CHAPTER 3

Psychometric Assessment

Programme

The psychometric tests which we planned to administer in an attempt to identify and quantify any significant sequelae of being born either too soon or too small are summarised in Table 3.1, together with the ages at which the various tests were to be administered. The numbers of children in each of our study groups to whom this programme was successfully administered, at successive ages, are shown in Table 3.2. The reductions in the numbers tested at six and seven years are accounted for by families moving away from the area.

TABLE 3.1

Programme of psychological tests

Name and nature of test	*Approximate age (years)*
Wechsler Preschool and Primary Scale of Intelligence (WPPSI)	6
Illinois Test of Psycholinguistic Abilities (ITPA)	5, 7
Skemp Test of Visual Concepts	6
Frostig Developmental Test of Visual Perception	5
Purdue Pegboard Test of Manipulative Dexterity	5, 6, 7
Birch tests of Intersensory Abilities	6, 7
Bender-Gestalt Test of Visuo-Motor Abilities	7
Ozeretsky Test of Motor Impairment	7
Harris Tests of Laterality	7
Schonell Test of Word Recognition	7
Holborn Test of Sentence Reading	7
Junior Eysenck Personality Inventory	7

The actual mean ages of the children tested at the three successive stages were 5 years 8 months, 6 years 6 months and 7 years 6 months: there were no differences between the groups, except that the mean age of the short-gestation group was only 5 years 7 months for the first round of testing. The standard deviations of the ages of the children in each group at the time of each round varied between 0.92 and 1.57 months.

TABLE 3.2

Numbers of children successfully tested

Approximate age (years)	*Random control*	*Short gestation*	*Rather light-for-dates*	*Very light-for-dates*
5	186[1]	59	67	73[2]
6	183	59	67	72
7	178	58	66	70

[1]One child in control group was too unco-operative for testing to be valid.
[2]One child in very light-for-dates group was deaf and unable to communicate.

Procedure

Each mother and child was brought to the clinical unit on which our study was based, where the child was tested first by the psychologist and then by one or more other members of the research team (*e.g.* psychiatrist, speech therapist). All observers were working 'blind', in the sense that they did not know either the group to which any child belonged, or each other's findings. With few exceptions each round of testing was completed in one session, the psychologist's examination normally lasting about 1¼ hours. In addition to the tests listed in Table 3.1, the psychologist rated each child's behaviour during testing for shyness, negativism and distractibility on a five-point unipolar scale extracted from the Minnesota Pre-School Scale (Goodenough *et al.* 1940). The psychiatrist who examined the child later during the same visit to the unit made his own independent rating, using the same scale. The two sets of results are described and discussed in Chapter 4.

Presentation of Findings

So as to present the results of the psychologist's examination in a manner which can be understood easily by the reader who has no specialised knowledge of psychometric tests, but is mainly interested in the general findings and the conclusions which can be drawn from them, we have adopted a simple, routine format for dealing with each test in the text. First there is a brief description of the test's purpose, the type of procedure involved and references to the published evidence concerning its validity and previous use; then the over-all results are given in the form of a table and the results for individual subtests (if any) are presented, usually in the form of a line diagram. So as to facilitate quick appreciation of the implications of these diagrams they are all drawn in such a way that a higher point on the vertical scale implies a 'better' result, in terms of the value judgments on which the test is based. Where the mean score for either abnormal group differs significantly from the controls by accepted standards, this fact is indicated. Only three sets of results are shown graphically in each diagram (those for the control and the two extreme groups); this is in the interests of clarity, and the results for the 'rather light-for-dates' group, if analysed separately, are described briefly in a footnote to the diagram.

So as to enable the more technically sophisticated reader to check the data in detail we have set out the numerical scores for the subtests, with statistical ranges and the actual results of tests of significance, in an Appendix (available on request from the Editors I.K. and D.S.).

Results of Psychometric Tests

Wechsler Preschool and Primary Scale of Intelligence (WPPSI)

This test is designed to measure global intellectual capacity and to produce separate measures of the verbal and performance components of intelligence in children aged 4½ to six years. It was developed as an extension to a younger age group of the Wechsler Intelligence Scale for Children: its description and standardisation were published by Wechsler (1963). Two reports of its administration to British populations have been published, in the Isle of Wight by Yule *et al.* (1969) and in Nottingham by Brittain (1969).

The over-all results are summarised in Table 3.3 and the results for the subtests are shown in Figure 2. The two extreme abnormal groups differ significantly from the controls in their mean verbal, performance and full-scale quotients. There are no significant differences between the mean quotients of the two extreme abnormal groups. The rather light-for-dates group's mean quotients are all higher than those of the very light-for-dates, but not significantly so. The patterns of the mean subtest scores show a remarkable similarity between the three groups (Fig. 2); although all mean scores in the two extreme abnormal groups are depressed as compared with the controls, there is no specific pattern of disability in particular subtests to distinguish the extreme groups either from the controls or from each other.

