Month of birth and its relationship to streaming in the primary school

Mckenzie, Ian. R.

UNIVERSITY OF DURHAM

M.ED. THESIS 1968.

IAN. R. McKENZIE.

MONTH OF BIRTH AND ITS RELATIONSHIP TO STREAMING IN THE PRIMARY SCHOOL

Abstract

Month of Birth and its Relationship to Streaming in the Primary School.

This study is concerned with an investigation into the relationship between month of birth and stream placement in the primary school. It is particularly concerned with the possibility that, where traditional streaming is implemented, there may be an under-estimation of the younger children in a school year age group.

Streaming is usually defined as "grouping according to ability with considerations of attainment", but, in practice, only attainment seems to be assessed adequately, and ability tends to be given less attention. In the traditionally streamed primary school, allocation is usually based on attainment level at the time of leaving the infant department. It is possible that some of the younger children in the year group, who have matured less intellectually, and who have had less time in the infant department to benefit from early formal tuition, may be under-estimated and placed in lower streams than their potential would warrant.

In the study 1000 children from 5 schools, 500 in the first year of the junior department and 500 in the fourth year, were investigated with respect to Month of Birth, I.Q., and Stream Placement.

Results showed that, although, in general, the children were successfully streamed, and although no birth months were superior with respect to intelligence, the younger children tended to be placed more readily in the lower streams. This was the case at first year level but not at fourth year level. Thus, although there was a tendency for early underestimation of the younger children of the school year group, this seemed to be rectified later to a great extent.

ACKNOWLEDGEMENTS

I should like to thank J. J. C. McCabe, B.Sc., M.Ed., A.B.Ps.S., for his helpful suggestions with respect to this study.

I should also like to thank education department officials, and members of the teaching profession, who have co-operated to make this study possible. In particular, I refer to the Director of Education of the local education authority selected for the study, and to the head-teachers and class teachers of the schools and classes directly involved.

I. R. McKENZIE

October, 1968

CONTENTS

PREFACE

The concern of this study is to investigate the possibility that some of the younger children of a school year age group may be undervalued at the primary school level when the traditional procedure of streaming, or grouping according to ability, is adopted.

It is distinctly possible that some of the younger children, who have matured less intellectually and who have had less time in the infant department to benefit from early formal tuition, may be underestimated on arrival at the junior department at the age of seven years. If junior departments adopt the procedure of streaming at this time, some of the younger children may be placed in lower streams than their true ability and potential would warrant, and a preponderance of younger children may be found in the lower streams.

Also, unless there is a great degree of flexibility and opportunity for movement between streams in the traditionally streamed junior department, the effect could be lasting, with the younger children still being undervalued at the end of their junior school career.

This study aims to investigate a sample of approximately 1000 junior school children, 500 seven year olds in their first year in the junior department, and 500 eleven year olds in their fourth year in the junior department.

Any evidence for the underestimation of younger children in terms of stream placement will be assessed and this will be done for both sub-samples.

Comparisons between the two sub-samples will also be made so that any lasting effects can be noted.

It is particularly stressed that the comparisons will be between two different groups of children. This is not a follow-up study.

General Introduction and Discussion

In this section the author summarizes the general position of streaming as he sees it, and in the next section surveys the relevant literature.

In discussions on Streaming in the Primary School the object has usually been to evaluate the desirability of the system or the results of the system in some way. Suggestions have been made as to the beneficial effect on both the brighter and duller children of being taught in homogeneous groups.

Alternatively claims have been made that brighter children benefit from the stimulation of their brighter colleagues when placed in such groups.

Discussion has centred around the relationship of streaming to the morale and the motivation of pupils, and it has also covered the social aspects.

Theoretically the position may be summarised as follows.

1. Concerning Social and Emotional Factors

Heterogeneous grouping is more natural in a democratic society and it is better for the social development of both the individual and the group at large. It is also less likely to be emotionally damaging to some individuals in the group and less likely to affect adversely the personality development of such individuals. The point here is that ego development may be affected, some persons perhaps over-developing their ego and over-evaluating themselves and others doing the opposite, with the added danger of the traumatic effect of classification changes.

As against this homogeneous grouping is said to be realistic, and it aids an individual's recognition and acceptance of his abilities, his limitations and his role in society. It may also be argued that, up to a point, it protects the individual psychologically as it prevents competition at an impossible level, the individual being grouped with others of the same level of ability and attainment. Also, as with the principle of proximity in perception, wide individual differences will be more easily noticed and emphasised when they are brought together in heterogeneous groups. Thus the psychological point concerning emotional damage and personality development could also operate against the heterogeneous group.

2. Concerning Motivational Factors

Those in favour of streaming stress the need for competition with peers of similar interest and ability and point out that this may be lacking in the heterogeneous group. They seem to be more concerned about positive motivational effects on the brighter pupil and perhaps less concerned about negative motivational effects on the duller pupil. They seem to work from an a priori assumption that competition itself is essential and beneficial, but perhaps ignore the fact that it can have differential effects on different members of the group. Individual members of a group can be affected by competition in different ways according to their own personality characteristics and also with respect to their standing and performance in the group.

Those opposed to streaming are probably not less interested or concerned about the progress of the brighter child, but they seem to be probably less anxious or worried about the possibility of inadequate performance by the brighter children when competition is at a minimum. They tend to take the view that the brighter pupil will have much the same success irrespective of the system or organisation. They seem to be more concerned about the duller children, particularly with respect to the possible negative effects of streaming on the duller child's level of motivation. Of course they do not regard competition itself to be so essential or beneficial, and often consider that cooperation within a group can motivate the individuals in that group more than direct competition between those individual members of the group. Thus there seems to be an underlying fundamental difference in attitude towards competition v. cooperation.

There has been very little done to test the relative effects of cooperation and competition in education but what has been done suggests that both have a part to play but that cooperation can have beneficial effects greater than had hitherto been thought. Maller (1929) found that competition generally tended to be a stronger motivating force than cooperation, but that many children in some circumstances are more strongly motivated by cooperation. As Craft's, Schneirla et al (1950) point out, Maller's cooperative situations in his study involved group competition, individuals in groups cooperating to set their performance against other groups. It is possible that cooperation may be more telling and effective when the group is working on a project for its own sake rather than to promote group prestige.

3. Concerning Practical and Technical Aspects

Those in favour of streaming claim that the teaching is more efficient when groups are homogeneous and that the resulting learning will be better. The range of ability is less extensive and lessons can be more easily presented to the group as a whole. Up to a point the same is said for the range of interests of the pupils. Certainly in terms of teaching organisation it is easier to operate with homogeneous groups as class teaching can be applied throughout, and less teacher contact time is wasted as the teacher has not got to divide his or her time between three or four sub groups. It is argued that teacher contact or direct teacher stimulation is reduced in heterogeneous groups as the teacher has to attend to three or four groups, and also lesson preparation becomes a much more complex and difficult business.

It is also claimed that discipline tends to suffer with heterogeneous grouping as, when class teaching is given, the extremes in ability become bored, and, when group work is invoked, some children take advantage of the lack of direct supervision.

Those opposed to streaming reject the claim that teaching is more efficient when homogeneous groups are arranged. It is claimed that some things, such as art and music, can be taught to a heterogeneous group as a whole, and that grouping and "setting" with other work can ensure that all pupils in a heterogeneous group make progress at their own level. Skilled preparation and organisation by the teacher can ensure that all work to their full potential. Those opposed to streaming deny that the interests of the bright and dull are so very different, particularly in the early school years, and they stress that the interests of the bright can often stimulate the dull. They thus claim that heterogeneous grouping does not retard the bright and that it may well assist the development of the dull.

Those opposing streaming would agree that, with a heterogeneous group, the teacher's task is a more skilled and exacting one, but they believe the teachers should accept this challenge. It is conceded that teacher contact and teacher stimulation at the class teaching level is not so operative with the heterogeneous group as with the homogeneous group, but it is claimed that this is a good thing. In the heterogeneous group children tend to be given more opportunity to work in groups and with some degree of initiative, and it is claimed that they actually benefit in terms of teacher contact and teacher stimulation, as when this does come round it is at a much closer and more personal level.

It is also conceded that discipline problems are inherent in the heterogeneous grouping where the teacher's attention may be concentrated in one direction, but it is claimed that this can be minimised by good preparation ensuring that every child is occupied in some way all the time. It is also claimed that working in groups in a social, cooperative manner will bring about a realisation of the importance of reciprocity in social relationships. This in turn will bring about socially responsible behaviour and self discipline at an earlier age, and it is considered that this is more important than traditional order under supervision.

One big difference between the two positions is that of underlying assumptions concerning accurate selection or allocation and the problem of ability constancy. One position accepts that the allocation is reasonably accurate and that there is a high degree of constancy in terms of ability and so subsequent performance. Streaming thus appears just and necessary.

The other position rather questions the accuracy of the allocation and doubts a high degree of constancy in ability. It suggests that constancy in performance may be partly due to a form of conditioning, andacceptance of a role given and an acceptance of standards set. To the supporters of this position streaming appears inaccurate, unnecessary, and unjust. The two positions are, of course, closely related to the opposite sides in the heredity v. environment debate. The positions described here tend to be extreme.

4. Summary and Conclusions

Thus in brief one side believes that heterogeneous grouping and the abolition of streaming will bring about greater social integration and less emotional damage to personality development, and at the same time maintain the educational standards of the bright and increase the educational standards of the dull. It also doubts the validity of the allocation in the present streaming system and resents the relative finality of this. It also places more accent on cooperation than upon competition.

The other side believes that homogeneous grouping or streaming should be retained. It fears that heterogeneous grouping would result in a lowering of educational standards, particularly for the brighter children, and it doubts that social integration would necessarily be improved by this system. It believes in a more tough-minded approach in which acceptance of abilities and limitations etc. is considered good training for life. There is satisfaction that the actual allocation is accurate and that over the years only a few corrections will be necessary. Great faith is placed in the value of competition as a motivating force.

It is worth noting that the two views are very similar to those taken by the two sides in the debate on comprehensive education, and, of course, this is not surprising as both debates deal with the issue of allocating and separating children. It would probably be fair to say that most educationalists supporting the comprehensive position have gone some way towards ascribing to the non-streaming philosophy. They have at least identified with the view that segregating and streaming children into different schools and buildings is socially undesirable, although many draw the line at this point and consider that streaming within comprehensive schools is necessary from the strictly educational point of view. Others take the position further, rejecting streaming of class groups as such, but accepting "setting" for individual school subjects.

In conclusion it can be noted that the theoretical positions on homogeneous v. heterogeneous grouping emanate from fundamental underlying attitudes, but that practical considerations and compromise are producing a middle of the road approach. It is not absolutely necessary to accept one position or the other, although this seems the case on first examination. It is quite possible to accept some of the values of one viewpoint and at the same time appreciate the weaknesses. It is possible to accept both positions, or at least not reject both outright, and to pose the question "How homogeneous?" or "To what extent heterogeneous?" To illustrate with extremes. Who would grade a normal sample of 240 three year old nursery children into twelve rigid ability streams? Or would place a fifteen year old low grade mongol of I.Q. 27 in the same teaching group as a fifteen year old boy whose I.Q. was 140+? Here is a question of deciding how much divergence of type or variance of ability is possible in a working group.

As the brightest and dullest become further apart as they grow older perhaps more divisions are necessary in the later school years.

Finally it may seem feasible to some to operate homogeneous grouping to some degree without making this rigid. Flexibility of approach, with constant reappraisal of the groups, could perhaps offset some of the more serious objections to streaming, with perhaps most important of all a reduction in the matter of obvious valuations made upon the various arbitrary groups.

It is worth noting that most educationalists accept the value of individual attention for a pupil, one teacher to one pupil, and, on first inspection, this may seem to support the streaming viewpoint, it being the ultimate in streaming. Yet few would deny that a child can only be truly educated if he also has interplay with others in a group. Thus perhaps an ideal situation could be postulated as follows. A heterogeneous classgroup, to promote social education and social cohesion, which includes homogeneous sub-groups for some instruction and which also includes individual attention for other instruction. As mentioned before when discussing the heterogeneous position this would necessitate teachers of very high calibre, and it is doubtful at present if there are sufficient such persons to go round, but the advent of programmed learning in the classroom could well make such a proposition feasible.

Thus in terms of homogeneous v. heterogeneous grouping, or streaming v. non streaming, the theoretical positions are fairly clear cut, being based on different philosophies, but in terms of practical commonsense approaches intermediate positions can be adopted. In assisting one to arrive at such a position a consideration of the empirical side may be useful. Experimental evidence testing any of the views of either side should be carefully weighed. Such evidence is dealt with in the next section.

A REVIEW OF THE LITERATURE ON STREAMING

A. Concerning the Effect on Educational and Intellectual Progress

Edmiston and Benefer (1949), Moyer (1924), Cook (1924), Purdom (1924), Gray and Hollingworth (1931) have all reported studies which indicate that streaming or homogeneous grouping, brings no statistically significant improvements in educational attainments. In fact in some instances the opposite is reported. For example Edmiston and Benefer found that the average gain in reading achievement for their eleven and twelve year old subjects was slightly greater for a wider I.Q. range grouping (41 points) than for a more narrow range grouping (29 points).

On the other hand Billett (1929), Sorenson (1948) and Barthelmess and Boyer (1932) have all reported experiments which show that streaming helps the educational attainment of both the dull and the bright child. Barthelmess and Boyer found that among equated groups of ten year old children attending either homogeneous or heterogeneous classes, those from the homogeneous classes achieved an average attainment gain of 12.8 months during a school year as against 10.4 months by those from heterogeneous classes.

Apart from that of Edmiston, Benefer and Sorenson, the work mentioned above was all done before the Second World War and it caused Raup to remark (1936)

"For every scientific claim made in support of homogeneous grouping there is an opposed claim made on grounds of research equally painstaking."

More recently, in the last decade, research into the problem has again been taken up, but, as yet, it has been insufficient, and certainly not conclusive enough, to advocate the complete adoption of one system or the other.

In 1961 Daniels published the results of an experiment in which he contrasted the development of pupils in two streamed schools, English primary schools, with that of two matched schools which were unstreamed. The investigation indicated that non-streaming may produce an improvement in intelligence and scholastic progress.

In the unstreamed school the average I.Q. had improved by four points at the end of the primary school period. In the streamed school the average I.Q. had improved by only one and a half points in the same time. Thus the unstreamed school had an increase of two and a half points over the streamed school. Similar results were obtained with attainment scores in Reading, Arithmetic and English, the effect on Reading and English being most operative in the early primary school years.

Daniels points out that the average increase in ability and attainment is achieved without noticeable "holding back" of brighter pupils, but rather a "pulling up" of the more backward.

No detailed account is given as to what relative extra help is given at these schools, help such as internal or external remedial teaching from class teachers, remedial teachers or psychologists etc., and so it is not possible to know if different amounts of such help are partly responsible for the "pulling up" of the more backward in certain schools.

Another criticism of this study is that, at the time it was carried out, 1957-61, the unstreamed schools would be under the influence of persons highly charged with enthusiasm in their roles as pioneers, and it is possible that the enthusiasm itself was largely responsible for the success. After initial enthusiasm had dulled somewhat, at a time when unstreamed schools were not so very new to the modern primary school system, it is possible that the same results would not have been forthcoming. The noted success could be the result of what is known in psychology as the "Hawthorne Effect". Referring to his unstreamed schools Daniels says he was "fortunate in being able to collect full and accurate test data from two three class entry junior schools whose heads did not stream the children because they felt it was educationally wrong to do so". The above statement would appear to leave him open to the criticism of the "Hawthorne Effect". However, Daniels study should certainly not be dismissed as invalid and it has set the pattern for controlled experimentation into the problem. More studies of its type are needed.

A more recently published work, in 1965, was that of Kellmer Pringle. Kellmer Pringle conducted a longitudinal study in which she compared the progress of children in two contrasting junior schools, and then followed up part of the sample in the secondary range. The two junior schools were different in orientation, method and organisation, but were both very good schools of their type. One was a traditional school with emphasis on class teaching, the other was more modern and progressive and it combined the project method and group work etc., with a more limited measure of class teaching. One big difference was that the traditional school was streamed throughout but the progressive school was unstreamed for the first three years. Children in both schools were given standardised tests throughout the four years. There were hardly any significant differences between the two schools with regard to academic progress, the only clear difference being that the traditional school children performed better on spelling tests.

After a more complex analysis of the follow up study some cautious tentative conclusions and generalisations were drawn.

1. A traditional approach (with streaming) to education may favour the development of mathematical ability in boys, while a progressive approach (without streaming) may favour its development in girls.

2. The traditional approach favours the brighter child while the progressive regime benefits the duller child.

3. Boys in general benefit more from a traditional framework, while girls in general benefit from the progressive environment.

Thus the position is not so simple as might be thought. Different methods and systems of organisation may be better or worse for different groups of children or different individual children. Kellmer Pringle's conclusion was that the progress of a child is the result of a complex interaction of potential, the particular school subject, teaching methods and school organisation, sex, and socioeconomic background.

Kellmer Pringle was measuring the effects of method as much as organisation but the method is in some sense determined by the organisation. Certainly the methods adopted by the two differently organised schools would be deemed appropriate by the respective advocates of the two main systems of organisation, streaming and non-streaming.

No attempt was made to control method and contrast organisation alone, but it is doubtful if this could ever be done, and even if it could, it is even more doubtful whether it would be valid, as by the nature of things the organisation and the teaching method tend to go hand in hand.

Even more recently, in fact at the very time of writing (1967), further evidence concerning the effects of streaming on educational progress has been published in the Plowden Report. It refers to its Manchester Survey. In the

1964 Manchester Survey it was shown that attainment in objective tests tended to be better in streamed schools. It also gave no support to the view that streaming has an adverse effect on children of low ability.

The Plowden Report also refers to the N.F.E.R.* cross sectional study of attainment in matched streamed and unstreamed schools. The results of this tended to show that the streamed schools did somewhat better than the unstreamed schools, although the differences were not great. The N.F.E.R. enquiry also showed that 'the system of streaming favoured girls, who are, age for age more mature than boys and more disposed to play "the good pupil role" and therefore to gain the approval of their teachers'.

It can be seen that the picture is no less confusing with the more recent work studies.

Daniels found the unstreamed schools to be generally better with regard to attainments whereas the N.F.E.R found the opposite. Kellmer Pringle found them generally much the same with the outcome very slightly in favour of the schools that were streamed. Daniels found that the dull benefited from being unstreamed and there was support from Kellmer Pringle on this point, but the N.F.E.R. study did not find that this was the case.

Kellmer Pringle found that girls benefited from the progressive type school with an unstreamed system, whereas the N.F.E.R. study claimed that it was the streaming system that favoured the girls.

An interesting criticism is made by Daniels (1955) in answering those who believed educational attainment was benefited by streaming and who directed experiments to prove the point. It is that "streaming, even for experimental purposes, introduces changes in teaching practices and syllabuses between the various groups which, inevitably, produce those very group differences which

*National Foundation of Educational Research (1962 Survey) are then used to justify streaming and so make cross comparisons impossible."

Wyndham (1934), and Cornell (1936) make the same point, and it is valid, but, of course, if differences are able to be introduced in experimental streamed groups because of the very nature and organisation of the group, and they are not able to be introduced in the controlled unstreamed groups because of the nature of those groups, and, if these differences are considered to be desirable, then they may legitimately be used to justify streaming.

Another interesting criticism of streaming in the primary school made by Daniels in 1955 is the suggestion that streaming has a differential effect on educational and intellectual progress. It is claimed that children in the higher streams progress as expected or above expectation but those in the lower streams progress at a rate below expectation.

Of course the bright and the dull do become further apart as they grow older, and two children with I.Q.'s of 80 and 120 have respective mental ages of four and six when their chronological ages are five years, these mental ages being eight and twelve when they are ten years old. However, this is not what is really meant by Daniels as such development would be within the realm of expectation. Rather is it implied that the child of I.Q. 120 may even raise his or her I.Q. to say 125 because of stimulation in a lively A stream, whereas the child whose I.Q. was 80 may even drop to 75 or 76 because of apathy and lack of stimulation.

This is more than likely when one considers the current position on the nature of intelligence. Taking the development of Hebb's (1948) original classification of Intelligence A and B we now have.

1. Intelligence A, the inherited genetical component, the potential at birth which is not measurable.

2. Intelligence B, the resultant of Intelligence A and environmental influence; that is Intelligence A after the environment has played its part in influencing its development. This is dependent on the degree of stimulation in the cultural setting. It is thought that the result is an actual change in the potential, and possibly even a change in the quality of the cortical apparatus.

3. Intelligence C, that which tests measure, the actual test result. We attempt to measure Intelligence B and finish up with Intelligence C, although this may be a fairly close approximation. The measure is never perfect because of the difficulty of sampling all forms of Intelligence B and because most widely based tests are contaminated with attainment and cannot be completely culture free.

It is certainly likely that we will be able to note differential effects on Intelligence C, the measured scores, which correlate with streaming arrangements. The obtained scores probably also reflect changes in Intelligence B. Actual potential may be lost forever with the dull because of poor stimulation in the early pre school years and during the important primary school developmental phase.

Following on from Daniels' 1955 suggestion concerning the differential effect streaming has on intellectual and academic progress, Douglas made the same point in 1964 and backed it up with evidence from his study.

After a three year follow up, children in the upper streams improved their scores of measured ability by an average of 0.17 points and those in the lower streams deteriorated by 0.49 points.

Douglas further noted that the less bright children in the upper streams improved relatively more than their brighter colleagues. This conclusion could, of course, be due to an artefact of the testing. Comparisons were made from one test at eight years to another at eleven years, the composition of the two tests not being the same. Also the distribution of scores in the second test may not have been the same as in the first test.

Douglas also noted that in the lower streams the brighter children show a greater average deterioration in test score than the dull children. He remarks that in the lower streams the relatively bright children are handicapped either by unsuitable teaching or lack of competition.

The above conclusions and remarks can be taken to support the cause of non streaming, as the implication is that the relatively brighter children of the lower streams would benefit from the competition and stimulation of the brighter children of the higher streams if heterogeneous groups were introduced. Also one might presume that the less bright of the higher streams would still receive, and benefit from, competition and stimulation from the very brightest if heterogeneous grouping was introduced.

