting, albeit tentatively, that, apparently, those who had been in the top fifth on one score at 11 made comparable progress in test scores in comprehensives and grammars. It was not left entirely clear, with the examination measures used, whether this conclusion could also be expresed in terms of examination results. There was one indication, from reliability-corrected regressions of numbers of 'O' level equivalents, that top scoring comprehensive pupils were not significantly different from grammar pupils who were equivalent to eleven. Without scaling this measure to allow for the fact that half the sample had no 'O' level, and for other distributional properties of the measure, it is not possible to rely on its implications. Those implications were, nevertheless, similar concerning the high scoring comprehensive pupils to the test score findings of the earlier study. They also resembled the pattern found in combined scales of English

results (despite the fact that this was a measure on which the overall average for comprehensive pupils was lower than that of the grammar+modern group). There is some justification for regarding numbers of 'O' level equivalents as a more suitable measure of exam performance than some other measures used in this work, and the analysis in question as the one which approached the sophistication of investigations of test scores at sixteen carried out for the earlier stage of the study. Even that indicator, though, is not as useful as the test scores, since the latter applied to everyone in the sample and were suitably distributed for the statistical techniques of the research. But there was weak support for the implications of results on test scores at sixteen in the lack of reliable evidence of differences in number of 'O' levels between top scorers in comprehensives and their equivalents in grammars.

On most measures of examination results, though, the apparent superiority of the grammar school performance relative to that in comprehensives and in secondary moderns held over the range of ability. This was evidenced by the lack of interactions. Where there were interactions, these had somewhat conflicting implications. Those in high scoring groups on tests at eleven seemed more likely in grammars than in comprehensives to pass five or more 'O' levels and to pass 'O' level maths or English. The secondary modern pupils of such high scores at eleven progressed relatively worse. Examination results on some dichotomous measures resembled the findings for truancy and parental satisfaction of the earlier study, in that the comprehensive pupils of high ability were 'worse' than their counterparts in the grammar+modern mixture. Those earlier observations were also based on dichotomous measures of outcomes (truants/not truants; satisfied/not satisfied) and the possibility suggests itself of a statistical artefact based on properties of the 'all-or-none' nature of the measure, affected by the lower intake of comprehensives when attainments were allowed for by grouping eleven year scores into fifths. On the criterion measure of pass/fail in mathematics at 'O' level, for instance, those in the top fifth at 11 in comprehensives had lower odds than the grammar+modern mixture. Results on this measure and on the analyses of proportions with five or more 'O' levels may be a result of lack of adequate allowance for eleven year attainment. Whether or not this was so, the results for the subgroups of comprehensive pupils did not suggest lower odds of five or more 'O' levels than those in grammar+moderns, among pupils in the top fifth. In that range, lower odds could have been expected, given the superiority of grammar pupils and the few secondary modern pupils. There were slender and not entirely consistent grounds for regarding the 'able' comprehensive pupils as lower than equivalent grammar pupils in the numbers of 'O' level equivalents they gained by sixteen (with the possible exception of those comprehensive pupils with the highest scores). If the question concerned the ablest children of all, however, it can not be answered empirically. This is because there were a number of grammar pupils at the top end of the combined score at eleven years old, that is, the ablest two per cent or so of the sample, for whom no comparison with any equivalent comprehensive pupils could be made, because no-one with such a combination of high scores went to comprehensives. (Thus, Figure 5 shows the grammar pupils drawn further to the right than any others.)

These details may seem slightly inconclusive to those who are impatient to know about 'bright' children. This is indicative of limitations to the measurement of school populations by means of examination results. To the extent that it

was possible to measure the relative progress of grammar and comprehensive pupils, the findings in terms of average numbers of 'O' level equivalents were not strong enough to contradict findings for test scores, published in 1980.

Problems in interpreting interactions with prior attainments

The difficulties in interpreting interactions stem from the fact that, in 1974, it was still true, as Monks had concluded six years earlier, that there were too many grammar schools to permit evaluation of comprehensive schools (Monks, 1968). Too many of the top ability pupils were creamed off to grammar schools to allow any inferences as to what true comprehensives would be like. The results here highlighted this. The average intakes at eleven of comprehensives were lower attaining and tended to be of lower social class than the average for grammars and moderns (and than the national average). And results at sixteen of the grammar pupils of highest scores at 11 could not be compared with anyone else's results, as no one who had been their equivalent at eleven went to a comprehensive or a secondary modern. This meant that something like the 'top one or two per cent' of the sample here were not included in the more sensitive comparisons, but they were included in and may have altered implications of comparisons of the 'top fifth' which looked at proportions passing exams or taking no exams.

The interactions complicate the account, and their interpretation is difficult. The fact that examination measures apply only to particular ranges of ability and that different types of schools, too, take different ranges of ability means that interactions, or a lack of statistically significant interactions, may not be translatable into educational evidence. One difficulty is that a lack of schooltype interaction ought to have implied an interaction between test score and the grammar+modern/comprehensive distinction, but this did not occur. Coupled with variation (perhaps systematic variation between types of school) in pupils put in for examinations, and the fact that some measures apply only to entrants for exams, this illustrates some of the difficulty in inferring anything from interactions. These are problems peculiar to examination results and some other measures used earlier in this research, and do not apply to test scores.

