ASPECTS OF INTELLIGENCE

Andrew M. Colman

Contents

1 Introduction: individual differences and the basics of intelligence testing 323
  1.1 What is intelligence? 324
  1.2 Correlation 325
  1.3 Galton and Binet 326
  1.4 The concept of the intelligence quotient 328
  1.5 The normal distribution 329
  1.6 Conversion to IQ scores 331
  1.7 Advantages of the statistical approach 332
      Summary of Section 1 332

2 Modern psychometric approaches to intelligence 333
  2.1 Wechsler scales 333
  2.2 British Ability Scales 336
  2.3 Raven's Progressive Matrices and the Mill Hill Vocabulary scale 337
  2.4 General intelligence and specific factors 338
  2.5 How accurate are IQ tests? 341
  2.6 Does intelligence decline in later life? 342
      Summary of Section 2 344

3 The nature-nurture controversy 344
  3.1 Heritability of IQ 345
  3.2 Separated identical twins 347
  3.3 Family studies 350
  3.4 Adoption studies 351
  3.5 Conclusions 353
      Summary of Section 3 354

4 Race and intelligence 354
  4.1 Racial admixture studies 356
  4.2 Racial crossing studies 357
  4.3 Test bias 358
  4.4 Environmental influences on IQ 360
      Summary of Section 4 361

5 Cognitive styles 361
  5.1 Field dependence–independence 361
  5.2 Convergence–divergence 363
  5.3 Concluding comments 365
      Summary of Section 5 366
      Further reading 366
      References 367
      Answers to SAQs 371

Further reading 366
References 367
Answers to SAQs 371
1 Introduction: individual differences and the basics of intelligence testing

People differ from one another, not only physically but also psychologically. Some people are self-assertive while others are submissive, some are persevering while others are fickle, some are outgoing while others are shy, some are trusting while others are suspicious, and so on. Allport and Odbert (1936) carried out a dictionary search and found 4500 distinct adjectives in the English language denoting psychological differences between people. Each of these adjectives relates to a personality trait, a more or less consistent pattern of behaviour that a person possessing the trait would be likely to display in certain circumstances. When we describe someone as shy, for example, we implicitly assume that he or she would probably show signs of embarrassment when meeting strangers.

The study of individual differences in psychology arises from the fact that different people do behave differently, in ways that are to some degree consistent and predictable, not only when meeting strangers, but also in many other circumstances. The ancient Greeks were the first to notice these differences and to attempt to explain them. Their doctrine of the four temperaments, which was widely accepted for many centuries, sought to explain individual differences in terms of the mixture of four fluids or humours in people's bodies. Optimistic or sanguine people were thought to have in their bodies a predominance of blood (sanguis), depressive or melancholic people an excess of black bile (melaina chole), short-tempered or choleric people an excess of yellow bile (chole), and apathetic or phlegmatic people a predominance of phlegm (phlegma). Although the physiological basis of the doctrine of the four temperaments was undermined by biological research during the Renaissance, the classification has survived in a modified form in some modern personality theories. This is discussed in greater detail in Chapter 8.

The first systematic study of individual differences using modern research methods was Francis Galton's study of intelligence in England in 1884. Since then psychologists have devoted particular attention to individual differences in intelligence and thinking. These differences have been more thoroughly investigated than any others in psychology, partly because of their uniquely important effects on people's educational prospects and prospects in life generally, and partly because of their controversial social implications.

In this chapter, some of the fundamental ideas and research findings related to individual differences in intelligence and thinking will be discussed. This will provide you with an introduction to one of the most important classes of individual differences in psychology.

ACTIVITY 1

Think of someone who seems to you to be very intelligent. List the person's qualities that led you to this judgement. Did you choose someone who is academically clever? How would you define intelligence? Are there other kinds of intelligence?
1.1 What is intelligence?

Intelligence can be defined informally as intellectual ability. A person who solves a difficult crossword puzzle quickly or gives the right answer to a tricky mathematical problem or gets a high score on an IQ (intelligence quotient) test is showing intelligent behaviour, and it is reasonable to infer that such a person is intelligent. Someone who does badly at the same tasks is not showing intelligent behaviour and may have a low intelligence, but the inference is uncertain in this case because other explanations are possible. Poor performance, even on an IQ test, might be due to tiredness, lack of interest or motivation, test anxiety, or many other causes apart from low intelligence.

Until fairly recently, psychologists who devised IQ tests tended to base their definitions of intelligence on their own preconceptions about intellectual ability and the types of behaviour associated with it. By the early 1920s there were almost as many different definitions of intelligence in the psychological literature as there were psychologists writing about intelligence. Some of the early expert definitions are shown in Box A.

In 1981, the American psychologist Robert J. Sternberg and his colleagues asked a large group of experts to rate many different kinds of behaviour according to how characteristic of intelligence they considered each one to be (Sternberg et al., 1981). A statistical technique called factor analysis, which will be explained in Section 2.4, was used to search for common themes.

---

**Box A  Early definitions of intelligence**

A famous symposium on ‘Intelligence and its measurement’ was published in the *Journal of Educational Psychology* in 1921. Fourteen experts gave their own informal definitions of intelligence, some of which may be paraphrased as follows:

1. The ability to carry out abstract thinking (L.M. Terman).
2. The ability to give responses that are true or factual (E.L. Thorndike).
3. The capacity to inhibit instincts, coupled with analytical ability and perseverance (L.L. Thurstone).
4. The ability to acquire abilities (H. Woodrow).
5. The ability to learn or to profit by experience (W.F. Dearborn).
6. The ability to adjust oneself to relatively new situations in life (R. Pinter).
7. The ability to adjust oneself to the environment (S.S. Colvin).
8. The capacity for knowledge and knowledge possessed (V.A.C. Henmon).

After 1921, many other informal definitions appeared in the psychological literature (see Miles, 1957, for an interesting summary and discussion).
Three factors which emerged from the analysis were: **verbal intelligence**, **problem solving** and **practical intelligence**. Sternberg interpreted these as the major **components of intelligence**, at least according to expert opinion.

Sternberg and his colleagues also showed that experts and non-experts have remarkably similar conceptions of intelligence. When ordinary people were asked to rate the same kinds of behaviour as the experts, there was almost complete agreement about how characteristic each one is of intelligence. In technical terms, the correlation between the two sets of ratings was 0.96, which is very high. The technique of correlation is often used in the study of intelligence and in other branches of psychology, so it is worth pausing briefly to explain it.

### 1.2 Correlation

The technique of **correlation** is simply a method of assessing whether, and to what extent, one measure varies together with another. Two measures that are related to each other are co-related (hence correlated), and the statistical technique for determining how strongly they are correlated is called **correlation**. As a child grows, for example, his or her arms and legs get longer, and this simultaneous change in the same direction is called **positive correlation**. Two unrelated measures that vary independently of each other are said to be **uncorrelated**. The length of a child’s arms and the amount of rain that falls in Singapore, for example, are uncorrelated. If high scores on one measure tend to go with low scores on the other, such as the age of a preschool child and the number of grammatical errors he or she makes, then there is a **negative correlation** between the two measures.

The usual **index of correlation** ranges from zero (0.00), for uncorrelated measures, up to +1.00 for a perfect positive correlation and down to −1.00 for a perfect negative correlation. The further a correlation is from zero, the more closely the two measures are related to each other, either positively or negatively.

The high positive correlation of 0.96 that Sternberg and his colleagues found shows that the experts and non-experts tended to rate the various kinds of behaviour very similarly, and this suggests that their conceptions of intelligence were very similar. There were two slight differences. First, only the experts considered motivation to be an important ingredient of intelligence; second, the non-experts attached more importance than the experts to social aspects of intelligence, such as the ability to make witty remarks and to understand jokes.

(SAQ answers are given at the end of the chapter) Try to estimate, on the basis of common sense, whether the correlation between each of the following pairs of measures is highly or moderately positive, close to zero, or negative:

(a) height and weight among adults;
(b) height and age among children;
(c) alcohol intoxication and manual dexterity.
1.3 Galton and Binet

Francis Galton, a cousin of the biologist Charles Darwin, constructed the world’s first intelligence test in England more than a century ago. He also carried out the first empirical studies designed to determine the extent to which differences in intelligence are (1) hereditary or (2) due to the different environments in which people grow up. In the firm belief that hereditary factors are overwhelmingly important, he founded the eugenics movement which aimed to improve the hereditary stock of the human population by selective breeding; that is, by encouraging intelligent people to have more children than less intelligent people. He was also the first psychologist to suggest that racial groups differ in innate intelligence.

Galton’s intelligence test was based on his theory about the mental processes involved in thinking, reasoning and problem solving. He believed that mental ability depends on the capacity to perceive subtle differences, so that, as he put it, ‘the more perceptive the senses are of difference, the larger is the field upon which our judgement and intelligence can act’ (Galton, 1883). Galton therefore assumed that measures of sensory discrimination, the ability to detect small differences through the sense organs, should be good tests of intelligence.

In 1884, at the International Health Exhibition in the South Kensington Museum, Galton set up a stall where visitors could have their mental abilities tested for 3 pence, and more than 9000 men and women took up his offer. This was the first and one of the largest studies of intelligence ever undertaken. It was also unusual, to say the least, that the subjects paid the researcher for their participation rather than the other way round. Galton’s tests measured reaction times (the speed with which people can react to signals), visual and auditory discriminations (the smallest differences in the lengths of lines and the pitches of musical notes that they can detect), touch sensitivity (the minimum distance between a pair of pinpricks that they can feel as separate rather than as a single pinprick), and various other sensory and motor functions (e.g. the maximum number of taps they can make with a stylus in a minute).

It soon became clear that Galton’s theoretical approach to the measurement of intelligence was misconceived. In particular, the American psychologist Clark Wissler reported in 1901 that the various sensory and motor tests did not correlate with one another when he tried them out on college students (Wissler, 1901). People who scored very highly on one test did not necessarily score highly on the others and vice versa. Galton’s assumption that all the tests measured the same general intellectual ability could not, therefore, be right. Worse still, Wissler reported that none of the tests correlated with students’ examination marks. On the other hand, marks in different subjects, such as mathematics and English, did correlate with one another, which was in line with the commonsense assumption that intelligence plays a part in academic achievement in all subjects. Wissler concluded that the sensory and motor tests were not good measures of intelligence, and Galton’s sensory–motor approach to the measurement of intelligence was abandoned in favour of direct tests of reasoning ability.
The first useful intelligence test was developed by Alfred Binet and Theodore Simon in France in 1905. Binet had been asked by the Minister of Public Instruction to find a means of detecting mentally subnormal children and classifying them as 'débiles' or 'morons' (mildly subnormal), 'imbeciles' (severely subnormal), or 'idiots' (profoundly subnormal), with a view to putting them in special classes or schools in which teaching could be geared to their abilities. Intelligence tests are still used for that purpose, although the terms used to denote mental handicap are now less offensive.