However, it is also possible to argue that some of the conclusions and remarks actually support streaming. For example if in the top streams the brightest improve least as dompared with their less bright colleagues, and if in the lower streams the relatively bright deteriorate more that their dullest colleagues, is this not clear evidence that the dull and mediocre of any group set the standard and pattern for that group and act as a brake upon the advancement of the bright? This, of course, is one of the major traditional arguments of the supporters of streaming. Also if it is remarked that in the lower streams the relatively bright children are possibly handicapped by unsuitable teaching, would the teaching not be even more unsuitable if the standard deviation of ability in the group was made even greater? If it was difficult to organise suitable teaching for a relatively homogeneous group would it not be even more difficult to organise and arrange it for a more heterogeneous group?

Most of the studies mentioned have dealt with the contrasting of general intellectual and academic development in the two types of system, but it is possible that one system or the other may favour progress in a particular subject. This is a possibility which must always be kept in mind and there is some measure of backing for the viewpoint from the work of Morris (1959). He found that, with regard to reading, less able children in particular benefit from being taught in classes made up of children of similar ability.

A most interesting study, published in 1966 by Thompson, points to the relative inefficiency of streaming and the need to restrict the practice of it. This study concerned the secondary school where differences in attainment within the same age group are obvious, and where the extensive range of ability is more easily noticed. If Thompson's assertions and claims are true for the secondary school they must be even more applicable to the primary school where attainment differences are not yet so pronounced and where the ability range, although more or less constant, has not yet developed to the stage where the brightest and dullest are so far apart, e.g. at 5 years the I.Q. 120 and I.Q. 80 are two years apart in terms of mental age; at 15 years they are 6 years apart in terms of mental age.

However, this study is not a controlled experiment with control and experimental groups and with null hypotheses agreed at the outset. Rather is it an investigation into what happened to individuals in a part of a year group in a comprehensive school, when this part of the year group was unstreamed. What happened was contrasted with what might have happened if Verbal Reasoning Quotients at 11+ had been used to stream the group. Periodical attainment ratings were noted after pupils had worked under non-streaming conditions and it was found that the predictive value of the Verbal Reasoning Quotient at 11+ to the later attainment ratings was very poor. Note was made of those figuring in the top thirty places of a merit list after one term, one year and three years. The results of the investigation certainly give support to the non-streaming thesis and suggest that an original streaming at 11+ based on the Verbal Reasoning Quotients would have been most inefficient.

The study is not completely invalid and the findings and assertions are probably largely true, but, apart from the criticism of lack of experimental design, the following criticisms must be levelled.

1. Some of the results may be due to inadequate, incompetent, subjective assessments by teachers in the school. The objective test is being validated against highly subjective internal assessments. It is even possible that bright children from poor backgrounds are being discriminated against in that teachers are "marking them down" without fully realising that they are doing so. Although not culture free the objective test is less prone to do this. The "marking down" process, if operating, could substantially reduce the correlation between the original valuation at 11+ and the subsequent valuation by the school.

2. Few streaming advocates would agree to streaming by Verbal Reasoning Quotient alone as is the comparative situation in the study. Also selection for groups by intelligence test alone is being validated against attainment later.

3. It can be argued from the results that the school has failed its brightest pupils and it is possible to argue that the system of organisation is responsible for this. Perhaps some of the brightest have been allowed to just "get by". Perhaps in a streamed group they would have been "pushed".

4. The investigation is restricted to just over one third of the total distribution (I.Q. 108-135) so generalisations on streaming and setting are hardly justified. In fact the group or block used for the unstreaming exercise is actually a streamed group. Indeed the main assessments are made on the constancy of appearance of about one quarter of this group in appearing in the top thirty places.

Nevertheless, despite the criticisms, it may be that Thompson's claims are very near the truth, and if so they would support Gatfield's (1958) assertion that there is in any case often a low degree of homogeneity in streamed groups.

Of course it is obvious that the unstreamed group will be even less homogeneous.

B. Concerning the Social Effects of Streaming

The literature dealt with so far has been concerned with the effects of streaming or non-streaming on education in its narrowest sense, that is in the intellectual and academic sphere. There has also been work dealing with the effects of streaming or non-streaming on education in a wider sense, and some has been particularly concerned with the social implications.

In 1951 Davis suggested that when selection is by ability there is a tendency to place middle class children in the top streams because they have learnt in their homes to use words with precision. Once there they receive continuing verbal training which maintains their initial superiority. On the other hand working class children may be first placed in lower streams because they lacked the initial stimulation at home, and then they may be further deprived by being given a relatively unacademic type of education at school.

This statement seemed extreme at the time but today most psychologists would agree that social bias is implicit in early selection by ability. Many advocate a stimulating nursery school system to offset early cultural deprivation.

The work of Douglas, published 1964, confirms the point made by Davis a decade earlier. With respect to this point Douglas states that the evidence points to the fact that streaming by ability reinforces the process of social selection. He goes on to say:

"Children who come from well kept homes and who are themselves clean, well clothed and shod, stand a greater chance of being put in the upper streams than their measured ability would seem to justify. Once there they are likely to stay and to improve in performance in succeeding years. This is in striking contrast to the deterioration noticed in those children of similar initial measured ability who were placed in the lower streams. In this way the validity of the initial selection appears to be confirmed by the subsequent performance of the children, and an element of rigidity is introduced early in the primary school system."

Jackson in 1964 and the Plowden Report in 1967 both concur with the above point of view.

Plowden notes the point that more middle class children are to be found in upper streams and fewer in lower streams than would be expected from their results in objective tests, and that a higher proportion of poor children are to be found in lower streams than their test scores warranted.

Plowden remarks: "How much of this placing was due to characteristics in the children which might have made them unsuccessful in an upper stream, how much to teachers' assumptions that clean and well kept children are abler, it is impossible to say."

Thus there does seem to be some case against streaming in so far as it has some undesirable social effects. In the first place there appears to be social bias in allocation to the streams and this in itself is unjust. Secondly, with children accepting a role or being conditioned to a level of response, the groupings are consolidated with the ultimate result of entrenched social division.

C. Concerning the Psychological Effects of Streaming

Following on from possible social effects of streaming is the question of possible psychological effects. In the review of the theoretical position at the beginning of this study the question of possible effects on emotional and personality development were discussed. There has been little work done on this but one study which attempted in some way to measure the psychological effects of streaming was that conducted by Rudd in 1958. Rudd applied attitude tests to the children in the various groups and also made use of sociometric techniques. Individual child studies were also completed.

Some interesting conclusions to the work were as follows:

1. Transfers of pupils between streams had traumatic effects both upon the pupils transferred and upon the streams to which they were transferred, but these effects were temporary.

2. The more lasting effects of transfer upon pupils were a highly individual matter, depending for their direction and strength upon the organisation of the psychological field of each pupil at the time of transfer.

3. The traumatic effects upon pupils tended to pass unnoticed by the teachers involved.

Level of morale is another psychological variable that could be affected by streaming. In 1961 Chetcuti published a paper relating to work on this aspect. Secondary school boys in streamed schools were studied in two respects; first, morale of the pupils as individuals; second, morale of the pupils as a group.

It was assumed that where there is high morale the individual feels self confident, accepts authority, feels accepted and appreciated, feels that he is receiving a fair amount of success, and participates freely in the activities of the group and feels that he belongs to it.

Attempts were made to measure this and the results between streams were compared. The test measures were in the form of group tests and questionnaires. These included a test to measure self confidence, a sociometric test to find the choices of children in six situations, a sentence completion test to measure acceptance of authority, a test of attitude towards school, and a test to measure feelings of being accepted by the teachers.

The main conclusion was that streaming tends to lower morale in the lower streams. It was also noted that the differences between high and low streams were most marked in the case of individual morale and not so clearly marked in group morale, although in every case lower stream boys rated their form lower than did higher stream boys.

Criticisms that can be levelled at this work are as follows:

1. In terms of design. There were no controlled comparisons between groups in streamed schools and unstreamed schools. It is possible that class differences in morale occur in non streamed schools perhaps because of such factors as form master influence, group sporting successes etc.

2. With respect to the definition of morale. It is difficult to reach general agreement as to what this is, and in particular the concept of acceptance of

authority as being an important aspect is something with which many will disagree. Certainly it very much depends upon what kind of authority is envisaged.

3. In respect to the attempts to measure morale. One seriously doubts whether a pencil and paper group test can adequately measure self confidence. It may be possible for a trained psychologist giving an individual personality test in a face to face situation to make some assessment of the level of self confidence, but it is doubtful if any group test can do this. Even more doubtful is the use of a sentence completion test to measure acceptance of authority. Secondary school children are sophisticated enough to "beat the test" in this situation. Also there is a query as to what is meant by acceptance of authority. Acceptance of authority on paper is not the same as acceptance of authority in practical situations. Different individuals can accept one more easily than the other. Some are more co-operative in this respect in the practical situation and others are more co-operative in the intellectual sense. There is also the question as to what kind of authority is being envisaged.

4. Finally, morale itself is based largely on the individual's personality and ego strength and it is affected by the whole environment. It is the resultant of all aspects of life, at home, socially, at play, at sport etc., and it is not just dependent upon academic success and grading. As the more successful pupils in a most wide and general sense tend to be placed in the higher streams, and, as the less successful tend to be placed in the lower streams, the morale of the individuals in the higher groups will be better than that of those in the lower groups regardless of the specific effect of academic work and streaming itself. It is interesting to note here Chetcuti's conclusion that the differences in morale between high and low streams were most marked in the case of individual morale. This tends to give weight to the above criticism.

Also, of course, in groups where individual pupils of personal high morale are numerous the resultant group morale will be high, and naturally the converse will apply.

Thus the results of the experiment could be said to be expected and are not necessarily due to the streaming.

Although this was a study in a Secondary School, it was felt that it was worth noting in the current discussion as little else has been done on the problem of morale in streamed or unstreamed schools.

D. CONCERNING THE VIEWS OF THE TEACHING PROFESSION ON STREAMING

In concluding the summary of the literature on streaming in the primary school it is worth noting the work done reporting professional views.

In 1961 Daniels published a study examining teachers' attitudes to streaming in the English primary school. His main findings were as follows:

1. The large majority of teachers believe that streaming is educationally sound and that it should be carried out on the basis of ability or of acholastic attainment, or some combination of the two.

2. A large majority of English primary school teachers believe that dull and backward children make the best progress when in groups of their own level. Most are so concerned as to this point that they are prepared to ensure that they are taught in small classes, even if this means increasing the size of the A and B classes.

3. The majority of English primary school teachers believe that streaming helps the brightest to make the best possible progress.

More recently there has been some evidence that professional opinion is less strongly in favour of streaming. In 1965 Butcher tested student teachers' attitudes to education. Using the Manchester Scales of Opinions about Education (Oliver and Butcher 1962), Butcher found that there were changes in educational opinion during training in the direction of increased naturalism, radicalism and tendermindedness. There was, however, some tendency towards reversal of attitude after experience of full time teaching.

McIntyre and Morrison (1967) found much the same with regard to teachers in training. The development of a more radical viewpoint of course correlates with a move towards non-streaming as the latter in some way helps to compose the former.

The Plowden Report also notes from its own enquiry that professional opinion is no longer so solidly behind streaming.

Finally, it might be noted that Coxe (1936) has stressed the effects which the opinions of the teachers involved in any experiment in this field can have upon that experiment. Attitudes of teachers towards streamed groups can produce attitudes in the children and so indirectly produce differences, or at least exaggerate basic differences.

E. CONCLUSIONS

There has been a fair amount of literature dealing with the subject of Streaming v. Non Streaming in the Primary School, and there has also been some experimental work. However, the controversy has not by any means been resolved as the results of some studies tend to contradict the results of others. Also, as has been shown, some aspects of the problem seem to have been well covered, but perhaps some aspects have not been adequately examined. It is hoped that this present study will in some way scientifically examine one aspect of the problem that has not received adequate attention.

THE PRESENT STUDY

This study was suggested by the observation of the author, while working in a School Psychological Service, that children in the lower streams of the primary school often tended to be young in respect of their school age group, and that some of these children displayed an intelligence level more typical of a higher stream. The study was devised after a pilot experiment had given a small measure of objective evidence suggesting that the above observation may in fact be an operative variable in the process of streaming in the primary school.

It deals with an aspect of streaming that so far seems to have been neglected or un-noticed. It is concerned with a possible defect in the usual system of streaming that, if demonstrated, would bring further distrust upon the system as it stands.

Actually since this study was started the tendency to find younger children in the lower streams has been noted and remarked upon by Butler, Pringle and Davies in the 1965 follow up of the 1958 National Cohort Study. The 1965 summary was prepared for the Plowden Committee and the above point was one of the findings mentioned in the 1967 Plowden Report.

In the geographical area studied streaming in the primary school is usually based, in the first instance, on the assessments made by infant departments on children being transferred to junior departments.

This assessment is mainly a matter of attainment and there is very little attempted assessment of ability or possible potential even at a subjective level. With the co-operation of class teachers, infant head teachers draw up a merit list of leaving pupils based on classroom performance in Reading, Writing and

Number. Also very little account seems to be taken of the actual length of time spent in the infant department and usually no age allowance, such as is inherent in a psychometric quotient, is considered. The result is that often the younger children within the school year group may be placed in a stream lower than their potential ability would indicate was suitable. Younger children may be penalised because they have had less tuition and because they may be relatively immature, nine, ten, eleven months etc. being a fairly significant development span at the age of seven years.

If a classification at seven years is considered necessary attainment is certainly important, but it should be considered in terms of opportunities that have been available for it to be acquired. Also ability should not be ignored.

Further to what is outlined above, the position is usually insufficiently corrected, and is often reinforced, as children pass through the junior department of the primary school. Transfers from stream to stream tend to be relatively few with respect to the number of children in a school year group, and, of course, there may be a gradual acceptance of the attitudes and standards of the group or stream in which the children are placed. Eventually there is often an identification with the role of A former, B former, etc.

If the assertations above are largely true then the position could be said to be both unjust and wasteful, and to warrant the attention of educational administrators.

THE AIM OF THIS STUDY IS TO INVESTIGATE THE RELATIONSHIP BETWEEN MONTH OF BIRTH AND STREAM PLACEMENT IN THE PRIMARY SCHOOL. IN PARTICULAR THE INTENTION IS TO ASSESS THE VALIDITY OF THE CLAIM THAT THERE IS MIS-PLACEMENT AND UNDERESTIMATION OF THE YOUNGER CHILDREN IN A SCHOOL YEAR AGE GROUP, AND TO INVESTIGATE THE POSSIBILITY THAT THIS EFFECT IS LASTING IN SO FAR AS IT PERSISTS INTO THE FINAL YEAR OF THE PRIMARY SCHOOL.

If an investigation is to be made into the relationship between streaming and month of birth, with special reference to the possible underestimation of younger children, then the following points must be shown to be true.

1. Firstly that the children in the study actually are being streamed. That an attempt is being made to stream the children of a year group and that, in general, the attempt is successful. If the children are not being streamed, or grouped according to ability, it would be pointless to study stream placement in any respect, let alone to study it in relation to another variable such as age.

   That there is an attempt to stream in every school in the study is without doubt as the intention was stated by all the headteachers, and all classes are named A, B, or C. Whether the aim to stream is, in general, successfully carried out can be ascertained by noting the relationship between I.Q. and stream placement.

2. Secondly it must be shown to be true that intelligence is evenly distributed throughout the sub-age groups of each school year group, and that children born in certain months have no significant advantage with respect to measured intelligence. A sub-age group is an arbitrarily formed

division of the school year group, containing all the children born in certain months or in a certain part of the year, e.g. September to December. If a school year group is divided into such sub-groups or categories statistical evaluation is simplified, particularly with respect to comparisons between the oldest and youngest children within the school year group.

If the children are evenly distributed throughout the sub-age groups there will be approximately the same numbers of children in each sub-age category for each general level of intelligence. For example, with the cases of above average intelligence, that is + 0.44 s.s., approximately one third should fall into the September to December sub-age group, approximately one third should fall into the January to April sub-age group, and approximately one third should fall into the May to August sub-age group. Similarly with those of average intelligence, + 0.44 s.s. to - 0.44 s.s., and for those of below average intelligence, that is - 0.44 s.s.

No sub-age group should be superior or inferior at any intelligence level, and such a position is confidently expected in this study. Nevertheless such a position must be shown to be true and the means of demonstrating it will be a testing of the relationship between I.Q. and month of birth.

Once these two points are established, that is that the children are, in general, being streamed, and that intelligence is found to be evenly distributed throughout the sub-age groups, without any sub-age group having an advantage with respect to intelligence, a valid examination of the relationship between age and stream placement can be made. Evaluation of the placement of the younger children can be attempted.

Thus the three main phases of the overall investigation must be.

1. The establishment of the fact that streaming is, in general, operating.
2. The establishment of the fact that intelligence is evenly distributed throughout the streams, with no sub-age group having any significant advantage.
3. An investigation of the relationship between date of birth and stream placement, with special note of the position of the younger children. An examination of the distribution of children from the various sub-age groups throughout the streams.

The questions raised by the three phases will be posed in the form of null hypotheses in the next section.

If points one and two above are confirmed, and if it is found that there is a significantly higher proportion of younger children in the lower streams, then the claim that there is misplacement and underestimation of the younger children in a school year group will receive some validation, certainly with respect to the study sample.

The study will also seek to compare first year junior school children with fourth year junior school children with respect to the above points, particularly point three. Thus any lasting effect of misplacement and underestimation will receive some form of measurement. It should be noted that the children of the study sample in the fourth year of the junior school are NOT THE SAME CHILDREN mentioned in the study as first year junior school children. This is not a follow up study, the same children being investigated when seven years old and then again when eleven years old. Such a long term measure was not practicable for the present study.

However, as there are approximately 1,000 children in the sample, 500 in the first year group and 500 in the fourth year group, it is considered that the pattern of results obtained for each school year group will be reflective of results in general where similar streaming is operating.

It will not be invalid to compare what should be a typical first year group with a typical fourth year group.

To place the investigation on a scientific and experimental basis, null hypotheses are to be formulated and tested.

A null hypothesis is based on the assumption that, in an experimental situation, whenever things are enumerated or measured, nothing but the laws of chance are operating. That is it is assumed that there are null correlations and no significant differences operating.

The Null Hypotheses to be tested are as follows.

1. That, in general, as a group, the first year junior school children in the sample were NOT streamed. That is the allocation to class groups was random, and was not according to ability.
2. That, in general, as a group, the fourth year junior school children in the sample were NOT streamed. That is the allocation to class groups was random, and was not according to ability.
3. That there is an even distribution of intelligence throughout the sub-age groups of the first year junior school children in the sample, and that no birth months have a significant advantage with respect to intelligence.
4. That there is an even distribution of intelligence throughout the sub-age groups of the fourth year junior school children in the sample, and that no birth months have a significant advantage with respect to intelligence.
5. That, incidentally, there does not seem to be any overall relationship between month of birth and intelligence. That is no birth months have a significant advantage with respect to intelligence.

6. That there is no significant relationship between month of birth (and thus age) and stream placement in the first year junior school sample, and that there is no tendency for the older children to be placed in higher streams, and the younger children to be placed in lower streams.
7. That there is no significant relationship between month of birth (and thus age) and stream placement in the fourth year junior school sample, and that there is no tendency for the older children to be placed in higher streams, and the younger children to be placed in lower streams.
8. That the distribution throughout the three streams, of children from the youngest sub-age group, will be similar for the two school year groups involved, that is for the children of the first year sample and the children of the fourth year sample.

   (For each school year the distribution would be even if no bias was operating. The hypothesis predicts that any bias found at first year level will also be found at fourth year level.)

Variables to be Noted and Assessed

In the present study it seems probable that the following variables would appear to be relevant.

1. **Original Streaming Procedure**

   In the study were the subjects, that is the children, actually streamed according to ability or rather attainment, or perhaps both?

   What procedure was adopted by the infant departments?

2. **The Relation between Ability and Attainment**

   What is understood by these concepts and particularly what is the attitude to these concepts of those concerned in the original streaming?

   What empirical relation has been found between the two concepts?

3. **Intelligence and Month of Birth**

   Is intelligence normally distributed in the same way throughout the months and seasons of birth or is it possible that births in certain months and seasons tend to produce more bright children or vice-versa more dull children?

   This is probably the most important variable needing investigation.

4. **Socioeconomic and Cultural Background**

   That is of the experimental subjects, the children in the sample. What effects could this have on the experiment?

5. **Conditioning to a Role**

   How much could this have played a part in the development of the experimental subjects?

6. <u>Differences Between Schools</u>

It is possible that there could be differences between the schools which could produce different results in the investigation.

Discussion of the Variables Noted

1. Original Streaming Procedure

Discussions were held individually with all the head teachers of infant departments whose ex-pupils were subjects of the study, and also with the head teachers of the junior departments whose schools the subjects now attended.

All the headteachers of junior departments received lists of pupils who had all been classified by infant department head-teachers, and these lists served to provide the basis of the original streaming in the junior school. The headteachers of the infant departments had all adopted the same procedure. After consultation with class teachers they had categorised children into A, B or C types. This was done in a completely subjective way without use of internal tests devised for the purpose, although all had access to classroom tests given by the teachers throughout the last school year. Certainly no objective tests were used. No separate consideration was made for ability as opposed to attainment, although headteachers claimed that, in the overall subjective assessment, note was made of children who could eventually improve. Similarly with young children, allowance in the overall assessment was made for some children, particularly for those whom they thought might improve their academic status in time. However, there was no scaled weighting allowed for age such as is typical of a psychometric test.

Thus although the procedure adopted is honest it is, perhaps, too subjective to be satisfactory. It is probably quite effective in assessing actual level of attainment, but inadequate in assessing ability and in making allowance for age.

Further tentative enquiries, beyond the scope of this present study, suggest that the procedure outlined above is fairly common. Very seldom do objective tests seem to be used, partly because of the administrative difficulties and partly because of the inadequacies of most group tests designed for the seven year old age group.

On arrival at some junior departments use is occasionally made of Schonell's Word Graded Reading Test, but this, of course, only tests attainment level in one aspect of reading. It does not give quotients or make allowances for age. In the present study the junior departments did not use it at all in the original streaming procedure. All relied solely on the classified lists from the infant departments.

Inadequate classification is a variable which could play a part in the present study and could be partly responsible for results it is anticipated may be found.

