

4 Method of Research

This study follows the careers of children who transferred to 20 multi-ethnic secondary schools at the age of 11 in the autumn of 1981. The central objective was to measure differences between schools in the outcomes they achieve, in academic and other terms, after taking full account of differences in the attainment and background of children at the point of entry. As a part of this objective, the study aimed to focus on the results the schools achieve with children belonging to racial minority groups. A second objective was to understand the reasons for school differences and if possible to describe processes underlying school success. A third objective was to describe the educational experience of children belonging to racial minority groups.

Outline of the method

The study was carried out in four local education authority areas: two in the South East, one in the Midlands and one in the North of England. Our original intention was to study five schools in each area, or 20 in all. In practice, substantive information is available for 19 schools for the first two years and for 18 schools up to the fifth year.

Over the first two years, the general plan was to collect information, essentially from staff, about the structure and organisation of the schools, the style of management, and the curriculum; and to collect detailed information about the individual children in the reference group, through tests of attainment, pupil questionnaires and a survey of parents.

In the first instance, information about the structure and organisation of the schools, the style of management and the curriculum was obtained by informal interviews with staff and by collecting documentation. Later, in the autumn of 1983, all staff in the study schools were asked to complete a questionnaire, and in addition staff with special responsibilities were asked to complete questionnaires about their area of responsibility. For reasons discussed below, response to these teacher questionnaires was poor and uneven. This meant that the objective of collecting quantitative information about school policies and practices could not be met. Some information about school characteristics can, however, be derived from information collected from the children.

In the third year, a special study was made of 'option choice', the structures and processes leading to a decision about the subjects that each child will study in the fourth and fifth years. In the fifth year, details of examination entries and results were collected.

The first two years

There are broadly six sets of information collected in the first two years from the individual children and their families.

1. The Rutter B2 score, an indicator of behavioural problems shown by the child.¹ The scores are based on questionnaires that were filled in by the primary school class teachers or head teachers some three to five months after the children had left primary school. The response from primary schools was somewhat uneven, and it was not always possible to follow up primary schools well outside the study areas. Nevertheless, the sample sizes are adequate for most of the study schools.
2. Attainment tests completed in the first half of the autumn term in 1981, that is, as soon as possible after the arrival of the children at the study schools. The following five test scores are available: maths, English comprehension, writing, verbal reasoning, numerical reasoning.
3. Attainment tests completed in the summer term of 1983, that is when the children had nearly completed their second year at the study schools. Three tests scores are available: English, maths, verbal reasoning.
4. The number of half days that the child was absent from school in each of the first two years.
5. A pupil questionnaire, completed by the children in the spring term of 1983. This covered a variety of subjects including specific kinds of encouragement and discouragement received from teachers, enthusiasm for school, participation in various school activities, friendship patterns, aspirations, languages spoken at home.
6. A survey of parents, in which personal interviews were attempted with one or other parent of each of the study children. This provides reliable information about the ethnic and socio-economic group of the family, their religion, about whether the parents are working and whether one or two parents are present in the household; it describes the extent and nature of contact between the parents and the schools; and it shows how satisfied parents are with the school, how they think their child is getting on, and what features of the school they criticise and praise.

The third year

The research carried out in the third year focused on the option choice processes. It aimed to establish what subjects it was decided the pupils should study and at what levels, and to describe the processes by which these decisions were made in the context of the relevant school policies, resources and structures. During the school year 1983/84, teachers involved in the option choice processes were interviewed, option choice booklets and other written materials were collected, and 2,273 pupils completed an 'option choice questionnaire' in class time a few days after they had finalised their subject choices for years four and five. In addition, the researchers sat in on 50 interviews between teachers, pupils and parents at two of the schools. Further details about the methods of research on option choices are contained in Chapter 12.

The fifth year

At the end of the fifth year information was collected about the public examinations for which the reference group were entered, and the results. In addition, a record was made, from the school registers, of the number of half days that each child was absent during the third and fourth years, and during the first two terms of the fifth year.