Binet worked out a commonsense, pragmatic way of measuring intelligence, in sharp contrast to Galton's earlier theoretical approach. Binet's was a commonsense approach in that it assumed that the best way to measure a child's intelligence is to test the child's ability to follow instructions, to exercise judgement and to solve a wide variety of problems generally associated with intelligence. The final version of the Binet–Simon scale contained fifty-four items arranged in order of difficulty, from following the movement of a lighted match with the eyes, through naming parts of the body and counting backwards from twenty, to working out what time a clock face would show if the hour and minute hands swapped places.

Without looking at your watch, try to work out what time a clock face would show if the hour and minute hand swapped places at 1.55.

Binet's method was pragmatic in that items were selected for the scale according to how well they correlated with an independent criterion known to be associated with intelligence. The problem, of course, was to find a suitable independent criterion for intelligence. You might think that it is impossible to find such a criterion, because if one existed there would be no point in constructing intelligence tests, but you would be wrong. Binet was the first to realize that age is an ideal criterion. It is an indisputable fact that as children grow older their ability to solve problems generally increases.

Binet and Simon administered the items of their intelligence test to representative groups of children of various ages (standardization samples). They found that on many items the older children performed better than the younger children. This was good evidence that these items were indeed measuring intelligence. On this basis they were able to establish test norms. The idea behind test norms is to provide standards of 'normal' or average ability against which other children's scores can be compared.

Binet and Simon established norms for each age group by recording the average number of items that children at each age level could pass. This enabled them to introduce the concept of mental age. If, for example, a 10-year-old child could pass only those items that 6-year-olds in the standardization sample could pass, then it was reasonable to infer that the child was functioning at a mental age of 6 and was therefore retarded by about 4 years. Binet defined 'morons' ('débiles') as people whose mental age had not increased beyond the norm for 11-year-olds, 'imbeciles' as those who had developed only up to a mental age of 5 years, and 'idiots' as those whose mental development stopped short at a mental age of 2 years.
Binet was fiercely critical of other psychologists, such as Galton, who considered intelligence to be a fixed and hereditary quantity. 'We must protest and react against this brutal pessimism', he wrote in 1909. 'With practice, enthusiasm, and especially method, one can succeed in increasing one’s attention, memory, and judgement, and becoming literally more intelligent than one was before' (Binet, 1909). He even went to the lengths of developing a programme of intellectual exercises to improve mental fitness in the way that physical exercises improve bodily fitness. These ‘mental orthopaedics’, as he called them, were designed to raise the intelligence of mentally retarded children.

(a) What evidence is there that Galton’s tests were not good measures of intelligence?
(b) What evidence is there that Binet’s tests do measure intelligence?
(c) How do Galton’s and Binet’s tests differ in content?

1.4 The concept of the intelligence quotient

The rest of Section 1 describes the techniques used for calculating IQ scores. If you find these difficult, read on and come back after finishing the chapter.

The idea of mental age led to the invention of the intelligence quotient, which is what the letters IQ stand for. In 1912 the German psychologist William Stern pointed out the obvious fact that a person’s mental age tells us nothing about his or her intelligence unless we also know the person’s actual (chronological) age (Stern, 1912). Imagine three 10-year-old children called Anne, Beatrice and Charles. Anne has a mental age of 7 years and is therefore obviously below average intelligence for her chronological age of 10. Beatrice has a mental age of 10 and is therefore of average intelligence. Charles has a mental age of 12 and so is above average intelligence. Stern hit upon the ingenious idea of dividing mental age by chronological age and regarding this quotient, which he called the intelligence quotient, as an index of intelligence.

In symbols, Stern’s quotient is the fraction MA/CA; that is mental age (MA) divided by chronological age (CA). The American psychologist Lewis Terman later introduced the abbreviation IQ for intelligence quotient and suggested multiplying Stern’s fraction by 100 to convert it to a percentage (Terman 1916). The revised concept of the IQ therefore is defined as:

\[ IQ = \frac{MA}{CA} \times 100. \]

According to this formula, Anne, with a mental age of 7 and a chronological age of 10, has an IQ of 7 divided by 10, multiplied by 100, which works out as 70. This means that her mental age is 70 per cent of her chronological age. Beatrice, whose mental and chronological ages are both 10, has an IQ of 10 divided by 10, multiplied by 100, which comes to 100, which is average for her age. Charles, whose mental age is 12 and whose chronological age is 10, has an IQ of 12 divided by 10, multiplied by 100, which comes to 120, so his mental age is 20 per cent higher than his chronological age. The most
important point to notice is that an IQ of 100 is average by definition, so IQs below 100 are below average and IQs above 100 are above average.

(a) Elizabeth, Andrew and William are all 5 years old. On an IQ test Elizabeth passes only those items that an average 4 year old in the standardization sample passed, Andrew passes only those that an average 5 year old passed, and William passes only those that an average 6 year old passed. Use the formula $IQ = \frac{\text{mental age (MA)}}{\text{chronological age (CA)}}$ multiplied by 100 to work out Elizabeth’s, Andrew’s and William’s IQ scores.

(b) Mark is 20 years old and Philip is 40 years old. They both pass only those items that an average 20 year old in the standardization sample passed. Using the formula, calculate their IQs. Does the answer seem fair to Philip when you think about it?

Although IQ is defined in most elementary psychology textbooks according to the mental age/chronological age formula shown above, it is seldom calculated that way today. The main reason for this is that the definition leads to absurdities when it is applied to adults. Beyond the age of about 17 or 18, people do not show any increase in intellectual ability, as measured by IQ tests. Mental age tends to stabilize, although in later life it may even decline (a point we will come to in Section 2.6). Average 40 year olds perform on IQ tests at about the same level as average 20 year olds. Yet, according to the old MA/CA formula, a 40 year old person performing at the level of a 20 year old has an IQ of 20 divided by 40, multiplied by 100, which comes out as an IQ of only 50. This would imply that the 40 year old is severely retarded or, in the new terminology of educational psychology, suffers from ‘severe learning difficulties’. To put it another way, the standard textbook definition of IQ is weighted unacceptably against age and is actually nonsensical when applied to middle-aged and elderly people.

1.5 The normal distribution

In 1939, David Wechsler introduced a purely statistical definition of IQ that avoids the problems of the old mental age/chronological age formula (Wechsler, 1939). His definition is used by virtually all contemporary psychologists who construct IQ tests. The basic idea is straightforward. An intelligence test is given to large samples of people from each age group in the population. Each person is given a test score according to how many items he or she answered correctly; the numbers of people with each possible test score are then counted. Figure 7.1 shows a typical pattern of test scores. For simplicity’s sake, the data in Figure 7.1 are taken from a small sample of twenty people given a test containing just ten items. In practice, much larger samples are needed, and the tests usually include more items.

Listed along the horizontal axis (line) of Figure 7.1 are test scores: 0 items correct, 1 item correct, and so on. Each square in the diagram represents a person who achieved the test score shown directly below it. The number of
squares above each test score represents the number of people who achieved that test score: two people scored 3, for example, and three people scored 4 on the test. Figure 7.1 represents the distribution of test scores in the sample. You will notice that there are more people with scores near the middle (test scores of 5 or 6) than there are with very low or very high test scores.

SAQ 5 In Figure 7.1:
(a) How many people achieved test scores of 5 or 6?
(b) How many people scored more than 7?
(c) How many people scored less than 3?
(d) What is the mean (average) test score of the sample? (Hint: add up the total scores in each column and divide the sum of these total scores by the number of people, i.e. twenty.)

When an intelligence test is given to a very large sample of people, it is found in practice (and there is a mathematical theorem that shows why it happens) that the distribution of scores closely resembles the bell-shaped normal distribution shown in Figure 7.2; the 'height' of the 'bell' represents the number of people with each test score. Here again, most of the test scores are close to the mean (average) (represented by the 'bulge' in the centre of the 'bell'). The further away a test score is from the mean, the fewer the people who achieve that test score.

The normal distribution is symmetrical: 50 per cent of test scores fall below and 50 per cent above the mean. Figure 7.2 also shows the percentage of test scores that fall within certain evenly spaced distances from the mean of the normal distribution. These distances are expressed in terms of a measure of variability called the standard deviation (abbreviated SD). Standard deviation
is a measure of the extent to which scores deviate from the average score. Scores that are one standard deviation from the mean are closer to the average than scores that are 2 standard deviations from the mean. Notice that 34.13 per cent of test scores fall within one standard deviation above the mean, and the same percentage of test scores fall within one standard deviation below the mean.

1.6 Conversion to IQ scores

By convention, the mean of a normal distribution of IQ scores is set at 100. As shown in Figure 7.2, the mean test score of 50 is accordingly converted to a mean IQ of 100 points. Every other test score can be converted to IQ points according to how many standard deviations above or below the mean it falls. Again by convention, each standard deviation is assigned a value of 15 IQ points. A test score of 40 (see Figure 7.2) is one standard deviation below the mean test score of 50, shown by an SD of \(-1\). To convert it to an IQ, we simply take the mean IQ of 100 and subtract 15 (one standard deviation) from it, giving a calculated IQ of 85 points. Similarly, a test score of 60 is one standard deviation above \(+1\) the mean test score of 50, so the equivalent IQ is 100 plus 15, which is 115 points. One last example: a test score of 20 is 3 standard deviations below \(-3\) the mean test score. The equivalent IQ is therefore 100 minus 45, which is 55 points.
In Figure 7.2, what is the equivalent standard deviation and calculated IQ of the following test scores: (a) 80; (b) 30; and (c) 70?

1.7 Advantages of the statistical approach

The distributions in Figures 7.1 and 7.2 each represent the test scores of a particular age sample. If they were scores of 8 year olds, for example, then the test score of any particular 8 year old could be converted to an IQ by the statistical method based on standard deviations outlined in Section 1.6.

There are two main reasons why this method of calculating IQ scores is preferable to the old formula of mental age divided by chronological age, multiplied by 100. First, it eliminates the absurdities that arise when the old MA/CA method is applied to adults. You will remember that 40 year olds who achieve the same test score as average 20 year olds end up with IQs of only 50 when the old MA/CA formula is used, although their performance is, in reality, average for their age. The statistical method involves comparing a person’s test score with the average for people in the same age group. This means that since the average scores of these age groups are the same, the test scores of both 20 year olds and 40 year olds will then be compared against the same mean.

The second advantage of the statistical approach is that an IQ score calculated in this way indicates a person’s intelligence relative to his or her peers in a way that is easy to understand. A person with an IQ of 130, for example, is in the top 2½ per cent for that age group, as is clear from Figure 7.2.

Summary of Section 1

- Intelligence can be defined informally as intellectual ability, although this can cover many different types of thinking, including verbal intelligence, problem solving and practical intelligence, which were interpreted by Sternberg as components of intelligence.
- The technique of correlation is used to measure the degree to which abilities are related to each other.
- Two measures that relate to each other are correlated; two unrelated measures that vary independently of each other are uncorrelated.
- The intelligence quotient (IQ) was originally defined by a formula, devised by Stern, comparing mental age and chronological age.
- These days, IQ is measured by comparing people’s performance on intelligence tests with norms for their own age groups, a purely statistical definition of IQ introduced by Wechsler in 1939.
Psychometrics (from the Greek psyche, meaning 'mind', and metron, meaning 'measure') focuses on the measurement of intelligence. Psychometric comparisons between people with different IQ scores have dominated the study of intelligence ever since the turn of the century. Research into the nature of intelligence and the psychological processes involved in thinking were rather neglected until comparatively recently, and some of the developments in this area will be discussed in Section 5.