Illinois Test of Psycholinguistic Abilities (ITPA)

This test is designed to detect specific abilities and disabilities in the field of communication. The experimental version which we used has nine subtests and was developed by McCarthy and Kirk (1961) from the model described by Osgood (1957). It aims to test the understanding of meaningful symbols and the automatic recall and manipulation of symbols and sequences of symbols; the processes of decoding (comprehension), encoding (expression) and association (the link between the two); and the auditory, visual, vocal and motor pathways used in communication. The individual subtests are described in the Appendix (available on request from the Editors). This test was administered at both five and seven years of age so that we could study the development of language over the period of infant school attendance.

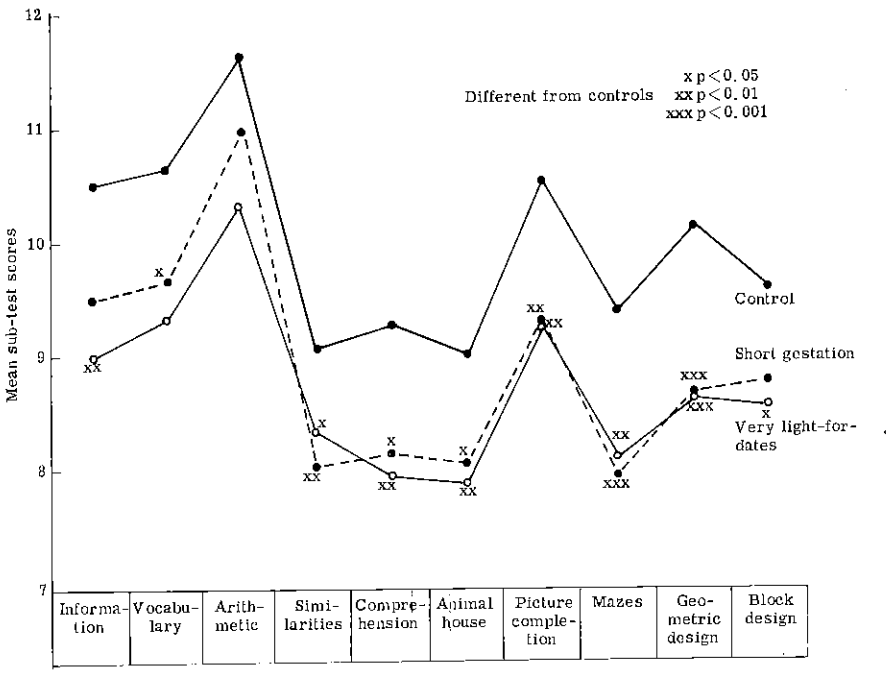

Fig. 2. Wechsler Preschool and Primary Test of Intelligence: mean subtest scores at age six years.

TABLE 3.3

Wechsler Preschool and Primary Scale of Intelligence at age five years

IQ	*Random control (a)*	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant differences* $p<0.05$	$p<0.01$	$p<0.001$
Verbal	101.0	95.5	95.1	93.8	a/b	a/c	a/d
Performance	98.2	90.3	94.4	89.9	—	—	a/b, a/d
Full-scale	99.7	92.3	94.3	91.3	a/c	a/b	a/d
standard deviation	16.2	14.6	15.7	16.7			

N.B. No significant differences between short-gestation and very light-for-dates groups.

TABLE 3.4

Illinois Test of Psycholinguistic Abilities at age five years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant differences* $p<0.05$	$p<0.01$	$p<0.001$
Over-all score	—0.20	—0.66	—0.58	—0.82	a/c	a/b	a/d
Language quotient	98.4	91.5	93.5	90.0	a/c	a/b	a/d
standard deviation	14.3	15.0	16.5	16.8			

N.B. No significant differences between short-gestation and very light-for-dates groups.

TABLE 3.5

Illinois Test of Psycholinguistic Abilities at age seven years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant differences* $p<0.05$	$p<0.01$	$p<0.001$
Over-all score	—0.51	—0.93	—0.84	—1.11	a/b	—	a/d
Language quotient	97.8	92.0	93.4	90.5	a/b	a/d	—
standard deviation	15.5	14.8	16.4	16.7			

N.B. No significant difference between short-gestation and very light-for-dates group.

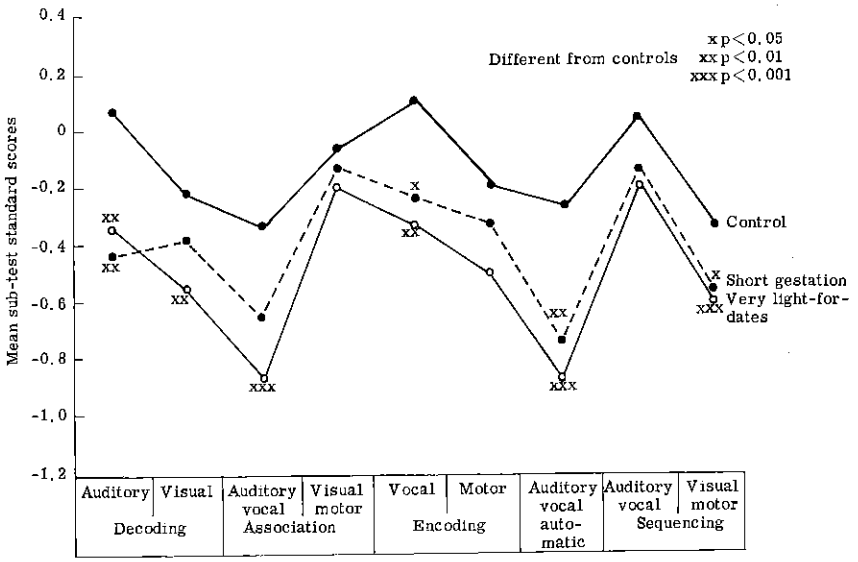

Fig. 3. Illinois Test of Psycholinguistic Abilities: mean subtest standard scores at age five years.