The difficulties encountered

From the beginning we found this an exceptionally difficult project to carry through. It is worth describing and discussing the difficulties we encountered, since this will help the reader to form judgements about the meaning and solidity of particular findings; also, this discussion will be useful for the planning of any future research in comprehensive schools.

Considerable difficulties were encountered in getting agreement from local education authorities to take part in the study. Two of the authorities which were originally approached agreed to take part (one of them after a long delay) but the other two refused and there was a refusal from a third authority before all four slots were filled. This greatly exacerbated what was perhaps our greatest problem, that is lack of time at the early stages of the project. It was not until the late summer of 1981 that we knew what the study areas would be, yet our schedule required that the study children should be tested in the first half of the autumn term. Consequently there was no opportunity to get to know the schools in a gentle and relaxed way; instead, we had to make immediate demands. The organisation of the testing inevitably took up a significant amount of staff time, and the chosen tests were inevitably imposed on the participating schools without any kind of consultation. Another consequence of the initial timing was that arrangements for contacting the feeder primary schools to ask them to complete the Rutter B2 test were not as efficient as they could have been if a system had been set up beforehand. This affected the response rates on this test and lengthened the recall period for the primary school teachers involved.

In any future project, we would plan to obtain the agreement of local education authorities and schools a full school year before any major research initiative (such as testing, questionnaires or structured observation) was scheduled to start. We would spend the preliminary year in getting to know the schools in a relaxed atmosphere, discussing the proposed research with them and securing their agreement, in broad terms, to the main elements of the programme.

A second difficulty was that the amount of researcher time available was too small in relation to the number of schools and research tasks within each school, given that there were substantial difficulties in obtaining cooperation. In some schools, particularly those in one of the four areas, very large amounts of time were taken up by discussions with teachers about the research that we wished to do, and indeed the research team probably spent more time on discussions of this kind than on anything else. Resources that would be adequate given ready cooperation become inadequate when cooperation is reluctant.

A third problem was that our research was intended to be exploratory. It was always our plan to work out the details of the research instruments to be used over the first year of the project. But this meant that the schools (and also the local education authorities) could not be told in advance exactly what we proposed to do. Later, when we were in a position to put forward specific plans (for example, for the parental survey and the pupil questionnaire) some of the schools felt they were being faced with something quite different from what they had expected. In a future project we would, if anything, over-state our demands to the schools at the outset and explicitly list each of the elements of the proposed research, even though none of them could be specified precisely.

A fourth set of difficulties was the specific features of the study and of the surrounding circumstances that made teachers and sometimes whole schools reluctant to cooperate. In some schools, there was considerable opposition from teachers to any testing of the children. Stated objections to tests were that all standard tests had been shown to have serious technical failings, that testing was historically associated with the view that intellectual abilities were genetically determined, that to test the children would have a

deadening and negative effect on them while the educational philosophy of the school was to be stimulating and encouraging, that standard tests could not be appropriate to the ability range of children or take into account the scope and emphasis of teaching in a particular school. A usually unstated objection was that the results might suggest that either the ability range of the intake to the school or the improvement in the children's performance over the first two years was below average. The schools were, of course, assured that they would not be identifiable in any published report, but teachers tended to think that the informed reader would be able to guess; and in any case, regardless of whether anyone would get to know the results, teachers naturally tended not to welcome tests in which they thought their children would do badly. From the fact that the project started with testing it was easy for teachers to conclude that it was to be an assessment of themselves and their children by a hostile outside body on dimensions that would not take account of what the school was trying to achieve. A specific anxiety was that children belonging to ethnic minority groups (especially those of West Indian origin) would achieve low scores, and that these results would be used to promote racist ideologies or policies not by the 'naive' researchers, but by others reading their report. These difficulties were the greater because several of the tests designed by NFER for the first wave (autumn 1981) were too difficult, on average, for the children in the study schools, who were less advanced than the populations on which NFER tests are standardised. Even when the overall level of difficulty of the tests is reduced, it remains hard or impossible to design a single standard test that adequately covers the full ability range. This difficulty is much more evident in the inner city schools covered in this study than in the schools generally used for test standardisation.