In 1916, Lewis Terman of Stanford University translated and adapted the Binet–Simon scale for use in the United States, and the Stanford–Binet scale, as it became known, was the prototype of nearly all subsequent IQ tests. As we saw in Section 1.3, in constructing tests for children, Binet had used the independent criterion of age as a basis for choosing suitable test items. This was a reasonable thing to do because older children are obviously better than younger children at following instructions, exercising judgement and solving problems. However age cannot be used as an independent criterion of intelligence in constructing tests for adults as we saw in Section 1.4, and the problem was one of how to select test items which would stretch the abilities of adults and thus reflect their true level of intelligence.

Psychologists solved this problem simply by extending or extrapolating the tests. Terman, for example, included a vocabulary subtest in his Stanford–Binet scale and used the method of extrapolation to make it applicable to adults. When testing children, he found that 6-year-olds could generally define common words like orange, older children could define less common words like lecture, and only teenagers could define uncommon words like philanthropy. He therefore extended or extrapolated this scale so that it could be used with adults by including very uncommon words like sudorific (which means sweat-producing). A similar method of extrapolation can be used with other types of test items. The commonsense assumption is that intelligent adults should be able to cope with more difficult items than less intelligent adults.

The Stanford–Binet scale is still used by researchers and educational psychologists. The most recent version of it, published in 1985, is unusual among IQ tests in so far as it is explicitly designed to minimize sexual and ethnic bias in its questions. There are many other intelligence tests available today, and the ones most often used are the Wechsler scales, the British Ability Scales, Raven’s Progressive Matrices and the Mill Hill Vocabulary scale, which are described below in Sections 2.1, 2.2, and 2.3, respectively.

2.1 Wechsler scales

The Wechsler scales are based on an adult IQ test originally developed by the American psychologist David Wechsler in 1939. There are now three different Wechsler scales for use with different age groups: the Wechsler Adult
Intelligence Scale (revised) or WAIS-R, the Wechsler Intelligence Scale for Children (revised) or WISC-R, and the Wechsler Preschool and Primary Scale of Intelligence or WPPSI. Each of these scales has to be administered individually by a trained tester and each yields three separate scores: verbal IQ, performance IQ and full scale IQ. Some simulated items of the type used in the verbal and performance subtests of the Wechsler Intelligence Scale for Children (revised) are shown in Box B.

The items that make up the Wechsler scales are remarkably diverse and many of them—for example, the questions about the writer of Tom Sawyer and the meaning of ‘flexible’ in Box B—seem to be measures of knowledge or memory rather than of thinking ability. Wechsler’s view was that knowledge reflects intelligence because intelligence is needed to acquire knowledge. It is certainly possible to criticize some of the items. Perhaps the most questionable items are in the Comprehension subtest which often asks for judgements based on general knowledge. Even so, each of the Wechsler subtests has been shown to be positively correlated with each of the others. This means that people who get high scores on one subtest tend to get high scores on the others, and vice versa. The correlations between various subtests of the Wechsler Adult Intelligence Scale (revised) are shown in Table 7.1.

Table 7.1 is called a correlation matrix. It shows the correlations between people’s scores on each subtest and each other subtest. The names of the subtests are listed down the side and in the same order across the top. The first column, for example, contains the correlations between scores on the Information subtest and scores on each of the other subtests. There is a dash rather than a correlation at the top of the first column, in the row labelled Information, because the correlations in the matrix are between scores on different subtests. The second row, labelled Digit Span, shows a correlation of 0.46 between the Digit Span and the Information subtest. The third row, labelled Vocabulary, shows a correlation between scores on the Vocabulary

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Information</th>
<th>Digit Span</th>
<th>Vocabulary</th>
<th>Arithmetic</th>
<th>Comprehension</th>
<th>Similarities</th>
<th>Picture Completion</th>
<th>Picture Arrangement</th>
<th>Block Design</th>
<th>Object Assembly</th>
<th>Digit Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Digit Span</td>
<td>0.46</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.81</td>
<td>0.52</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>0.61</td>
<td>0.56</td>
<td>0.63</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Comprehension</td>
<td>0.68</td>
<td>0.45</td>
<td>0.74</td>
<td>0.57</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Similarities</td>
<td>0.86</td>
<td>0.43</td>
<td>0.72</td>
<td>0.56</td>
<td>0.8</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>0.52</td>
<td>0.37</td>
<td>0.55</td>
<td>0.48</td>
<td>0.52</td>
<td>0.54</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>0.50</td>
<td>0.37</td>
<td>0.51</td>
<td>0.46</td>
<td>0.49</td>
<td>0.50</td>
<td>0.51</td>
<td>0.47</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Block Design</td>
<td>0.40</td>
<td>0.43</td>
<td>0.32</td>
<td>0.56</td>
<td>0.48</td>
<td>0.51</td>
<td>0.34</td>
<td>0.47</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Object Assembly</td>
<td>0.39</td>
<td>0.33</td>
<td>0.41</td>
<td>0.42</td>
<td>0.40</td>
<td>0.43</td>
<td>0.52</td>
<td>0.40</td>
<td>0.63</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Digit Symbol</td>
<td>0.44</td>
<td>0.42</td>
<td>0.47</td>
<td>0.45</td>
<td>0.44</td>
<td>0.46</td>
<td>0.42</td>
<td>0.39</td>
<td>0.47</td>
<td>0.38</td>
<td>—</td>
</tr>
</tbody>
</table>

(Source: data adapted from Wechsler, 1981, p. 46)
CHAPTER 7 ASPECTS OF INTELLIGENCE

(Answers in brackets)

**Verbal subtests**

**Information**
- How many wings does a bird have? (Two)
- What is steam made of? (Vaporized water)
- Who wrote Tom Sawyer? (Mark Twain)

**Vocabulary**
- What is a desk? (A writing table)
- What does 'flexible' mean? (Easily bent)

**Arithmetic**
- Sam had three pieces of candy and Joe gave him four more. How many pieces of candy did Sam have altogether? (Seven)
- Three women divided eighteen golf balls equally among themselves. How many golf balls did each person receive? (Six)
- If two buttons cost 15p, what will be the cost of a dozen buttons? (90p)

**Comprehension**
- What should you do if you see someone forget his book when he leaves a restaurant? (Run after him with it)
- What is the advantage of keeping money in a bank? (Security)
- Why is copper often used in electrical wires? (It is a good conductor)

**Similarities**
- In what way are a lion and a tiger alike? (They are both animals)
- In what way are a saw and a hammer alike? (They are both tools)
- In what way are an hour and a week alike? (They are both periods of time)

**Performance subtests**

**Picture completion**
The subject points out the missing elements in several line drawings; for example, a wheel missing from a pram.

**Picture arrangement**
The subject arranges a number of pictures into a sequence so that they tell a meaningful story.

**Block design**
The subject arranges painted wooden blocks to copy designs formed by the examiner.

**Object assembly**
The subject fits a number of shapes like jigsaw pieces together to make recognizable objects.

(Adapted from The Psychological Corporation, 1974)
and the Information subtests of 0.81 and a correlation between the Vocabulary and Digit Span subtests of 0.52. There are no correlations in the top right of the matrix because they would merely repeat what is in the bottom left. The correlation between Information and Digit Span is missing from the first row, for example, because it is the same as the correlation between Digit Span and Information shown in the second row.

If there were no relation between the Block Design subtest (a non-verbal test of abstract problem solving) and the Information subtest (a verbal test of knowledge), to pick two of the most dissimilar subtests, then the correlation between these two would be zero. In fact, Table 7.1 shows that it is 0.50, which is higher than the correlation of 0.47 that exists between the heights and weights of adult men in Britain (Knight, 1984, p. 29). Most of the correlations between the Wechsler subtests are above 0.50, and this is surprising and inexplicable to anyone who doubts the existence of general intelligence, suggesting as it does that there is a global mental ability underlying performance on all the subtests, as we shall see in Section 2.4.

In Table 7.1, what are the correlations between the following subtests:
(a) Vocabulary and Information?
(b) Vocabulary and Comprehension?
(c) Comprehension and Information?
(d) Object Assembly and Block Design?

2.2 British Ability Scales

The British Ability Scales (BAS), first published in 1979, are a set of twenty-three tests designed to measure an even wider diversity of mental abilities than are the Wechsler scales. The BAS yields visual IQ, verbal IQ and general IQ scores. It can be used with children and adolescents up to 17 or 18 years of age. Its subtests can be classified according to the following six mental processes: speed of information processing, reasoning, spatial imagery, perceptual matching, short-term memory, and retrieval and application of knowledge. Speed of information processing, which is supposed to underlie performance on all other subtests, is the most unusual component of the BAS. Its presence reflects the influence of modern information-processing approaches to intelligence, to which you will be introduced in the chapters in Part V.

The various subtests of the BAS, like those of the Wechsler scales, each correlate significantly with each of the others, and general IQ scores derived from the BAS correlate quite highly with independent measures of scholastic and academic attainment (Elliott, 1983, chapters 9 and 10). Correlations between general IQ and school tests of mathematical attainment range from 0.62 to 0.81 for children of different ages. Correlations between general IQ and tests of reading attainment range from 0.67 to 0.72. Lastly, correlations between the British Ability Scales' general IQ and 'O' level (now GCSE) grades in various subjects range from 0.24 (English literature) to 0.63.
(mathematics). These correlations suggest once again that different measures of intellectual performance tend to be related and that there is a general intellectual ability underlying IQ test scores.

2.3 Raven’s Progressive Matrices and the Mill Hill Vocabulary scale

Raven’s Progressive Matrices, first published in 1938 and revised several times since then, is a non-verbal test designed to measure abstract reasoning ability through the use of meaningless geometric diagrams. SAQ 9 asks you about one of the easiest items from the test.

In Figure 7.3, which of the numbered diagrams at the bottom would complete the arrangement at the top?

The Raven’s Progressive Matrices test is designed to cover a very wide range of mental ability and to be usable with subjects irrespective of age, sex, nationality, or education. Scores on Raven’s Progressive Matrices correlate

![Figure 7.3 An easy item from Raven's Progressive Matrices](Source: Raven et al., 1983, item A5)
highly with those on other intelligence tests. Correlations with the Wechsler scales, for example, range from 0.54 to 0.86, and correlations with measures of educational attainment, such as school grades and examination results, are generally between 0.20 and 0.60. Many psychologists believe that Raven’s Progressive Matrices provide the purest available measure of general intelligence, uncontaminated by cultural and educational influences. But research to be discussed in Section 4 casts doubt on that belief.