Social class

As well as looking at different school groups over different ranges of ability, the study explored the patterns for different social classes, but there were few observations to qualify the findings corrected for social class. The slender evidence relevant to the theory that bright working class children were assisted more by the selective system than by comprehensives failed to support that theory.

Subgroups of comprehensives

Were there differences between categories of comprehensives? It has been made clear that, once due allowance was made for the contribution of social class, sex, parents' interest and attainment and ability before the start of secondary school, to general 'O' level and CSE attainment of this kind, there was no consistent difference between grammar+modern pupils taken was a whole and

comprehensive pupils. Beyond that, however, the relative position of sixteen-year-olds in different types of school was less clear, as seen in the findings for those in comprehensives of various origins and dates.

Pupils in the grammar+modern group were in between different subgroups of comprehensive pupils, as might be expected from the lack of difference between them and the average for comprehensives, in terms of the numbers of 'O' level equivalents they obtained. In proportions getting five or more 'O' level equivalents, though, grammar+modern pupils were lower scoring than any comprehensive subgroup. This was the only measure, apart from proportions with no exams, which showed overall differences reaching high levels of statistical significance. Some comprehensive groups, namely early, ex-grammar and ex-secondary modern comprehensives, were clearly relatively likely to contain pupils with five or more 'O' level passes. It seemed also the case in early ex-grammar and former secondary modern comprehensives that fewer people took exams than among equivalent grammar+modern pupils.

Inconsistencies in the relations between early and other comprehensives, in other findings as in examination results, and the implications of results of interactions, were further deterrents to generalisations. It is possible, however, to draw one implication from findings for early comprehensives. Exam results were closer to consistency with pupils' plans for the future, with patterns of intended school-leaving age and perhaps with plans for advanced courses of study, rather than being related to findings on attainment tests or on truancy, staying away from school or teachers' ratings on behaviour. But it was not possible to characterise the exam performance of early comprehensive pupils as consistently different from that of recent comprehensive pupils.

It seemed ex-grammars may have conferred a slight advantage, along with ex-secondary moderns, in exam progress relative to eleven year old score. Interactions detracted from the likely interpretation that findings for ex-grammar schools were linked with the greater prevalence of ambitions for advanced study in those schools. It is possible to speculate that the method of reorganisation assoicated with converting single schools into comprehensives had advantages, but this finding is specific to relatively few schools at that particular period of comprehensive re-organisation. The reason for exploring these subgroups was that the comprehensive sample could have been affected by selection between comprehensives. It does not seem that the differences in examination results between older and newer comprehensives, and between those formed from a grammar school or a secondary modern, and the purpose-built or amalgam comprehensives, were simply predictable from differences in intakes.

Transitional pupils' examination results

The study distinguished pupils in comprehensives throughout their secondary years from 'transitional' pupils. Results in some respects suggest that being in school while it was being redesignated 'comprehensive' was associated with disturbance to examination results. The lower examination attainment of transitional pupils was not, apparently, due to any reduction in exam entries among the low attaining pupils, if that is indicated by the proportions with no exams. Transitional pupils had relatively low proportions with no exams.

(Moreover, the lower exam results of transitional pupils overall do not tell the full story, as, on some interactions in analyses of proportions, their top fifth pupils were quite high attaining. It has been seen that this has to be treated with caution and might relate to inadequacies in 'proportions' analyses.) It is also possible that pupils in transitional schools made up for passing fewer exams by 1974 by taking more later on, after sixteen years old.

If reorganising was temporarily detrimental to examination results, it has to be pointed out again how specific these results are to the period studied. Areas which changed over to comprehensives later might have done so without this temporary disruption to exam performance. 'Reorganisation' as a longer term establishment of all-through comprehensive intake may be achieved with more stability in examination performance.

Other results of 'transitional' pupils

The poorer performance of transitional pupils, though it lends itself to comparatively easy interpretation, would not be predicted from their results on test scores, where they resembled the grammar+modern pupils. If exam results were lowered as a result of disrupted staffing or accommodation or changed syllabus caused by reorganisation, this was not part of a classic picture of disturbance. There was no more truancy or bad behaviour in the transitional group, relative to earlier attainment, than in the grammar+modern combination. The closest similarity between exam results and any other measure for this group was with the plans for advanced study. Transitional pupils had the lowest odds of wanting advanced courses and planned as early a leaving age as the grammar+modern group. These measures may imply, as this project suggested in 1980, that some 'children found school particularly difficult during the year or so of comprehensivisation'.