2. <u>The Relation between Ability and Attainment</u>

The concepts of ability and attainment have been well recognised for a long time in the field of education. Briefly the concept of attainment is concerned with the development and acquisition of educational standards set by peer groups and the society in which one exists, and a level of attainment indicates to what extent the standards have been aquired. The concept of ability is concerned with the potential to achieve these standards. Even theoretically it can be appreciated that there could never be a perfect correlation between ability and attainment as there are so many intervening social and personality variables.

Extreme cases help to illustrate the difference between the two concepts. On the one hand there is the individual who has a relatively high degree of ability but a poor level of attainment. A clever individual who has never been encouraged and who has never become well motivated, who has had poor social and cultural support, and who may even have had grossly inadequate attendance at a place of formal education. On the other hand there is the individual of only moderate ability whose attainments are relatively good, and who is now known as an "over-functioner". This is an individual who has been encouraged and well supported and who has developed a high level of motivation. Here the attainments are in advance of what would be predicted from the individual's age and level of ability.

Although the concepts of ability and attainment are fairly clearcut, the measurement of them in any individual of group of individuals is not so simple. The measurement of attainment is the easier function. An individual has either reached a set standard or he has not reached it, although, of course, it is true that the level of ability very much determines the attainment level that can be reached. The measurement of ability is more complicated. Tests of ability are available but it is difficult for them to be devised without the involvement of some attainment. The current psychometric position concerning ability and attainment tests is one not so much of different base and type but rather of different emphasis and degree. The ability test aannot help involving some measure of attainment, but it attempts to minimise the influence of attainment and to measure the powers of the intellect.

High and positive correlations have always been found between ability and attainment, e.g. Pidgeon and Yates (1956) (1960), Thorndike (1931) etc., and this is not surprising, but that is not to say in any selection or allocation procedure the one can be assessed and the other ignored. One is dealing with individuals and some of these individuals will be the ones who disturb a perfect correlation. They will perform better on one test than the other. For a fuller picture assessments on both ability and attainment should be obtained.

There are arguments for using only attainment assessments when allocating individuals to homogeneous groups and this procedure is particularly reasonable when children are older and need a basic attainment level to cope with the work envisaged for that group. However, at the age of seven years there has not been equal time and opportunity for the reaching of standards set by the group as a whole, and for this reason it would seem to be wrong for attainment to be assessed to the exclusion of ability. Both should be given a fair weighting.

It would appear that in our study the assessment of ability by the head-teachers has been rather inadequate, although all headteachers understood that ability and attainment were different, and this could affect the actual stream placement of any child in the sample. However, it is probably the younger children of the age group who are affected most, as their ability has had less time to mature and flourish and to bring about a high level of the more easily assessed attainment.

If it is found that younger children predominate in the lower streams the reason will be the original streaming procedure with too trusting a belief in the correlation of ability and attainment at the early age of seven years. That is unless it is found that children born in certain months, or a certain season, are less intelligent than the other children in the sample.

3. Intelligence and Month or Season of Birth

The question is posed as to whether intelligence is evenly distributed in the same way throughout the months and seasons of birth. If it is found that births at one time of the year tend to include more bright children than usual then it would be reasonable to expect more numbers from this birth group in the higher streams. Similarly, if more dull children were found in the birth group one would expect to find more children of that group in the lower streams. With particular reference to the present study, if it was found that the midsummer births, the younger children in the sample, tended to be duller as a group, then any tendency for them to be found in high numbers in the lower streams would be only to be expected. It is therefore important that this matter be closely checked and scrutinised, and comparisons not be made on the mere assumption that intelligence will be evenly distributed in the same way throughout the months or seasons of birth. The assumption seems reasonable but it must be seen to be true for the sample in the study.

Butler, Pringle and Davies in the 1958 National Cohort Study were confident that month of birth was not likely to affect the even distribution of intelligence as all the children selected for their sample were born in one week in March. The same applies to the sample of Douglas (1964) whose sample were also all born in one week in March.

However, there has been some work suggesting that month or season of birth can in some way affect intelligence. Results have been somewhat contradictory.

In 1941 Fitt concisely summarised the literature of the 1930's which reported small but consistent differences in ability according to month of birth. Summer and autumn births were found to correspond with greater ability than winter and spring births. In the northern hemisphere these are May to October and November to April respectively. The studies ranged from dull to bright subjects, children and adults. Although the actual months showing highest or lowest scores tended to vary, the overall seasonal effect was remarkably consistent. However, many of these earlier studies did not use psychological tests or objective criteria, and those that did used techniques which would now be considered outmoded.

Fitt considers the time of conception to be critical rather than the actual time of birth, with following seasonal changes in the pregnancy period being important.

Pintner and Forlano (1933) took a sample of 17,500 New York schoolchildren and divided them into four groups of equal size on basis of I.Q. High I.Q. High Average, Low Average, and Low I.Q. They found the three lower groups to have the same seasonal distribution of births, with a pronounced maximum of frequency in February or March and a minimum in May or June. A minor maximum was found in August or September and a minor minimum in October or November. The distribution for High I.Q. s was found to be different, with a minimum in mid-winter to a maximum in August and September, and with a secondary maximum in April.

The problem was tackled in an interesting way but there was no investigation to discover the mean I.Q. s of the children born in the different months or seasons, and consequently there was no investigation to see if these groups were significantly different. Such a procedure may have been more satisfactory.

Huntington (1944) conducted a similar study but his peak for very high I.Q.'s of 130+ was in March and April. The curves or distributions for the lower I.Q. groups were similar to those obtained by Pintner and Forlano in their study.

Huntington points out that seasonal distributions of births cannot be rightly understood until we take into account many factors such as climate, diet, percentage of first births, social customs and standards of living.

Knobloch and Pasamanick (1958) showed that a significantly high proportion of intellectual subnormals are born in the first three months of the year. Knobloch and Pasamanick suggest the effect of summer heat on the pregnant woman, at the time when the embryonic cortex is being organised, as a possible reason for the results obtained. Of course this work is with subnormals only but it makes possible the inference that this birth group as a whole will tend to be less intelligent because of the high proportion of sub-normals. This might not follow as there could also be a high proportion of highly intelligent individuals in the group thus making the mean I.Q. similar to that of other groups born at different times of the year. Of course, the standard deviation of the I.Q.'s of the group would be greater than that of other groups.

Orme (1962), with adult subnormals, noted the relationship between I.Q. and Birth Month or Season of Birth.

His results were as follows:

	Summer Birth	Autumn	Winter	Spring
I.Q. 55-69	29	19	12	18
I.Q. 40-54	13	17	20	20

On a chi-square test this is significant at the one per cent level of confidence and it is in the direction to support the claim that, within the subnormal range, children born in winter seem to be a little less intelligent than children born in summer. However, it is worth noting that in terms of the frequencies in this sample there are actually more summer birth sub-normals than winter birth sub-normals.

Orme used the Raven's Coloured Progressive Matrices to obtain his measure of I.Q. and he makes considerable claims for this, considering it to be an efficient measure to use with intellectually sub-normals, and a measure that is not contaminated by culture, education, etc. He even claims that the test is one of the most homogeneous measures of "what is variously called "g" factor, fluid ability, and non verbal performance". It is difficult to reconcile this claim as "g" factor or fluid ability are not the same thing as non-verbal performances. He also claims that the test is possibly one of the best measures of Hebb's Intelligence A, but, of course, Hebb's Intelligence A is by definition not measurable.

It is also worth noting that the test only samples one aspect of intelligence. It is a non-verbal test involving perceptual and spatial ability. Even the test manual suggests that, if a measure of general ability is wanted, the test should not be used by itself but in conjunction with the Mill Hill Vocabulary Test. Nor, incidentally, does it give results in terms of I.Q. but actually percentile ranks.

Of course, the extravagant claims made for the Coloured Progressive Matrices do not in themselves invalidate the results of Orme's small study.

In a study published in 1963 Greenberg claimed to have found a significant association between low climatic temperatures seven to eight months before birth and the incidence of mongol births to young mothers. This is, in some respects, a little evidence to suggest a reversal of the claim that winter and spring births are least favourable to intelligence and summer and autumn births are more favourable to intelligence.

In another study, in 1963, Orme again found adverse performance with winter and spring births as compared with summer and autumn births. Again the study was with sub-normals in the 40-69 I.Q. range, and again the Coloured Progressive Matrices Test was used. The most obvious climatic variable is that of temperature and it was hypothesised that seasonal effect on intelligence would be paralleled by the effect of monthly temperatures above or below that month's average temperature. It was further agreed that the effects would be restricted to a certain period during development if that period was specially critical for cortical development and intelligence. For each individual the month of birth and the preceding nine months were examined.

The only obvious conclusion that could be made by Orme was that summer and autumn births were best for intelligence, and the reason for this was thought to be the gradual increase of temperature throughout the pregnancy from the middle period when the development of intelligence is critical.

Williams (1964) completed a study which rather contradicts the view propounded by Orme, that is, that winter and spring births tend to affect intelligence adversely as compared with summer and autumn births. Williams found an undue preponderance of summer born children in special schools for the educationally sub-normal. This was a study involving 265 children drawn from E.S.N. Schools. Williams noted that the results could be partly due to educational organisation, such as differential entry to infant school, but that they could also be due to other factors such as the actual age group position (in time of year) with direct relation to the intra-uterine development etc. He noted that the phenomena seems to have a bigger impact on the special schools. Williams was not able to clearly separate the age group effect from the term of entry effect but he seems to think that the former is more responsible for the high number of summer birthdays in the E.S.N. sample.

It may be that a certain type of fetus may be more predisposed to intra-uterine injury or damage in any case, and that the temperature standards or variations of certain months may be more liable to affect mothers and so bring about further risk for those passing through an important developmental phase at that time. This would account for there being more children of a certain birth group in an E.S.N. sample.

As has been demonstrated the evidence on intelligence and season of birth is rather muddled and contradictory, and most of the work that has been done has been restricted to the sub-normal range. An investigation of the issue with respect to the brighter end of the intelligence distribution is that of Ojha, Kelvin and Lucas (1966). They investigated the problem with university students using the A.H.5 Intelligence Test and they found no evidence of a relationship between season of birth and intelligence.

An interesting study on a related topic is that by Johns (1962) noting the age factor in reading retardation. The youngest children of school year age groups, that is those born May to August, tended to be (as a group) more retarded in reading than the older children of the year group. This result was not attributed to a lower intelligence of the summer birth children but rather to the shorter infant training they had received.

It could be argued from this that, if such a difference persisted until the late junior school age, the system of streaming had failed, and it could be further argued that it had even reinforced the disadvantage of shorter infant training.

4. Socioeconomic and Cultural Background

This has an influence on the stream in which a child is placed. Douglas (1964) and others have confirmed this point. However, in the present study it would be reasonable to assume that the background variable will be randomly spread through all sub-age groups, and that the variable would be controlled in investigating month or season of birth and its relation to streaming in the primary school.

In any event to test the control of the variable would involve an impractical addition to the work of the study. An assessment of each individual child's background would have to be made and a complicated analysis would have to be undertaken. Tests would have to be implemented to see if there was any difference between good, average, and poor background children in terms of the proportion or number coming from each sub-group.

Of course it is possible that younger children from poorer backgrounds could suffer from a cumulative effect greater than might be expected. To test this would be extremely difficult, involving perhaps linear estimates of probabilities and expectations according to combined background group and sub-age group. This would be a major study in itself. However, it will be interesting to note if there is any difference between a relatively high social background school in the study and a relatively low social background school in the study with respect to the proportion of younger children in the lower streams.

5. Conditioning to a Role

It is not possible to note the effect of this in this study, but it is known that it can happen and it may be that streaming encourages it.

If it is found that the younger children of the school year age group still tend to occupy a disproportionate number of low stream places in the fourth year of the junior school, then it can be assumed that conditioning to a role has played some part. This will be particularly so if the younger children are found to be as intelligent as the other children in the sample.

6. Differences Between Schools

It is possible that there could be differences between the schools which could produce different results in the investigation.

These could come about because of differences in the general socio-economic backgrounds of the various schools. As already mentioned it will be interesting to note if there are any differences between schools of high socioeconomic background and low socioeconomic background with respect to the proportion of younger children in the lower streams.

Differences between schools could also result from differences in the attitudes of headteachers and teachers, particularly with respect to the idea of streaming itself. Toleration of movement from stream to stream, involving extra administrative work, could also be an important factor.

The Pilot Experiment

This was carried out in February 1965 at a junior school in an industrial town in the north-east of England. The school takes its pupils from a wide range of social and cultural backgrounds but the majority of the children come from lower middle class and upper working class homes. On an overall assessment the social background of the pupils would be rated as above average. The same assessment could be made as to the intellectual calibre of the pupils. On the intelligence test in the experiment the average intelligence quotient was above average of that for the population at large.. The average I.Q. was 108 on the Moray House Picture Intelligence Test.

The school could be described as a happy school, not subject to a great deal of repression. However, it is also a school with definite aims and standards in which some pressure is brought to bear, especially upon children in the top streams.

Each school year group is divided into three streams, according to a merit list prepared by infant head-teachers, and although there is movement from one stream to another this is restricted to about four promotions per year from C to B, and B to A, demotions being similar in number.

The headteacher and most members of staff are quite happy with the streaming system and they believe the implementation is reasonably fair and accurate. This is consistent with the finding of J. C. Daniels (1961) in assessing the attitude of teachers to streaming.

In the pilot experiment all the first year children of the school were selected as subjects.

The dates of births and stream placements of all the children were noted.

On one day all the children in the sample were tested on the Moray House Picture Intelligence Test I for Seven Year Olds.

After marking, scoring, checking and conversion of raw scores into standardised scores, the material was tabulated. Eventually the data was used to construct contingency tables in order to test the relationship between the following variables.

1. Month of Birth (And Consequent Age) v. Level of I.Q.

2. Month of Birth (And Consequent Age) v. Stream Placement

Chi square values were computed for both contingency tables.

With respect to the first relationship tested, Month of Birth and Consequent Age v. Level of I.Q. the chi square value obtained was well below the point of significance. The obtained value was 4.12 whereas that required for a five per cent level of confidence was 9.49. Thus as the chi square value was so small it can be claimed that there is no significant difference in intelligence between children in the sample born in different parts of theyear. The younger children within the school year group were no less intelligent than the older children, intelligence being defined in terms of level or quotient.

It is worth noting here that older children of the same I.Q. as other younger children would be slightly more mature in terms of mental age, but this involves the main point at issue that some allowance should be made for age, otherwise some bright but young children may be under-estimated and condemned to a lower standard of education than is warranted.

Incidentally, with respect to the first relationship investigated, the trend in the chi square analysis was in favour of the younger children in the year group, that is in respect of high I.Q. Of course, as stated above, this trend did not reach anything like a level of significance.

With respect to the second relationship investigated, Month of Birth and Consequent Age v. Stream Placement, a highly significant value of chi square was obtained. The chi square value obtained was 25.79 and the value needed for a five per cent level of confidence was 9.49, and for a one per cent level of confidence 13.28.

Month of Birth (And Consequent Age) v. Level of I.Q.

	Sept.1956 Dec. 1956	Jan.1957 Apr.1957	May 1957 Aug.1957	
I.Q. 116+	13	12	10	35
I.Q. 95-115	21	12	15	48
I.Q. 94-	4	7	9	20
	38	31	34	103

Chi Square = 4.12 Not Significant.

For 4 degrees of freedom a Chi Square Value of 9.49 is needed at the five per cent level of confidence.

Month of Birth (And Consequent Age) v. Stream Placement

	Sept.1956 Dec. 1956	Jan.1957 Apr.1957	May 1957 Aug.1957	
A Stream	24	10	2	36
B. Stream	9	10	16	35
C Stream	5	11	16	32
	38	31	34	103

Chi Square = 25.79 Highly Significant - beyond the one per cent level of confidence.

A Maxwell Chi Square Trend Analysis showed a definite trend towards the older children in the sample being placed in higher streams and the younger children being placed in lower streams. A more detailed study of the contingency table revealed that the greatest chi square value for an individual cell was that for the oldest children in the A stream. Actually in terms of original cell frequencies 24 out of 38 of the children in the oldest sub-group in the school year were in the A stream.

Contrasting these two results it would seem that, although there was no significant difference between month or season of birth and intelligence, and although intelligence was evenly distributed between the sub-age groups within the school year, there was a significant difference in the distribution of places in the higher streams, with the younger children being more readily placed in the lower streams.

It would appear that the younger children had been under-estimated and allowance had not been made for their lack of opportunity and their relative immaturity which would be both cortical and social. Streaming had probably been implemented in terms of attainment and level of maturity when leaving the infant department.

THE EXPERIMENT PROPER

A similar procedure to the pilot experiment was adopted but it was conducted on a larger scale.

Description of the Sample

Five junior schools were included in the sample and all the first year and fourth year children were subjects in the experiment. All five junior schools were situated in an industrial town of the north east of England. The socio-economic backgrounds of all the schools were rather mixed but the average rating for the different schools was different.

Two schools, A & B, both built about 1910 but quite pleasant to work in, could be said to be above average with regard to socioeconomic background. They drew more children from high status residential areas than the other three schools. Included among the parents were a fair number of professional and clerical workers etc. and approximately half the children came from owner occupier homes. School B was the school which had been used in the Pilot Experiment.

One of the five schools, C, was very much average with regard to home background. Almost all the children came from homes on a modern council estate, and the school itself was only seven years old. Most of the parents were skilled or semi-skilled workers with a fair number of casual labourers amongst them. Almost all the children were well dressed and well kept, and shortage of money did not seem to be a problem. This school was by far the most homogeneous with regard to socioeconomic background.

The other two schools, D & E, could be said to be below average with regard to socioeconomic background. Most of the parents were semi-skilled workers or labourers and each school had in attendance the children of several problem families. Most of the children were well dressed and well cared for but some were obviously neglected and were living near or below the poverty line. Most of the children lived in older private property and the vast majority of the parents were tenants and not owner occupiers. Both schools themselves were built before the turn of the century and are situated in older parts of the town. However, both have been modernised to some extent, with new flooring etc., and both are in good decorative order.

The above assessments of socioeconomic background were not obtained by detailed formal methods. They were the result of discussions with headteachers, health visitors and school welfare officers, together with the author's own assessment based on his experience as a social worker and teacher in the town for the past fifteen years.

In all five schools the children are divided into three streams, and, as mentioned before in an earlier section, this is done originally on the basis of a classification made by the infant departments. All schools claimed that, on an average, two children were moved up and two down from stream to stream at the end of each term, up to the end of third year.

Three schools had only two movement times once the children were in their third year and all the schools admitted that there was little promotion or demotion in the final junior school year, as this was thought to be unsettling.

All five headteachers were in favour of streaming, two being very strongly in favour. One of these was the headteacher of a school of above average socioeconomic background, school B, and one was headteacher of a school of below average socioeconomic background, school E. The two headteachers from the other above average socioeconomic background school, school A, and the other below average socioeconomic background school, school D, both thought streaming was the best procedure but they were not so hostile to the idea of non-streaming. Perhaps significantly, the headteacher of the school of average and rather homogeneous socioeconomic background, school C, was the most receptive to non-streaming. He was convinced that there had to be streaming at some stage but he had already come to a decision, following talks with a local inspector of schools, to unstream his first two year groups at the beginning of the next academic session.

The Experimental Design and Procedure

The subjects were all the first year and fourth year children in all five schools. The information required for each subject was Date of Birth, Present Stream, and I.Q.

Once this was obtained it would be possible to ascertain statistically whether children were generally being streamed according to ability, whether intelligence was evenly distributed throughout the months or seasons of birth, and also whether a disproportionate number of young children were being placed in the lower streams.

The information would also enable comparisons to be made between first year and fourth year children with regard to younger children being placed in the lower forms. It would also allow comparisons to be made between schools.

Once collected the data would be tabulated and placed in contingency tables to enable the testing of the null hypotheses outlined in a previous section of this study.

These were as follows:

1. That, in general, as a group, the first year junior school children in the sample were NOT streamed. That is the allocation to class groups was random, and was not according to ability.
2. That, in general, as a group, the fourth year junior school children in the sample were NOT streamed. That is the allocation to class groups was random, and was not according to ability.
3. That there is an even distribution of intelligence throughout the sub-age groups of the first year junior school children in the sample, and that no birth months have a significant advantage with respect to intelligence.
4. That there is an even distribution of intelligence throughout the sub-age groups of the fourth year junior school children in the sample, and that no birth months have a significant advantage with respect to intelligence.
5. That, incidentally, there does not seem to be any overall relationship between month of birth and intelligence. That is no birth months have a significant advantage with respect to intelligence.
6. That there is no significant relationship between month of birth (and thus age) and stream placement in the first year junior school sample, and that there is no tendency for the older children to be placed in higher streams, and the younger children to be placed in lower streams.

7. That there is no significant relationship between month of birth (and thus age) and stream placement in the fourth year junior school sample, and that there is no tendency for the older children to be placed in higher streams, and the younger children to be placed in lower streams.
8. That the distribution throughout the three streams, of children from the youngest sub-age group, will be similar for the two school year groups involved, that is for the children of the first year sample and the children of the fourth year sample.
(For each school year the distribution would be even if no bias was operating. The hypothesis predicts that any bias found at first year level will also be found at fourth year level.)

To test the Null Hypotheses the following relationships would have to be investigated.

1.	With the First Year Children	-	Level of I.Q.	v	Stream Placement
2.	" " Fourth " "	-	" " "	v	" "
3.	" " First " "	-	" " "	v	Birth Month (and thus age)
4.	" " Fourth " "	-	" " "	v	" " " " "
5.	" " Total Sample of Children	-	" " "	v	" " " " "
6.	" " First Year Children	-	Stream Placement	v	Birth Month (and thus age)
7.	" " Fourth " "	-	"	v	" " " " "

8. The distribution of the youngest children throughout the three streams in Fourth Year v the distribution of the youngest children throughout the three streams in First Year.

<u>The Proposed Statistical Treatment of the Data</u>

A. For the relationships 1 to 7 mentioned above the application of a Chi Square Test of Significance. Three by three contingency tables will be arranged with the variables concerned being divided as follows.

1. Level of I.Q.

High I.Q.	- A	Standard	Score	of	+ 0.44
Average I.Q.	- "	"	"	"	+ 0.44 to -0.44
Low I.Q.	- "	"	"	"	-0.44

2. Stream Placement

A Stream

B Stream

C Stream

3. Month of Birth (and thus age)

Born September to December (inclusive)

" January to April (inclusive)

" May to August (inclusive)

In addition Maxwell's Chi Square Trend Analysis may be used where necessary. (Maxwell 1961).