The Mill Hill Vocabulary scale, published in its original form in 1944 as a verbal companion to Raven’s Progressive Matrices, consists of a list of eighty-eight words divided into two sets of forty-four. The subject’s task is to explain the meanings of the words or (in an alternative form of presentation) to select the correct synonym for each word from a list of six alternatives provided. Most children with a mental age of 5 years can explain the meanings of the first few words (cap, loaf, unhappy), and between the ages of 5 and 16, children of average intelligence can usually define about three additional words per year. The most difficult words in the list, which only a small minority of adults can define, include recondite, exiguous and minatory. The correlation between scores on the Mill Hill Vocabulary scale and Raven’s Progressive Matrices is 0.75, which is remarkably high considering that they are two utterly different ways of measuring intelligence.

There are many other intelligence tests in common use. The most popular are paper-and-pencil, multiple-choice group tests that can be administered to large numbers of subjects at the same time. The types of test items most often featured in these tests, apart from those already described, are analogies, odd-one-out items, and number sequences. Examples of each type are given in SAQ 10.

Can you solve the following (difficult) intelligence test items? They are similar to ones found in paper-and-pencil group IQ tests.

(a) Analogies
C is to X as F is to:
(i) M (ii) U (iii) L (iv) V (v) H.

(b) Odd-one-out
Which of the following is the odd one out?
(i) goal (ii) line (iii) sore (iv) hurt (v) wand.

(c) Number sequences
What is the missing number in the following sequence?
0, 2, 8, 18, ...

2.4 General intelligence and specific factors

Scores on different IQ tests correlate highly with one another, partly because a new test is not considered valid unless it correlates with existing tests. To that extent, validating a new IQ test is rather like hoisting oneself with one’s own bootstraps. But that is not the whole story. The most impressive finding to emerge from decades of research into intelligence is the way in which
different types of tests and subtests correlate with one another no matter how
different they are in content. People who score highly on one test generally
score highly on others, and people who score badly on one tend to score badly
on the rest. It is difficult, in fact, to devise any set of intellectual problems that
are not more easily solved by people with high IQs than by those with low
IQs. People who do not believe that IQ tests measure intelligence are hard put
to explain this well-established fact or to point to anything that does indicate
intelligence in commonsense terms but does not correlate with IQ scores.

Most contemporary psychologists interpret the fact that different IQ tests and
subtests correlate with one another, although not perfectly, as evidence that
they must all be measuring a single common factor, which the British
psychologist Charles Spearman (1927) called $g$ (for general intelligence). This
implies that people are either intelligent, average, or unintelligent across the
board, depending on their levels of general intelligence.

A rival interpretation championed by the American psychologist Louis Leon
Thurstone (1938) focused on the other side of the coin—the fact that the
correlations between IQ subtests are not perfect and that certain groups of test
items usually correlate more highly with one another than with items in other
groups of tests. Thurstone believed that these clusters of correlated items
indicate the existence of seven independent primary mental abilities. He
labelled these Number, Word Fluency, Verbal Meaning, Memory, Reasoning,
Spatial Perception and Perceptual Speed. His assumption was that
intelligence is made up of independent abilities, so that a person may score
highly on some groups of test items and poorly on others. A person with good
Word Fluency, for example, would be expected to score well on all the items
which test that particular ability, but the same person may or may not score
highly on items which test the Perceptual Speed ability.

The statistical technique of factor analysis has been used in an attempt to
settle the argument about general intelligence versus specific abilities. The
details of factor analysis are complicated, but the underlying idea is easy to
grasp. Factor analysis begins with a correlation matrix such as the one shown
in Table 7.1. A mathematical technique is then used to scan the matrix for
clusters of high correlations, indicating groups of tests or subtests that are all
attributable to a common underlying factor. The purpose of factor analysis is
to reduce the matrix of correlations to the smallest possible number of factors
that can account for all the correlations in the matrix. This process is
analogous to that of reducing all the colours of the rainbow to the three
primary colours—red, green and blue—of which they are all composed. If all
the tests or subtests measure the same general intellectual ability, then only
one main factor ought to come out of the analysis. This would be evidence in
favour of Spearman's concept of general intelligence ($g$). If, on the other hand,
the various tests or subtests measure several independent abilities, then
several factors should emerge from the analysis.

When factor analysis is applied in an effort to decide objectively whether
intelligence is a single general ability or a set of independent specific factors,
the results are ambiguous to say the least. It is not difficult to see why this is
so. If Spearman was right, and all tests and subtests measure $g$ with varying
degrees of accuracy, then there should always be very high correlations
between all intelligence subtests, especially between those subtests which are the best measures of intelligence.

If, on the other hand, Thurstone was right and different groups of items measure relatively independent primary mental abilities, there should be clusters of high correlations between subtests measuring the same factor, and correlations close to zero between subtests measuring different factors. One might, for example, find high correlations between subtests involving visual ability, and high correlations between subtests involving verbal ability, but zero correlations between visual and verbal subtests.

In practice, the correlation matrix which shows correlations between IQ subtests usually turns out like the correlations between the WAIS-R subtests shown in Table 7.1. It is difficult to decide whether the pattern confirms Spearman's theory of a general factor or Thurstone's theory of independent primary mental abilities. As you can see from the answers to SAQ 8, Vocabulary, Information and Comprehension seem to form a cluster of high correlations (all over 0.68), as do Object Assembly and Block Design (0.63). But the correlations between subtests belonging to different clusters are nowhere near zero as Thurstone's theory demands; the lowest correlation in the whole table is 0.33. Factor analysis cannot produce clear-cut results with messy data like these. When applied to the Wechsler scales, it usually generates three factors, labelled Verbal Comprehension, Perceptual Organization and Memory (Wechsler, 1981, p. 50). But these factors are not independent because correlations between subtests belonging to different groups are substantial, as we have seen.

The evidence that has accumulated since the 1930s, when Thurstone's views were most influential in the United States (though not so much in Britain), strongly supports the theory of general intelligence. Without accepting that theory in some form it is hard to explain the significant correlations that exist between all intelligence tests and subtests, no matter how diverse. But Thurstone was right in thinking that different people with the same IQ can have very different intellectual strengths and weaknesses.

Modern IQ tests are often designed so that separate scores can be calculated for different aspects of intelligence. The Wechsler scales, as we have seen, measure verbal IQ, performance IQ and full scale IQ. The British Ability Scales measure verbal IQ, visual IQ and general IQ. People with identical general IQ scores can, and often do, have quite different patterns of ability, which may tell us more about them than do their global IQ scores (the average of all their subtest scores). For example, some people are excellent at verbal thinking but rather weak at visual thinking. It is a mistake, however, to believe that the different mental abilities are completely independent. The notion of an overall general intelligence, represented by global IQ, and the rival notion of specific abilities, represented by various factors such as verbal or visual ability, undoubtedly both contain a portion of the truth.
2.5 How accurate are IQ tests?

The accuracy of an IQ test, or of any measuring instrument for that matter, consists of two main ingredients: reliability and validity. The reliability of a test is the consistency and stability with which it measures. If a tape measure is a reliable measure of length, then it should show consistency in results if we measure objects with different parts of the tape. It should also show stability in the sense that it should give the same results if we measure the same objects with it on two separate occasions. If the tape measure is not consistent and stable, perhaps because it tends to stretch, then it is not a reliable measure. The same idea applies to psychological tests such as IQ tests. If an IQ test reliably measures general intellectual ability, then scores on some test items should correlate highly with scores on the other items. People's scores on the test on two separate occasions should also correlate highly.

There are three main techniques for establishing the reliability of tests. The first establishes split-half reliability, which shows the test's consistency. Researchers give the test to a large group of people and record the score that each person gets on the odd-numbered test items and the even-numbered items. The researchers then work out the correlation between these two sets of scores, each of which represents half of the items. If all the test items measure the same general ability, then people who do well on the odd-numbered items should tend to do well on the even-numbered items, so the correlation should be high. If the correlation is low, with many people scoring very differently on the two halves of the test, then the test is not reliable.

The second technique determines the parallel-form reliability of a test. Some tests, such as the Stanford-Binet scale and the Mill Hill Vocabulary scale, come in two equivalent forms each containing similar kinds of items. In these cases, the parallel-form reliability of the test can be assessed by giving both forms of the test to a large group of people and working out the correlation between their scores on the two forms. Once again, the correlation will be high if the test is reliable.

The third technique establishes test-retest reliability, which is an index of the test's stability. Researchers give the test to a large group of people on two separate occasions and work out the correlation between the two sets of scores. If the test is reliable, then people who score highly on the first occasion should score highly on the second, and those who score badly on the first occasion should score badly on the second. In other words, the correlation between the two sets of scores should be high; if it is not, then the test is not reliable. Good IQ tests have very high reliabilities. The split-half and test-retest reliabilities of the Wechsler Adult Intelligence Scale and the British Ability Scales (general IQ) are both 0.95.

The second ingredient of a test's accuracy is its validity. A test is said to be valid if it measures what it purports to measure. There are many ways of assessing a test's validity, of which the most important are various types of criterion validation. In the case of IQ tests, criterion validation involves giving the test to children of different ages or different levels of educational attainment. If an IQ test is a valid measure of intelligence, it ought to correlate
significantly, although not perfectly, with measures of educational attainment. It follows that children who do well at school should tend to score higher on IQ tests than those who do badly at school. This expectation has been confirmed by numerous careful research studies (Jensen, 1980; Mackintosh and Mascie-Taylor, 1985). Correlations between IQ scores on the one hand and marks in 'O' level (now GCSE) examinations, or scores on school tests of reading attainment or mathematics on the other, are typically found to lie between 0.40 and 0.70. If IQ were a perfect predictor of educational attainment as measured by such tests, then the correlations would be 1.00; if IQ were quite unrelated to educational attainment, then the correlations would be zero. The fact that the correlations lie somewhere between these two extremes suggests that IQ tests are valid measures of intelligence, and also that intelligence contributes substantially to scholastic performance, but that the two are not identical.

Another method of validating an IQ test is by correlating scores from it with scores from other IQ tests of known validity. This method, called congruent validation, ensures that the new test measures whatever the other tests measure. A new IQ test would be considered suspect if it gave scores that did not correlate highly with those from established IQ tests.

It is important to understand that a test can be reliable but not valid. A test that purports to measure intelligence, for example, may in fact simply be measuring reading ability. Such a test could be very reliable, in the sense that it gives consistent and stable scores, but it would not be a valid intelligence test because it would be measuring reading ability rather than intelligence. On the other hand, a test that is known to be valid must be reliable, because it cannot measure accurately what it is supposed to measure unless it is consistent and stable.

### 2.6 Does intelligence decline in later life?

It is worth commenting briefly on the long-term stability of intelligence. Young children's intellectual ability increases as they grow older, but this increase levels off when people reach the age of about 17 or 18 years. IQ scores indicate intellectual ability relative to the average for people of the same age. The question naturally arises as to whether people's IQs then remain stable over long periods of time. IQ tests are not reliable for very young children, under about 5 years of age, so it is hardly surprising that IQs at that age do not correlate very highly with IQs in later life, even though the correlations are well above zero. From about the age of 7 years, IQ scores tend to stabilize: people's IQs at the age of 7 correlate about 0.70 with their own IQs at the age of 17.