(v) Reservations

Reservations about examination results as measures

The studies of exam results here are by no means exhaustive of the countless aspects of examination performance which it would be possible to measure. Different measures might have had different implications. Though it is tempting to try to list alternative approaches, it is left to readers to suggest their own. The examination measures used were a selection, albeit the most obvious ones. Most of the continuous measures (numbers of 'O' levels, etc.) are unsatisfactory measures for the statistical techniques used, though the control for eleven year attainment which the techniques permitted was relatively sophisticated. Proportions were statistically legitimate measures but control for eleven year attainments was in some respects less sensitive. The drawbacks and limitations of examination results as measures of school progress discussed here are those relevant to research measurements; this report is not concerned with the fitness of examinations as indicators of the educational progress of individuals. For the purpose of comparing schools, there are so many defects to exam measures that they are potentially misleading. Chapter III of this report has highlighted the arguments against using exams, as well as some strategies

using exams, as well as some strategies adopted for offsetting those drawbacks and some of the improvements, such as scaling of the measures, which could be carried out by further work and might even alter conclusions.

There is an appearance of meaningfulness to examination results which may convince readers who are unfamiliar with NCDS tests, for instance, that examinations provide more convincing evidence. It is particularly easy to slip into the error of regarding the intake correction here as actually adjusting individual's grades. Thus Wrigley (1981), in his criticism of this study's correction of test scores for prior attainments and social class, ridiculed the notion that anyone would allow him the benefit of a few extra grades in his school exams on account of his humble origins. Such a comment serves to highlight the danger of confusion surrounding researchers' attempts to assess the role of type of schooling. The danger is particularly acute when, for an individual school leaver, his or her examination results are publicly comprehensible and have 'market value' as qualifications, whereas there is no such familiar value attached to a score on an NCDS test. It is important not to confuse that with the relative validity of test scores, as against examinations, as indicators of national averages. The latest stage of this study has put forward some difficulties in using examination results to measure large groups. It would be unwise to rely heavily on any single measure of examination performance, and the test scores used in the previous stage of the work are better material for use with the techniques employed. In Chapter III, it is made clear that interactions, particularly, may not be the best way to answer questions about examinations. One way to judge the value of analyses, however, is to assess the extent to which the statistical model used in each analysis provided a fit of the data, and how good an explanation of the factors that go into examination performance is provided by school differences here (see Appendix S).

Reasonable comparisons using examination results present many difficulties, not least that findings depend to some extent on the measure used to characterise examination performance. They are also affected by the exact nature of the allowance made for extraneous factors influencing the results. No study in this field could be definitive, therefore. The account of examination results here should be regarded as a preliminary investigation of examination results measures, basing tentative conclusions on the evidence so far. There are plenty of suggestions throughout the text for further analyses which could be carried out. Apart from that, the evidence points to the difficulty of drawing conclusions from examination results. There is a risk in generalising from performance in any one subject, to any one level or with any particular type of measure or sample of children. Even if the measures of examinations here were adequate as measures of examinable progress, it would be wrong to judge schools solely by that, particularly as exams have restricted access. Other explorations of this study revealed that the types of school under consideration varied greatly in the number of pupils going in for various examinations, after allowance for earlier attainments. Judgements of schools based on examination results will depend on the extent to which these factors can be taken into account. Yet they must be allowed for, and interpreters of exam results will therefore have to be fully aware of pitfalls such as are demonstrated here.

The overall approach of this study to measuring examination performance may perhaps be relevant to the now compulsory publication at Local Authority level of local examination statistics. At least, it has been possible to point out a few vital considerations in attempting to allow for intake differences. What the study underlines is the fact that examinations are in many ways unsatisfactory as measures of large numbers of pupils.

A particular feature of the examination measures should be noted here, which is that they are restricted in the case of 'O' levels and CSE's to exams taken by 1974. This again is a traditional bias and may artificially reduce the level of performance in schools (perhaps those with new sixth forms), where 'O' levels can be taken later. Ryrie (1981), writing of Scottish schools more recently, notes 'evidence of a slight but interesting change in the character of the fifth year in recent times [17-year-olds, in Scotland]: more young people have apparently been staying on with a view to sitting or resitting 'O' level grades than was previously the case'. If such changes were affecting schools by 1974, it may be that examination performance in less traditional schools was underestimated by this research. In particular, the lower examination results in transitional schools may have been a function of this restriction. Table 9 in Chapter III suggests that reorganisation may have delayed rather than prevented examination work.

Exams are not the only or best way to measure schools. Research such as this can promote overconcentration on exams, indeed on those of high attainments at eleven. Gray (1981a), for instance, points out that 'in a system where the average leaver obtains barely two 'O' levels and assorted CSE awards ... there is a distinct danger that by emphasising the primacy of examination qualifications, we produce models of effective schools which are inadequate to the needs of a substantial minority of pupils'. It is hoped that, by relating exam results to the earlier findings of this study, on test scores, plans for the future and behaviour, this report broadens the picture that would otherwise be wholly based on exams.

In that context, to permit comparisons with other research on selective and nonselective schools and to assess comprehensive pupils by the criteria, amongst others, of 'grammar school performance', though questionable, was part of the task required. Nevertheless, considerations such as those above encouraged the use of a variety of different measures, so that the collection of observations should be relevant to the whole spectrum of pupils.