At the same time it should be emphasised that the attitude to tests varied very sharply between schools and also between teachers within schools. Some schools and teachers were strongly opposed, and indeed one school withdrew from the project at a very early stage entirely because of teachers' objections to the first wave of tests. On the other hand, some schools were strongly in favour of testing (some of these were themselves in the habit of testing the children at regular intervals). For these schools, the testing element of the project was in its favour, not least because it would provide further data about the children.

Two other factors were important in causing a reluctance to cooperate. The fact that the research had some connection with racial or ethnic minority groups was a further difficulty. Teachers are in any case liable to be defensive, but they are particularly sensitive to accusations of racism. Although they understood the objectives of our research, teachers still feared that the published findings might stimulate unreasonable attacks. Also, many teachers think that ethnic and racial differences are best ignored, that to make distinctions between ethnic groups is to show prejudice, that the enlightened are 'colour-blind'. A research project that is explicitly concerned with distinct ethnic groups is digging up something that they would prefer to remain buried.

The other cause of reluctance was the insecurity felt by teachers at the time of the research because of the contraction and reorganisation of secondary schools in response to falling rolls. In a number of schools, teachers felt they were under threat, and in a number of ways this feeling was justified by the facts. Not many teachers have actually lost their jobs, but two schools in our sample have been affected by amalgamations, and even where schools have remained intact, some teachers have had to move from one school to another. This climate of uncertainty is not conducive to happy cooperation with outside researchers.

Details of the method

Selection of areas and schools

For the reasons discussed in the last chapter, we are inclined to think that future research on school differences should be designed to test specific theories that seek to explain why some schools are more successful than others. However, this study was not designed to test a specific theory. If it had been, then it would have been best to limit the study to a single, homogeneous area and to choose a set of schools that resembled each other as closely as possible in most respects, but varied with respect to factors expected to have an important influence. Since, by contrast, we were planning a more open and exploratory study, it was important that the areas and schools should encompass as much variety as possible. For example, we stated in the proposal ‘We do not plan to hold the area constant, because the factors associated with success may vary between one type of area and another, and this is something which the study should try to take into account’. The proposal went on to argue that there should be among the chosen schools a variety of ethnic mixes, social class profiles and geographical locations. More specifically, ‘We hope that the following would be covered:

areas in London, the Midlands and the North of England;
conurbations and free-standing towns;
inner city and suburban areas;
working-class areas and areas with a strong middle-class element;
schools with high, medium and relatively low proportions of pupils belonging to minority groups in aggregate;
schools containing both Asians and West Indians, and schools containing only Asians and only West Indians;
Moslems, Hindus and Sikhs, and Asians speaking various sub-continental languages.’

In the event, all of these conditions were met by the actual selection. The proposal went on, ‘Of course, all the different combinations of these variables cannot be covered within a sample of twenty schools.’ This does, indeed, turn out to be a serious problem. For example, most of the Bangladeshi pupils in the sample are concentrated in a single girls school. The overall proportion of Afro-Caribbean pupils in the sample is rather low, and there is no school with a very high proportion belonging to this group. Past American research has suggested that school success may be related to the proportion of the school population belonging to racial minority groups. Within a sample of 20 schools, it is not possible to explore the effects (if any) of the more detailed ethnic balance: for example, the effect of having a high proportion of Asians belonging to the same linguistic and religious group, as opposed to a high proportion belonging to different specific groups.

A further objective was to ensure, as far as possible, that some of the selected schools were ones thought to be successful, and others less so. This, together with the other requirements, was explained to the participating authorities, which in each case responded by suggesting a set of five schools. In one area, two changes were made to the original list when the research team pointed out that only schools thought to be fairly successful had been included. In that same area, a replacement school was included at an early stage (and before the first wave of tests) when it looked as though one of the original set might drop out (it did drop out a little later).

In short, the schools were selected purposively, to meet certain criteria, essentially by the participating authorities, though some checks were available to the research team.