Does intelligence decline in middle age? There is a great deal of evidence that average intelligence, at least as measured by the 'pure' reasoning tests and subtests, is highest among 18–30 year olds and significantly lower in older groups. On Raven's Progressive Matrices, for example, average 20 year olds solve about 45 of the 60 matrices correctly, average 40 year olds solve about 37, and average 60 year olds solve only about 28 (Raven, Court and Raven,
A similar progressive decline with age has been reported on the performance scales of the WAIS-R (Wechsler, 1981) and on other non-verbal tests and subtests (Bates and Schaie, 1976). These data used to be accepted stoically as evidence that intelligence declines quite precipitously after about 30 years of age. It is now known that the findings arise from something quite different and less alarming.

Most of the findings showing a decline in intelligence with age have been based on studies in which groups of people of different ages were compared. These are called cross-sectional studies because they focus on a cross-section of the population—comprising, that is, different age groups—at a particular time. Recent research (e.g. Schaie, 1983) has used longitudinal studies in which researchers tested and retested the same group of people at different times rather than different age groups at the same time. In other words, they followed the IQ test performance of one group of people as they got progressively older. (People of the same age who are tested repeatedly in a longitudinal study are called an age cohort.) In the longitudinal studies, a decline in intelligence was found only after the age of about 60.

Why, then, do 50-year-old people score lower in cross-sectional studies than 30-year-olds? The reason seems to be that younger generations are more intelligent than their elders, or at least better than their elders at solving 'pure' reasoning problems in IQ tests. When a person born in 1930, for example, is tested and retested later in life, there is little decline in IQ. But that person's IQ is likely to be slightly lower than that of someone born 20 years later. A comparison with height is instructive. Young adults are taller, on average, than members of their parents' generation, because people today tend to grow taller. The reason why 50 year olds are generally shorter than 30 year olds is not that they have begun to shrink with age but that they belong to a shorter age cohort, and there is a similar cohort effect with intelligence.

The evidence is quite unambiguous. Like physical stature, average levels of intelligence, at least as measured by non-verbal tests like Raven's Progressive Matrices, have been rising steadily over the past half century or so, not only in Britain, but also in Belgium, Canada, France, New Zealand, the Netherlands, Norway, the United States and probably Japan (Flynn, 1987; Lynn and Hampson, 1986). This must be the reason, or at least a substantial part of the reason, why older cohorts do less well in the tests. Even the decline after the age of 60, which shows up not only in cross-sectional but also in longitudinal studies, may have nothing to do with age as such. It may be due to the tendency for intelligence to decline, probably as a side-effect of illness, in people who are close to death. Age cohorts over 60 are bound to include a minority who are suffering from various kinds of illnesses, and the decline in average intelligence after the age of 60 may be due entirely to the decline in intelligence in that minority group.
Summary of Section 2

- Modern intelligence tests include the Wechsler scales, the British Ability Scales, Raven's Progressive Matrices and the Mill Hill Vocabulary scale.
- A correlation matrix shows the correlations between scores on several tests or subtests.
- Factor analysis is a statistical technique for analysing clusters of correlations to identify underlying factors.
- Spearman's theory of general intelligence assumes a single \( g \) (general intelligence), reflected by high correlations between all tests and subtests. Thurstone's theory emphasizes several clusters of correlations associated with primary mental abilities.
- The reliability of a test depends on its being a consistent and stable measure, as demonstrated by split-half, parallel-form and test-retest correlations.
- The validity of a test is the extent to which it measures what it purports to measure. Criterion validation involves correlating IQ scores with measures of educational attainment and with other IQ tests. Congruent validation ensures that the new test measures whatever the other tests measure.
- While cross-sectional studies indicated a steep decline in IQ for older age groups, longitudinal studies indicate that intelligence does not decline as much as was formerly believed.

The nature–nurture controversy

For over a century a controversy has raged in psychology over the relative contributions of heredity (nature) and environment (nurture) to intelligence. Lewis Terman and other early American psychometricians believed intelligence to be a fixed and innate mental capacity. Many of them were members of eugenic societies which campaigned in favour of sterilization laws to prevent people with very low IQs from having children, and against immigration from southern, central and eastern Europe. A number of sterilization laws were passed in the United States, beginning with the Act for the Prevention of Idiocy in Pennsylvania in 1905. In the state of Virginia alone, where a similar law was passed, over 7000 people with low IQs were compulsorily sterilized between 1924 and 1972. In 1912 a psychologist called Henry Goddard administered IQ tests to immigrants at the receiving station in New York harbour and reported, incredibly, that most of the Italians, eastern
Europeans and Jews were ‘feebleminded’ (Goddard, 1913). The infamous Immigration Restriction Act of 1924 limited the number of Jewish refugees from Nazi Germany who were able to escape to the United States, as a result of which hundreds of thousands perished in concentration camps after being refused entry into the United States. The writings of Terman, Goddard and other American psychometricians, while not a direct cause of these deplorable measures, lent intellectual credibility to them (Fancher, 1985; Gould, 1981; Colman, 1987).

In Britain, Cyril Burt, later to be knighted for his services to educational psychology, advised the Consultative Committee on Secondary Education in 1938 that intelligence is a fixed and hereditary factor which can be measured reliably by the time a child is 11 years old. His advice was incorporated into the Education Act of 1944 which introduced the ‘eleven-plus’ examination based largely on IQ tests. Children who failed the eleven-plus were denied entry to grammar schools which prepared pupils for higher education and were sent instead to academically inferior secondary modern schools, where they remained even if their intellectual or scholastic ability improved dramatically in later years.

Is it true that a person’s intelligence is a fixed and hereditary characteristic, as eye colour is, or is it a variable characteristic traceable to environmental influences? Stated in that way, the question is misleading for two main reasons. First, there is no necessary connection between the fixedness of a characteristic on the one hand and its genetic or environmental origins on the other, as the question implies. An example to illustrate this is an illness called phenylketonuria. This is a purely hereditary disease caused by a single recessive gene and associated with severe mental retardation. Nevertheless, the effects of this hereditary disease can be prevented by a single environmental change: by eliminating phenylalanine, which is present in food proteins, from the patient’s diet. This shows that a hereditary characteristic need not be fixed and unmodifiable. It is also the case that a fixed, unmodifiable characteristic need not necessarily be hereditary. This applies, for example, to mental retardation resulting from oxygen deprivation at the time of birth.

Second, it is meaningless to pose the heredity-versus-environment question about intelligence as an either/or issue. In order to have any human intelligence a person needs a brain, and it is impossible to have a brain without inheriting the necessary genes and acquiring the food, oxygen, parental care and other environmental requirements for its growth.

3.1 Heritability of IQ

The question about which psychologists argue is slightly more complicated: to what extent are differences in IQ between people genetically determined? In technical terms, the question is, ‘what is the heritability of IQ?’ The heritability of a characteristic in a specified population is the proportion of the variability of that characteristic that is attributable to genetic differences between individuals. The heritability of IQ is therefore the proportion of the
variability—that is, the differences between individual IQ scores—that is attributable to genetic differences between people in the population.

Two hypothetical limiting cases will help to fix the concept of heritability and clarify the point at issue. Imagine first a perfectly egalitarian utopia in which no one is advantaged or disadvantaged relative to anyone else with regard to any environmental factors that might affect IQ. In that society any IQ differences observed between people, and therefore also the variability of IQ, would necessarily be due entirely to genetic factors (nature), so the variability of IQ would be 100 per cent. Now imagine a science-fiction dystopia (the opposite of utopia) populated entirely by clones who share identical genes. In that case, since there are no genetic differences between people, any observed variance in IQ would necessarily be due entirely to environmental factors (nurture) and so the heritability of IQ would be zero.

If education influences IQ, and it is hard to believe that it does not, then any increase in equality of educational opportunity is likely to lead to the paradoxical result that hereditary factors become increasingly important in producing differences between people. Because heritability is defined as the proportion of the IQ differences due to the heredity, then if differences due to environmental factors play a smaller role, it follows that the remaining proportion due to hereditary factors automatically becomes relatively larger. In fact, there is empirical evidence to show that this has happened. A study of Norwegian twins born in the 1930s, 1940s and 1950s confirms that, as equality of educational opportunity increased in that country, the estimated heritability of intelligence increased correspondingly (Heath et al., 1985).

How can the heritability of IQ be estimated? The problem would be easily solved if researchers could hold either environmental or hereditary factors constant as in the hypothetical utopia and dystopia discussed above. Experimental biologists routinely perform similar experiments on plants to determine the heritability of various characteristics, but it is manifestly impossible to settle the IQ question by such methods. Psychologists are limited to indirect methods of estimating the heritability of IQ, and the results are highly controversial.

Researchers have used three main methods for estimating the heritability of IQ: studies of separated identical twins, family studies and adoption studies. Galton pioneered two of these methods in a primitive way and reported his results in a book entitled Hereditary genius; an inquiry into its laws and consequences (Galton, 1869). His family study consisted of an examination of the family trees of 415 highly distinguished people in various walks of life. He found that their blood relatives were also eminent much more often than would be expected by chance, and, furthermore, that their close relatives were more often eminent than their more distant relatives. For example, 48 per cent of their sons, 7 per cent of their grandsons, and only 1 per cent of their great-grandsons were eminent. Genius thus seemed to run in families rather like red hair, which was already known to be hereditary, and Galton concluded that differences in mental ability are also largely hereditary.
Galton found that intelligence seemed to run in families, and he inferred from this that differences in intelligence are mainly due to heredity. Can you think of any other explanation for his finding?

Galton's adoption study capitalized on the common practice of nepotism in the Roman Catholic Church, whereby popes used to adopt young boys from ordinary backgrounds and raise them in their homes as 'nephews'. Galton argued that these boys had similar environmental advantages to those enjoyed by the natural sons of other eminent men but none of the genetic advantages: 'the social helps are the same but the hereditary gifts are wanting' (Galton, 1869). His investigation showed that the popes' 'nephews' seldom attained eminence in later life, although almost half the natural sons of other eminent men did so. He interpreted these findings as further evidence that intelligence is largely hereditary.

Galton's family and adoption studies were quaint and interesting but hardly convincing. Sections 3.2, 3.3 and 3.4 are devoted to a review of the evidence that has accumulated since then. For over a century psychologists have tried to resolve the nature-nurture question in relation to IQ, using both refinements of Galton's methods and studies of separated identical twins. However, all this research effort has produced no very definite conclusions.

3.2 Separated identical twins

As explained in Chapter 5, section 5.3, identical twins are formed when a single ovum, fertilized by a single sperm, splits in two and grows into two separate individuals. The twins have identical genes, which means that they are the same sex and very similar in all hereditary characteristics. Non-identical twins develop from two separate ova fertilized by two separate sperms at about the same time; they may be the same or opposite sex and share about half their genes in common, just like ordinary siblings (brothers and sisters). Both identical and non-identical twins are normally brought up together and share a common environment.