Reservations about 'intake correction'

(a) General remarks

The findings here, then, have to be seen as the product of choices of outcome measures. They are also dependent on choices of what factors to take into account, of how to allow for intake differences. Social class (father's occupation) was explored, for example, but the study did not take into account the material circumstances of pupils' homes (as indicated by income, housing tenure, crowding, or free school meals, for instance). These factors might well have been influential in truancy (Tibbenham et al., 1977; Halsey et al., 1980; Galloway, 1982). There are any number of other possible influences that

could have been taken into account. Further work could usefully extend the analysis here to relate progress in secondary schools to progress in primary school, between seven and eleven years old. Above all, the results here should underline the importance of a longitudinal approach to measuring outcomes of schooling. However, it is always arguable in defence of the necessarily limited allowance for extraneous factors that, once several important background factors have been taken into account, others make only modest contributions, statistically speaking, to explaining outcomes. Here, social class and parents' interest in schooling may have stood in for material home circumstances.

Nevertheless, just because allowance for intake has permitted isolation of the secondary school element, there is no justification for ascribing causal links to the relation between school group and performance. All such a study can do is remove some obvious extraneous factors and report any associations between school category and results. It will be clear that not all extraneous factors are included in the intake-corrections; technical constraints, as well as the complexity of educational influences, precluded anything like complete allowance. Because of that, and because of limitations mentioned next, it would be a mistake to regard even 'intake-corrected' observations as necessarily a function of school factors.

(b) Balance of intake

There were all sorts of other factors that may have contributed to results. The detail of findings for various ranges of eleven year score frequently suggested likely interpretations in terms of the dissimilar intakes of schools and pupils' peer-groups. These would be influences over and above the individual's eleven year score. In very few analyses (all in the earlier stage of the project) was the social composition of the school allowed for, and in none was it really possible to allow for ability mix, as was discussed in the earlier publication. Yet, it is a commonplace that the intake of a school will create an environment beyond what is allowed for here in the controls for characteristics of pupils at the time of intake. Rutter and his co-authors, for instance, judge that 'The presence of a relatively high concentration of pupils in the upper ability groups may work to the advantage not only of those pupils themselves, but also of their peers. In a similar way, a largely disadvantaged intake might depress outcomes in some cumulative way, over and above the effects of a disadvantaged background on the individual pupil' (Rutter et al., 1979). The pupils here were in groups of schools that differed in the mixture of children they took at that time.

Some of the more detailed observations on the age range of schools, for instance, prompt the speculation that the balance of intake in a school, or sample of schools, was a powerful influence, even after the attributes of individual pupils were allowed for. Those comprehensive schools which formed part of an upper school system and started at 12 or 13 years old, for example, seemed both to have more advantages in attainments at eleven years old and to produce examination results comparable to those of grammar schools at sixteen. If these were more likely to be part of a Leicestershire system in which only those schools starting at 12 or 13 had pupils going on to sixth forms, perhaps they had a more favourable 'balance' of intake. Their average intakes were

closer to those which would be expected in 'fully comprehensive' schools; this study does not reveal whether a more advantageous balance, as a function of the 'selective' nature of such schools relative to 11-16 comprehensives, also contributed to results. The balance of intake argument, though, is not supported by the findings for transitional schools, which had higher scoring pupils at intake than comprehensives and seemed to produce relatively lower exam results.

(c) The weights in the grammar+modern combination

In the comparisons above between the combined grammar+modern 'selective' group and the comprehensive pupils, allowing for initial differences at eleven years old was more reasonable than other approaches. The statistical allowances here may offset some initial lack of comparability. But the comparison of comprehensive pupils even then is relative to a non-equivalent grammar+modern sector. The weight of grammar school pupils in the 'selective' group meant that the comparison did not correspond to anything like a 'real' selective versus nonselective comparison. The proportion of grammar pupils is too high in that grammar+modern mixture. This is part of the artificiality of attempting to compare 'selective' and 'nonselective' pupils in the mixture of schools which coexisted in 1974. The comprehensives were comprehensive in name only. We cannot know which way the results would go if the proportion of secondary modern pupils had been more like the four-fifths majority of full selection. Appropriate proportions could be created by reweighting, and other researchers could undertake that (though such reweighting requires many questionable assumptions: that the interaction effects can be taken in, that certain measures cannot be reweighted as they apply only to limited ranges, that the comprehensive sample does not need reweighting). It was explained in Chapter I that, without such reweighting, the ratio of one-third grammar to two-thirds secondary modern makes the comparison with comprehensives peculiar. For example, 27 per cent of the grammar+modern group is in the top fifth, as against 13 per cent of comprehensive pupils; but how should such groups be reweighted? There are as many questionable assumptions in reweighting as in using the actual numbers of pupils as they happened to be.

(vi) Overview: some possible implications and limitations of the work

The implications of findings are not entirely consistent and conclusions are therefore not easy to draw. It can be said, however, that there is no evidence here for or against a policy of selection or a policy of comprehensive schooling. With those provisos, it is also clear that no-one could draw the conclusion from these data that the introduction of comprehensive schooling was a mistake. Though this study reveals that selection may have made a difference, in a number of ways, it did not show comprehensives as 'failing'. To make that claim, it would have been necessary to show that their pupils had lower averages overall than the combination of grammars and secondary moderns, and this was not the case. It is therefore possible, despite all the reservations, to give an answer to the oversimple question, 'Do comprehensive pupils do better or worse on public exams?' The answer appears to be 'No'. The answer, that no evidence of a schooling difference appeared, derives its meaning chiefly from its negation of claims for which there is an equal lack of evidence, such as that 'comprehensive pupils are worse at exams'. Ill-founded though such claims are, it is worth posing against them the sort of evidence which it is possible to

obtain, and the reasons why there is no evidence on that level for such generalisations. From the evidence here, which, it will be remembered, concerns an earlier period of educational history, it appeared that the particular combination of grammar and secondary modern schools which there happened to be did not seem to enhance examination performance, relative to the schools then in existence which were called 'comprehensive'.