The over-all results at five years are summarised in Table 3.4 and the results for the subtests are shown in Figure 3. In both cases the range of possible scores is —3 to +3, with a mean of 0. The mean quotients of the abnormal groups again differ significantly from the control group, and again the differences are more significant in the case of the very light-for-dates than the short gestation group. There is again no significant difference between the two extreme abnormal groups. Again the rather light-for-dates quotient is higher than the very light-for-dates, but not by a significant amount. The patterns of the subtest scores of the two extreme abnormal groups again resemble each other remarkably closely, and this time to a greater degree than they resemble the pattern of the controls. The greatest relative depression of the extreme abnormal group's scores is seen in the two auditory-vocal subtests; but this hardly amounts to a significant pattern of disability relative to the controls. There are two subtests, visual-motor association and auditory-vocal sequencing, in which there are no significant differences between the scores of the groups: in the case of the former, clinical impressions formed during its administration strongly suggested that chance played an important part in determining the child's score at the age of five years.

For purposes of comparison, the over-all results at the age of seven years are summarised in Table 3.5 and the results for the subtests are shown in Figure 4 to the same scale as in Figure 3. Again the quotients of the two extreme abnormal groups differ significantly from the controls, but the actual differences are rather less than they were at five years; and at this age the rather light-for-dates group no longer differs significantly from the controls. As at five years, the patterns of the subtest scores of the two extreme groups resemble each other, but this time they show no such gross depression below the controls in the two auditory-vocal subtests as they did previously. In the visual-motor association subtest there is this time a significantly

worse score by the very light-for-dates group as compared with the short gestation group, suggesting that, though possibly too difficult for all groups at the age of five, this subtest may have presented a specific difficulty to the very light-for-dates group at the age of seven. There are no other significant differences between the two extreme abnormal groups. In vocal encoding and auditory-vocal sequencing there were no significant differences between any of the groups.

Skemp Test of Visual Concepts

No adequately standardised tests of visual conceptualization in children of this age were available to us: but we thought it worthwhile to administer the test devised by Skemp (1967) since, although it had not at the time been adequately assessed for reliability or validity, there were sufficient children in our random control group to provide norms for purposes of statistical comparison with the abnormal groups. The test provides a total score of mental abilities assessed through the visual channel and consists of three subtests:—

(1) Memory Cards, a test of short visual memory;
(2) Paired Associates, a visual learning task involving the association of geometric-type symbols with coloured patterns; and
(3) Visual Concepts, in which the child must without guidance develop the concept of removing the non-exemplar from a series (*i.e.* choosing the 'odd man out').

An account of the background data available to us is given in the Appendix.

The results are summarised in Table 3.6, where again the total scores of the abnormal groups, but particularly of the very light-for-dates, are significantly depressed. There is also a significant depression of the very light-for-dates group in all subtests (Fig. 5).

Frostig Developmental Test of Visual Perception

This test was designed to investigate specific skills in the area of visual perception (Frostig *et al.* 1963) and was standardised for children aged four to seven years. It consists of five subtests:—

(1) Eye-motor co-ordination, a test of accuracy of control over a pencil when drawing between fixed points;
(2) Figure-Ground, a test of ability to perceive geometric shapes against an increasingly complex background;
(3) Form Constancy, a test of ability to recognise geometric shapes under varying conditions of size, shading, texture, etc.;
(4) Position in Space, a test of ability to recognise reversals and rotations of figures; and
(5) Spatial Relationships, a test of copying simple forms and patterns, using dots as guide points.

The over-all results are summarised in Table 3.7, and the results of the subtests are shown in Figure 6. Once again, the quotients of the two extreme abnormal groups differ significantly from the control group: once again, the patterns of their subtest scores are similar to the controls, so that there does not seem to be any evidence of a specific disability in either of the extreme abnormal groups. Each of these groups differs significantly from the controls in four out of the five subtests.