This more refined technique notes any linear trend in the proportions in the cells of the contingency tables, and it readily indicates what is actually contributing greatly to the chi-sqaare value.

It sub-divides the overall chi-square value for inspection. It is applicable to contingency tables where the classification categories fall into a natural order, for then it is possible to search for trends in the data. In particular a component of chi-square due to a linear trend can be separated out and tested for significance.

When the categories of the variables have a natural order it may be assumed that there is a continuous variable underlying them, and the variables can be treated as if they were quantative variables, numerical values being allotted to the categories. Regression methods can be used for partitioning the overal chi-square value, so that trends can be examined statistically. Moreover, since a trend, or regression line, is based on just one degree of freedom it is possible that, although the overall chi-square value is not significant, the trend may be.

B. For relationship 8, mentioned above. In order to see if any correction, or change by fourth year, occurred with respect to the distribution of younger children throughout the various streams, the following two procedures were considered to be necessary.

1. For the youngest sub-age groups within first year and fourth year (born May to August) a Chi-Square Test and Trend Analysis for School Year v Stream Placement.

Stream	A	B	C
1st Year Youngest Age Group			
4th Year Youngest Age Group			

2. (a) A test of proportions between 1st Year Youngest Age Group Children in Form A and 4th Year Youngest Age Group Children in Form A.

   (b) A test of proportions between 1st Year Youngest Age Group Children in Form B and 4 th Year Youngest Age Group Children in Form B.

   (c) A test of proportions between 1st Year Youngest Age Group Children in Form C and 4th Year Youngest Age Group Children in Form C.

C. As a check to see if there were any significant differences between the different schools, (in particular any differences with respect to the three main relationships being investigated, viz. I.Q. v Stream Placement, I.Q. v Birth Month, and Birth Month v Stream Placement) the following procedures will be adopted.

1. The means and standard deviations of the I.Q. distributions to be compared by simple inspection and possibly statistical tests of difference.

2. The construction of contingency tables for each individual school with respect to Level of I.Q. v Stream Placement, I.Q. Level v Birth Month, and Birth Month v Stream Placement. The comparison of the chi-square values between the different schools. This can be done for the first year and fourth year children.

3. The noting and contrasting of the proportions of younger children in the lowest forms, i.e. C forms, between the different schools. This can be done for first year and fourth year children.

It is noted that in order to prevent this study from growing to an unwieldy size, work on comparisons between schools will be kept to a minimum. Most of the comparisons will be made at the level of simple inspection and a complex matrix of statistical procedures will not be undertaken. Where inspection suggests any interesting differences statistical work can be pursued.

It is anticipated that there will be significant differences between some of the schools with respect to the mean I.Q.s and possibly even the standard deviations, as the schools serve areas of different socioeconomic background, but more interesting will be the question as to whether these different schools will produce similar or different patterns and distributionsof data.

The Procedure for Obtaining the Data

As already stated the information required for each subject was Date of Birth, Present Stream, and I.Q. The first two were easily obtained from class registers. As the total sample included approximately one thousand children it was not possible practically to obtain I.Q. assessments for all the children on individually administered tests such as the Stanford Binet Intelligence Scale or the Wechsler Intelligence Scale for Children. Accordingly group intelligence tests had to be used. In May 1966 all the children in the fourth year of every junior school in the town concerned in the experiment were given, in a classification examination, a Moray House Group Test, of the verbal reasoning type. Although too heavily loaded with verbal items, this type of test is widely accepted as a reasonable measure of a child's intelligence at the age of eleven years. The results of this testing were used to provide the I.Q. data for the fourth year children in the sample.

All the fourth year children were given the test on the same day, and so, obviously, different individuals administered the testing of the fourth year children of the experimental sample. This last point matters little as instructions are brief and well standardised, and the testing situation is such that the personality and the teaching skill of the tester play no significant part. The actual test administered was the M.H. 77 and it was standardised on a sample of 65,872 children in 1964. It has a mean of 100 and a standard deviation of 15. Two measures of reliability were implemented in the standardisation of the test.

(a) A measure of internal consistency, calculated by Ferguson's method on a sample of 201 children, gave a correlation coefficient of 0.976.

(b) A coefficient of equivalence and stability of M.H. 77 with M.H. 76 was calculated from the scores of the complete year group of 2,270 children in a certain area. The interval between testings was fourteen days and the obtained coefficient was 0.954.

The M.H. 77 test itself is made up of the usual group test material such as similarities, analogies, series, reasoning problems etc.

In June 1966 all the first year children in the sample were tested on the Moray House Picture Intelligence Test 1. This test is specifically designed for seven year olds, and, as implied by the name, it consists of picture items and does not at all depend on reading ability. The test is the most comprehensive of its kind, having nine sub-tests and a total of 100 items. This compares favourably with the N.F.E.R. Picture Intelligence Test which has only 60 items.

The sub-tests deal with acting on instructions, noting the object that does not belong to a group, and the completion of pictures by selecting the correct missing part. They deal with picture absurdities, that is the noting of the picture in the group that is absurd, with the ordering and sequence of picture series, and with the selection of the reversal of a given picture or figure from a group of like pictures or figures. Other sub-tests deal with the selection from a group of drawings a part that a given object always has, with a completion of a series of picture story analogies, and with the completion of a diagrammatic series.

The test is an interesting test for children of seven years and it seems to motivate them well and to hold their attention, However, it is a test which is difficult to administer because of the great number of instructions to be given and the number of examples to be worked or taught by the tester. Although the instructions are standardised word for word the approach is very much a teaching one. For example the involved instructions should not be read in a stilted manner but should be spoken rather than read, and there should be emphasis on good rapport. Also some of the example work involves the eliciting of answerssfrom the class.

As this testing situation involves the personality, the teaching skill, and the testing expertise of the tester, it was thought that all the testing should be done by one person, in this case the author himself. In this way the variable of tester would by held constant. Otherwise fifteen different groups would have been tested by fifteen different people.

Of course, with this procedure not all the children could be tested at the same time. However, all were tested within a period of three days. With the administration of the Moray House Picture Intelligence Test 1, it was considered that holding the tester variable constant was more important than having the same time of testing for all subjects.

Following the testing all the papers were marked and checked and all totals and conversions were double checked. This, together with the fieldwork, involved a tremendous ammunt of work as appmoximately 500 papers with 4,500 sub-tests containing a total of 50,000 items had to be scored and checked.

The Moray House Picture Intelligence Test standardisation is based on work done in 1943 with 8,107 children, and the conversion tables calculated from the norms give a mean of 100 and a standard deviation of 15. This is the same as the mean and standard deviation of the M.H. 77. Thus if detailed comparisons of I.Q. scores between first year groups and fourth year groups were wanted, the matter would be simple, as the scores are directly comparable. The reliability of the M.H. Picture Intelligence Test was calculated by the Ferguson method and the coefficient quoted in the manual is of the order of 0.96.

It is true, of course, that the two tests, M.H. 77 and M.H. Picture Intelligence Test, are of a different type. The M.H. 77 test is mainly concerned with verbal ability and it necessarily involves reading ability. The M.H. Picture Intelligence Test is more in the nature of a performance test and it involves no reading ability.

However, the tests are applicable to and suitable for the two main age groups involved in the project, that is seven year olds and eleven year olds. Each in its place probably gives as reasonable an estimate of intelligence as it is at present possible to obtain from group tests. The content and testing procedure is different for the two tests, and this means comparisons between scores on the two tests have to be made with caution. However, the content and testing procedure for each test is what is applicable to each age group and each is necessary for the best estimate of intelligence at each age.

To obtain the best estimate of intelligence at each age is more impobtant for comparisons than merely ensuring the tests are of similar type. Of course if it had been possible to have them of similar type as well this would have been ideal.

The Arrangement of the Data

To apply the chi-square test to the relationships being investigated, the relevant variables had to be sub-divided. As mentioned previously the variables were sub-divided into three parts, and three by three contingency tables were constructed. A tripartite sub-division was decided upon because the variable of Stream Placement naturally was of this pattern. All the subjects fall into one of the three categories, A stream, B stream, or C stream. The evidence here was from distinct, discrete categories. It was easily obtained and it was ready for allotment to cell frequencies in the contingency tables.

With regard to the variable of Month or Season of Birth (and thus age) the obvious way of dividing the variable into three parts was to allot a four month span to each sub-divison. The eldest children, September to December births inclusive, were the subjects placed in the first sub-division. January to April births inclusive were placed in the second sub-division. The subjects born towards the end of the academic year, May to August inclusive, were placed in the third sub-division. As the dates of birth had all been noted it was only a simple clerical task to place subjects in the right category and note the cell frequencies in the contingency tables.

The data for the two variables just mentioned was easily obtained and arranged and it involved no direct measures with the children. This was not the case with the variable of I.Q. Level. The fourth year junior school children were all given the M.H. 77 Test and the first year junior school children were all given the M.H. Picture Intelligence Test. After marking and checking and conversion to I.Q.'s etc. each individual child had a score, and a distribution of scores was obtained.

As the means and standard deviations of the standardised tests are the same, comparisons should be simple but it is possible that one or both of the standardisations are dated or are not quite applicable to the sample in the study. For instance the sample as a whole could be slightly above or below average, with respect to intelligence, or possibly the M.H. Picture Intelligence Test standardisation, completed in 1934, could be a little out of date, not being quite applicable to the first year children in the sample. With respect to the latter the sample mean and standard deviation could be somewhat different from that of the original standardisation.

Individual scores, and group measures such as means and standard deviations, should be suitable for comparisons within each year group where the same test was being used throughout. However, in order to make comparisons between years it may be as well to convert all scores into standard scores. Comparable continuums of ability for each year group will thus be available.

This last measure, conversion to standard scores, will also make the task of sub-dividing the I.Q. variable very much easier and much more exact. Approximately one third of the population fall above the standard score of 0.44, one third fall below a standard score of -0.44, and one third fall between a standard score of +0.44 and a standard score of -0.44. These cut-off points, at +0.44 and -0.44 standard scores, seem to be the most reasonable ones to use in sub-dividing the variable into approximately three equal parts. Using the mean I.Q. and the standard deviation of I.Q. for the total first year sample, the I.Q. levels at the cut-off points of +0.44 and -0.44 standard scores can be established. To obtain the I.Q. score at +0.44 standard scores the equation would be as follows, with X denoting the required I.Q. score.

$$0.44 = \frac{X - \text{Mean}}{\text{Standard Deviation}}$$

Once the I.Q. levels at the +0.44 and -0.44 standard score points have been found the tripartite sub-divison of the I.Q. variable is a simple matter.

The same procedure can be adopted for the fourth year sample, the mean and standard deviation of this I.Q. distribution being used to enable a threeway subdivision of the fourth year I.Q. variable.

Although in the original standardisations there were no differences between the means and standard deviations of the two tests used with the first year children and the fourth year children, differences may be found in this study. Possible reasons for this have been outlined above, and it could transpire that the cut-off points in terms of I.Q. could be different for the two main groups. The use of standard scores based on the two main sub-samples, first year children and fourth year children, enables valid comparisons to be made.

RESULTS

A.

The Investigation of the I.Q. Variable

1. The Means and Standard Deviations

With First Year Children - test used being the M.H. Picture Intelligence Test.

For the Total First Year Sample in all Five Schools Combined.

The mean I.Q. was 110.849. (110.9)

The standard deviation was 15.69. (15.7)

For School A (High socioeconomic background).

The mean I.Q. was 117.219.

The standard deviation was 13.52.

For School B (High socioeconomic background).

The mean I.Q. was 115.75.

The standard deviation was 13.63.

For School C (Average socioeconomic background).

The mean I.Q. was 107.83.

The standard deviation was 14.54.

For School D (Low socioeconomic background).

The mean I.Q. was 107.0.

The standard deviation was 15.89.

For School E (Low socioeconomic background).

The mean I.Q. was 106.089.

The standard deviation was 16.8.

With Fourth Year Children - test used being M.H. 77.

For the Total Fourth Year Sample in all Five Schools Combined.

The mean I.Q. was 103.6.

The standard deviation was 17.98.

For School A (High socioeconomic background).

The mean I.Q. was 109.7.

The standard deviation was 16.39.

For School B (High socioeconomic background).

The mean I.Q. was 109.7.

The standard deviation was 17.5.

For School C (Average socioeconomic background).

The mean I.Q. was 98.65.

The standard deviation was 16.7.

For School D (Low socioeconomic background).

The mean I.Q. was 100.7.

The standard deviation was 14.2.

For School E (Low socioeconomic background).

The mean I.Q. was 98.1.

The standard deviation was 19.6.

(For further detail see Appendix A)

These results tend to show that, as a whole, the sample is a little above average with regard to intelligence. Both the mean I.Q. of the first year sample and the mean I.Q. of the fourth year sample were higher than the mean I.Q.'s of the original samples used in the standardisation of the respective tests M.H. Picture Intelligence Test and M.H. 77. The mean I.Q. of the total group of fourth year children was three points above the mean of the 1964 standardisation sample. This difference is significant at the one per cent level of confidence and the three points difference probably reflects the true deviation of the total project sample from the population at large. There is little reason to suspect that the first year schoolchildren in the project sample are genuinely different in terms of intelligence from the fourth year schoolchildren in the project sample, particularly as there has been little population movement in the last few years in the local areas involved. The fact that the first year sample's mean I.Q. is so very much higher than the mean I.Q. of the standardisation sample (10 points) is best explained in terms of dated norms. The M.H. Picture Intelligence Test was standardised in 1943 and the norms are no doubt now out of date. Seven year old children probably perform better on this test now because of factors such as earlier social maturity, the influence of nursery education, better infant teaching with better methods and smaller classes than of twenty five years ago, the influence of mass media in early education, and possibly test sophistication and increased confidence in a testing situation.

The M.H. 77 Test, used with the fourth year sample, was standardised only two years before its use in the project and the norms are much more likely to be accurate. The deviation of the fourth year sample from the population at large is probably indicative of a genuine deviation of the total sample.

As the distribution statistics obtained for the samples in our study differ from the original statistics of the standardisation samples, and particularly as the difference is greater from one set of distribution statistics than for the other, it is imperative that use be made of standard scores based on the relevant data. Comparisons between sub-samples, that is between first year schoolchildren and fourth year schoolchildren, would be extremely difficult without this procedure.

The standard scores will be based on the data of the relevant individual distributions, that is the mean and standard deviation of the I.Q. distribution of the first year children in the sample (tested on M.H. Picture Intelligence Test), and the mean and standard deviation I.Q. of the fourth year children in the sample (tested on M.H. 77.)

The actual working of the standard scores follows this discussion on the obtained I.Q. means and standard deviations.

In some ways the I.Q. means and standard deviations obtained for the different schools are as might be expected. The relative differences between the different schools in the sample were more or less predictable. The two schools A and B, having children from predominantly high socioeconomic backgrounds, obtained similar mean I.Q. s. These mean I.Q. s were significantly higher than the means for the other three schools. This result applied to both tests, that for the first year sample and that for the fourth year sample.

Perhaps surprisingly School C, considered to be of average socioeconomic status and drawing on a good working class population from a modern estate, scored only at the same level as the two schools D and E, designated as catering mainly for low socioeconomic backgrounds. However, it is interesting to note that this school, whose population is considered to be rather homogeneous with respect to background, did have a smaller standard deviation on the M.H. Picture Intelligence Test than those obtained by schools D and E.

Also as expected schools D and E obtained similar group scores to one another, and they tended to score lowly as compared to schools A and B. The now well known sociological maxim, that children from areas of high socio-economic background tend to obtain above average I.Q. s and that children from areas of low socioeconomic background tend to obtain I.Q. s of below average, was amply borne out in this study.

2. Conversion to Standard Scores and the Division of the I.Q. Variable

In the original planning of the statistical work of this study it was considered that the conversion of I.Q. scores into standard scores would be a desirable procedure. After the results had been obtained for the first year and fourth year I.Q. distributions it was considered that conversion to standard scores was essential, particularly if comparisons were to be made between the year groups.

with the First Year Group.

The mean I.Q. was 110.9 and the standard deviation was 15.69. For any individual a standard score could be obtained by the formula $Z = \frac{(M - X)}{S}$

where Z = the standard score,

where M = the mean I.Q.

where X = the individual I.Q.,

where S = the standard deviation of the I.Q. distribution,

and where (M - X) means the <u>difference</u> between M and X,

$$\text{e.g. } Z = \frac{(111 - 120)}{15.7} = \frac{+9}{15.7} = +0.6369$$

$$\text{or } Z = \frac{(111 - 80)}{15.7} = \frac{-31}{15.7} = -2.0 \text{ approx}$$

As the main statistical work with the intelligence variable requires that variable to be sub-divided into three parts the following procedure was adopted.

Investigation of areas under the normal distribution curve (Cambridge Elementary Statistical Tables, 1953) showed that approximately one third of the population fall above + 0.44 standard deviations or standard scores, that one third fall between + 0.44 standard deviations and - 0.44 standard deviations, and that one third fall below - 0.44 standard deviations. Thus two arbitrary I.Q. cut off points were obtained approximating to these points on the normal distribution curve. The I.Q. cut off points were obtained in the following way.

$$Z = \frac{(M - X)}{S}$$

where X is the I.Q. point required.

$$\therefore (Z)(S) = (M - X)$$

$$\therefore (Z)(S) + M = X$$

With the First Year Group this is as follows.

$$0.44 = \frac{(110.9 - X)}{15.69}$$

$$\therefore (0.44)(15.69) = 110.9 - X$$

$$\therefore 6.9036 = 110.9 - X$$

$$\therefore X = 110.9 - 6.9$$

$$= 104$$

The I.Q. score of 104 is therefore the lower cut off point in the I.Q. distribution of the first year children.

In obtaining the upper cut off point

$$0.44 = \frac{(X - 110.9)}{15.69}$$

$$\therefore X - 110.9 = (0.44)(15.69)$$

$$\therefore X = 6.9036 + 110.9$$

$$= 117.8$$

$$\text{i.e. } X = 118 \text{ approximately.}$$

Thus the first year I.Q. distribution can be divided into three parts using the I.Q.'s of 104 and 118 as division points. Obviously I.Q.'s above 118 would be placed in the top or above average category, and I.Q.'s below 104 would be placed in the bottom or below average category. I.Q.'s between 118 and 104 would be placed in the middle or average category. With regard to I.Q.'s that actually were 118 or 104 an arbitrary decision had to be made as to placement.

To maintain balance either both had to be placed in the end categories (above or below average) or both had to be placed in the middle category of average. It was decided that the most reasonable procedure, causing the least distortion of any kind in a three way division, was placement in the middle or average category.

Thus the first year I.Q. distribution was divided into approximately three equal parts as follows.

Above Average ...	+ 0.44 standard scores	- I.Q. 119 and above.
Average	+ 0.44 standard scores to	
	- 0.44 standard scores	- I.Q. 104 - 118 inclusive.
Below Average ...	- 0.44 standard scores	- I.Q. 103 and below.

With the Fourth Year Group

The mean I.Q. was 103.6 and the standard deviation was 17.98. The same procedure was adopted with the fourth year group as with the first year group.

Using the same system it was ascertained that the fourth year I.Q. distribution could be divided into approximately three equal parts as follows.

Above Average ..	+ 0.44 standard scores	- I.Q. 112 and above
Average	+ 0.44 standard scores to	
	- 0.44 standard scores	- I.Q. 96 - 111 and inclusive
Below Average ..	- 0.44 standard scores	- I.Q. 95 and below.

3. <u>An Assessment of the Normality of the I.Q. Distributions.</u>

This is concerned with assessing whether the I.Q. distributions of the individual schools in the study are normally distributed, that is that the I.Q. scores in terms of frequencies will correspond to the normal (Gaussian) curve. The I.Q. scores should tend to be normally distributed about their mean. Certainly this seems to be the nature of things, and, as noted by many psychologists, for example Vernon (1960), there is a definite tendency for human abilities to be normally distributed. If a large unselected group of children is tested, and the numbers obtaining each score, or I.Q., are plotted, the graph usually approximates to the symmetrical, bell-shaped curve known as the normal distribution curve. Most individuals score near the mean and fewer and fewer individuals score as either extremes are approached.

In practice the graph tends to be a little irregular unless the numbers are very large, and, of course, if the group has been specially selected it will tend to become skewed rather than symmetrical. This could be the case with some of the schools in the study, as the pupils of schools having non-average socioeconomic backgrounds could be considered as specially selected groups. Certainly the high socioeconomic background schools may tend to produce distributions which are slightly negatively skewed as compared to the distribution of all schools combined. Similarly the low socioeconomic background schools may tend to produce distributions which are slightly positively skewed as compared to the distribution of all schools combined.

The following table shows the I.Q. distributions for all schools in both years.

The distribution of each individual school can be examined to see if it is normally distributed about its own mean. It can also be examined in relation to the mean and distribution of the whole, combined, school year group.

Table Illustrating the I.Q. Distributions

At First Year Level

I.Q. Range	A	B	C	D	E	Total School Year Group
	Individual Schools (with frequencies)					
129.5 - 139.5	21	14	2	6	8	51
119.5 - 129.5	29	33	21	18	17	118
109.5 - 119.5	26	27	19	21	23	116
99.5 - 109.5	15	19	25	27	21	107
89.5 - 99.5	8	7	15	14	23	67
79.5 - 89.5	4	2	3	7	6	22
69.5 - 79.5	0	1	4	6	7	18
59.5 - 69.5	0	1	1	1	2	5
Totals	103	104	90	100	107	504
Mean I.Q.	117.2	115.7	107.8	107.0	106.0	110.8
Standard Devs.	13.0	13.6	14.5	15.8	16.8	15.7

At Fourth Year Level

Individual Schools (with frequencies)

I.Q. Range	A	B	C	D	E	Total School Year Group
139.5 - 149.5	3	3	1	0	1	8
129.5 - 139.5	6	11	3	1	2	23
119.5 - 129.5	23	17	7	5	19	71
109.5 - 119.5	24	24	18	15	12	93
99.5 - 109.5	22	25	18	16	11	92
89.5 - 99.5	14	10	27	18	18	87
79.5 - 89.5	9	8	20	9	19	65
69.5 - 79.5	3	6	8	6	18	41
59.5 - 69.5	1	1	4	0	4	10
Totals	105	105	106	70	104	490
Mean I.Q.s.	109.7	109.7	98.6	100.7	98.1	103.6
Standard Devs.	16.4	17.5	16.7	14.3	19.6	17.9

At the level of inspection an examination of the twelve individual I.Q. distributions indicates that the distributions tend to be normally distributed about their own means. This is the case with all schools at both first year level and at fourth year level, although perhaps this is not so, completely, with schools A and B at first year level. Here there seems to have been insufficient "headroom" or "ceiling" with a resulting "bunching" in the highest category. The same applies to the Total Group.