In so far as conclusions can rely on the overall lack of difference between the grammar+modern combination and comprehensives, it poses the crux of the problem for any 'selective system', underlined by these findings. For, even though grammar schools produced better exam results, selection as an organisation requiring grammars and secondary moderns did not seem, overall on balance, to achieve more in the way of examinations than the average for comprehensives, when due allowance was made for initial attainment and background. No statement that comprehensives were worse or better could be founded on these data.

There are a number of theories - not all mutually compatible - concerning the differences between a selective and a comprehensive system. They are based on the theoretical underpinning of the two systems, as they would be ideally, and not on empirical eyidence. This study has highlighted the lack of empirical evidence for many claims that are made about different kinds of schools.

It may be suggested, for example, that selection for grammar school enhances the academic performance and opportunities of the minority who are selected. There is no way of testing that theory in these data; the comparison of grammar pupils and other children was possible, because grammar pupils still existed. Because they did exist, however, the intake of comprehensive schools was closer to that of secondary moderns, and did not cover the full range over which grammar pupils existed. There did appear to be an advantage in selection, though, overall, to those who were selected, the grammar school pupils.

There are other theoretical questions relevant to the investigation. One view of comprehensives as making a 'grammar school' type of education more widely available, implies that a test of comprehensives would be proportions obtaining a 'grammar school' qualification of five or more 'O' levels. Taking that as the criterion, it seemed that, given attainment at eleven, comprehensives were as likely as the selective grammar+modern combination to promote the acquisition of five or more 'O' level equivalents. When the three school groups were separated and analysis was confined to the top fifth at eleven, results confirmed that odds of five 'O' levels were lower in comprehensive than in grammar schools, while secondary modern pupils' results were lower still*. On another measure of 'grammar school education', i.e. the numbers of 'O' level equivalents, the very highest scorers at eleven in comprehensives may have fared about as well as their equivalents in grammars, but this did not hold for all the 'top fifth'. Because comprehensives include children of lower initial attainment than grammar schools, it may be believed that they must therefore fail to cater for bright children. From the data here, there is little reliable evidence to support or

* Though see the Statistical Appendix for possible effects of adding .5 to scores.

detract from that theory. The difficulty in obtaining reliable, clear evidence has been shown; but, in so far as it is possible to draw conclusions, the only eliability-corrected interaction suggested that 'able' children in comprehensives differed slightly (with the possible exception of the very top scorers at eleven) in numbers of 'O' level equivalents from similar pupils in grammars. The pattern in English grades and the lack of interactions means that averages for those high scoring at eleven seemed lower in comprehensives than grammars, and higher in comprehensives than secondary moderns. The implications of the evidence from the earlier work on test scores of 'able' children doing as well in comprehensives as in grammars and better than in secondary moderns had been that there was no evidence of an advantage in selection; those in comprehensives did as well and as badly as under selection. These examination measures were less revealing. It seemed from both test scores and examination results, though, that those who under a selective system went to secondary moderns were hampered, if anything, relative to their progress in comprehensives.

It did not appear from this study that the relatively lower examination performance in secondary moderns stemmed from any inability of secondary moderns to develop the idea of pupils going in for some examinations. It seemed that secondary moderns had taken up CSE examinations with as much speed as comprehensives. Another hypothesis, that well-established patterns in secondary moderns of catering for pupils with non-academic interests would mean larger proportions of non-examinees, was therefore not supported. Exam entry, it appeared, had been revised with the advent of CSE's - and, perhaps even in its first year, with the raising of the school leaving age. Despite taking examinations, the secondary modern pupils were 'under-achieving' in the examinations which they did take.

The grammar school pupils were, before corrections to data, at the sort of advantage which Gray (1981a) identifies as accruing to pupils from going to a 'good' comprehensive rather than another comprehensive school. Gray likens this advantage to the 2 'O' levels or so which might appear from the Oxford Mobility Study to be the result of going to a public/independent school. The latter, however, showed no gap between grammar and direct-grant/HMC schools, after allowing for social composition of schools, in line with Douglas' findings (Douglas et al., 1968; Halsey et al., 1980). Intake-corrected data here suggested that, for the half of pupils who were obtaining 'O' level equivalents, the average advantage attributable to being in a grammar rather than a comprehensive was less than two thirds of an 'O' level equivalent. The average advantage of being sent to a comprehensive rather than a secondary modern appears to work out at a little over a third of an 'O' level. The theoretical justifications for one or other system often fail to emphasise that the existence of grammar schools meant the simultaneous existence of secondary moderns. Here, as well as comparing the three school types, the project has given equal emphasis to comparing the combination of grammar and secondary moderns with comprehensives. According to the theory of selection, the hypothesis could have been entertained that the overall average for grammar and secondary modern pupils would be 'better' in examinations than the comprehensive on average. This could be rejected. Since many of the arguments for selection focussed on its value for those predicted as potentially excellent at

examinations by eleven years old, the absence of improved performance overall among the 'selective' combination is noteworthy. In terms of the actual allocation of pupils to schools in this sample, moreover, it might be seen as unlikely that comprehensive pupils would equal the combination of grammars and secondary modern, given that at this particular period, between 1969 and 1974, comprehensives co-existed with a considerable degree of selection.