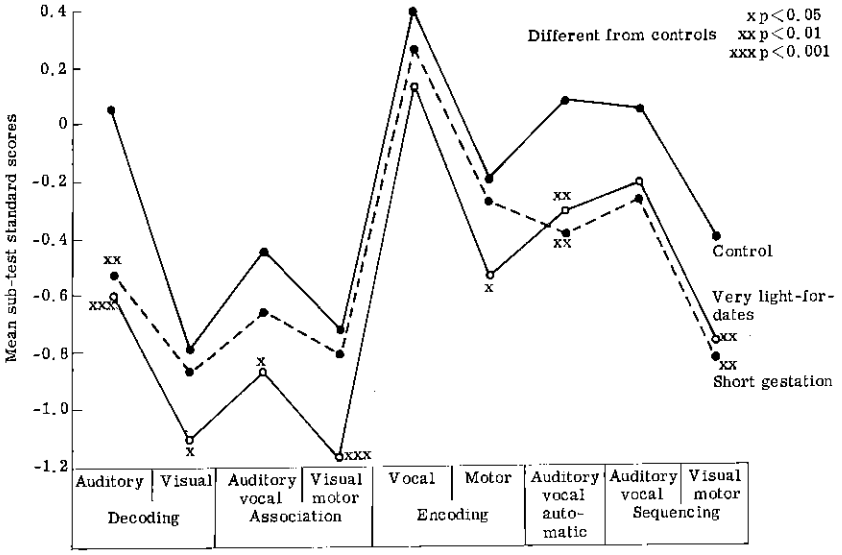

Fig. 4. Illinois Test of Psycholinguistic Abilities: mean subtest standard scores at age seven years.

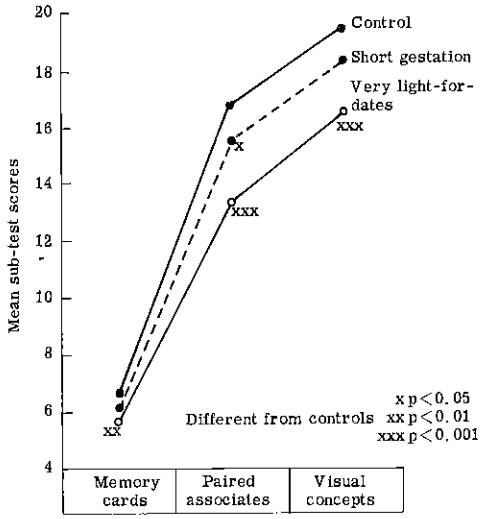

Fig. 5. Skemp Test of Visual Concepts: mean subtest scores at age six years.

TABLE 3.6

Skemp Test of Visual Concepts at age six years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant differences* $p<0.05$	$p<0.01$
Total score	34.0	32.4	33.4	29.9	a/b, b/d, c/d	a/d
standard deviation	5.2	5.6	5.8	7.0		

TABLE 3.7

Frostig Developmental Test of Visual Perception at age five years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant differences* $p<0.05$	$p<0.01$	$p<0.001$
Perceptual quotient	102.7	96.1	100.5	95.5	c/d	a/b	a/d
standard deviation	12.3	13.5	12.5	14.2			

N.B. No significant differences between short-gestation and very light-for-dates group.

TABLE 3.8

Intersensory tests: mean total scores of three subtests at six and seven years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant differences* $p<0.05$	$p<0.01$	$p<0.001$
Total score at six years	17.0	15.2	15.4	14.7	—	a/c	a/b, a/d
Total score at seven years	18.5	17.7	17.8	17.1	a/d	—	—

N.B. No significant differences between short-gestation and very light-for-dates groups.

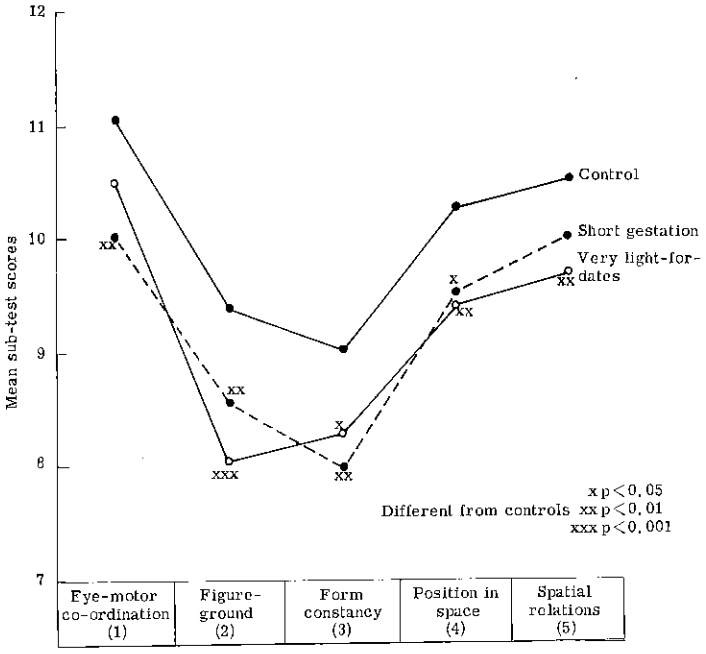

Fig. 6. Frostig Developmental Test of Visual Perception: mean subtest scores at age five years.