Neither the schools nor the children in them are a representative sample of a wider population. However, it may be reasonable to assume, subject to confirmation, that the *relationships* observed hold good in a more general way.

Behaviour scale

One of the strategies of the research is to allow for differences in the conditions with which the schools have to cope when making comparisons between them. This of course leads to the need to measure some of these conditions, and it is important to note that the way in which the conditions vary between one school and another is of great interest in itself, and may have important implications for policy, quite apart from the need to take these variations into account when making other comparisons between schools.

Schools may differ in the proportion of the children entering them who have emotional or behavioural problems. As in *Fifteen Thousand Hours* we tried to take account of this factor by asking primary school teachers to fill in the B2 questionnaire for each of the study children. The 26 items in this questionnaire 'were chosen to cover the main common emotional and behavioural problems of children as they might be seen in a school setting, and the wording was designed as far as possible to provide descriptions of overt behaviour which required the minimum of inference by the teacher'. Various studies by Rutter and his colleagues have shown that the scores from this questionnaire are correlated with independent measures of delinquency and psychiatric disturbance.²

At our early visits to the schools, in the autumn of 1981, we established lists of the new intake of children showing the primary school from which each one had come. We then wrote to the head teachers of these primary schools, enclosing one blank B2 questionnaire for each study child who had been at the primary school, and asking the head to have the questionnaires filled in by the teacher who knew most about the child. The returns were fairly slow to come in: the latest ones came in, after reminders, late in the spring term of 1982. Questionnaires were completed for 1,763 children altogether. There are three schools in which the scores are available for less than 50 children (the actual number of responses in these three schools is 45, 27 and 4). For the remaining schools, the sample sizes are adequate.

Tests of attainment

We made the assumption that secondary schools might have a number of functions, and we therefore hoped to find a number of criteria of success corresponding to these various functions. A good level of scholastic attainment among children belonging to all ethnic groups, both boys and girls, was seen as one important criterion of success among a number of others.

There are two kinds of difficulty in obtaining useful measures of attainment. The first is that secondary schools may differ widely in the attainment of the children entering them at the age of 11, so it is difficult to distinguish the success of the school in improving attainment from the prior characteristics of the children it has to deal with. Secondly, it is very hard to obtain agreement among all of those involved (teachers, schools and parents) as to what kinds of attainment are important and how they may best be measured or assessed. One criterion that certainly has relevance in the world outside of school is success in public examinations, but it only becomes available when the children have completed their fifth year.

Our approach to the problem was to have the children take standard tests at the beginning and end of their first two years at the study schools. These tests concentrated on the basic

skills of reading, writing and maths; this is justified by the assumption that improvement in these skills is associated with progress in many other subjects. Tests of verbal and non-verbal reasoning were also included, but chiefly to help us analyse the differences between intakes to the various schools (and not to serve as indicators of progress). By this approach we are able to overcome the two difficulties we have mentioned to a considerable extent. The first-year test scores provide us with the possibility of 'controlling' for the differences in the attainment of the children at the time of entry to the schools. By concentrating on basic skills we give ourselves the best chance of obtaining agreement that what is being measured is relevant to a wide variety of educational methods and philosophies.

The children were tested in September-November 1981 and in July 1983 on norm-referenced, standardised tests. In the first year, the following tests were used.

1. NFER verbal reasoning test of 25 items.
2. NFER non-verbal reasoning test of 25 items.
(These tests were specially commissioned and drawn from items contained in the NFER 'LEA and schools item bank'.)
3. NFER reading comprehension test, maximum score 15. This test comprised three passages with associated questions testing extraction of information, information brought to bear on the text by the reader and appreciation of the style or purpose of the passage.
4. NFER writing test, marked on a score of 1-7. This was a specially commissioned test of free writing, assessed by 'rapid impression marking'.
5. NFER maths test of 30 items. This was a specially commissioned test with items drawn from the 'LEA and schools item bank', after maths advisers in the four study areas had been consulted as to the balance of items on various subjects.

In the second year, the following tests were used.