Identical twins are very occasionally separated soon after birth and raised in different homes. The importance of these rare cases is that these twins share only their genes, not their environments, in common. Any similarity in their IQs beyond what is to be expected by chance would seem, on the face of it, to be attributable to their genetic identity. Some psychologists regard studies of separated identical twins as the most convincing data available concerning the heritability of IQ, but the results are, in reality, curiously inconclusive. Only four such studies have so far been published, partly because suitable cases are extremely hard to find. The results, in descending order of the sizes of the studies, and of the correlations reported, are summarized in Table 7.2.

By far the largest and most influential of the published studies, and the one reporting the highest correlation, was carried out by Sir Cyril Burt of University College London between the 1940s and 1960s. It is also the only one purporting to show that the home environments of the twins were quite
Table 7.2  Studies of the correlation between the IQs of separated identical twins

<table>
<thead>
<tr>
<th>Study</th>
<th>Number of twin pairs</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burt (1955, 1958, 1966)</td>
<td>53</td>
<td>0.86</td>
</tr>
<tr>
<td>Shields (1962)</td>
<td>37</td>
<td>0.77</td>
</tr>
<tr>
<td>Newman, Freeman and Holzinger (1937)</td>
<td>19</td>
<td>0.67</td>
</tr>
<tr>
<td>Juel-Nielsen (1965)</td>
<td>12</td>
<td>0.62</td>
</tr>
</tbody>
</table>

different. If separated identical twins are raised in similar home environments, then the correlation between their IQs is hard to interpret because it might be due to environmental similarities rather than to the genetic similarities. Burt reported that the correlation between the twins’ home environments, which he rated on a scale from ‘higher professional’ to ‘unskilled’, was close to zero, which suggests that the massive 0.86 correlation between their IQ’s must be due to genetic differences alone, implying that the heritability of IQ is 86 per cent.

In 1974, the American psychologist Leon Kamin drew attention to certain anomalies in Burt’s data which seemed to strain the laws of probability (see Box C), and in 1979, Leslie Hearnshaw’s official biography of Burt proved beyond doubt that the data were faked. Burt simply fabricated IQ scores to prove his hereditarian point, and he even went to the lengths of publishing papers in psychological journals with fictitious co-authors.

The second largest study (Shields, 1962) reported a correlation of 0.77 between the IQs of thirty-seven pairs of separated identical twins in England. This study, unlike Burt’s, can be criticized on the grounds that the twins’ environments were quite similar (Kamin, 1974). In more than half the pairs, the ‘separated’ twins were raised in different branches of the same family; often one member of the pair stayed with the mother while the other was raised by a grandmother or aunt. If these pairs are removed from the sample, the correlation in the remaining thirteen pairs drops to 0.51. But in several of these cases the twins’ environments were still similar because both twins attended the same school, played together, and so on. Shields’ data therefore overestimate the heritability of IQ to an unknown degree because the correlation between the twins’ IQ scores might be due chiefly to their similar environments.

The next largest study (Newman et al., 1937) reported a correlation of 0.67 between the IQs of nineteen pairs of separated identical twins in the United States. This study contains three serious flaws. First, there is abundant evidence that the twins’ environments were similar, as in the Shields study (Kamin, 1974). Second, the researchers used a flawed method of recruiting subjects for their study. To avoid accidentally including non-identical twins in their sample, they excluded all pairs who said in reply to a preliminary questionnaire that they were at all dissimilar in appearance or behaviour. This biased the sample from the outset in favour of twins who were extremely similar, because twins who were noticeably dissimilar were not included.
Three years after Burt's death, Kamin (1974) pointed out several 'numerical impossibilities' in Burt's data. Kamin's most devastating discovery was the existence of inexplicable identical correlations in the data. Burt claimed to have tested his twins on both individual and (presumably less reliable) group IQ tests. He initially reported data from twenty-one pairs of separated identical twins. By 1958 he claimed to have tested over thirty pairs, and by 1966 his sample had purportedly grown to fifty-three pairs. He reported the correlations as shown in Table 7.3.

<table>
<thead>
<tr>
<th>Date</th>
<th>Number of twin pairs</th>
<th>Group test</th>
<th>Individual test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1955</td>
<td>21</td>
<td>0.771</td>
<td>0.843</td>
</tr>
<tr>
<td>1958</td>
<td>30+</td>
<td>0.771</td>
<td>0.843</td>
</tr>
<tr>
<td>1966</td>
<td>53</td>
<td>0.771</td>
<td>0.863</td>
</tr>
</tbody>
</table>

The fact that the correlations, with one exception which may be a typing error, remain identical to three decimal places with different sample sizes, strains the laws of probability and the credulity of the reader. Kamin concluded that 'the numbers left behind by Professor Burt are simply not worthy of our current scientific attention' (Kamin, 1974).

This is a serious methodological error in a study in which it is precisely the degree of similarity that is at issue. The third flaw is the confounding of IQ with age. The researchers measured the twins' IQs with the original 1916 Stanford–Binet scale, which is notorious for its faulty standardization for different age groups. A properly standardized IQ test should yield average scores of 100 for all age groups because an IQ of 100 is average by definition. However, the 1916 Stanford–Binet scale yields higher average scores for younger than for older age groups (see the discussions in Sections 1.4 to 1.6). The correlation between the twins' IQs in the study by Newman et al. may therefore be due, at least in part, to their identical ages rather than their identical genes.

The last and smallest study (Juel-Nielsen, 1965), carried out in Denmark on twelve pairs of separated identical twins, reported a correlation of 0.62. Juel-Nielsen measured the twins' IQs with a Danish translation of the Wechsler Adult Intelligence Scale (WAIS) which had not been standardized on a Danish sample, and this resulted in IQ being confounded with age as in the study by Newman et al. (Kamin, 1974). A second flaw was that, in many cases, the twins' environments were similar, as in the Shields study.
3.3 Family studies

A child inherits half its genes from each of its parents and therefore shares half its genes with each of them. It follows from this that non-identical twins and ordinary siblings share about 50 per cent of their genes in common, a grandparent shares about 25 per cent of genes with each grandchild, and so on. In family studies these different degrees of genetic relatedness are compared with IQ correlations between the corresponding relation pairs. The underlying assumption is that, if the heritability of IQ is high, then the IQ correlations between different categories of relatives should decrease in an orderly fashion as the degree of genetic relatedness decreases.

The most influential family study (Erlenmeyer-Kimling and Jarvik, 1963) reported a remarkably orderly pattern of correlations between numerous different categories of relatives. Some of the data, as summarized in tabular form by Jensen and Eysenck, are shown in Table 7.4.

Table 7.4 Correlations between IQ scores of different categories of relatives

<table>
<thead>
<tr>
<th>Relationship category</th>
<th>Average IQ correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identical twins raised together</td>
<td>0.87</td>
</tr>
<tr>
<td>Identical twins raised apart</td>
<td>0.75</td>
</tr>
<tr>
<td>Siblings raised together</td>
<td>0.55</td>
</tr>
<tr>
<td>Siblings raised apart</td>
<td>0.47</td>
</tr>
<tr>
<td>Parent–child</td>
<td>0.50</td>
</tr>
<tr>
<td>Non-identical twins (opposite sex)</td>
<td>0.49</td>
</tr>
<tr>
<td>Second cousins</td>
<td>0.16</td>
</tr>
<tr>
<td>Unrelated</td>
<td>-0.01</td>
</tr>
</tbody>
</table>


Most of the correlations correspond uncannily closely to the proportions of genes that each of the categories of relatives share in common. It is now known that the data, which were obtained from fifty-two separate research reports, are indeed too good to be true (Kamin, 1974; Paul, 1985; Colman, 1987). One example will suffice to show how unreliable the data are. Look at the correlation of 0.47 for siblings raised apart. It is based on three published investigations, one by Burt and a mythical co-author of his whom he called Howard, which reported correlations of 0.46, 0.34 and 0.23. The average of these three numbers is not 0.47, as shown in Table 7.4, but 0.34, which is way below the theoretical figure of 0.50 that corresponds to the proportion of genes siblings share in common. There are numerous similar errors and distortions, all in the direction of making the data appear to support the simple genetic theory. To make matters worse, the data include many correlations fabricated by Burt to support that theory.

Erlenmeyer-Kimling and Jarvik’s data were discredited in the mid 1970s, but they continue to be cited in textbooks as evidence for the heritability of IQ. Of the nineteen textbooks of genetics published between 1978 and 1984 that
devoted more than half a paragraph to the heritability of IQ, twelve cited the

data as evidence, and in many of these books it was the only evidence cited
(Paul, 1985). One of the books included a dramatic insert on ‘Scientific fraud:
the case against Sir Cyril Burt’, but also stated that ‘the measured heritability
of IQ is relatively high’ and gave the Erlenmeyer-Kimling and Jarvik data,
which incorporate Burt’s fraudulent figures, as evidence for this conclusion

More reliable family studies have been done in recent years. To translate the
findings into numerical estimates of heritability, mathematical models have to
be used, and these models include a large number of assumptions which some
critics consider unacceptable. They all assume, for example, that identical
and non-identical twins raised together experience the same environmental
background, and that living with a very intelligent or a very unintelligent
sibling has no direct effect on a child’s IQ. The evidence from the best family
studies suggests that the heritability of IQ probably lies somewhere in the
range of 30–60 per cent (Henderson, 1982), but many psychologists consider
the evidence to be unconvincing.

3.4 Adoption studies

Children who are adopted early in life share none of their genes with their
adoptive parents but about 50 per cent with their natural mothers who
relinquished them soon after they were born. It is reasonable to assume that, if
IQ differences are due mainly to environmental factors, then the IQs of
adopted children should correlate more highly with those of their adoptive
parents than with those of their natural mothers with whom they have never
lived. If IQ is highly heritable, on the other hand, then roughly the reverse
should be found. The IQs of adopted children should correlate moderately
highly with those of their natural mothers, who donated ‘nature’ in the form
of half their genes, and not at all with the IQs of their adoptive parents, who
provided ‘nurture’.

Two classic adoption studies (Burks, 1928; Leahy, 1935) reported correlations
of only 0.13 and 0.18, respectively, between the IQs of adopted children and
those of their adoptive parents. These correlations are virtually zero and are
certainly much lower than those found between the IQs of children raised by
their natural parents and the IQs of those parents, which suggests that genes
are important. But the results are hard to interpret because the IQs of the
adoptive parents in both studies were mostly high and closely bunched
together (Kamin, 1974, chapter 5; Eysenck and Kamin, 1981, chapter 15). Why
is this a problem? The answer is that it may have artificially depressed the
correlations. The point is that, if the adoptive parents are all very intelligent,
then there is no scope for correlations throughout the range from low to high
IQs. To understand this statistical point, consider personal income and diet as
measured by calorie intake. There is undoubtedly a high correlation between
these two things in the world’s population as a whole. But it would probably
all but disappear if only well-off people in Britain were included in the
sample, because well-off people are, for the most part, equally well fed. The
same thing may have happened in the Burks and Leahy studies. The adoptive
parents were nearly all highly intelligent.
Skodak and Skeels (1949) were the first to examine both the IQs of adopted children and those of their natural mothers. They found the correlation to be 0.44 in sixty-three children adopted before the age of 6 months. For decades this finding was considered the most powerful of all the evidence showing the heritability of IQ to be high, because there seemed to be no possible environmental explanation for such a substantial correlation between the IQs of children and the IQs of their mothers from whom they were separated in early infancy.