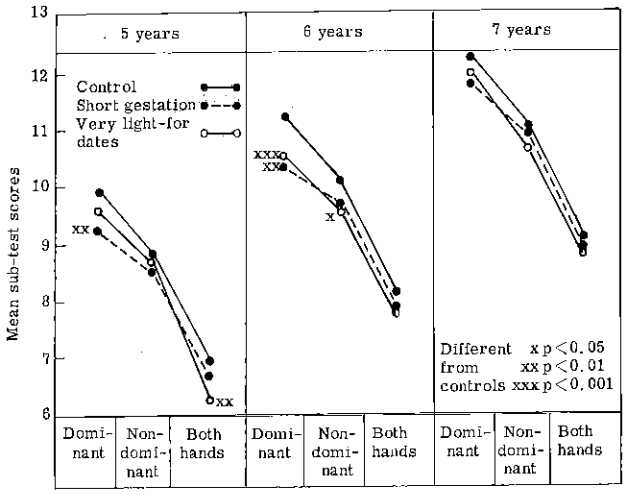

Fig. 7. Purdue Pegboard Test of Manual Dexterity: mean subtest scores at ages five, six and seven years.

Purdue Pegboard

This test is intended to measure manual dexterity and speed of working, independently of educational level. It was evaluated in children down to the age of six years by Rapin *et al.* (1966). Following a brief period of instruction, the child is given three 30-second trials, one with the left hand, one with the right and one with both hands together (Science Research Associates 1948). The total numbers of metal pegs (or of pairs when both hands are used) successfully placed in the right holes in the board at each trial are scored separately. It was also noted which hand the child preferentially used for holding a pencil in making a drawing and this was recorded as dominant for the purpose of expressing the results in Figure 7 (no over-all score is calculated in this test, so there is no table of results). The test was administered as part of each round of our assessment programme so as to observe changes with age. It is immediately apparent that there was progressive improvement in the scores of all groups at successive ages. Although the scores of the two extreme abnormal groups are all lower than those of the controls, relatively few of the differences are significant, and none is significant at the age of seven years.

Intersensory Tests

These tests were developed by Birch and his associates for investigating the development of intersensory relationships in brain-damaged children. We employed three subtests:—

(1) Auditory-visual integration. The child is presented auditorily with a series of long and short tapping sounds, like a morse code letter, and is asked to identify the sequence he has heard from a set of three slightly differing series presented visually on a card. The task is to convert an auditory stimulus into a visual type of response (Belmont *et al.* 1968).

(2) Haptic-visual integration. The child is seated with his hands extended through a screen which prevents his seeing them. He is presented in turn with eight geometric wooden shapes from the Seguin form board, and after exploring each by touch alone he is required to identify it from a visual array depicting all eight figures. The task is to convert a haptic (touch) stimulus into a visual type of response, on the basis of differences in shape (Birch and Lefford 1964).

(3) Haptic-visual discrimination. The child is required to investigate by touch alone (as in the previous subtest) a series of five graded metal discs of different sizes, and to indicate which particular disc he is handling from among a duplicate set which is within his field of vision. The task again is to convert a haptic (touch) stimulus into a visual type of response, but this time on the basis of differences in size.

These three subtests were administered as part of the second and third rounds of our assessment programme, again in the hope of observing changes with age. The over-all scores at six and seven years of age are summarised in Table 3.8 (p. 26), and the results of the subtests are shown in Figure 8. The significant difference between the total scores of the two extreme abnormal groups and the controls seen at the age of six years had almost disappeared by the age of seven years. Also, the lower scores of the two extreme abnormal groups in all subtests at six years, significant in five out of the

six possible comparisons, had almost disappeared by the age of seven years when the only two significantly reduced scores were in the very light-for-dates group.

Bender-Gestalt Test of Visuomotor Ability

This test was designed by Bender (1938) to detect difficulty in visuo-motor performance, as a possible indicator of cerebral dysfunction on an organic basis. It involves reproducing designs by direct copying. Norms for children aged from five to 10 years were established by Koppitz (1964).

The results in our population at the age of seven years are summarised in Table 3.9; since there are no subtests there is no corresponding Figure in this case. It should be noted that, because errors rather than successes are recorded, a higher numerical score in this test indicates a less satisfactory performance. It is clear that there is a very significant impairment of performance in the very light-for-dates group as compared with the control and rather light-for-dates groups. The score of the short-gestation group is intermediate and does not differ significantly from either the control or the very light-for-dates group. Taken at its face value, this finding suggests that the generally impaired performance of the very light-for-dates group in all the tests reported so far may be explained on the basis of organic brain damage.

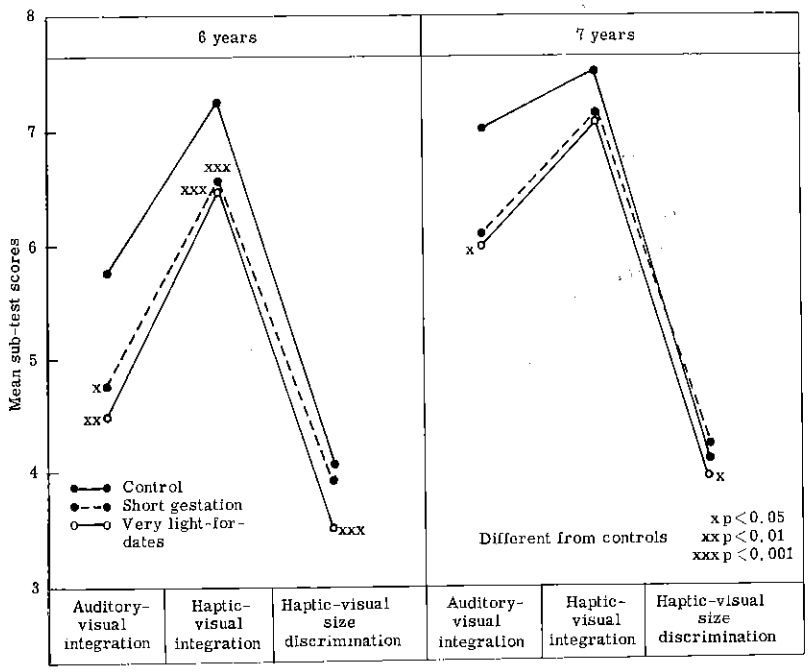

Fig. 8. Intersensory tests: mean subtest scores at ages six and seven years.