In that context, it is surprising that it did appear that some aspects of the underachievement and early school leaving associated, according to earlier research in this field, with selection, may have been mitigated by the existence, even within a selective system, of comprehensive schools. Uncertain though examination findings were, they suggested that, among those of higher scores at eleven, comprehensive pupils or subgroups of comprehensive pupils equalled or even surpassed the average for the grammar+modern group. Such results are not wholly unexpected, though, in view of the suggestive findings of smaller-scale or local research studies, including those of Ross et al. (1972), who found more examinations taken and a later leaving age than average in the twelve comprehensives studied.

It did not seem likely, from these results, that pupils' exam attainments were differing systematically in a way that would permit the conclusion that going to comprehensives, as against the 'alternative' grammar+modern combination, explained examination performance. Clearly, a large component of exam performance was related to preexisting attainment. Then, there was great variety between different subgroups. There was variety in the degree to which, after one or two measures of preexisting attainment were allowed for, another second or third measure (test score, social class) had a different role. These differences were such that it was impossible to be sure about high ability pupils. Various measures of examination performance did not have entirely similar implications. Certain measures of exam performance were used that, under selection at least, were more or less confined in application to the 'able' group in grammar schools, such as whether or not pupils obtained five or more 'O' levels. Overall averages for all but two measures implied that the examination results of the grammar+modern group and the comprehensive pupils' results did not differ. The superficially inconsistent details of examination results, however, illustrate some of the limitations of using examination results to measure national average characteristics of schools. The lack of an interaction (or lack of difference between the grammar+modern and comprehensive groups) over the range of initial ability is difficult to reconcile with the general lack of school type interactions and with the higher relative odds for some comprehensive subgroups than for the grammar+modern groups among higher scorers. This might be a function of relatively small numbers in each subgroup. Because of such inconsistences, though, it was necessary to conclude that there was no sign that one or other schooling organisation was 'better' or 'worse'. The only other conclusion to which the findings did seem to point was that, as far as exam results were concerned, on average, being in a school in the process of transition to comprehensive status was detrimental. This is unlikely to be a problem for as many schools now as it was between 1969 and 1974, the years under consideration. But there is no evidence that the apparent effect on exam results of the unsettling transition lasted, in the results from recently comprehensivised schools, classified as comprehensives before the study children arrived.

The main focus of this stage of the project has been examination achievement during compulsory school years in the different school groups. Only speculation is possible concerning work done after sixteen-plus examinations (though some mention is made in Appendix H of sixth form work). The fact that the study is child-based is a reminder that examinations measure pupils, and are not designed for measuring schools or even aggregates of children. The strength of the study was its ability to look at a national picture, with longitudinal data. Those data were carefully controlled to derive a reasonable comparison with allowance for differences which existed before secondary school, rather than crude averages. Even with NCDS data and with advanced statistical techniques, it was not possible to reach any certain conclusion. The difficulty of using examination results to measure schools and applying intake corrections to exam measures meant that no study of this question could be definitive. A study such as this is not an assessment of schools, still less of systems, but rather a description of how one system, the education system in England, worked in 1974, when the types of schools differed in their intakes. The child-based, as against school-based nature of the work prevents any investigation, unfortunately, of how performance relates to 'balance' of intake. This is not to imply that schools simply reflect their intakes. Intakes and performance reflect resources, teachers and headteachers, of course. The allowance for intake did not 'compensate for' the fact that schools vary in other ways, but this project was not concerned with such variations.

Nor do such empirical data allow investigation of the effect of 'balance of schools'. There are no empirical data on what the examination performance of this sample would have been like in various schools if these children had been allocated to different schools in different numbers.

The averages for each type of school may suggest that it is appropriate to interpret them as showing how pupils would have done if they had been in, say, a grammar as against a comprehensive or a secondary modern. However, it is a crucial function of the fact that these are empirical data that the averages show how the pupils in grammar schools fared, given that those other pupils were in comprehensives and secondary moderns. They demostrate how pupils in comprehensives did, given that the proportion of pupils in grammars was 26 per cent and that 42 per cent went to secondary moderns. They are not indications of how a child who went to a comprehensive would have fared if she had gone to a grammar, but of how she would have fared as long as some other 'equivalent' child went to a comprehensive instead. Comprehensive pupils were being compared with a group which, by containing the selected, defined the comprehensive intakes. The chances of being in the comprehensive group were not, then, independent of who was in the 'selective' group. Had they been so, the intakes to comprehensives would have been higher-attaining at eleven, and, presumably, schools would have responded to their intakes.