1. NFER verbal reasoning test EF 90 items.
2. NFER maths attainment test EF 60 items.
(Among this range of reasoning and maths tests produced by NFER - Nelson, running from an A test for 7-8 year olds to an EF test for 11-13 year olds, are those most commonly used by local authorities for their testing programmes.)
3. Edinburgh reading test (total score 155, with sub-tests of skimming, vocabulary, reading for facts, points of view, comprehension). The Edinburgh tests are the second most commonly used reading tests among local education authorities.

There has, of course, been much criticism of tests and of testing procedures, particularly where ethnic minority pupils are concerned. Much of this discussion is concerned with arguments or ideologies about the origins of individual differences. Where the focus is on effective schooling, many of the common criticisms of testing become much less relevant. Schools and education authorities are the major users of the extensive testing facilities that have grown up to service their needs; in practice, therefore, tests of the kind that we have used are seen by the educational system as useful and relevant measures of performance.

We carried out an item analysis of the first year reading and reasoning tests to check whether any of the items might be 'culturally biased'. The conclusion was that some of the items in the verbal reasoning and reading tests produced significantly poorer responses from ethnic minority pupils than from the rest. This item analysis and the problem of cultural bias is discussed at greater length in Chapter 9, which introduces the results.

Analysis of progress in attainment

Progress in attainment has been analysed in two separate time segments: up to the end of the second year, and from the end of the second year to the fifth year public examinations. In the analysis for the first two years, we consider the scores obtained in maths and reading at the end of the second year after controlling for the scores obtained at the beginning of the first year. In the analysis of the third to fifth year, we consider the results obtained in fifth year examinations after controlling for the maths or reading test score results obtained at the end of the second year. There are two reasons for adopting this approach. First, we find that the ethnic minorities show a different pattern of progress over the two periods, and it seems important to bring this out. Second, the reading test used at the end of the second year was superior to the one used earlier, so there is a strong case for preferring the second-year test scores as a control for the analysis of examination results.

Attendance

In the research proposal we suggested high attendance levels as a criterion of success for two reasons. First, whatever are the benefits to be derived from school, children cannot have access to them unless they attend, so the attendance level is a measure of the 'opportunity to benefit'. Secondly, we wanted to resist the tendency to see school entirely as a means to an end – a way of getting credentials or skills – rather than an enjoyable and fulfilling experience in itself. The attendance level is a behavioural indicator of enthusiasm for the school among children and parents; this behavioural measure is supplemented by attitudinal measures in the pupil and parental surveys.

The number of half days on which the child had been absent in each of the two school years was counted up from the school registers. The information was recorded for each individual child in the study group, so that the attendance record of a particular child can be related to that child's test scores, family background or other characteristics.

Pupil questionnaire

The children completed the pupil questionnaires in class during one period in the spring term of 1983. An earlier draft of the questionnaire was piloted during the previous term in two northern schools unconnected with the study. About 120 questionnaires were completed in the pilot. In general, the pilot showed that the questionnaire was well received by children and teachers, but the results caused us to make some detailed changes to wording and layout.

Classes were supervised by teachers while the questionnaires were completed. Teachers were encouraged to help any children who had difficulty in reading or understanding questions; in practice, the questionnaires were completed in a more sociable atmosphere in one of the areas than in the three others. The questionnaire was provided in only one, English-only version. Children whose first language was not English and who had difficulty in understanding the questions were given special help by teachers, and with this help most of them were able to cope. There was, however, a significant number of children who could not complete the questionnaire at one school having a high proportion of Bangladeshi children who had arrived in Britain fairly recently.

The information from the pupil questionnaire can be related to the other information about the particular child and the family, but much of it can also be used to derive variables about the school as a whole.

Survey of parents

Our first intention was to carry out a postal survey of parents known to speak English, followed by personal interviews (using the same questionnaire) with parents thought not to speak English or who had not responded to the postal questionnaire. Two postal pilots were carried out in two northern towns unconnected with the study, the first with a longer and the second with a much shorter and simpler version of the questionnaire. Both of these pilots produced very low response rates, and the postal method therefore had to be abandoned. Although postal surveys of parents have been successful in some areas, it seems that they will not work for the parents of children in inner city schools. By contrast, pilot personal interviews showed that most parents were willing to be interviewed and found the questions acceptable, even though teachers often believed that they would object to them.