Ozeretsky-Stott Test of Motor Impairment

The original test described by Ozeretsky was intended to assess the whole range of motor abilities: the modification described by Stott (1966) was specifically designed to identify impairment of motor abilities of a type presumed to be due to neurological damage in children aged from five to 16 years. He eliminated all the items in the original test which showed differences between boys and girls and also pruned away the tasks which he considered likely to be affected by complicating factors such as mental ability, height, muscular strength and prior learning experience. The version which we used yields only a single over-all score of general motor impairment, but this represents the results of assessing a wide range of fine and gross motor skills, including steadiness and control during movement of the whole body, balance while immobile, co-ordination of activities involving arms and hands, and manual dexterity.

The mean over-all impairment scores of our study groups are shown in Table 3.10. Again there is no Figure since there are no subtest scores. As in the previous test, a numerically higher score denotes a less satisfactory performance because impairments are recorded, not successes. The performance of the very light-for-dates group is again obviously and significantly impaired, but in this case the short-gestation group's score is even worse (though there is no statistically significant difference between these two extreme abnormal groups).

Harris Tests of Laterality

So as to compare the extent to which lateral dominance had developed and was appreciated by children in different groups at the age of seven years, we administered a number of the tests described by Harris (1958). We found few differences between the groups and have shown in Table 3.11 the results of the only three tests in which some difference was obvious — and even then this difference was only statistically significant in one test (for knowledge of right and left). The three tests included in the Table can be briefly described as follows:—

(1) Hand dominance: a composite rating based on the results of two tasks, which assess (*a*) the extent to which one hand is preferentially used for activities such as winding a watch and cutting with scissors, and (*b*) the extent to which mirror reversals and errors of legibility occur when the child is required to write digits with both hands simultaneously without being able to see what he is doing.

(2) Knowledge of right and left: testing the child's ability to identify correctly his own right hand, left ear, etc.

(3) Concordance of dominance of hand (as rated in (1) above) with eye, rated for both monocular and binocular vision.

This is the first type of test in our series in which the very light-for-dates group does not differ significantly from the controls at any point; it is only the short-gestation group whose performance is significantly different (in a direction which must be considered inferior) because of a high incidence of hesitancy in correctly identifying the right from the left parts of the body.

TABLE 3.9

Bender-Gestalt Test of Visuo-motor Ability at age seven years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant differences* $p<0.05$	$p<0.01$	$p<0.001$
Total error score	5.09	5.81	5.06	6.73	—	c/d	a/d
standard deviation	3.32	2.84	3.31	3.86			

N.B. No significant difference between the error scores of the short-gestation and very light-for-dates or control groups.

TABLE 3.10

Ozeretsky/Stott Test of Motor Impairment at age seven years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant differences* $p<0.05$	$p<0.01$	$p<0.001$
Total motor—							
impairment score	1.85	3.93	2.41	3.67	—	—	a/b, a/d
standard deviation	2.72	4.70	3.57	4.27			

TABLE 3.11

Harris Tests of Laterality at age seven years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant difference* $p<0.01$
	%	%	%	%	
Over-all hand dominance					
strong right	51.1	44.8	51.5	54.3	
strong left	1.1	5.2	1.5	2.9	—
remainder	47.8	50.0	47.0	42.8	
Knowledge of right and left					
normal	54.5	44.8	50.0	48.6	
hesitant	9.0	27.6	10.6	10.0	a/b
confused	36.5	27.6	39.4	41.4	
Hand/eye combination					
right hand/right eye	48.9	41.4	51.5	55.7	—
remainder	51.1	58.2	48.5	44.3	

TABLE 3.12

Reading Achievement Tests at age seven years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant difference p<0.05*
Schonell quotient	93.8	92.3	96.7	91.4	—
standard deviation	18.9	17.8	19.7	19.4	
Holborn quotient	93.2	90.8	94.7	90.0	—
standard deviation	17.9	16.9	17.8	17.8	

N.B. No significant differences between any two groups in either test.