Perhaps surprisingly, there is an environmental explanation, and it is called selective placement. Children are not delivered to their adoptive homes by storks. They are usually sent there by adoption agencies who used, before the Second World War, to place most children in well-off homes. Since the war, as adoption has become more popular, the agencies have generally tried to place each child with adoptive parents whose intellectual and educational level roughly matches those of their natural mothers. In the Skodak and Skeels study, children of natural mothers with high IQs had often been placed with college-educated adoptive parents in wealthy homes which probably provided good environments for the development of their IQs, while children of low-IQ natural mothers had generally been placed with less well educated adoptive parents in poorer homes (Kamin, 1974, chapter 5). What follows from this is that the correlation between the children's IQs and those of their natural mothers may have been due to selective placement in environments related to their natural mothers' IQs rather than to hereditary factors.

More recent adoption studies have avoided some of the pitfalls of earlier research by focusing on parents who have raised both adopted and natural children in their homes. All the children are brought up in the same environment by the same parents, so any significant difference between the parent–child IQ correlations for the adopted children and the natural children in these households would seem to be attributable to the sole difference that only the natural children share genes in common with their parents.

The results of two large American studies using this improved research design are shown in Table 7.5.

Table 7.5 Correlations between IQs of adoptive mothers and IQs of their natural and adopted children

<table>
<thead>
<tr>
<th>Mother–Child correlation</th>
<th>Natural children</th>
<th>Adopted children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scarr and Weinberg (1977)</td>
<td>0.34</td>
<td>0.29</td>
</tr>
<tr>
<td>Horn, Loehlin and Willerman (1979)</td>
<td>0.20</td>
<td>0.22</td>
</tr>
</tbody>
</table>

In Scarr and Weinberg's (1977) Minnesota study, the parents and their natural children were white and the adopted children were mostly black or of mixed race. The mother–child IQ correlation was slightly higher for the natural children, but the difference was not statistically significant (which means that it could easily have been due to chance). In Horn et al.'s (1979) Texas Adoption Project, as it is called, there was a slight difference, which was also non-significant, in the opposite direction. Both sets of results suggest that the
heritability of IQ is low because of the similar mother–child correlations regardless of whether or not the genes were the same.

Further evidence pointing to this conclusion comes from families in both studies with more than one adopted child and a natural child or more than one natural child and an adopted child. The correlations in these families between pairs of natural children (with half their genes in common), pairs of adopted children (genetically unrelated) and natural–adopted pairs (genetically unrelated) were all very similar. None of the correlations differed significantly from the others, although the first of these should surely have been higher than the others if the heritability of IQ is high.

In both studies correlations were also reported between the adopted children's IQs and those of their natural mothers. In the Minnesota study the correlation between the IQs of the adopted children and the estimated IQs of the mothers who relinquished them was 0.32. But this is probably attributable to selective placement because the correlation was later shown to be almost the same (0.33) when the IQs of children unrelated to the mothers but living in the same adoptive homes were substituted in the calculation (Kamin, 1974, chapter 5). In other words, the natural mothers' IQs correlated just as highly with the IQs of the other children in the adoptive homes as with the IQs of their own children. This is surprising and can be explained only by selective placement. In the Texas study the correlation was 0.31 between the IQs of adopted children and those of their natural mothers, and it was 0.19 when children sharing the same environment but none of the mothers' genes were substituted in the calculation. In this case, selective placement explains most of the original correlation, though not all of it.

### 3.5 Conclusions

Most people, including psychologists, seem to hold strong views about the heritability of intelligence. The research evidence has signally failed to settle the question once and for all or to provide a half-way convincing estimate of the relative contributions of heredity (nature) and environment (nurture) to individual differences in IQ scores. Studies of separated identical twins, family studies and adoption studies are all fraught with different but seemingly insuperable problems.

Any estimate of heritability is at best only a very rough approximation, and it applies only to a single population at a single point in time. The amount of environmental variation differs from one society to the next, and it even changes over time within the same society. A consequence of this is that the heritability of IQ is itself a variable quantity that can never be pinned down and carved in stone as a universal and eternal verity. Precise estimates give a false air of scientific exactitude because they are not based on properly controlled experimental studies. Controlled experiments on heritability are impossible to perform on human subjects for practical and ethical reasons. The research that has been done is non-experimental and inadequately controlled, and it is foolhardy to overinterpret its findings. The data come from studies that are always imperfect, often seriously flawed and occasionally downright fraudulent.
Depending on which studies you consider best, which you are willing to ignore, and which simplifying assumptions you are willing to accept, you can arrive at an estimate of the heritability of IQ anywhere from zero to about 70 per cent. There is another, probably wiser, option, and that is simply to suspend judgement on this issue. It is certainly not necessary to hold a view about the heritability of IQ because, whether it is high or low or zero, the degree to which we can influence people’s IQs by suitable environmental means is an entirely separate (and much more important) question. Whatever the heritability of IQ, improvements in education, housing, social environments, and nurture in general can only be of benefit to society as a whole and all the individuals within it.

Summary of Section 3

- Attempts to measure the proportion of the differences in IQ attributable to heredity have used studies of separated identical twins, family studies and adoption studies.
- The apparently high correlations that have been recorded between the IQ scores of separated identical twins are not necessarily due to the twins’ identical genes. One alternative explanation is the fact that, although separated, the twins are often raised in similar environments. There are cases also, where figures may have been doctored in a fraudulent manner.
- The reported correlations due to kinship in families are ambiguous, because genetic and environmental influences are difficult to separate.
- Correlations between the IQs of adopted children and those of their natural mothers may be due to selective placement of the children in home environments similar to those of their natural parents.
- In all these studies it is very difficult to sort out the degree to which differences in intelligence are due to heredity (nature) as opposed to environmental factors (nurture).

4 Race and intelligence

An acrimonious debate about race and intelligence began when Arthur Jensen of the University of California, Berkeley, published a paper entitled ‘How much can we boost IQ and scholastic achievement?’ (Jensen, 1969). Hans Eysenck of the University of London summarized Jensen’s ideas in a book entitled Race, intelligence and education (Eysenck, 1971). The point at issue in this debate is the interpretation of the evidence showing that black Americans score about 15 points lower, on average, on IQ tests than do white
CHAPTER 7 ASPECTS OF INTELLIGENCE

Americans (Shuey, 1966; Scarr, 1981). Also, people of West Indian descent in Britain score between 5 and 13 points lower, on average, on IQ tests than do members of the indigenous white population (Mackintosh and Mascie-Taylor, 1985; Mackintosh, 1986). Jensen and Eysenck both argued that the differences are due mainly to genetic differences between the black and white races, and their writings generated angry emotional reactions. Literally hundreds of criticisms were published, and there were calls for Jensen’s dismissal from the American Psychological Association. Students began disrupting his classes, and his university had to hire a bodyguard to protect him. Reactions to Eysenck in England were only slightly less extreme.

Before examining the evidence on this question, it is worth mentioning two theoretical points that are relevant to it. The first is that modern techniques of genetic analysis have shown only about 7 per cent of human genetic differences of all kinds to be attributable to racial differences (Rose et al., 1984, pp. 119–27). This suggests that race, like beauty, is only skin deep and that race differences in intelligence are therefore unlikely to be due to genetic factors. The second theoretical point is that some 70 per cent of Afro-Americans and Afro-Caribbeans have at least one white ancestor, and 15 per cent have more white ancestors than black; the ancestry of black Americans and West Indians is, on average, 20 per cent white (Reed, 1969). These facts also make it unlikely that there are any very marked inborn psychological differences between the black and white populations. The question is empirical, though, and there is evidence that bears more directly on it.

Jensen and Eysenck argued that race differences in IQ are mainly due to genetic factors, and they based this conclusion largely on evidence which they believed showed that the heritability of IQ is high. They reasoned that, if IQ differences in the general population are due mainly to heredity, then IQ differences between the black and white groups within the population must also be due mainly to heredity. The argument is now known to be invalid, as the biologist Richard Lewontin demonstrated with the following famous analogy (Rose et al., 1984, p. 18). Two handfuls of seed are taken from the same sack and planted in separate plots, which differ from each other only in that the second plot is deficient in nutrients. The plants in the first plot grow tall and those in the second are stunted. This is illustrated in Figure 7.4.

![Figure 7.4 Lewontin's analogy. Within each plot, genetic differences make some plants grow taller than others. Because of environmental differences (nutrients), however, the plants in plot A grow taller than the plants in plot B.](plotA.png)
Examine Figure 7.4 carefully. Then think about Lewontin’s analogy and try to anticipate where it is leading.

(a) Will the difference between the plants in the two plots be due to genetic factors, environmental factors, or both?

(b) There will also be differences between individual plants within each plot. Will these within-group differences be due to genetic factors, environmental factors, or both?

(c) What can you conclude from the answers to (a) and (b) about race differences in intelligence?

The average difference between the two groups of plants will be due entirely to environmental factors (nutrients), because the seeds come from the same sack and are therefore genetically the same, on average. On the other hand, differences between plants within each plot will be due entirely to genetic factors because the environment is identical for all seeds in the same plot. In other words, heritability is 100 per cent within each group, but the average difference between the two groups is due entirely to environmental factors. Lewontin’s analogy shows that, even if the heritability of IQ is very high, the differences between racial groups may be entirely attributable to environmental factors.

The argument based on heritability is fallacious, as shown above. However, we should examine the empirical evidence for the genetic or environmental origins of the black–white IQ gap. This comes from racial admixture and crossing studies, to which we can now turn.

4.1 Racial admixture studies

Racial admixture studies capitalize on the fact, mentioned earlier, that black people have varying amounts of white ancestry. If the black–white IQ gap is mainly due to genetic factors, then obviously black people with a great deal of white ancestry, and therefore many ‘white’ genes, should have higher average IQs than those with no white ancestry. It is difficult to estimate a person’s proportion of white ancestry accurately, so the results of racial admixture studies are not absolutely conclusive.

One of the earliest and best studies (Witty and Jenkins, 1936) focused on ninety-one black children with the highest IQs of all the many thousands in the Chicago public school system. When the researchers classified these high-IQ black children according to proportion of white ancestry by examining their family trees, they found no evidence that they had more white ancestry than did a comparison group of ordinary black Americans. The results showed, for example, that 14.3 per cent of the high-IQ children had predominantly white ancestry, compared to 14.8 per cent of the comparison group. If genetic factors are at the root of the black–white IQ gap, then there should have been substantially more children with predominantly white ancestry in the high-IQ group. In fact, the distribution of white ancestry was remarkably similar in both groups of children, and the brightest child in the
sample, a girl with an astonishing IQ of 200, was one of the ones who had no known white ancestry at all.