However essential it may be to allow for intake differences, measuring exams cannot take into account that schools are not static, but can change rapidly. When parents judge a school on its past record of examinations, they are also reckoning up the experience and expertise of the school. But the differences between school types here, which existed even after correction for individual pre-existing attainment, indicate that schools are more than pupils. It could be inferred that schools differ in their promotion of skills, that pupils respond to each other and to schools, and that schools respond to intakes. This

prompts the warning that, because it is tied to a particular historical stage, but also because it concerns national averages, a study such as the one reported here cannot enlighten any parent about which school would suit their children best. It would be a mistake to try and calculate the relative chances of a known or imagined pupil in one or other type of school; the average chances explored here are national averages and may not apply to paricular cases. Any individual case would be better informed by researching the relatively few schools and options available to a particular child, with more up-to-date and local knowledge than can be offered by a national study.

It would be wrong for other fundamental reasons, to take these observations to imply that all would be well if there were grammar schools for all. The grammar schools would not have existed without the secondary moderns. The grammar school average has to be judged alongside the secondary modern average, though it is not part of this study to speculate about how the resources for each type of school were interdependent. The study could only compare the apparent progress of grammar pupils with what they would have done in secondary moderns or in comprehensives if other children had gone to grammar schools. The results of comparing grammars and secondary moderns may not be unexpected, even though, in theory, the comparison could not be made, as the same children could not have been allocated to one type of school if they were suited to the other type. Had grammar schools not put their pupils at an advantage in academic attainments and examination results, though, relative to what would have happened if they had gone to secondary moderns and other pupils had gone to grammars, the tripartite apparatus would have been seen as a failure. One of the implicit questions with which this study started was whether,given the same bunch of children, comprehensives would produce the same result. As the schools did not have the same assortment of children, the study had to rely on statistical techniques which compensate for nonequivalence of samples. But that could not provide an answer to whether comprehensives would produce the same results as the grammars and secondary moderns, given the same children, because that is not amenable to empirical scrutiny.

Similarly, the comparison of the grammar+modern combination and comprehensive pupils could be interpreted as showing what pupils would be like if they were in one or other system. All that empirical data can demonstrate, however, is what pupils would be like if they were in comprehensives and those other pupils were in that mixture of grammars and secondary moderns. This does not detract from the fact - indeed, this project strems from the assumption - that it is a real question whether comprehensive schools, during the period of reorganisation of the first half of the 1970's, were catering for their pupils, in many aspects of attainment and plans, as well as in public examinations.

Arguments for and against the influence of selection on examination results could not properly be assessed, when selection continued to affect all the pupils in the study and made the intakes of comprehensives of lower attainment on average than the average for the grammar+modern combination. This was not, therefore, a test of two systems, the selective and the comprehensive, still less a test of the ideals of those systems, but simply a presentation of some information on the operation of the English secondary school system at a particular historical stage in its piecemeal reorganisation along comprehensive lines.

The somewhat conflicting results broadly confirmed the lack of evidence, either of triumph or disaster, which the earlier report of this project described concerning the comprehensive schools of the era. Nothing in this report should be taken as suggesting what would happen if there were a fully comprehensive system, nor what would happen if that had many years to become established, as the selective system had by 1974. It should be clear that the task here, to describe progress in comprehensive and selective schools, to the limited extent that progress can be judged with examination results, was almost impossible. The test scores at sixteen of the same pupils provided a more suitable indicator of relative progress than examinations. Even so, complete allowance for extraneous factors, even with data from NCDS, best suited to such an investigation, was impossible in the educational context of the 1970's. Nevertheless, the data used were the best empirical evidence available on the allocation of pupils to schools at that time and on the progress of pupils in those schools. They serve to show which of the theories mentioned in this chapter form questions to which no empirical answer is possible. Any evaluation of progress in comprehensives would need to find comprehensive intakes, just to study their examination results, let alone to test the aims of comprehensive schooling.

References:

Baldwin, R. (1979) Performance at GCE Advanced level in maintained schools (National Council for Educational Standards).

Baldwin, R. (1981) Secondary schools 1965-79 (National Council for Educational Standards).

Bellaby, P. (1977) The Sociology of Comprehensive Schooling (Methuen)

Benn, C., and Simon, B. (1972 second edn.); "Half Way There: Report on the British Comprehensive School Reform" (Penguin Education)

Coleman, J. (1966) "Equality of Educational Opportunity" (Washington D.C., U.S. Department of Health, Education and Welfare)

Conklin, M.E. and Dailey, A.R. (1981) "Does consistency of parental educational encouragement matter for secondary school students?" Sociology of Education Vol. 54 (Oct), 254-62

Cox, C. and Marks, J. (1980) "Real Concern" (Centre for Policy Studies).