TABLE 3.13

Junior Eysenck Personality Inventory at age seven years

	Random control (a)	*Short gestation (b)*	*Rather light-for-dates (c)*	*Very light-for-dates (d)*	*Significant differences p<0.05*
Extraversion score	16.7	16.5	16.2	16.6	—
Neuroticism score	12.2	12.7	13.0	12.6	—
Lie score	7.6	8.2	7.9	8.5	a/d
standard deviation	2.8	2.5	2.6	2.2	

Reading Achievement Tests: Schonell and Holborn

These two tests of the ability to read aloud differ in that the Schonell Test presents the child with single words (Schonell and Schonell 1965) and the Holborn Test presents him with sentences (Watts 1970). The over-all results are summarised in Table 3.12, and are very surprising in that this is the first of our series of tests in which there are no significant differences between the scores of any of our groups. We can only speculate that these two tests may not be sensitive enough to detect impairment of performance at the age of seven years. It is noteworthy that our whole population was underachieving in comparison with the standardisation samples originally used for developing the two tests, and all to a very similar extent. In terms of 'reading age', from which the quotients are calculated, our children's performance was an average six to nine months behind their chronological age.

Junior Eysenck Personality Inventory

This inventory was introduced by Sybil Eysenck in 1965 as an application to the child population (aged seven to 16 years) of the well-known inventory designed to evaluate in adults the two major personality factors of extroversion/introversion and neuroticism/stability (Eysenck and Eysenck 1964). There is also a Lie Scale, designed to measure the extent of any tendency to self-misrepresentation. Although the manual provides norms for children down to the age of seven years, we feel bound to express some reservations about the validity of the results at this age because of the children's difficulty in understanding some of the questions. However, since the mean ages of the children in our comparison groups were the same, we feel justified in making comparisons between their mean scores.

The results we obtained are summarised in Table 3.13, and it is clear that there are no significant differences between the groups in the main factors tested, but a (just) significantly higher Lie Score was recorded for the very light-for-dates group compared with the controls.

Discussion

The results of the battery of psychological tests look like a clear *prima facie* case for concluding that the performance of both extreme abnormal groups was inferior to that of the control group. In no single test, or even subtest, was either group's score significantly superior to that of the control group: and in the great majority of the tests the majority of their scores, even in subtests, was significantly inferior. Furthermore, the patterns of the differences for the two abnormal groups are remarkably similar, with only occasional significant differences between their subtest scores and only one whole test in which they differ significantly (to the disadvantage of the light-for-dates as compared with the short-gestation group). The tests in which there was a significant difference between one of the extreme abnormal groups and the control group, or between the two extreme groups themselves, are listed in Table 3.14. If we compare these lists with that of the tests administered (Table 3.1, p. 19) we can identify the tests in which no significant differences were found between either of the extreme abnormal groups and the controls. There were none at either five or six years of age; and at the age of seven years there were only the Purdue pegboard test of

manipulative dexterity and the two tests of reading achievement (Schonell and Holborn). We have not listed the individual subtest scores in which there were significant differences between the two extreme abnormal groups, but have referred to them in discussing each test. A large number of subtests was administered, and it must be remembered that a significant difference would be expected to occur by chance in one of every 20 subtests.

However, we cannot yet attribute the differences we have listed directly to the abnormalities of intra-uterine growth in which we are interested. We must try to allow for the effects of associated factors before reaching any such conclusions.

The first factor which we feel it necessary to investigate is the possibility that the mean gestational age difference of 37 days between the short-gestation and control groups may account for a significant proportion of the inferior performance of the former. We have looked into this possibility in the case of two tests in which the postnatal age is used to calculate a quotient, which were administered at the age of 5½ years (when the 37 days represents a bigger proportion of the child's age than later). Two correction procedures could be adopted. We could make a correction to each individual child's age, but since the standardisation range for screening purposes at this age is 5 years 3 months to 5 years 9 months for the ITPA and 5 years 6 months to 5 years 8 months for the Frostig Test, the mean correction of 37 days could make no significant difference. A more stringent procedure would be to make an over-all mean correction, but this would only raise the ITPA quotient from 91.5 to 93.2 and the Frostig from 96.1 to 97.9, neither of which is a significant change ($p<0.05$). We have therefore decided not to make any age correction in our analyses.

But the possible 'confusing' effect of associated biological and social factors clearly cannot be ignored, even at this preliminary stage of our analysis. We have therefore selected one factor of each type, which we consider likely to prove significant and interesting on the evidence already available to us, and we have investigated the effect of each of these factors upon examples of each of the different types of assessment, the results of which are reported in this and the next three chapters.

From the biological factors listed in Table 2.2 (p. 12), we have selected the child's sex as suitable for our purpose. In addition to the generally accepted differences between the performance of boys and girls, which we mentioned in discussing that Table, we have evidence that boys may be more vulnerable than girls to the later effects of an adverse perinatal factor such as breech delivery (Neligan *et al.* 1974). The well-known finding of a higher neonatal mortality rate among boys than girls tends to favour the possibility that this differential vulnerability has an organic rather than a social basis. Moreover, in the sphere of physical growth Tanner (1970) has shown that females recover more rapidly than males from the effects of transient adverse factors. The results of investigating the sexes separately are shown in Tables 3.15 and 3.16 in which we have not subdivided the light-for-dates group in order to avoid the problems resulting from very small numbers of children in some of the cells.