It certainly appears that the other I.Q. distributions are normally distributed about their means, although with school E at fourth year level there is an irregularity almost sufficient to suggest a bi-modal distribution. This is probably just a chance result, and perhaps, partially, it is due to the arbitrary category limits. Perhaps there were several borderline cases which did not fall into an expected (expected in terms of a normal distribution) category, but just fell into a neighbouring category. This would cause distortion to the normal distribution. If the arbitrary category limits were changed slightly a different, more normal, distribution might be obtained.

For example if 6 cases from category 3 crossed the borderline to category 4, 4 cases from category 4 crossed the borderline to category 5, 4 cases from category 8 crossed the borderline to category 7, 7 cases from category 7 crossed the borderline to category 6, and 5 cases from category 6 crossed the borderline to category 5, a normal distribution would be formed.

A tabulated representation illustrates the point.

School E

	Categories								
	1	2	3	4	5	6	7	8	9
Actual Distribution	1	2	19	12	11	18	19	18	4
After Suggested Changes	1	2	13	14	20	20	16	14	4

Examination by simple inspection suggests that, at first year level, schools A and B are rather negatively skewed, and schools C, D, and E are slightly positively skewed. At fourth year level schools A and B again seem to be negatively skewed but perhaps to a lesser extent than at first year.

level. Schools C, D and E again seem to show some positive skewing at fourth year level, but this is very slight and the distributions can be described as reasonably normal.

The above assessments have been made at the level of simple inspection but it is possible to test statistically a distribution for normality. Guilford (1956) describes a method of obtaining the frequencies that would be expected if a distribution (of given frequencies, mean and standard deviation) was normal. He then shows how these expected frequencies can be assessed in relation to the observed frequencies of the distribution. A form of the Chi-Square Test can be applied, so that a value can be obtained which will indicate whether the actual obtained distribution is significantly different from a normal distribution.

Chi-square values for each interval are obtained and these are totalled to produce an overall value. In assessing the overall value the degrees of freedom allowed are the number of intervals involved minus three. One degree of freedom has been lost in computing the mean, a second in computing the standard deviation, and a third is allowed for N, the size of the sample.

The formula for obtaining the chi-square values for each interval is as follows.

$$\frac{(fo - fe)^2}{fe}$$

where (fo - fe) is the difference between the observed frequencies and the expected frequencies.

Also where $fe = y \left(\frac{i.n.}{s.d.} \right)$

Further detail of the formula can be found in Appendix A.

It was decided that the statistical test described should be applied to the following two distributions.

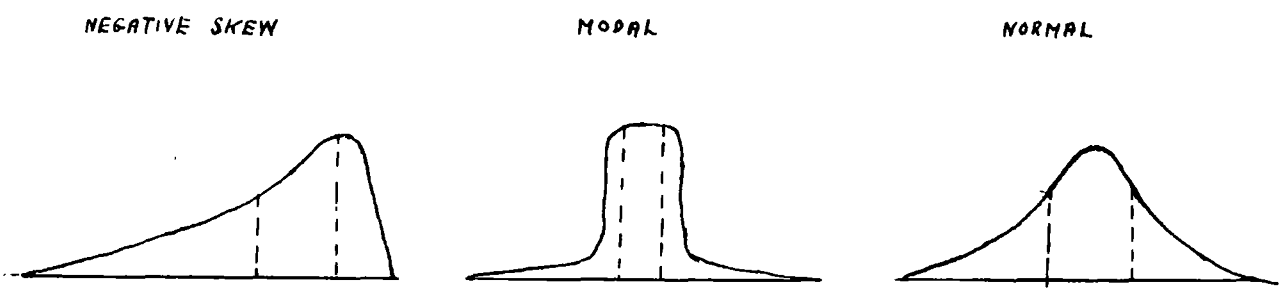
1. The I.Q. Distribution of the Total First Year Sample in all Five Schools Combined.
2. The I.Q. Distribution of the Total Fourth Year Sample in all Five Schools Combined.

The decision to apply the tests was based on the fact that the two distributions mentioned are used in the division of the I.Q. variable. The author considers that when a variable is being divided into three parts of equal size, and these parts are to be classified as upper, middle and lower divisions, it is desirable, although not absolutely necessary, that the variable be normally distributed. Obviously a score distribution of almost any form can be divided easily into three equally sized parts, in order of merit, by merely counting a third of the cases from the top and a third from the bottom, and in the case of grouped data interpolating into the appropriate category. However, if the distribution is not normal there is less satisfaction with the arbitrary division points, and there is less confidence in the placing of the individuals near the borderlines. Some divisions may not seem to be so greatly different, and there is less confidence in readily classifying the divisions as upper, middle and lower.

For instance, when a distribution has a large, negative, skew, the higher division point in a three way division will probably fall into a category having a large proportion of the frequencies. That is it will be at the highest point of the curve. Many cases will be "bunched" around the division point, and there can be little confidence that a fair number of the cases on one side of the division point will be so very different from a fair number on the other side.

With a distribution which has a very large proportion of its cases in the category known as the mode, it is even possible that the two division points could both be in this category. If this was the case one could not feel so confident about the value of dividing such a distribution into three parts of equal size, and classifying them as upper, middle and lower. Of course, the point of absurdity is reached when a distribution is such that almost every case obtains the same score. A three-way division would then be a nonsense.

With a normal distribution no division point is at the highest point of the curve, and the three parts can be more obviously and clearly classified as upper, middle and lower divisions. The three sketches below illustrate the points being made.



Although the normal distribution curve most readily lends itself to a confident three-way division, and although some abnormal curves are not suitable for this purpose, other curves that depart from normality may be divided with a fair degree of confidence.

This could apply to some flat curves. That is curves that are even and uniform but with little height as compared to the curve of a normal distribution. Such distributions are evenly spread over the normal range but with a less than normal grouping around the mean. It could also apply to tri-modal curves if such were ever found.

There is only great lack of confidence in a division when many cases of similar ability are "bunched" around a division point, such as in the case of a very highly skewed distribution.

Using the Chi-Square Test of a Normal Distribution the results for the two school year groups were as follows.

1. For the Total First Year Sample of all Five Schools Combined.

The obtained chi-square value was 12.65.

With 3 degrees of freedom, after regrouping, this is just significant at the one per cent level.

2. For the Total Fourth Year Sample of all Five Schools Combined.

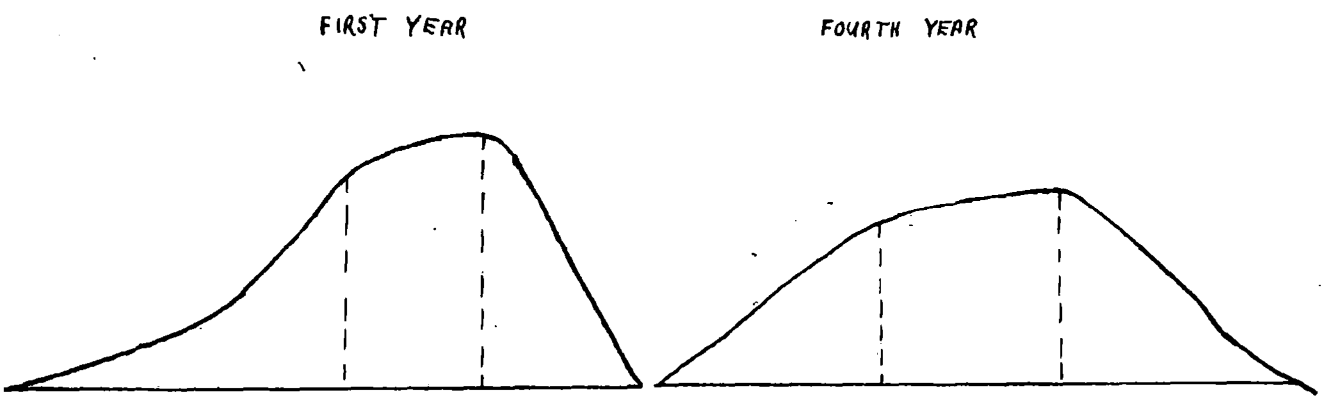
The obtained chi-square value was 10.73.

With 4 degrees of freedom, after regrouping, this is significant at the five per cent level.

(For further detail see Appendix A.)

These results indicate that both distributions are not completely normal. If they had been completely normal, low, insignificant chi-square values would have been obtained. However, the chi-square test is a very stringent one and small deviations from normality can produce a significant result. Actually in the case of the fourth year distribution the deviation of one class interval was sufficient to produce a significant chi-square value.

It can be seen from inspection that the curves of both school year distributions are not grossly abnormal or distorted. There is no excessive "bunching" at the division points. The curves are as follows.



Most important in the division of the I.Q. variable is the fact that standard scores have been used. As this is the case, and as the distributions are not grossly abnormal, it can be assumed that a reasonable division of the variable has been made.

As the standard scores are based on the total year distributions there should be reasonably even numbers in each division of upper, middle and lower intelligence, for the total year groupings, but some of the schools with skewed distributions will have more cases in an end category. For instance schools A and B will probably have more cases in the + 0.44 s.s. category. However, this does not mean that the cases in that end category, or in any other category, will not be evenly distributed between the sub-age groups. Unless there is some unknown birth-intelligence factor operating, there should be as many May to August births as September to December births in the + 0.44 s.s. category. This should apply to each intelligence division or category.

Although intelligence is not perfectly normally distributed in the study sample this does not mean that, at all levels, it will not be evenly distributed throughout the sub-age groups. Approximately the same numbers should be found for September to December births, January to April births, and May to August births at any of the three general levels of intelligence. The chi-square test of the relationship between I.Q. level and month of birth, examined in a following section, should be able to assess how far this is true. If it is not true it will indicate that there is some special influence operating, such as some birth months showing a consistent superiority in intelligence.

B.

The Investigation of the Relationship between I.Q. and Stream Placement

The aim of this investigation was to establish that, in general, as a group, the children in the sample were streamed. For all statistical tests the data pertaining to the two variables was arranged in three by three contingency tables. Chi-square values were computed. High, significant, chi-square values would indicate that there was a relationship between I.Q. and stream placement, and that, without doubt, the children had been streamed.

To reach significance at the five per cent level of confidence a chi-square value of 9.49 is necessary when four degrees of freedom apply. To reach significance at the one per cent level of confidence a chi-square value of 13.28 is necessary when four degrees of freedom apply.

With First Year Children.

For the Total First Year Sample in all Five Schools Combined.

The chi-square value was 140.8 This is very highly significant.

For School A

The chi-square value was 53.6 This is very highly significant.

For School B

The chi-square value was 33.6 This is highly significant.

For School C

The chi-square value was 30.6 This is highly significant.

For School D

The chi-square value was 18.06 This is still well beyond the one per cent level.

For School E

The chi-square value was 48.6 This is very highly significant.

With Fourth Year Children.

For the Total Fourth Year Sample in all Five Schools Combined.

The chi-square value was 385.5   This is very highly significant.

For School A.

The chi-square value was 100.3   This is very highly significant.

For School B.

The chi-square value was 121.2   This is very highly significant.

For School C.

The chi-square value was 90.4   This is very highly significant.

For School D.

The chi-square value was 46.2   This is very highly significant.

For School E.

The chi-square value was 93.7   This is very highly significant.

For the Complete Sample of First Year and Fourth Year Children Combined.

The chi-square value was 526.3   Again this is very highly significant.

(For further detail see Appendix B.)

Obviously there was no need to proceed with Maxwell's test here, and the above results confirm beyond any doubt that, in both school years, in all schools, the children, in general, as a group, were being streamed.

C. THE INVESTIGATION OF THE RELATIONSHIP BETWEEN LEVEL OF I.Q. AND MONTH OF BIRTH (and thus sub-age group)

The aim of this investigation was to establish that intelligence is evenly distributed throughout the sub-age groups of the children in the sample, and to ascertain that no birth months have a significant advantage with respect to intelligence. For the statistical testing the data was arranged in three by three contingency tables. Chi-square values were computed. As with the preceding work chi-square values of 9.49 and 13.28 were necessary if the five per cent and one per cent levels of confidence were to be reached.

Of course here the hypothesis is that no significant differences will be found and no significant chi-square values will be obtained.

With First Year Children.

For the Total First Year Sample in all Five Schools Combined.

The chi-square value was 2.78. This is not significant.

For School A.

The chi-square value was 0.833. This is not significant.

For School B.

The chi-square value was 3.427. This is not significant.

For School C.

The chi-square value was [illegible].143. This is not significant.

For School D.

The chi-square value was 7.213. This is not significant.

For School E.

The chi-square value was 1.015. This is not significant.

With Fourth Year Children.

For the Total Fourth Year Sample in all Five Schools Combined.

The chi-square value was 7.117. This is not significant.

For School A.

The chi-square value was 8.779. This is not quite significant at the five per cent level of confidence.

For School B.

The chi-square value was 9.658. This is just significant at the five per cent level of confidence.

For School C.

The chi-square value was 3.707. This is not significant.

For School D.

The chi-square value was 2.957. This is not significant.

For School E.

The chi-square value was 2.635. This is not significant.

For the Complete Sample of First Year and Fourth Year Children Combined.

The chi-square value was 2.18. This is not significant.

(For further details see Appendix C)

No significant chi-square values were found for the tests done with the first year children. With respect to the fourth year children one of the six chi-square tests, that for School A, produced a value which was just short of significance at the five per cent level of confidence. Another one of the six fourth year tests, that for School B, produced a value that did just reach this level of significance. The other four tests did not approach a level of significance.

A Maxwell trend analysis was applied to the fourth year contingency tables of Schools A and B. This did clarify matters to some extent as it did show that there were not significant linear trends in the contingency tables. This applied to both School A and School B and the respective scores were 0.12 and 1.47. Of course a linear trend was not really expected as this would entail a gradual, ordered, improvement or deterioration of I.Q., correlating with an ordered succession of birth months from September to August. One definite deviation from expectation by one sub-age group could produce a significant overall chi-square for the contingency table, and such a deviation might suggest that a particular sub-age group had an advantage or a disadvantage with respect to intelligence.

The contingency table of School A (Fourth Year) showed a different I.Q. distribution for the sub-age group whose birth months were January to April; a distribution different to those of the other two sub-age groups represented in the same contingency table. The distribution for this sub-age group was near normal, whereas for the other two sub-age groups, the distributions were negatively skewed. The mean I.Q. was also significantly lower.

It is rather interesting to note that with School B (Fourth Year), the chi-square value reached significance at the five per cent level of confidence mainly because of the contribution of the same sub-age group, namely January to April births. However, on this occasion the deviation was in the opposite direction. The distribution of I.Q. for the sub-age group in this contingency table was more negatively skewed than the distribution obtained for the other two sub-age groups. This time the mean I.Q. for that sub-age group was significantly higher than the means of the other two sub-age groups.

School A

		Month of Birth		
		Sept. to Dec.	Jan. to April	May to Aug.
Level	+0.44 s.s.	22	10	20
of	+0.44 to -0.44	7	15	10
I.Q.	-0.44 s.s.	6	9	6

School B

		Month of Birth		
		Sept. to Dec.	Jan. to April	May to Aug.
Level	+0.44 s.s.	14	20	16
of	+0.44 to -0.44	11	9	15
I.Q.	-0.44 s.s.	12	2	6

Clearly the evidence we have collected here indicates that intelligence seems to be evenly distributed throughout the sub-age groups. It supports the hypothesis that no sub-age group seems to have any advantage with respect to intelligence. No linear trends were noted in the data, and the two relatively high chi-square values, near to the point of significance at the five per cent level of confidence, tended to cancel each other out as regards postulating a superior or inferior sub-age group.

In any event as thirty-nine chi-square tests are being done in this study one would expect two of them to reach the five per cent level of confidence by chance alone. Thus a significant result such as that obtained for School B (fourth year) should not be totally unexpected.

Finally it should be observed that the most reliable chi-square test will be that in which the cell frequencies are at their highest. That is when the total data on the two variables being examined is grouped and arranged for testing. It is worth noting that when this was actually the case the chi-square value was low and nowhere near significance (i.e. 2.18).

D. THE INVESTIGATION OF THE RELATIONSHIP BETWEEN MONTH OF BIRTH (AND THUS AGE) AND STREAM PLACEMENT

This part of the investigation is the most crucial of all, as it is directly concerned with the main object of the study. It will reveal whether children from the different sub-age groups within a school year group are evenly and randomly distributed throughout the streams, or whether they are unevenly distributed, that is not distributed at random. The actual facts of the stream allocation will be revealed regardless as to whether these facts can be justified or not.

If there is a random allocation to streams, and if no sub-age group is at an advantage or a disadvanta e as to stream placement, then the contingency table should produce low, non-significant, chi-square values. If significant values are found it will suggest that the children born in the differing sub-age groups are not evenly distributed throughout the streams and that some bias is operating. If so some sub-age group, or groups, will be at an advantage with regard to stream placement, and some other, or others, will be at a disadvantage. Any trends in the contingency tables will be noted. The significant points are again 9.49 and 13.23.

With First Year Children.

For the Total First Year Sample in all Five Schools Combined.

The chi-square value was 37.39. This is highly significant.

For School A.

The chi-square value was 10.6. This is significant at the five per cent level.

For School B.

The chi-square value was 18.2. This is highly significant, being beyond the one per cent level of significance.

For School C.

The chi-square value was 10.77. This is significant at the five per cent level but not the one per cent level.

For School D.

The chi-square value was 14.37. This is significant beyond the one per cent level.

For School E.

The chi-square value was 2.14. This is clearly not significant.

With Fourth Year Children.

For the Total Fourth Year Sample in all Five Schools Combined.

The chi-square value was 4.38. This is not significant at the five per cent level of confidence, and it does not even approach it closely.

For School A.

The chi-square value was 9.06. This is almost significant at the five per cent level.

For School B.

The chi-square value was 4.98. This is not significant.

For School C.

The chi-square value was 4.57. This is not significant.

For School D.

The chi-square value was 7.65. This is not significant.

For School E.

The chi-square value was 1.34. This is nowhere near significance.

With the exception of one school, School E, the first year children from the different sub-age groups do not seem to be randomly spread throughout the streams. There appears to be some bias operating, with one of the sub-age groups being at an advantage with respect to stream placement and one being at a disadvantage. The degree of bias is quite appreciable as two schools produce chi-square values significant at the five per cent level of confidence (when the relationship between Month of Birth and Stream Placement is investigated) and another two schools produce values significant at the one per cent level of confidence. The chi-square value for the whole first year group is significant at the one per cent level of confidence. An inspection of the contingency tables having significant chi-square values shows that the trend or bias is always in the same direction. The first sub-age group tends to have greater representation in the A streams than would be expected by chance. This first sub-age group consists of children born in the months September to December inclusive, that is the oldest children in the sample. Conversely the third sub-age group consisting of the youngest children in the sample, those born from May to August inclusive, tended to have undue representation in the C streams and relatively poor representation in the A streams.

The contingency table for all the first year children combined illustrates the above mentioned position of trends quite well, and the distribution of frequencies is typical of the results noted for first year children in all the schools except School E.

Total First Year Group Birth Month v Stream Placement

		Birth Month		
		Sept. - Dec.	Jan. - April	May - Aug.
Stream Placement	A	92	62	44
	B	48	60	71
	C	25	40	62
The Resulting Chi-Square Values				
		11.42	0.04	9.41
		1.92	0.10	1.07
		6.62	0.01	6.78

The overall chi-square value was 37.4.

The trend, with a tendency for older children to be placed in higher streams and younger children to be placed in lower streams, can easily be seen if the two end distributions are examined i.e. the distributions of stream placement for those born September to December and those born May to August. Inspection of the frequencies in the four corner cells of the contingency table, and noting of the contributions of these cells to the overall chi-square value, further demonstrates the trend (see the table above).

The trend can be seen so clearly by inspection that a Maxwell Chi-Square Trend Analysis hardly seems necessary. Nevertheless this was done as a statistical measure. The outcome was that the trend described was highly significant. The results for the contingency table shown, that is the one relating to the total sample of first year children from all five schools combined, were as follows:

	Degrees of Freedom	Chi-Square Value
Due to Linear Regression	1	35.2
Due to Departure from Linear Regression	3	2.2
Overall Value	4	37.4

As the distribution of cell frequencies in the contingency table dealing with the total first year sample was so similar to the distributions found in the contingency tables dealing with schools A, B and D, the Maxwell analysis would be similarly significant for these individual schools. As the obtained Maxwell result was so highly significant it was not deemed necessary to repeat such a similar computation for each individual school. The overall chi-square values are more than sufficiently significant and the trend directions are obvious.

The contingency table for the first year children in School C produced an overall chi-square value of 10.76, which was significant at the five per cent level, but the degree and direction of trend was not quite so obvious at the level of simple inspection. A Maxwell test was therefore carried out and it had positive results. The portion of the overall chi-square value due to linear regression amounted to 8.288. Allowing one degree of freedom for regression this was significant at the one per cent level. Confirmation is given that the same trend found operating in schools A, B and D is also found operating in School C.

Although the contingency table dealing with the first year children in School E did not produce a significant overall chi-square value, it is possible that there could still be a trend within the table which could prove to be significant. Simple inspection of the distribution of frequencies did not suggest that this would be so, and a statistical check showed that a Maxwell test could not possibly be significant.

In a three by three contingency table there are four degrees of freedom appropriate to the overall table, and the significance of the obtained chi-square value is in relation to these degrees of freedom. However, in assessing a trend within a table the Maxwell test estimates the portion of the overall chi-square value that is due to linear regression. In assessing the significance of this portion of the chi-square value only one degree of freedom is appropriate. For one degree of freedom a chi-square value of 3.84 is needed for significance to be reached at the five per cent level of confidence.