We were able to obtain agreement to carry out the survey from the schools and education authorities only after long and in some cases difficult negotiations. The survey raised the difficult ethical issue of whether and in what circumstances the schools should be prepared to let outside researchers have the home addresses of the children (without a list of home addresses, the survey could not be undertaken). The schools in the Midlands and the North eventually agreed to release the addresses when they were satisfied with the form of the questionnaire. This was naturally subject to stringent guarantees that parents could refuse to be interviewed if they wished and that the information they provided would not be disclosed to the schools, the education authorities or any third party. The schools and authorities in the South East were more cautious in their approach and in a number of cases they required that parents should first give their permission for their addresses to be released, or be given an opportunity to decline a request for them to be released. This meant that the schools had to organise a mail-out to parents prior to the compilation of the address lists for the survey. Of course, this considerably reduced the number of interviews that it was possible to achieve in the schools concerned.

Senior staff in a number of the schools thought that parents would regard a number of the proposed questions as offensive or intrusive and that the survey would lead to complaints from parents and would cause friction between the families and the school. A number of questions which would have been useful, and which would not have caused offence in the vast majority of cases, as experience of many other surveys has shown, had to be deleted at the request of the schools. In the event, in the course of carrying out the survey we encountered only a very few informants who were abusive or plainly upset by the questioning; there was an appreciable but small number of refusals to be interviewed; complaints from the parents to the schools about the interviews were not a significant problem, and the problems were no greater in the areas where the addresses had been released without prior permission than elsewhere.

Bilingual versions of the questionnaire were produced in English with Urdu, Punjabi, Gujarati and Bengali. We tried to establish beforehand from the school the linguistic group to which each Asian family belonged and then sent an interviewer who could speak the appropriate language. Interviewers were specially recruited for this survey and worked under the supervision of a specially employed fieldwork supervisor in the Midlands and Northern areas and under the supervision of a member of the research team in the two southern areas. The bulk of the interviews were carried out between June and October 1983, though a few were carried out in November and December.

One interview was carried out in each family. A majority of the interviews (68 per cent) were with the mother, while 29 per cent were with the father and the balance (3 per cent) with some other relative.

Numbers

There were about 3,100 children who attended the 20 study schools for all or part of the first two years of the study. By the fifth year, there were 18 study schools left, and information was collected about the 2,426 pupils then on the rolls of these 18 schools. Table 4.1 shows the number of pupils for which various items of information are available. It should be remembered, however, that when analyses make use of information from various sources, the sample sizes are considerably reduced, as pupils can only be included if the information from all relevant sources is available.

Presentation of findings

Part II (*Children, parents and schools*) introduces the results of the research carried out in the first two years. It shows the results on each topic individually and explores some of the inter-relationships. Part III considers differences between schools in terms of progress in attainment up to the end of the second year, and in terms of outcomes other than attainment. Part IV presents the results of the study of subject option choice carried out in the third year. Part V presents the results of the fifth year examinations, then analyses differences between schools in terms of examination results, after allowing for attainment at the end of the second year. Part VI sets out the main conclusions.

Presentation of tables

Percentages are rounded to the nearest whole number. The asterisk (*) denotes a value less than 0.5 per cent and which is therefore too small to be rounded up to 1. The sign for nil is the dash (-).

Notes

1. The Rutter B2 behaviour scale was designed 'as a screening instrument to pick out children with possible emotional or behavioural difficulties and there is good evidence from a variety of studies that it does this well. Different teachers tend to rate the same children in a fairly comparable fashion and the questionnaire scores generally agree with more detailed individual diagnostic psychiatric assessments.' This quotation is from Rutter et al. (1979). Further details about the development and testing of the B2 scale are contained in Rutter (1967), Rutter et al. (1970) and Rutter et al. (1975).
2. See note 1 above.