The finding of such a significantly better performance by the boys than the girls in the control group, in terms of both over-all intelligence and language quotients, is rather unexpected. Among 150 Isle of Wight children aged five years, Yule *et al.* (1969) found no over-all sex differences in the WPPSI; but among 60 Nottingham

TABLE 3.14

Tests in which there was a significant difference between groups

Age (years)	*Short gestation cf. Control*	*Very light-for-dates cf. Control*	*Very light-for-dates cf. Short gestation*
5	ITPA Frostig Purdue-dominant	ITPA Frostig Purdue-both	
6	WPPSI Skemp Purdue-dominant Intersensory	WPPSI Skemp Purdue-dominant -non-dominant Intersensory	Skemp
7	ITPA Ozeretsky Laterality (knowledge of)	ITPA Intersensory Bender-Gestalt Ozeretsky Junior Eysenck Lie Score	

TABLE 3.15

Wechsler Preschool and Primary Scale of Intelligence related to sex of child at age six years

	Random control (a)	*Short gestation (b)*	*Light-for-dates (c)*	*Significant difference* $p<0.05$	$p<0.01$
Full-scale IQ of					
Boys (1)	103.0	92.1	93.9	—	a/b, a/c
Girls (2)	96.1	92.6	91.8	—	—
Significant differences					
$p<0.05$	—	—	—		
$p<0.01$	1/2*	—	—		

*Difference between boys and girls.

TABLE 3.16

Illinois Test of Psycholinguistic Abilities related to sex of child at age five years

	Random control (a)	*Short gestation (b)*	*Light-for-dates (c)*	*Significant difference* $p<0.05$	$p<0.01$
Language quotient of					
Boys (1)	100.7	90.3	92.9	a/c	a/b
Girls (2)	95.9	93.2	90.6	a/c	—
Significant difference					
$p<0.05$	1/2				

TABLE 3.17

Wechsler Preschool and Primary Scale of Intelligence related to social class at age six years

	Random control (a)	*Short gestation (b)*	*Light-for-dates (c)*	*Significant difference* $p<0.05$	$p<0.01$
Full Scale IQ of					
classes I and II (1)	115.8	(97.5)*	104.6	a/c	—
class III (2)	97.6	97.0	94.3	—	—
classes IV and V (3)	99.1	87.2	86.4	—	a/b, a/c
Significant differences					
$p<0.05$	—	2/3	1/2, 2/3		
$p<0.01$	—	—	—		
$p<0.001$	1/2, 1/3	—	1/3		

TABLE 3.18

Illinois Test of Psycholinguistic Abilities related to social class at age five years

	Random control (a)	*Short gestation (b)*	*Light-for-dates (c)*	*Significant difference* $p<0.05$	$p<0.01$
Language quotient of					
classes I and II (1)	110.6	(94.0)*	98.2	—	—
class III (2)	96.6	94.5	93.1	—	—
classes IV and V (3)	98.1	88.2	87.1	—	a/c
Significant differences					
$p<0.05$	—	—	2/3		
$p<0.01$	1/2, 1/3	—	—		

*There were only 2 children in Social Classes I and II in the short gestation group.

children, Brittain (1969) found a slight but consistent difference in mean subtest scores in favour of the boys. In the ITPA, Mittler and Ward (1970) found no over-all sex differences among 100 Manchester four-year-olds, "although girls show a clear tendency towards higher language scores"; and in American children, McCarthy and Kirk (1963) reported that "the battery as a whole did not markedly favour either sex, or at best favours girls slightly". However, the suggestion that boys may be more vulnerable than girls to harmful effects of the abnormalities of intra-uterine growth in which we are interested appears to be confirmed in two ways:—

(1) The significant gradient in favour of the control boys is not seen in any of the four comparisons that can be made within the abnormal groups.

(2) The mean scores of the children in the abnormal groups are significantly worse than those of the controls in all four instances in which the comparisons can be made among the boys, but in only one instance among the girls.

From the social factors listed in Tables 2.3 (p. 12) and 2.4 (p. 13), we have selected the social class derived from the father's occupation as most suitable for our present purpose. Although open to criticism on the grounds of being too non-specific,

this factor was described by Birch *et al.* (1970) as "perhaps the most effective single criterion available and the one most frequently used in socio-medical research". If effects can be demonstrated in this way, then the more specific and meaningful components of the global classification can be sought. The results of our analyses, again in terms of both intelligence and language quotients, are summarised in Tables 3.17 and 3.18. It appears that there are more significant differences between the scores of social-class groups (nine altogether) than between those of intra-uterine growth groups (four), but the cumulative effect of the two factors is such that there is nearly 30 points difference in the WPPSI Test between random control children in classes I and II and very light-for-dates children in classes IV and V, and over 20 points difference in the ITPA.

Our general conclusions concerning the results reported in this chapter are that the two abnormalities of intra-uterine growth in which we are primarily interested are both associated with impaired performance over a wide range of cognitive and sensorimotor functions at the age of five to seven years, but the adverse effects of being born too small are more severe in degree (though similar in kind) than the effects of being born too soon. That there is an organic basis for the impaired performance of the very light-for-dates group is suggested by the consistent finding of a gradient between the control, the rather and the very light-for-dates groups. But the modifying effects of the child's sex, and the even more striking effects of the social class of origin, make it clear that relevant associated biological and social factors will need to be allowed for in some sort of multivariate analysis before reaching any firm conclusions concerning the question of primary interest.