As the chi-square value for the whole table was only 2.14 it is obvious that even if almost all of that amount was due to linear regression it would not be significant. Incidentally the computed chi-square value due to linear regression in this table was only 1.49. This is clearly not significant.

In summarising the position with the first year children it can be said that in schools A, B, C and D, there is a definite tendency to place the older children in the upper streams and the younger children in the lower streams. Only in School E does there seem to be unbiased allocation to streams, with the younger children having the same opportunity as the older children of being placed in a higher stream.

With the fourth year children the position of stream placement seems to be different to some extent. Children from the different sub-age groups seem to be more evenly spread between the streams. The older children do not seem to be at such an advantage and the younger children do not seem to be at such a disadvantage.

The contingency table investigating the relationship between month of birth and stream placement for the total fourth year sample produced an overall chi-square value of 4.38. This is not significant and it does not even approach the five per cent level. Individual schools B and C obtained similar non-significant values of 4.98 and 4.57 respectively. School E produced the particularly low and non-significant value of 1.34. The value for School D was 7.65 but this is still well short of significance at the five per cent level. Only School A produced a chi-square value approaching significance. The value was 9.06 and this was just short of significance at the five per cent level.

Inspection of the contingency tables for the Total Fourth Year Sample, School B and School C did not indicate possible significant linear trends, except perhaps in the case of School C. Nevertheless, a statistical check with the Maxwell test was carried out for all three tables. As expected no significant trends were found for the Total Fourth Year Sample and for School B, and the computed chi-square values due to linear regression were as follows:

For the Total Fourth Year Sample 3.1. Not quite significant at the five per cent level.

For School B, 0.22. Nowhere near significance.

However, School C obtained a value of 3.87 which was just significant at the five per cent level, and the direction of the trend indicated that the older fourth year children in this school do have a better chance of being placed in a higher stream than do the younger children. There was obviously little point in attempting a Maxwell test with School E as the overall chi-square value for that contingency table was nowhere near significance for even one degree of freedom. Even if almost all the obtained value was due to linear regression it still would not be significant.

Schools A and D both had overall chi-square values large enough to make a Maxwell test a necessary check, although even with these two schools inspection of the contingency tables did not readily indicate linear trends.

For School A a score of 2.059 was obtained with the Maxwell test. This does not reach the five per cent level for one degree of freedom, and it can be concluded that there is no significant linear trend in this contingency table.

School D obtained a score of 0.19 with respect to linear regression and this is nowhere near significance.

Thus with the fourth year sample the position is found to be somewhat different to that found with the first year sample. With the fourth year sample hardly any significant relationship was found between month of birth and stream placement. The contingency tables investigating the relationship between these two variables were almost all non-significant. None of the overall chi-square values, from the contingency tables dealing with the total fourth year sample and all the individual schools, were found to be significant. The use of the Maxwell test only led to the discovery of one significant linear trend. That was with School C and the obtained value only just attained significance at the five per cent level.

The following table shows how the first year sample and the fourth year sample compare with regard to the relationship between month of birth and stream placement. Chi-square values from the respective contingency tables are given. For four degrees of freedom values of significance at the five per cent level and the one per cent level are 9.49 and 13.28 respectively.

	1st Year Sample		4th Year Sample	
School A	10.6	S (5)°	9.06	N.S.
" B	18.2	S (1)°	4.98	N.S.
" C	10.77	S (5)°	4.57	N.S.°
" D	14.37	S (1)°	7.65	N.S.
" E	2.14	N.S.	1.34	N.S.
Total Year Group	37.39	S (1)°	4.38	N.S.

S (5) Significant at the five per cent level

S (1) " " " one " " "

N.S. Not significant. Does not reach the five per cent level

° Shows a significant linear trend

The table readily illustrates the following points:

1. All the fourth year chi-square values are much lower than their counterparts of the first year. This is even true for School E which obtains non-significant values for both the first year sample and the fourth year sample. This suggests a correction to some extent of the placement bias found with the first year sample.

2. School E is the only school completely unaffected by a bias in the placement of children from different sub-age groups to the various streams.

3. School C was the only school which showed significant bias of placement with both the first year children and the fourth year children, although the significant trend noted in the fourth year only just reached significance at the five per cent level.

Further contrast between the first year children and the fourth year children in terms of stream placement will be dealt with in the next section.

(For further details relating to this section, see Appendix D)

E. THE INVESTIGATION OF CHANGE FROM FIRST YEAR TO FOURTH YEAR, WITH RESPECT TO THE DISTRIBUTION OF YOUNG CHILDREN THROUGHOUT THE STREAMS

It has already been noted that in stream placement at the first year level the different sub-age groups are not randomly spread, and a bias seems to operate against the youngest children. At the fourth year level this bias is not so obvious and it is not statistically significant. In this section it is aimed to measure differences between the first year and fourth year groups as to this bias. It was originally hypothesised that there would be no difference, and that the proportions of young children in the various streams would be the same for both school years. It was particularly thought that the proportion of young children in the lower streams would be the same in fourth year as in first year.

The evidence of the last section suggests that the hypothesis will not be substantiated, but two measures will be adopted to assess whether, with the youngest children, the proportions in the different streams are the same for the two school years.

1. For the youngest sub-age group (May to August birthdays), the relationship between stream placement and school year (1st year or 4th year), the contingency table was as follows:

		Stream A	Stream B	Stream C		
Year Group	First Year	44	71	62	-	177
	Fourth Year	54	63	42	-	159
		98	134	104	-	336

The overall chi-square value for the table was 4.389. This is not quite significant at the five per cent level of confidence.

The result suggests that the first year distribution of young children throughout the streams is not statistically different from the fourth year distribution of young children. However, closer inspection of the table shows that, although the proportions in the three first year cells are not so widely different from the proportions in the three fourth year cells in terms of plus or minus, the distributions for the two school years are quite different. There is a reversal of trend from first year to fourth year. Of the children under consideration, that is the youngest sub-age group, more were placed in C streams than in A streams at the first year level. At the fourth year level more were placed in A streams than in C streams. This revesal of trend is quite definite, and it is certainly possible that the table could contain a significant chi-square value due to linear regression.

Accordingly, a Maxwell Chi-Square Trend Analysis was done with respect to the table and a significant result was obtained with respect to linear regression.

The sub-division of the chi-square value for the table was as follows:

	Degrees of Freedom	Chi-Square Value	
Due to Linear Regression	1	4.376	S.
Due to departure from Linear Regression	1	0.013	N.S.
Overall Value	2	4.389	N.S.

This shows that the linear regression in the table is significant and this means that the first year and fourth year distributions can be regarded as different. Undoubtedly a change seems to have occurred by fourth year with respect to the distribution of younger children throughout the various streams.

2. A second measure was adopted to assess any possible changes in stream allocation from first year to fourth year with respect to the youngest sub-age group.

This measure was a short series of tests of proportions.

a. The first test of proportions was between the first year youngest sub-age group children in stream A and the fourth year youngest sub-age group children in stream A.

As the children in the fourth year sample are not the same children who appear in the first year sample the appropriate test will be a test of difference between uncorrelated proportions. The formula suggested by Fisher (1950) was used for this test. The formula is as follows:

$$\bar{Z} = \frac{p_1 - p_2}{\sqrt{\bar{p}\,\bar{q}\left(\frac{N_1 + N_2}{N_1 N_2}\right)}} \quad \text{where } \bar{p} = \frac{N_1 p_1 + N_2 p_2}{N_1 + N_2}$$

$$\text{and where } \bar{q} = 1 - \bar{p}$$

The test gives a score which can be interpreted in terms of the standard measurement $\bar{Z}$, appropriate to large normal samples.

The result of the first test was a $\bar{Z}$ of 1.90. This is just about significant at the five per cent level of confidence, the required value being 1.96. This suggests that the proportion of young children placed in the A streams differs between first year and fourth year. The proportion of young children in A streams in first year is 0.2486 and the proportion of young children in A steams in fourth year is 0.3396. The difference between these proportions is just about significant at the five per cent level.

b. The same test was applied to the proportion of youngest sub-age group children placed in B streams in first year as against the proportion of them placed in B streams in fourth year.

For this test a $\bar{Z}$ value of 0.09 was obtained. This is obviously not at all significant. The two proportions for young children in the B streams of first year and fourth year were 0.401 and 0.396 respectively. These are much the same, certainly not significantly different.

Of course, if a bias is operating at either first year level or fourth year level, or both, it will not be apparent by the examination of the proportions in the B streams, and contrasting the first year and fourth year with respect to the proportions of young children in the B streams will illustrate little. If a sub-group is being undervalued or overvalued in some way, and proportions are the measurement criteria, it will be the proportions found at the ends of the distribution that will provide the important evidence.

Thus in contrasting the two school year groups with regard to the proportions of young children in the various streams, it will be to the A and C streams that we will look to obtain our main evidence.

c. The same statistical test was applied to the proportion of young children placed in C streams in first year as against the proportion of young children placed in C streams in fourth year.

For this test a $\bar{Z}$ value of 1.39 was obtained. This does not reach the five per cent level of confidence and it could be said that the proportions of young children in the C streams of first year and fourth year are not significantly different. The actual proportions were 0.35 and 0.283 for first year and fourth year respectively. However, the probability of such a difference between proportions (one direction or the other) being obtained by chance is only about eighteen in a hundred, and the probability that the difference will operate in one expected or predicted direction to such an extent is only nine in a hundred.

The difference in proportions is only just about significant between the first year A streams and the fourth year A streams, (i.e. the proportion of younger children in these streams), and the difference in proportions between the first year C streams and the fourth year C streams does not reach significance. However, one vital point must not be overlooked. That is the fact that a contrasting study revealed not only differences in proportions from one year to another, but also a change as to which had the larger and smaller proportions. The two school years were contrasted and the following points were noted.
At first year level the A streams contained the smaller proportion of young children, but at fourth year level the A streams contained the larger proportion of young children.

With the C stream contrast the converse was noted.
At first year level the C streams contained the larger proportion of young children, but at fourth year level the C streams contained the smaller proportion of young children.

For the youngest sub-age group the frequencies pertaining to school year and stream placement were as follows:

Stream	A	B	C	
1st Year	44	71	62	177
4th Year	54	63	42	159

It can be seen that at first year level 44 children from the sub-group of 177 were found to be in A streams as compared to 62 out of 177 in C streams. At fourth year level 54 out of 159 were found in A streams as compared with 42 out of 159 in C streams.

Thus although the differences in the proportions may not be so significant from one school year to another, the fact that there is a reversal of the smaller-larger proportions is most important. Also when only the children (from the youngest sub-age group) placed in either an A or a C stream were considered as the basic group for comparison, significant proportional differences were found between first year and fourth year. At first year level 0.4152 of this group were in A streams and, of course, 0.5848 were in C streams. At fourth year level 0.5624 were in A streams and, of course, 0.4376 were in C streams. A test of the difference between proportions in A streams at first year level and fourth year level produced a $\bar{\bar{Z}}$ of 2.09. This is significant at the five per cent level of confidence. Naturally a similar reciprocal result was obtained with respect to this in C streams.

The evidence above of the noted differences and changes in proportions, and the evidence from the Maxwell trend analysis done for the youngest sub-age group with respect to stream placement versus school year, seems to be sufficient to warrant the acceptance of the view that there have been changes in stream placement from first year to fourth year. Change has occurred with respect to the distribution of younger children throughout the various streams.

(For fuller statistical data see Appendix B)

F. DIFFERENCES BETWEEN THE SCHOOLS

In contrasting the different schools and noting differences between them the following observations were made:

1. The mean I.Q.s. of the two schools whose pupils were mainly considered to be of high socioeconomic background (Schools A and B) were significantly higher than the means of the three other schools. The pupils of the three other schools were classified as being mainly of average (School C), and below average (Schools D and E), socioeconomic background. This result was more or less expected but the amount of difference was perhaps a little surprising. On both tests, M.H.T. (Pic I) and M.H. 77, the difference was exceptionally high to the order of 10 I.Q. points. It was interesting to note that the school considered to be of average socioeconomic background, School C, scored no better than the two schools considered to be of below average socioeconomic background. This school was also thought to be the most homogeneous with respect to background but, in fact, it had much the same standard deviation of I.Q. as the two schools of high socioeconomic background. There was a slight tendency for the schools of low socioeconomic background to obtain slightly larger standard deviations of I.Q. than those obtained by the schools of higher socioeconomic background.
2. With respect to streaming, the allocation of children to three groups in terms of ability, it is clear that all schools were actually operating this system. The results of the investigation which examined the relationship between I.Q. and stream placement showed that the children, in general, as a group, were being streamed. There were no differences between the schools with respect to this.

3. Intelligence seemed to be evenly distributed throughout the sub-age groups and no particular sub-age group was at an advantage with regard to intelligence. This was the case for all schools and there were no differences between the schools in this respect.

4. When the relationship between month of birth and stream placement was investigated some differences between the schools were revealed. The general tendency was for the children from the different sub-age groups to be unevenly distributed throughout the streams at first year level, with the younger children being disproportionately allocated to the lower streams. This tendency was clear and definite for Schools A, B, C and D, but it was not present in School E. It would appear that in School E greater skill had been applied to the allocation procedure at the original streaming. Possibly the head teacher of the infant department, or the head teacher of the junior department, or both, had been more sophisticated in approach and more thorough in the application of streaming procedures. Certainly some consideration seems to have been given to age and the result was that the younger children were more evenly distributed throughout the streams.

By fourth year the general tendency was for the bias operating against the younger children to be rectified to some extent. Generally the children from the different sub-age groups were more evenly distributed in fourth year. This was so with Schools A, B and D. With School E the even distribution obtained at the original streaming continued. After an investigation by the Maxwell test it was ascertained that School C still showed a definite trend in fourth year similar to that found in first year. That is that the distribution of the different sub-age groups throughout the streams was not even, and that older children tended to be more readily placed in higher streams, with younger children being more readily placed in lower streams.

Thus we might conclude that the differences between the schools with regard to stream placement were as follows. School E distributed the children from the different sub-age groups evenly throughout the streams for both the first year and the fourth year. No bias for or against older or younger children seemed to operate at all. Schools A, B and D did not show an even distribution at the first year level but they did seem to have relatively even distributions by fourth year. School C seemed to operate a biased allocation at first year level and to maintain this bias into the fourth year. Surprisingly this was the school of the headmaster who was most receptive to the idea of non-streaming.

An implied difference between the schools might be that Schools A, B, D and also E, managed to maintain a genuine mobility between streams, but that School C did not manage this to a sufficient degree. Mobility between streams would allow the bright younger children in the lower streams to move to higher streams as they matured and drew level with their older colleagues. It would also allow dull older children to revert to lower streams when they were overtaken.

Although mobility between streams is an implied difference between schools it is the most reasonable explanation to account for the evidence. Mobility between the streams could account for the distribution changes noted at the fourth year level.

5. With regard to the placement of children from the youngest sub-age groups in C streams the following proportions for the different schools were noted:

At first year level School A 0.42

" B 0.41

" C 0.24

" D 0.39

" E 0.30

The proportions for Schools C and E are lower than the proportions for Schools A, B and D, but in the case of School C this was merely because the C class was very small, fourteen pupils in all. This was a remedial group of fourteen combined with a second year remedial group of fourteen, making a junior remedial class of twenty-eight. This meant that only a small proportion of the youngest sub-age group could be placed in that form.

At fourth year level School A 0.20

" B 0.29

" C 0.38

" D 0.11

" E 0.27

The proportion for School C is significantly higher than the proportions for the other schools. This is to be expected when it is remembered that School C was the only school found to maintain a bias against the youngest children at the fourth year level.

It can also be seen that Schools A, B, D and E all reduced the proportion of young children in the C stream from year one to year four, but School C actually increased its proportion from first year to fourth year.

School D has a lower proportion than any other school at fourth year level, but this is not too reliable a result as the total frequencies for that school with respect to the youngest sub-age group in fourth year only come to nineteen.

(for fuller data see Appendix F)

CONCLUSIONS

The most reliable conclusions in a social science study are those based on hypotheses agreed on at the outset. Investigations are organised in such a way that the main questions to be asked are posed in the form of null hypotheses, which the subsequent evidence will either accept or reject. When at all possible the acceptance or rejection will depend on a statistical evaluation being made in accordance with generally accepted standards and limits of confidence.

Accordingly for the main conclusions of this study one must turn to the null hypotheses postulated before the manipulation of any data.

At the expense of perhaps giving the impression of repetition in the write up of this study it was decided that it would be to the readers' advantage to have the null hypotheses re-stated as they are discussed.

The first null hypothesis was as follows.

That, in general, as a group, the first year junior school children in the sample were NOT streamed. That is the allocation to class groups was random, and not according to ability.

The statistical evidence opposing this statement was so overwhelming that the null hypothesis must be rejected. High chi-square values were obtained for all schools on the relationship between I.Q. and stream placement, and the trend in all contingency tables was for bright children to be placed in A streams and dull children to be placed in C streams.

The second null hypothesis was similar to the first but this time it applied to the fourth year sample instead of the first year sample.

Again the evidence opposing the hypothesis was overwhelming and it must be rejected. Exceptionally high chi-square values were obtained and the same trends were noted.

Without doubt both the children of the first year sample and the fourth year sample had, in general, been streamed.

The third null hypothesis was as follows.

That there is an even distribution of intelligence throughout the sub-age groups of first year junior school children in the sample, and that no birth months have a significant advantage with respect to intelligence.

The statistical evidence was overwhelming in support of the null hypothesis, and the hypothesis must be accepted without reservations. The contingency tables testing the relationship between the variables of I.Q. and month of birth gave low chi-square values, showing that there is no significant relationship between the two.

The fourth null hypothesis was similar to the third but on this occasion the children being referred to are fourth year rather than first year children.

On balance the statistical evidence gives support to the null hypothesis, certainly sufficiently for the hypothesis to be accepted rather than rejected.

There was a small amount of evidence from two schools that perhaps intelligence may not be evenly distributed throughout the sub-age groups or the months of birth. However, on closer inspection (see the discussion in the results section) the evidence from the two schools tended to cancel out. In any case the greatest part of the evidence relating to this fourth year investigation of I.Q. level and month of birth gave support to the null hypothesis, and it is because of this the said hypothesis is accepted.

The fifth null hypothesis was as follows.

That, incidentally, there does not seem to be any overall relationship between month of birth and intelligence. That is no birth months have a significant advantage with respect to intelligence.

This hypothesis was postulated as it could be incidental to the main study without introducing complications, and it could be a useful addition to the evidence accumulated on this subject. (i.e. evidence accumulated by other research workers in other studies.)

The evidence obtained again gives great support to the null hypothesis, and again it is accepted with confidence.

The sixth null hypothesis was as follows.

That there is no significant relationship between month of birth (and thus age) and stream placement in the first year junior school sample, and that there is no tendency for the older children to be placed in higher streams, and the younger children to be placed in lower streams.

The overall evidence is such that the null hypothesis must be rejected, but with some reservation. Four of the five schools did produce significant relationships between the two variables month of birth and stream placement and these four all exhibited the tendency to place the older children in the higher streams and the younger children in the lower streams. However, one school did not produce a significant relationship between the variables and did not show the tendency mentioned.

Thus although the null hypothesis is rejected, and although the indications are that biased stream placement related to age is general, it is noted that such bias is not unavoidable.

The seventh null hypothesis was similar to the sixth, only on this occasion the school year group being referred to is the fourth year not the first year.

The overall evidence concerned with the seventh null hypothesis is such that the hypothesis can be accepted.

In the fourth year the relationship between month of birth and stream placement is not significant and it only nearly approaches significance in one school. The tendency to place older children in the higher streams and younger children in the lower streams is not so obvious, and in only one school out of the five is there any evidence for this trend.

With the fourth year sample it would appear that the children from the different sub-age groups are randomly spread throughout the streams, and the older children no longer seem to be at an advantage as regards stream placement and the younger children no longer seem to be at a disadvantage.

The eighth null hypothesis was as follows.

That the distribution throughout the three streams, of children from the youngest sub-age group, will be similar for the two school year groups involved, that is for the children of the first year sample and the children of the fourth year sample.

The evidence pertaining to this null hypothesis was such that the hypothesis was rejected.

A change seems to have occurred by fourth year with respect to the distribution of younger children throughout the various streams. Detailed examination of a contingency table dealing with the distributions of the two school years showed that they were different. The table showed a significant trend. This trend actually indicated a reversal of the biased form of stream placement from first year to fourth year.

Consideration of the proportions of younger children from each year group in the respective streams A, B and C produced further evidence for the rejection of the null hypothesis. The proportions from one year to another remained much the same with stream B, but with streams A and C there were changes beyond expectation. At first year level more children were placed from the youngest sub-age group in C streams than in A streams, and this difference was quite appreciable. However, at fourth year level not only had the proportions evened out but it was actually found that more children from this group were placed in A streams than in C streams.

The overall evidence thus definitely rejects the eighth null hypothesis.

Having accepted or rejected the null hypotheses it is now possible to discuss the implications involved. Further conclusions may be obtained in a logical manner rather than in the empirical manner used in accepting or rejecting the null hypotheses.

From the study of the first year sample and the appropriate null hypotheses we have obtained the following main facts.

1. That children, in general, are being streamed, that is grouped according to ability.

2. That month of birth and level of I.Q. are not significantly related, and that no birth months have any advantage, or superiority, with respect to intelligence.

As points one and two above are established it follows that all sub-age groups should be randomly distributed throughout the streams. If this is not the case the streaming is inefficient and possibly unjust, with probably inadequate assessment of ability at the original streaming being at fault. If any sub-age group is disproportionately represented in any stream this will indicate bias.

Our third main finding from the study of the first year sample was as follows.

3. Children from the different sub-age groups are not randomly distributed throughout the streams, and a disproportionate number of children from the youngest sub-age group are placed in the C streams.

The conclusion from this must be that streaming, or grouping according to ability, is inefficient at the first year level, and that there is a bias operating in favour of the older children and against the younger children.

The questions now arise as to whether streaming remains inefficient and whether it retains its bias.

The answers to these queries, in terms of the present project, are to be found in the study of the fourth year sample. A study of the null hypotheses relating to the fourth year sample showed the following points to be true.

1. That children, in general, are being streamed in terms of ability.

2. That month of birth and I.Q. level are not significantly related, and that no birth months have any advantage with respect to intelligence.