Cox, C.B. and Dyson, A.E. (eds.) (1969) "Black Paper Two: The Crisis in Education" (London; The Critical Quarterly Society, October)

Department of Education and Science: Statistics of Eduaction (1974) Vol. 2, Tables 4,6, 22 (1970) Vol. 2, p.viii

Douglas, J.W.B. (1964); "The Home and the School" (MacGibbon and Kee/Panther, 1967)

Douglas, J.W.B., Ross, J.M., and Simpson, H.R. (1968) "All our Future: A longitudinal study of secondary education" (Peter Davies)

Eggleston, J. (1967) "Some environmental correlates of extended education in England" Comparative Education 3 (2) March 1967 reprinted in Eggleston, J. (ed.) (1974) "Contemporary research in the sociology of schooling" (Methuen)

Eggleston (1977) "The Ecology of the School" (Methuen)

Floud, J. and Halsey, A.H. (1957) British Journal of Sociology Vol. 8, pp. 33-39 "Intelligence tests, social class and selection for secondary schools".

Floud, J. (ed), Halsey, A.H. and Martin, F.M. (1956); "Social Class and Educational Opportunity" (Heinemann)

Fogelman K. (1976) "Britain's Sixteen-Year-Olds" (NCB)

Fogelman, K., Tibbenham, A and Lambert L. (1981) Absence from school, in Berg, I. and Hersov, L. (Eds) "Problems of School Attendance and Refusal"

Ford, J. (1969) "Social Class and the Comprehensive School" (RKP)

Fraser, E.D. (1959) "Home, Environment and the School" (University of London Press)

Galloway, D. et al. (1982) "Schools and Disruptive Pupils" (Longman)

Ghodsian, M., Fogelman, K., Lambert, L. and Tibbenham, A. (1980) "Changes in behaviour ratings of a national sample of children" British Journal of Social and Clinical Psychology Vol. 19 (No. 3) 247-256

Glossop, J.A. and Roberts, C. (1980) "An Exploratory Study of Examination Policy Differences and Performance in Three Comprehensive Schools", Education Review Vol. 32, No. 1 pp. 67-85

Goldstein, H. (1981) "Research on Comprehensive Schools: a review of 'Progress in Secondary Schools'" in Forum Vol. 23 No. 3 Summer 1981

Gray, J. (1981a) "A Competitive Edge: examination results and the probable limits of secondary school effectiveness"; Educational Review, Vol. 33 No. 1.

Gray, J. (1981b) Review of Progress in Secondary Schools, British Educational Research Journal Vol. 7 No. 1 1981

Halsey, A.H., Heath, A. and Ridge, J. (1980) "Origins and Destinations" (OUP)

Heath, A. and Clifford, P. (1980) "The seventy thousand hours that Rutter left out"; Oxford Review of Education, Vol. 6 pp. 3-19

Hopkins, A. (1978) "The School Debate" (Penguin)

Husén, T. (1979) "The School in Question: a comparative study of the school and its future in Western society" (OUP)

Ives, R. (1980) Examination Results of NCDS in relation to test scores and social class (NCB, unpublished)

Jackson, B. and Marsden, D. (1962) "Education and the Working Class" (RKP)

Lacey, C. (1981) Review in Times Higher Education Supplement, 16 January 1981

Mitchelmore, M. C. (1981) "Reporting Student Achievement: How Many Grades?" British Journal of Educational Psychology., Vol. 51 June pp. 218-223

Monks, T.G. (ed) (1968) "Comprehensive Schools in Action" (NFER)

Morrison, A. and McIntyre, D. (1971) "Schools and Socialization" (Penguin Education)

"Robbins Report" (1963) Committee on Higher Education Report Cmnd 2154 (HMSO London)

Ross, J. M. and Chanan, C. (1972) "Comprehensive Schools in Focus" (NFER)

Rutter, M., Tizard, J. and Whitmore, K. (eds.) (1970) Education, Health and Behaviour (Longman)

Rutter, M., Cox, A., Tupling, C., Berger, M. and Yule, W., (1975) British Journal of Psychiatry Vol. 126 pp. 493-509.

Rutter, M., Maughan, B., Mortimore, P. and Ouston, J. (1979) "Fifteen Thousand Hours; Secondary Schools and their Effects on Children" (Open Books)

Ryrie, A. (1981) "Routes and Results: a study of the later years of schooling" (Hodder and Stoughton for the SCRE)

Schools Council (1980) "Exams Brief" (Schools Council)

Soares, A. and Soares, L. (1969) "Self perceptions of culturally disadvantaged children"; American Educational Research Journal Vol. 6 pp. 31-45

Steedman, J. (1980) "Progress in Secondary Schools" (NCB)

Steedman, J., Fogelman, K. and Hutchison, D. (1980) "Real Research" (NCB, mimeo)

Szamueli T., (1969) in Cox, C.B. and Dyson, A.E. (eds.) op. cit.

Tenth Report of the Expenditure Committee (1977) Session 1976-77, "The Attainments of the School Leaver" (HMSO)

Wilson, J.M. (1982); "The Accuracy of 'A' level Forecasts" Educational Research Vol. 24 No. 3 June 1982 pp. 216-221

Wright, N. (1977) "Progress in Education" (Croom Helm)

Wrigley, J. (1981) "The NCB Report - Some Problems" in Dancy J. (ed.) Perspectives 6, "Children's Progress in Secondary Schools: The National Children's Bureau Report". (School of Education, University of Exeter).