Again, as the above points are established, it follows that all sub-age groups should be randomly distributed throughout the streams, and no sub-age group should be disproportionately represented in any stream. If this is not so it can be assumed that the streaming is still inefficient at the fourth year level, and that the bias is still operating.

However, the third main finding from the study of the fourth year sample was as follows.

3. Children from the different sub-age groups are randomly distributed throughout the streams, and that no sub-age group is disproportionately represented in any stream.

The conclusion from this must be that streaming, or grouping according to ability, is efficient at the fourth year level, and that much of the bias operating against the younger children has been eliminated.

A comparison study between the school year groups, and relating to the eighth null hypothesis, supports this last point.

Thus the main conclusions of the study must be that, although the streaming tends to be inefficient at the first year level, and although at this level the younger children seem to be at a disadvantage, by the time the fourth year level is reached the streaming is quite efficient, and the bias against the younger children seems to have disappeared.

The reason for the above noted change can only be implied, but the most reasonable assumption seems to be that there has been sufficient mobility between the streams in the intervening years. Bright younger children who were placed in streams below their appropriate potential will have had the opportunity to move up to higher streams, and, of course, some of the older children, who were perhaps over-estimated at first year level, will have moved to lower streams.

It should be pointed out that the study also gives sufficient evidence to show that inefficient streaming can continue right through to the fourth year of the primary school, and that the bias against younger children can persist. This was demonstrated for one of the five schools, and it should serve as a warning against rigid streaming with lack of movement from one stream to another.

One final conclusion from the study is that it seems possible to eliminate bias, and to stream efficiently at the outset, if sufficient careful measures are taken, albeit these measures be largely subjective. This was demonstrated in the case of one school in the sample, and the point was appreciated when differences between schools were noted.

Strictly speaking the conclusions derived from the study are only applicable to the study sample, and, with a high degree of confidence, to similar samples such as could be found in the industrial towns of North-East England.

However, as the study sample is large, and as it contains a fairly wide cross section of social strata, the conclusions are probably applicable to, and the results typical of, most industrial and urban areas where traditional streaming is operating. One must avoid generalising beyond the data, but some guarded general inferences should be drawn, and some possible general indications should be noted, from a study of this type.

A General Discussion

This study aimed to investigate the relationship of month of birth to stream placement in the primary school, and in particular it aimed to study the possibility that some of the younger children of a school year age group may be underestimated in a traditional streaming system.

Younger children may be underestimated in non-streamed schools if teachers do not make allowances for age differences. However, the underestimation is more formal and obvious if it happens in a streaming system. In a non-streamed school an underestimated child may not be fully aware of the estimation made about him and he can still set his level of aspiration at the higher standards of the group. The underestimated child who is placed in a C stream can hardly fail to note the estimation made of him, and he no longer has high standards in front of him. Thus it is underestimation in the streamed system in which we are interested. In particular the study is concerned with the underestimation of young children as it is probably they who are most at risk.

The study does not set out to oppose or support streaming. It is merely trying to ascertain whether month of birth is significantly connected with streaming, and if in particular the streaming tends to formalise any under-estimation of the younger children. It seeks to assess whether the mere accident of month of birth is a factor involved in determining streaming, and it aims to question the validity of this if established. Streaming as a concept or practice is not being opposed, but the danger of unjust underestimation because of age is underlined. One of the dangers of streaming a school year group is assessed as scientifically as possible. The study is successful in so far as it does show that the younger children are at risk to some extent, certainly at the first year junior school level. However, it is quite encouraging that the study does also show that, if streaming is not too rigid, and if there is sufficient mobility between streams, the younger children will not be underestimated and unfairly placed at the end of their primary school career.

As it is desired that the project should not be too unwieldy, only the one aspect of streaming is closely studied, that being the relationship of age or month of birth to streaming. However, this does not mean that other aspects of streaming are not considered to be important. Social background as a variable influencing stream placement is appreciated together with such factors as cultural support, position in the family, size of the family, previous attendance at a nursery school, absence from school during the critical infant period, and quality of teaching and stability of staffing in the infant school. These factors are all recognised as having an influence on the stream placement of any child. However, in this project little account could be taken of them and no experimental control could be implemented. It is submitted that this is not so important for the present study and that lack of control of these variables does not invalidate the study. It is reasonably assumed that the influence of these variables will be randomly spread through all age groups, and to a great extent through all school and class groups, although it has been recognised that the general socioeconomic backgrounds of the various schools will have some effect and perhaps help to produce inter-school differences.

The possible effects that streaming can produce are also appreciated, particularly social and psychological effects. These effects were discussed fully in an earlier section of this study, the section reviewing the literature on streaming. This study was not designed to measure these effects, but because this was not done the conclusion should not be reached that there is a lack of appreciation of their importance. The reason why such effects were not measured is that the finite limits of the study enforced a strict discipline of approach, with a concentration on the main aim of the study. However, it was noticed that one possible serious effect did not manifest itself in this study. This is the effect noticed by Douglas (1964) that there can be a differential effect on the I.Q. because of streaming. Placement in an A stream tends to assist the further development of intelligence with a corresponding rise in I.Q., whereas placement in a C stream tends to hinder development with a corresponding fall in measured scores. If this had operated to any great extent the disproportionate number of older children who were placed in A streams would have improved their I.Q. ratings, and the disproportionate number of younger children who were placed in C streams would have deteriorated in their I.Q. ratings. This would have had the overall effect of producing differences in the I.Q. distributions of the two groups, with a higher distribution of scores coming from the older children. This did not happen despite the fact that the verbally loaded test at fourth year would suit those who had received early verbal stimulation in the atmosphere of an A stream. At first year level the oldest sub-age group (September to December births) obtained a mean I.Q. of 112.1 as compared to the mean I.Q. of 111.8 for the youngest sub-age group (May to August births) (Test M.H.T. Pic. I.) At fourth year level the oldest sub-age group obtained a mean I.Q. of 104.1 as compared to the mean I.Q. of 104.3 for the youngest sub-age group. (Test M.H.77.) None of these differences are at all significant. No I.Q. difference could be demonstrated between the two sub-age groups although the oldest sub-age group had been exposed to relatively more A stream influence, at least during the early junior school stage, than had the youngest sub-age group. Also, of course, the lowest sub-age group had been exposed to more C stream influence than had the oldest sub-age group.

Perhaps the bad effects of streaming are not so disastrous as the opponents of the system would have us believe, but it is probably true to say that flexibility is necessary if streaming is to be implemented. This flexibility should be threefold.

1. In the general approach to the problem. Different forms of streaming should be investigated and assessed with respect to advantages and disadvantages. Different forms of streaming may suit different circumstances, and, with this in mind, controlled experimentation should be carried out. Perhaps even uncontrolled local experimentation may help to achieve a flexible approach to the whole question of streaming. Experimental work could be attempted with the idea of unstreaming for the first two years of the junior school career, streaming only being introduced in the third and fourth year.

There could also be experimental work on the Plowden (1967) suggestion that two parallel forms could be introduced to replace the present A and B streams but that the traditional C stream should continue as a slower learning group. This is something of a compromise attempting to "get the best of both worlds."

2. Flexibility in terms of allocation procedures. The allocation should not merely be based on an attainment level at a particular point in time. There should also be regard for ability, allowance for age, and consideration of previous opportunity to acquire certain academic standards. A more comprehensive and flexible approach is needed, taking account of all factors and all relevant information.

3. There should also be flexibility in terms of movement between the streams. The present study seems to give support to the idea that if mobility between streams is maintained some of the disadvantages of streaming can be offset.

These points should be considered at all administrative levels when future policy on streaming in the primary school is discussed.

New forms of primary school organisation are being tried out at the moment. For example the family grouping in some infant departments, where siblings of several years age difference are in the same class, and intraclass grouping in unstreamed primary schools. The intraclass grouping system has operated for a while in some schools in the U.S.A. Frandsen (1961) suggests that the sub-groups may be involved in different subjects or they may work at different levels of abstraction. The children in a group may have similar or complementary abilities. Occasionally each child may select whichever group he wishes to join, and it is thought that social as well as academic needs should be considered when groups are arranged.

These forms of organisation, together with such ideas as open plan schools with no rigidly defined classrooms, will undoubtedly have some influence on streaming in the future, but at present these systems are in the experimental stage.

Without doubt two barriers to the development of further unstreaming in the primary schools are the present size of primary school classes, and the short supply of high calibre teachers trained to deal with the extra intricacies of organisation and method necessary for success with the unstreamed group.

Perhaps the day will arrive when the teacher will have a relatively small class enabling more time to be spent with small groups and individuals. The time may also come when the teacher will be relieved of time wasting clerical activities, and when he or she will be aided to some extent by mechanical teachers in the form of programmed teaching machines. Unstreaming may then become more practicable and acceptable to all, but until then streaming procedures will still be widely used.

It is thus sensible to examine the dangers which can beset the present traditional system of streaming in the primary school. In this study an attempt has been made to underline one of those dangers. This was the danger that stream placement and month of birth can be significantly related, with older children within a school year group being at an advantage and younger children being at a disadvantage. It may be a source of surprise and relief to some that, although this danger was clearly demonstrated in this study, the suspicion that it would have a lasting effect received little support.

APPENDICES

APPENDIX A

The investigation of the I.Q. Variable

The Means and Standard Deviations

With First Year Children - test used being M.H.T. (Pic I).

The Total First Year Sample in all Five Schools Combined

I.Q.	Frequencies	
129.5 - 139.5	51	
119.5 - 129.5	118	
109.5 - 119.5	116	
99.5 - 109.5	107	
89.5 - 99.5	67	
79.5 - 89.5	22	
69.5 - 79.5	18	
59.5 - 69.5	5	Total = 504

The mean I.Q. = 110.849

The standard deviation = 15.69.

School A (High socioeconomic background)

I.Q.	Frequencies	
129.5 - 139.5	21	
119.5 - 129.5	29	
109.5 - 119.5	26	
99.5 - 109.5	15	
89.5 - 99.5	8	
79.5 - 89.5	4	
69.5 - 79.5	0	
59.5 - 69.5	0	Total = 103

The mean I.Q. = 117.219. The standard deviation = 13

School B (High socioeconomic background)

I.Q.	Frequencies
129.5 - 139.5	14
119.5 - 129.5	33
109.5 - 119.5	27
99.5 - 109.5	19
89.5 - 99.5	7
79.5 - 89.5	2
69.5 - 79.5	1
59.5 - 69.5	1 Total = 104

The mean I.Q. = 115.75. The standard deviation = 13.63

School C (Average socioeconomic background)

I.Q.	Frequencies
129.5 - 139.5	2
119.5 - 129.5	21
109.5 - 119.5	19
99.5 - 109.5	25
89.5 - 99.5	15
79.5 - 89.5	3
69.5 - 79.5	4
59.5 - 69.5	1 Total = 90

The mean I.Q. = 107.83. The standard deviation = 14.54

School D (Low socioeconomic background)

I.Q.	Frequencies
129.5 - 139.5	6
119.5 - 129.5	18
109.5 - 119.5	21
99.5 - 109.5	27
89.5 - 99.5	14
79.5 - 89.5	7
69.5 - 79.5	6
59.5 - 69.5	1 Total = 100

The mean I.Q. = 107. The standard deviation = 15.87

School E (Low socioeconomic background)

I.Q.	Frequencies
129.5 - 139.5	8
119.5 - 129.5	17
109.5 - 119.5	23
99.5 - 109.5	21
89.5 - 99.5	23
79.5 - 89.5	6
69.5 - 79.5	7
59.5 - 69.5	2 Total = 107

The mean I.Q. = 106.089. The standard deviation = 16.8

With Fourth Year Children - test used being M.H.77

The Total Fourth Year Sample in all Five Schools Combined

I.Q.	Frequencies	
139.5 - 149.5	8	
129.5 - 139.5	23	
119.5 - 129.5	71	
109.5 - 119.5	93	
99.5 - 109.5	92	
89.5 - 99.5	87	
79.5 - 89.5	65	
69.5 - 79.5	41	
59.5 - 69.5	10	Total = 490

The mean I.Q. = 103.601. The standard deviation = 17.98

School A (High socioeconomic background)

I.Q.	Frequencies	
139.5 - 149.5	3	
129.5 - 139.5	6	
119.5 - 129.5	23	
109.5 - 119.5	24	
99.5 - 109.5	22	
89.5 - 99.5	14	
79.5 - 89.5	9	
69.6 - 79.5	3	
59.5 - 69.5	1	Total = 105

The mean I.Q. = 109.7. The standard deviation = 16.39

School B (High socioeconomic background)

I.Q.	Frequencies	
139.5 - 149.5	3	
129.5 - 139.5	11	
119.5 - 129.5	17	
109.5 - 119.5	24	
99.5 - 109.5	25	
89.5 - 99.5	10	
79.5 - 89.5	8	
69.5 - 79.5	6	
59.5 - 69.5	1	Total 105

The mean I.Q. = 109.76. The standard deviation = 17.5

School C (Average socioeconomic background)

I.Q.	Frequencies	
139.5 - 149.5	1	
129.5 - 139.5	3	
119.5 - 129.5	7	
109.5 - 119.5	18	
99.5 - 109.5	18	
89.5 - 99.5	27	
79.5 - 89.5	20	
69.5 - 79.5	8	
59.5 - 69.5	4	Total = 106

The mean I.Q. = 98.65. The standard deviation = 16.7

School D (Low socioeconomic background)

I.Q.	Frequencies
139.5 - 149.5	0
129.5 - 139.5	1
119.5 - 129.5	5
109.5 - 119.5	15
99.5 - 109.5	16
89.5 - 99.5	18
79.5 - 89.5	9
69.5 - 79.5	6
59.5 - 69.5	0 Total = 70

The mean I.Q. = 100.7. The standard deviation = 14.26

School E (Low socioeconomic background)

I.Q.	Frequencies
139.5 - 149.5	1
129.5 - 139.5	2
119.5 - 129.5	19
109.5 - 119.5	12
99.5 - 109.5	11
89.5 - 99.5	18
79.5 - 89.5	19
69.5 - 79.5	18
59.5 - 69.5	4 Total = 104

The mean I.Q. = 98.15. The standard deviation = 19.6

It should be noted that both intelligence tests used in the study had conversion tables which had I.Q. score limits of 70 to 140. However, although the highest possible I.Q. score on the M.H.T. (Pic I) was 140, a score of 150 could be obtained by extrapolation on the M.H. 77. Both had similar lowest possible I.Q. scores of 60 by extrapolation. Accordingly, when the arbitrary categories of the distribution were arranged an extra category at the top end was made for the M.H. 77 (fourth year) distribution of scores and eight cases were found to score at 140+. Extrapolation from the conversion table could produce scores up to 150.

The M.H.T. (Pic.I) produced a "bunching" in the top category at first year level for schools A and B. This is the resuBt of the M.H.T. (Pic I) not having sufficient "ceiling", or opportunity for spread at the top for the oldest bright children in the study sample.

The Chi-square Tests for a Normal Distribution used the following formula.

$$\text{Chi-square} = \frac{(fo - fe)^2}{fe}$$

where fo = observed frequencies
fe = expected frequencies
and where (fo - fe) = the differences between fo and fe

The appropriate degrees of freedom are N intervals - 3.

The formula for obtaining the fe values is $fe = y\left(i\ \frac{N}{S.D.}\right)$

where N = the total number of frequencies in the distribution.

where S.D. = the standard deviation of the distribution.

where i = the class interval of the distribution.

and where y = the height of the ordinate at Z. (This is at the mid-point of each interval)

In turn Z is obtained from the formula $\frac{x}{S.D.}$

where x = X - M, X being the mid-point of each interval of the distribution in turn, M being the mean of the distribution, and X - M being the deviation of the interval mid-point from the mean.

The Chi-Square Test for a Normal Distribution at First Year Level.

M = 110.8 S.D. = 15.7 N = 504

	X	x (i.e. X–M)	Z $\left(\text{i.e.} \frac{x}{S.D.}\right)$	y	fe	fo
129.5 - 139.5	134.5	+23.7	+1.510	0.127	43	51
119.5 - 124.5	124.5	+13.7	+0.873	0.275	92	118
109.5 - 119.5	114.5	+ 3.7	+o.236	0.390	126	116
99.5 - 109.5	104.5	- 6.3	-0.401	0.370	120	107
89.5 - 99.5	94.5	-16.3	-1.039	0.240	78	67
79.5 - 89.5	84.5	-26.3	-1.675	0.100	33	22
69.5 - 79.5	74.5	-36.3	-2.312	0.028	10	18
59.5 - 69.5	64.5	-46.3	-2.951	0.006	2	5

Regrouped Frequencies for the Chi-Square Test

	fo	fe	(fo-fe)	$(fo-fe)^2$	Chi-Sq. values
129.5 - 139.5	51	43	8	64	1.4
119.5 - 129.5	118	92	26	676	7.5
109.5 - 119.5	116	126	10	100	0.8
99.5 - 109.5	107	120	13	169	1.4
89.5 - 99.5	67	78	11	121	1.5
79.5 - 89.5) 69.5 - 79.5) 59.5 - 69.5)	45	45	0	0	0.0
					Total 12.6

With 3 degrees of freedom this is just significant at the one per cent level. Thus the distribution is not completely normal.

The Chi-square Test for a Normal Distribution at Fourth Year Level

M = 103.6 S.D. = 17.98 N = 490

	X	x (i.e. X-M)	Z (i.e. $\frac{x}{S.D.}$)	y	fe	fo
139.5 - 149.5	144.5	+40.9	2.274	0.030	8	8
129.5 - 139.5	134.5	+30.9	1.718	0.091	25	23
119.5 - 129.5	124.5	+20.9	1.162	0.200	55	71
109.5 - 119.5	114.5	+10.9	0.606	0.332	93	93
99.5 - 109.5	104.5	+ 0.9	0.050	0.398	109	92
89.5 - 99.5	94.5	- 9.1	0.506	0.351	95	87
79.5 - 89.5	84.5	-19.1	1.062	0.225	64	65
69.5 - 79.5	74.5	-29.1	1.618	0.108	30	41
59.5 - 69.5	64.5	-39.1	2.174	0.038	11	10

Regrouped Frequencies for the Chi-Square Test.

	fo	fe	(fo-fe)	$(fo-fe)^2$	Chi-Sq. values
139.5 - 149.5 } 129.5 - 139.5 }	31	33	2	4	0.120
119.5 - 129.5	71	55	16	256	4.600
109.5 - 119.5	93	93	0	0	0.000
99.5 - 109.5	92	109	17	289	2.800
89.5 - 99.5	87	95	8	64	0.660
79.5 - 89.5	65	64	1	1	0.016
69.5 - 79.5 } 59.5 - 69.5 }	51	41	10	100	2.640
					Total = 10.736

With 4 degrees of freedom this is just significant at the five per cent level. Thus the distribution is not completely normal.

Appendix B

The Investigation of the Relationship between I.Q. and Stream Placement With First Year Children

The Frequencies in the contingency tables, and the resulting chi-square values, were as follows. (All tables have four degrees of freedom).

The Total First Year Sample in all Five Schools Combined

		Stream A	Stream B	Stream C
	+0.44 S.S.	115	57	8
I.Q.	+0.44 to -0.44	69	64	39
	-0.44 S.S.	14	58	80

Overall chi-square value = 140.8

With four degrees of freedom, significant well beyond one per cent level.

School A

		Stream A	Stream B	Stream C
	+0.44 S.S.	36	14	3
I.Q.	+0.44 to -0.44	5	19	8
	-0.44 S.S.	1	4	13

Overall chi-square value = 53.6

Highly significant; well beyond one per cent level.

School B

		Stream A	Stream B	Stream C
	+0.44 S.S.	24	20	4
I.Q.	+0.44 to -0.44	12	11	18
	-0.44 S.S.	0	1	14

Overall chi-square value = 33.59

Highly significant; well beyond one per cent level.

School C

		Stream A	Stream B	Stream C
	+0.44 S.S.	19	6	0
I.Q.	+0.44 to -0.44	18	10	2
	-0.44 S.S.	4	19	12

Overall chi-square value = 30.6

Highly significant; well beyond one per cent level.

School D

		Stream A	Stream B	Stream C
	+0.44 S.S.	17	9	1
I.Q.	+0.44 to -0.44	18	12	8
	-0.44 S.S.	7	12	16

Overall chi-square value = 18.06

Significant; beyond one per cent level.

School E

		Stream A	Stream B	Stream C
	+0.44 S.S.	19	8	0
I.Q.	+0.44 to -0.44	16	12	3
	-0.44 S.S.	2	22	25

Overall chi-square value = 48.6

Highly significant; beyond one per cent level.

With Fourth Year Children

Thefrequencies in the contingency tables, and the resulting chi-square values, were as follows. (All tables have four degrees of freedom).

The Total Fourth Year Sample in all Five Schools Combined

		Stream		
		A	B	C
	+0.44 S.S.	153	25	0
I.Q.	+0.44 to -0.44	41	89	14
	-0.44 S.S.	0	58	110

Overall chi-square value = 385.5

Very highly significant.

School A

		Stream		
		A	B	C
	+0.44 S.S.	39	13	0
I.Q.	+0.44 to -0.44	4	24	4
	-0.44 S.S.	0	3	18

Overall chi-square value = 100.3

Very highly significant.

School B

		Stream		
		A	B	C
	+0.44 S.S.	39	11	0
I.Q.	+0.44 to -0.44	1	27	7
	-0.44 S.S.	0	0	20

Overall chi-square value = 121.2

Very highly significant.

School C

		Stream A	Stream B	Stream C
	+0.44 S.S.	25	0	0
I.Q.	+0.44 to -0.44	14	17	1
	-0.44 S.S.	0	18	31

Overall chi-square value = 90.4

Highly significant.

School D

		Stream A	Stream B	Stream C
	+0.44 S.S.	18	1	0
I.Q.	+0.44 to -0.44	13	11	2
	-0.44 S.S.	0	12	13

Overall chi-square value = 45.1

Highly significant.

School E

		Stream A	Stream B	Stream C
	+0.44 S.S.	32	0	0
I.Q.	+0.44 to -0.44	9	10	0
	-0.44 S.S.	0	25	28

Overall chi-square value = 93.7

Highly significant.

The Complete Sample of First Year and Fourth Year Children Combined

		Stream A	Stream B	Stream C
	+0.44 S.S.	268	82	8
I.Q.	+0.44 to -0.44	110	153	53
	-0.44 S.S.	14	116	190

Overall chi-square value = 526.3

Appendix C

The Investigation of the Relationship between Level of I.Q. and Month of Birth (and thus sub-age group)

With First Year Children

The Total First Year Sample in all Five Schools Combined

		Month of Birth		
		Sept. to Dec.57	Jan. to April 58	May to Aug.58
	+0.44 S.S.	54	63	63
I.Q.	+0.44 to -0.44	63	48	61
	-0.44 S.S.	48	51	53

Chi-square value = 2.78. Not significant.

School A

		Month of Birth		
		Sept. to Dec.57	Jan to April 58	May to Aug.58
	+0.44 S.S.	21	17	15
I.Q.	+0.44 to -0.44	12	9	11
	*0.44 S.S.	6	7	5

Chi-square value = 0.833. Not significant.

School B

		Month of Birth		
		Sept. to Dec.57	Jan. to April 58	May to Aug.58
	+0.44 S.S.	11	16	21
I.Q.	+0.44 to -0.44	15	13	13
	+0.44 S.S.	5	3	7

Chi-square value = 3.4 Not significant.

School C

		Month of Birth		
		Sept. to Dec.57	Jan. to April 58	May to Aug.58
	+0.44 S.S.	8	9	8
I.Q.	+0.44 to -0.44	8	10	12
	-0.44 S.S.	10	8	17

Chi-square value = 2.14. Not significant.

School D

		Month of Birth		
		Sept. to Dec.57	Jan.to April 58	May to Aug.58
	+0.44 S.S.	7	13	7
I.Q.	+0.44 to -0.44	16	8	14
	-0.44 S.S.	12	16	7

Chi-square value = 7.2 Not significant.

School E

		Month of Birth		
		Sept.to Dec.57	Jan.to April 58	May to Aug.58
	+0.44 S.S.	7	8	12
I.Q.	+0.44 to -0.44	12	8	11
	-0.44 S.S.	15	17	17

Chi-square value = 1.815 Not significant.

With Fourth Year Children

The Total Fourth Year Sample in all Five Schools Combined

		Month of Birth		
		Sept.to Dec.54	Jan.to April 55	May to Aug.55
	+0.44 S.S.	66	56	56
I.Q.	+0.44 to -0.44	36	53	55
	-0.44 S.S.	60	60	48

Chi-square value = 7.117 Not significant.

School A

		Month of Birth		
		Sept. to Dec.54	Jan.to April 55	May to Aug.55
	+0.44 S.S.	22	10	20
I.Q.	+0.44 to -0.44	7	15	10
	-0.44 S.S.	6	9	6

Chi-square value = 8.7. Not quite significant at five per cent level.

Maxwell Analysis

Due to linear regression	0.123	1 d.f.	N.S.
Due to departure from linear regression	8.655	3 d.f.	S.(5%L)
Overall value	8.77	4 d.f.	N.S.

School B

		Month of Birth		
		Sept.to Dec.54	Jan.to April 55	May to Aug.55
	+0.44 S.S.	14	20	16
I.Q.	+0.44 to -0.44	11	9	15
	-0.44 S.S.	12	2	6

Chi-square value = 9.58. Significant at five per cent level.

Maxwell Analysis

Due to linear regression	1.477	1 d.f.	N.S.
Due to departure from linear regression	8.381	3 d.f.	S.(5%L)
Overall value	9.858	4 d.f.	S.(5%L)

School C

		Month of Birth		
		Sept.to Dec.54	Jan.to April 55	May to Aug.55
	+0.44 S.S.	10	8	7
I.Q.	+0.44 to -0.44	6	12	14
	-0.44 S.S.	17	16	16

Chi-square value = 3.707. Not significant.

School D

		Month of Birth		
		Sept.to Dec.54	Jan.to April 55	May to Aug.55
	+0.44 S.S.	9	6	4
I.Q.	+0.44 to -0.44	6	12	8
	-0.44 S.S.	8	10	7

Chi-square value = 2.957. Not significant.

School E

		Month of Birth		
		Sept.to Dec.54	Janlto April 55	May to Aug.55
	+0.44 S.S.	11	12	9
I.Q.	+0.44 to -0.44	6	5	8
	-0.44 S.S.	17	23	13

Chi-square value = 2.635. Not significant.

For the Complete Sample of First Year and Fourth Year Children Combined

		Month of Birth		
		Sept. to Dec.	Jan. to April	May to Aug.
	+0.44 S.S.	120	119	119
I.Q.	+0.44 to -0.44	99	101	116
	-0.44 S.S.	108	111	101

Chi-square value = 2.18. Not significant.

Appendix D

The Investigation of the Relationship Between Month of Birth (and thus age) and Stream Placement.

With First Year Children

The Total First Year Sample in all Five Schools Combined

		Month of Birth		
		Sept.to Dec.57	Jan.to April 58	May to Aug.58
	A	92	62	44
Stream	B	48	60	71
	C	25	40	62

Chi-square value = 37.9. Highly significant.

Maxwell Analysis

Due to linear regression	35.2	1 d.f.	S.(1%L)
Due to departure from linear regression	2.2	3 d.f.	N.S.
Overall value	37.4	4 d.f.	S. (1%L)

School A

		Month of Birth		
		Sept.to Dec.57	Jan.to April 58	May to Aug.58
	A	20	15	7
Stream	B	13	13	11
	C	6	5	13

Chi-square value = 10.6. Significant at five per cent level.

School B

		Month of Birth		
		Sept.to Dec.57	Jan. to April 58	May to Aug. 58
	A	20	8	8
Stream	B	4	12	16
	C	7	12	17

Chi-square value = 18.2 Significant at one per cent level.

School C

	Month of Birth		
	Sept. to Dec. 57	Jan. to April 58	May to Aug. 58
A	15	16	10
Stream B	10	7	18
C	1	4	9

Chi-square value = 10.77. Significant at five per cent level.

Maxwell Analysis

Due to linear regression	8.29	1 d.f.	S.(5%L)
Due to departure from linear regression	2.48	3 d.f.	N.S.
Overall value	10.77	4 d.f.	S. (5%L)

School D

	Month of Birth		
	Sept. to Dec. 57	Jan. to April 58	May to Aug. 58
A	23	12	7
Stream B	7	16	10
C	5	9	11

Chi-square value = 14.37. Significant at one per cent level.

School E

	Month of Birth		
	Sept. to Dec. 57	Jan. to April 58	May to Aug. 58
A	14	11	12
Stream B	14	12	16
C	6	10	12

Chi-square value = 2.14. Not significant

Maxwell Analysis

Due to linear regression	1.49	1 d.f.	N.S.
Due to departure from linear regression	0.65	3 d.f.	N.S.
Overall value	2.14	4 d.f.	N.S.

With Fourth Year Children

The Total Fourth Year Sample in all Five Schools Combined

		Month of Birth	
	Sept.to Dec. 54	Jan. to April 55	May to Aug. 55
Stream A	72	68	54
Stream B	49	60	63
Stream C	40	42	42

Chi-square value = 4.38 Not significant.

Maxwell Analysis

Due to linear regression	3.0	1 d.f.	N.S.
Due to departure from linear regression	1.38	3 d.f.	N.S.
Overall value	4.38	4 d.f.	N.S.

School A

		Month of Birth	
	Sept. to Dec. 54	Jan. to April 55	May to Aug. 55
Stream A	21	9	13
Stream B	8	16	16
Stream C	6	9	7

Chi-square value = 9.06. Not quite significant at five per cent level.

Maxwell Analysis

Due to linear regression	2.06	1 d.f.	N.S.
Due to departure from linear regression	7.00	3 d.f.	N.S.
Overall value	9.06	4 d.f.	N.S.

School B

		Month of Birth	
	Sept. to Dec. 54	Jan. to April 55	May to Aug. 55
Stream A	12	16	12
Stream B	13	11	14
Stream C	12	4	11

Chi-square value= 4.98. Not significant.

Maxwell Analysis

Due to linear regression	0.02	1 d.f.	N.S.
Due to departure from linear regression	4.96	3 d.f.	N.S.
Overall value	4.98	4 d.f.	N.S.

School C

		Month of Birth	
	Sept. to Dec. 54	Jan. to April 55	May to Aug. 55
Stream A	16	14	9
Stream B	9	12	14
Stream C	8	10	14

Chi-square value = 4.57 Not significant.

Maxwell Analysis

Due to linear regression	3.87	1 d.f.	S.(5%L)
Due to departure from linear regression	0.70	3 d.f.	N.S.
Overall value	4.57	4 d.f.	N.S.

School D

		Month of Birth	
	Sept. to Dec. 54	Jan. to April 55	May to Aug. 55
Stream A	9	15	7
Stream B	9	5	10
Stream C	4	9	2

Chi-square value = 7.65. Not significant.

Maxwell Analysis

Due to linear regression	0.02	1 d.f.	N.S.
Due to departure from linear regression	7.63	3 d.f.	N.S. (but almost at 5%L)
Overall value	7.65	4 d.f.	N.S.

School E

		Month of Birth	
	Sept. to Dec. 54	Jan. to April 55	May to Aug. 55
Stream A	14	14	13
Stream B	10	16	9
Stream C	10	10	8

Chi-square value = 1.34. Not significant.

APPENDIX E

For the Youngest Sub-Age Group. (May to August birthdays)

The Relationship between Stream Placement and School Year

		Stream A	Stream B	Stream C	
Year Group	First Year	44	71	62	- 177
	Fourth Year	54	63	42	- 159
		98	134	104	- 336

Chi-square value = 4.389 d.f. = 2

Not quite significant at the five per cent level of confidence.

Maxwell Analysis

Due to linear regression	4.376	1 d.f.	S (5%L)
Due to departure from linear regression	0.013	1 d.f.	N.S.
Overall value	4.389	2 d.f.	N.S.

Detail of Maxwell Analysis (with reference to the above example)

When there exists a natural order amongst the categories in a classification it may be assumed that there is a continuous variable underlying them. Under this assumption it is possible to quantify the variable by allotting numerical values to the categories. With the above example the following values were nominated.

		x variate -1	x variate 0	x variate +1
y variate	+1	44	71	62
	0	54	63	42

From a frequency distribution of y' the sum of squares of the y' values about their mean is 83.74.

From a frequency distribution of x' the sum of squares of the x' values about their mean is 201.89

Also $S_y^2 = 0.2492$ and $S_x^2 = 0.6009$

From a frequency distribution of (x' - y') the sum of squares of the (x' - y') values about their mean is 255.96

From a frequency distribution of (x' + y') the sum of squares of the (x' + y') values about their mean is 315.3

$$byx = \frac{(83.74 + 201.89 - 255.96)}{2(201.89)} = 0.0735$$

$$bxy = \frac{(83.74 + 201.89 - 255.96)}{2(83.74)} = 0.1772$$

$$Vbyx = \frac{0.2492}{201.89} = 0.001234$$

$$Vbxy = \frac{0.6009}{83.74} = 0.007176$$

$$\text{Chi-square for } byx = \frac{(byx)^2}{Vbyx} = \frac{(0.0735)^2}{0.001234} = 4.376$$

$$\text{Chi-square for } bxy = \frac{(bxy)^2}{Vbxy} = \frac{(0.1772)^2}{0.007176} = 4.376$$

∴ Amount of chi-square due to linear regression = 4.376

<u>A Test of Proportions Between 1st Year Young Children in Stream A and 4th Year Young Children in Stream A (May to August birthdays)</u>

Using the formula

$$\bar{z} = \frac{|p_1 - p_2|}{\sqrt{\bar{p}\,\bar{q}\,\frac{N_1 + N_2}{N_1 N_2}}} \qquad \text{where } \bar{p} = \frac{N_1 P_1 + N_2 P_2}{N_1 + N_2}$$

$$\text{and where } \bar{q} = 1 - \bar{p}$$

1st Year 44 children out of 177 in Stream A ∴ $p_1 = 0.2486$

4th Year 54 children out of 159 in Stream A ∴ $p_2 = 0.3396$

$N_1 = 177$ $N_2 = 159$

∴ $\bar{p} = 0.2917$

and $\bar{q} = 0.7083$

∴ $\bar{z} = 1.832$

<u>A Test of Proportions Between 1st Year Young Children in Stream B and 4th Year Young Children in Stream B (May to August birthdays)</u>

Using the same formula as before.

1st Year 71 children out of 177 in Stream B ∴ $p_1 = 0.4012$

4th Year 63 children out of 159 in Stream B ∴ $p_2 = 0.3960$

$N_1 = 177$ $N_2 = 159$

$\therefore \bar{p} = 0.39$

and $\bar{q} = 0.61$

$\therefore \bar{z} = 0.09$

A Test of Proportions Between 1st Year Young Children in Stream C and 4th Year Young Children in Stream C (May to August birthdays)

Using the same formula as before.

1st Year 62 children out of 177 in Stream C $\therefore p_1 = 0.35$

4th Year 42 children out of 159 in Stream C $\therefore p_2 = 0.283$

$N_1 = 177$ $N_2 = 159$

$\therefore \bar{p} = 0.3155$

and $\bar{q} = 0.6845$

$\therefore \bar{z} = 1.39$

A Test of Proportions Between 1st Year Young Children in Form A and 4th Year Young Children in Form A (When only children placed in either A or C Streams are considered)

Using the same formula as before.

1st Year 44 children out of 106 in Form A

4th Year 54 children out of 96 in Form A

$N_1 = 106$ $N_2 = 96$

$\therefore \bar{p} = 0.4851$

and $\bar{q} = 0.5149$

$\therefore \bar{z} = 2.09$

APPENDIX F

Inter-School Differences can be seen in the following table.

Schools	A	B	C	D	E
1st Year Mean I.Q.	117.2	115.7	107.8	107.0	106.0
4th Year Mean I.Q.	109.7	109.7	98.6	100.7	98.1
1st Year Stan.Dev.I.Q.	13.5	13.6	14.5	15.8	16.8
4th Year Stan.Dev.I.Q.	16.3	17.5	16.7	14.6	19.6
1st Year - I.Q. v. Stream Placement (Chi-square values given)	53.6(1%)	33.5(1%)	30.6(1%)	18.0(1%)	48.6(1%)
4th Year - I.Q. v. Stream Placement (Chi-square values given)	100.3(1%)	121.2(1%)	90.4(1%)	45.1(1%)	93.7(1%)
1st Year - I.Q. v. Mon. of Birth (Chi-square values given)	0.8(N.S.)	3.4(N.S.)	2.14(N.S)	7.2(N.S.)	1.8(N.S.)
4th Year - I.Q. v. Mon. of Birth (Chi-square values given)	8.7(N.S.)	9.5(5%)	3.7(N.S)	2.9(N.S.)	2.6(N.S.)
1st Year - Month of Birth v. Str. Place. (Chi-square values given)	10.6(5%)	18.2(1%)	10.7(5%)*	14.3(1%)	2.1(N.S.)
4th Year - Month of Birth v. Str. Place. (Chi-square values given)	9.0(N.S.)	4.9(N.S.)	4.5(N.S.)*	7.6(N.S.)	1.3(N.S.)

* Shows a significant linear trend.

BIBLIOGRAPHY

BARTHELMESS, H.M., and BOYER, P.A. 1932.
An Evaluation of Ability Grouping. Journal of Educational Research, Volume 26.

BILLETT, H.O., 1929.
A Controlled Experiment to Determine the Advantages of Homogeneous Grouping. Educational Research Bulletin XII. University of Ohio.

BUTCHER, H.J., 1965.
The Attitudes of Student Teachers to Education. British Journal of Social and Clinical Psychology.

BUTLER, N.R., PRINGLE, M.L.K., DAVIES, R. 1965.
The National Child Development Study (1958 Cohort). 11,000 Seven Year Olds. Publisher Longmans and Green.

CHETCUTI, F. 1961.
A Study of the Morale of A Stream and C Stream Pupils in Secondary Schools. Educational Review. University of Birmingham.

CRAFTS, L., SCHNEIRLA, T. et al. 1950.
Recent Experiments in Psychology. Publisher McGraw-Hill.

COOK, P.R., 1924.
A Study of the Results of Homogeneous Grouping of Abilities in High School Classes. 23rd Year Book. Public School Publishing Co., Bloomington, Illinois.

CORNELL, E.D., 1936.
Effects of Ability Grouping. Published Studies. 35th Year Book. Public School Publishing Co., Bloomington, Illinois.

COXE, W.W., 1936.
The Grouping of Pupils. 35th Year Book. Public School Publishing Co., Bloomington, Illinois.

DANIELS, J.C. 1955.
Spurring the Intellect. Times Educational Supplement. 29.7.65.

DANIELS, J.C. 1961.
The Effect of Streaming in the Primary School. British Journal of Educational Psychology, Vol. 31, 1961. I. What Teachers Believe. II. A Comparison of Streamed and Unstreamed Schools.

DAVIS, A. 1951.
Intelligence and Cultural Differences. Publisher University of Chicago Press.

DOUGLAS, J.W.B., 1964.
The Home and the School. Publisher Macgibbon and Kee.

EDMISTON, R.W. and BENEFER, J.C. 1949.
The Relationship between Group Achievement and Range of Abilities within the Group. Journal of Educational Research.

FISHER, R.A., 1950.
Statistical Methods for Research Workers. Publisher Oliver and Boyd.

FITT, A.B., 1941.
Seasonal Influence on Growth. Function and Inheritance. New Zealand Council for Educational Research, Series No.17. (London) Oxford University Press.

FRANDSEN, A.N., 1961.
Educational Psychology. Publisher McGraw-Hill.

GATFIELD, R.A., 1958.
An Experimental Investigation into Certain Aspects of Streaming in a Primary School. Thesis - University of Southampton.

GRAY, H.R. and HOLLINGWORTH, L.S. 1931
The Achievement of Gifted Children Enrolled or not Enrolled in Special Opportunity Classes. Journal of Educational Research, Vol. 24.

GREENBERG, R.C. 1963.
Two Factors Influencing the Births of Mongols to Younger Mothers. Medical Officer No.109.

GUILFORD, J.P., 1956.
Fundamental Statistics in Psychology and Education. Publisher McGraw-Hill.

HEBB, D.O., 1948.
The Organisation of Behaviour. Publisher Wiley, New York.

HUNTINGTON, E. 1944.
Season of Birth and Fame. Journal of Genetical Psychology 1944, Vol. 63.

JACKSON, B. 1964
Streaming: An Educational System in Miniture. Publisher Routledge and Kegan Paul 1964.

JOHNS, E.M. 1962
The Age Factor in Reading Retardation. Research Studies No. 24, University of Leeds Institute of Education.

KNOBLOCH, H. and PASAMANICK, B. 1958
Seasonal Variation in the Births of Mentally Deficient. American Journal of Public Health No. 48.

LINDLEY, D.V. and MILLER, J.C.P. 1953
Cambridge Elementary Statistical Tables. Publisher Cambridge University Press.

McINTYRE, D. and MORRISON, A. 1967
The Educational Opinions of Teachers in Training. British Journal of Social and Clinical Psychology.

MALLER, J.B. 1929
Co-operation and Competition. An Experimental Study of Motivation. Teachers College Contributions to Education 1929, No.384.

MAXWELL, A.E. 1961
Analysing Qualitative Data. Publisher Methuen & Co. Ltd.

Moray House Picture Test 1
Publisher University of London Press.

Moray House Verbal Reasoning Test 77.
Publisher University of London Press.

MORRIS, J.M. 1959
Reading in the Primary School. National Foundation for Educational Research. Publication No. 12. Publisher Newnes.

MOYER, E.L. 1924
A Study of the Effects of Classification by Intelligence Tests. 23rd Year Book. Public School Publishing Co., Bloomington, Illinois.

A Cross Sectional Study of Attainment in Matched Streamed and Unstreamed Schools. National Foundation for Educational Research Survey 1962.

OJHA, A.B., KELVIN, R.P. and LUCAS, C.J. 1966
A Note on Season of Birth and Intelligence. British Journal of Educational Psychology 1966, Vol. 36.

OLIVER, R.A.C. and BUTCHER, H.J. 1962
The Manchester Scales of Opinions about Education. Teachers Attitudes to Education. British Journal of Social and Clinical Psychology.

ORME, J.E. 1962
Intelligence and Season of Birth. British Journal of Medical Psychology No. 35.

ORME, J.E. 1963
Intelligence, Season of Birth and Climatic Temperature. British Journal of Psychology 1963, No. 54.

PIDGEON, D.A. 1960
Ability and Attainment of Children at Three Age Levels. British Journal of Educational Psychology.

PIDGEON, D.A. and YATES, A. 1956
The Relation between Ability and Attainment. An Examination of Current Theory. National Foundation for Educational Research No. 8.

PINTNER, R. and FORLANO, G. 1933
The Influence of Month of Birth on Intelligence Quotients. The Journal of Educational Psychology.

PLOWDEN REPORT 1967
Children and their Primary Schools. Publisher H.M.S.O.

PRINGLE, K. and GOOCH, M. 1965
Four Years On. National Bureau for Co-operation In Child Care.

PURDOM, I.L. 1924

The Value of Homogeneous Grouping. Publisher Warwick and York. Baltimore.

RAUP, R.B. 1936

Some Philosophical Aspects of Grouping. 35th Year Book, Public School Publishing Co. , Bloomington, Illinois.

RAVEN, J.C.

Standard Progressive Matrices. Publisher H.K. Lewis & Co.

RUDD, W.G.A. 1958

The Psychological Effects of Streaming By Attainment. British Journal of Educational Psychology, Vol. 28, 1958.

SORENSON, H. 1948

Psychology in Education. Publisher McGraw-Hill.

TERMAN, L.M. and MERRILL, M.A. 1961

Stanford-Binet Intelligence Scale (3rd Revision) Publisher G. Harrap & Co.

THOMPSON, D. 1966

Streaming in the Secondary School. Educational Review. University of Birmingham.

THORNDIKE, E.L. 1931

Human Learning. Appleton Century Crofts - New York.

VERNON, P.E. 1960

Intelligence and Attainment Tests. Publisher University of London Press.

WECHSLER, D. 1949

Wechsler Intelligence Scale for Children. Publisher The Psychological Corporation, New York.

WILLIAMS, P. 1964

Date of Birth, Backwardness and Educational Organisation. British Journal of Educational Psychology 1964, Vol. 34.

WYNDHAM, H.S. 1934

Ability Grouping. Melbourne University Press.

INDEX