Sandra Scarr and her colleagues carried out the most carefully controlled racial admixture study yet reported (Scarr et al., 1977). They used biochemical methods based on blood samples to estimate racial ancestry in a sample of Minnesota schoolchildren and found no relationship between racial ancestry and IQ scores on any of the four separate tests that they used. These findings confirm those of Witty and Jenkins (1936). They suggest strongly that the IQ difference between black and white Americans does not stem from genetic causes.

### 4.2 Racial crossing studies

**Racial crossing studies** focus on black, white and mixed-race children who happen to be raised in similar environmental circumstances. If genetic factors are mainly responsible for the IQ difference between the races, then the white children should have higher average IQs than the mixed-race children, and the mixed-race children should have higher average IQs than the black children. If environmental factors are all-important, on the other hand, then black, mixed-race, and white children raised in similar circumstances should have similar average IQs.

The problem, of course, is to find children from the different race groups raised in truly similar environmental circumstances. The most obvious objection to this line of research is that it is difficult or impossible to raise black, white and mixed-race children in comparable environments because racism is itself an important environmental factor to which only the black and mixed-race children are likely to be exposed. If they score lower than the white children on IQ tests, critics of the research can always argue that the supposedly equal environments actually favoured the white children. Interestingly, this objection turns out to be beside the point because, when they are raised in similar circumstances, the black and mixed-race children do not, in fact, score lower than the white children. In all three racial crossing studies that have been published, children from the different racial groups were found to have remarkably similar average IQs. The results of these studies are shown in Table 7.6.

<table>
<thead>
<tr>
<th>Study</th>
<th>Average IQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Black</td>
</tr>
<tr>
<td>Eyferth (1961)</td>
<td>96</td>
</tr>
<tr>
<td>Tizard et al. (1972)</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>106</td>
</tr>
<tr>
<td>Scarr and Weinberg (1976)</td>
<td>97</td>
</tr>
</tbody>
</table>

Note: The children in the Tizard et al. study were given three separate IQ tests.
The children in Eyferth’s (1961) study were Besatzungskinder (occupation children), the illegitimate offspring of sexual liaisons between American occupation troops and German women after the Second World War. Some of their American fathers were black and some were white. All the children were raised by their white mothers in post-war Germany where anti-black prejudice was, and is, comparatively rare. The mixed-race children, who inherited half their genes from their black fathers, and the white children, whose genetic make-up was entirely white, differed by only a single point in average IQ.

Barbara Tizard and her colleagues (Tizard et al., 1972) gave young children living in ‘residential nurseries’ in England three separate IQ tests. It is clear from Table 7.6 that the black and mixed-race children scored consistently higher, on average, than the white children, though the differences were very small and non-significant.

The third racial crossing study, carried out by Sandra Scarr and Richard Weinberg in the United States (Scarr and Weinberg, 1976), was the Minnesota inter-racial adoption study already discussed in a different context in Section 3.4. The black, mixed-race and white children in the study were all living in white adoptive homes. Scarr and Weinberg pointed out that the black children were not really raised in the same environmental circumstances as the others. They were adopted later in life and had less well educated adoptive parents than the mixed-race and white children. The mixed-race and white children were found to have very similar average IQs, but the black children scored 12 to 14 points lower, on average, than the other groups (see Table 7.6). Only the mixed-race and white children were raised in truly comparable environmental circumstances, according to the researchers, and the similar average IQs of these two groups suggest that environmental rather than genetic factors account for IQ differences between the races.

### 4.3 Test bias

Many people, especially non-psychologists, believe that it is hardly surprising that black people get lower average scores on IQ tests than white people. The reason they usually give is that the tests are probably biased. Let us examine this popular explanation a little more closely.

It is certainly true that most conventional IQ tests were constructed by white, middle-class, male psychologists, standardized on white samples, and validated by correlating their scores with tests of scholastic attainment in white schools. (The Wechsler scales, the British Ability Scales, and the newest version of the Stanford–Binet scale are exceptions: at least their standardization samples contained members of ethnic minorities.) But these facts, in themselves, do not prove that the tests are biased. The claim that IQ tests are biased, if it means anything intelligible, implies that they underestimate the true intelligence of black people.

It is not clear what this claim amounts to in the absence of any measure of true intelligence apart from IQ tests themselves. It might mean that IQ tests underestimate the ability of black people to excel at school examinations and other activities usually assumed to require intelligence. There is a great deal
of evidence to show that IQ tests are not biased against black people in that sense. IQ tests predict scholastic attainment just as well for black as for white children in the United States (Jensen, 1980). In Britain also, two studies based on large, random samples of schoolchildren showed that IQ scores correlate highly with school tests of mathematics and reading ability among both West Indian and white 11-year-old children. If anything, the West Indian children do slightly worse on the school tests than their IQ scores suggest they should do (Mackintosh and Mascie-Taylor, 1985; Mackintosh 1986).

It could be argued that IQ tests are biased against black people in a different sense. Perhaps they fail to take account of various forms of educational and general environmental deprivation that cause black people to do badly in IQ tests and also in school examinations and all other activities requiring intelligence. This is a misleading argument. It is equivalent to saying that unemployment statistics are racially biased because they show that black people in Britain are twice as likely to be unemployed (or in the case of Pakistanis and Bangladeshis three times as likely) as white people.

Of course educational and general environmental deprivation adversely affect the development of intelligence. Poverty, overcrowding, poor housing and amenities, and similar environmental factors are all associated with relatively poor performance on IQ tests (Mascie-Taylor, 1984) and a disproportionate number of black people certainly grow up in educationally and environmentally deprived circumstances. It is these inequalities rather than any bias in the tests themselves which are responsible for this poor performance.

To avoid accusations of test bias, psychologists have constructed a number of culture-fair tests which depend to a minimum degree on information and language. Although no test is completely culture-free, some are clearly more cultural-fair than others. Raven’s Progressive Matrices, which was described in Section 2.3, is the best known and most respected of the culture-fair tests. Perhaps surprisingly, the evidence shows that black people in the United States and Britain often perform worse on culture-fair tests than on culturally loaded tests such as the Mill Hill Vocabulary scale. Jensen interpreted this as further evidence that the black–white IQ gap is due mainly to genetic factors. He argued that, if the IQ gap is due to genetic influences, it should show up more markedly on culture-fair tests (Jensen, 1969, p. 81).

Three studies carried out independently in Britain cast doubt on this interpretation, and on the whole notion of culture-fairness. They all showed that Asian children who had recently immigrated into Britain scored badly on Raven’s Progressive Matrices and on other supposedly culture-fair tests, but that those who had lived in Britain for several years scored as well as white children (Ashby, Morrison and Butcher, 1970; Sharma, 1971; Mackintosh and Mascie-Taylor, 1985). In all three studies, the Asian children’s average IQ scores rose by 15 to 20 points after they had lived in Britain for 5 years, although their actual intelligence could hardly have changed so dramatically in such a short time. These findings decisively dispel the myth that Raven’s Progressive Matrices and similar tests are even remotely culture-fair. They suggest, on the contrary, that people from certain cultural backgrounds are at a disadvantage on these tests until they learn new ways of approaching IQ tests.
4.4 Environmental influences on IQ

The National Child Development and Child Health and Education studies in Britain (Mackintosh and Mascie-Taylor, 1985; Mackintosh, 1986) investigated environmental factors such as father's occupation, family size, family income, and neighbourhood quality that might influence children's IQs. The researchers tried to match each West Indian child with a white child in similar environmental circumstances in order to see whether the IQ gap persisted when the environmental factors were equalized in this way. Some of the West Indian children were so deprived that the researchers could find no suitable white counterparts for them among more than 10,000 children randomly selected from the whole population for each study, but among the successfully matched pairs the IQ gap was indeed greatly reduced. In the National Child Development Study, an 11-point average difference between West Indian and white IQ scores decreased to a 5-point difference, and in the Child Health and Education study, a 9-point gap decreased to 3 points. Similar results have been reported in the United States and elsewhere. Blau (1981), for example, compared 500 white and 500 black American children and found that by statistically eliminating environmental factors, such as socio-economic status and family structure, the IQ gap was reduced from about 15 to 3 IQ points.

Do findings like these prove that black–white IQ differences are due to environmental rather than genetic factors? Unfortunately they do not, and to believe that they do is to commit the so-called sociologist's fallacy (Mackenzie, 1984). The reason is that the artificial suppression of environmental differences can obscure real genetic differences. To see this clearly, consider one environmental factor, socio-economic status, a little more closely. It is possible, at least in principle, that high intelligence helps people to get good jobs and salaries and may therefore partly determine socio-economic status. If this is so, then researchers who compare black–white pairs of similar socio-economic status, or equalize this environmental factor statistically, are bound to find smaller IQ differences between the races. Intelligent people, both black and white, would still end up in the higher-status jobs and less intelligent people, both black and white, would end up in lower-status jobs, so people of similar socio-economic status would tend to have similar IQs.

Black people in the United States and Britain undoubtedly score lower, on average, than white people on conventional IQ tests. It is worth noting that members of some ethnic minorities, such as Jews and Orientals (Chinese and Japanese), in the United States score higher, on average, than the general white population, and the same is probably true in Britain. Psychologists do not fully understand the reasons for any of these racial differences, although they seem to be due chiefly to environmental differences between the groups. The suggestion by Jensen, Eysenck and others that the differences are largely due to genetic factors seems hardly credible in the light of the evidence that has accumulated since the early 1970s, from racial admixture and racial crossing studies in particular.
Summary of Section 4

- Jensen and Eysenck have argued that average differences in IQ scores between members of different racial groups are due to genetic differences between the races.
- Findings from racial admixture and racial crossing studies show that black, white and mixed-race children brought up in similar environments have similar IQ scores.
- There is no evidence that IQ tests are biased against black people in Britain or the United States.
- Some of the environmental factors that cause black people to score lower on IQ tests than white people have been identified, but they are not fully understood.

5 Cognitive styles

We have concentrated so far on differences in overall intelligence as measured by IQ tests. However, people differ from one another, not only in their overall levels of intellectual ability but also in the way they think. Different people tend to display distinctive styles in the way in which they think, remember, perceive and generally process information, and these are known as cognitive styles. These styles affect people's functioning in virtually all areas of their lives. Psychologists have discovered numerous cognitive styles which appear to be quite distinct from general intelligence. Research in this area deals with the manner in which people think rather than their ability to think. The various cognitive styles that have been identified relate to different modes of mental functioning rather than to different levels of mental ability.

Cognitive styles develop slowly and cannot easily be modified by experience or training (Messick, 1976; Morrison, 1988). They seem to be deeply rooted in personality structures, and some of them may even be biologically based (Witkin, 1976). Research into cognitive styles has its roots in cognitive psychology and personality theory. Most of the techniques used to investigate them are therefore based on the methods of experimental and clinical psychology, in contrast to the psychometric methods used in the study of intelligence. In Sections 5.1 and 5.2, two of the most important cognitive styles will be discussed: field dependence–independence and convergence–divergence.

5.1 Field dependence–independence

Field dependence–independence is a cognitive style discovered accidentally by the American psychologist Herman Witkin in 1949, which has been extensively studied since then (Witkin, 1949, 1976; Witkin et al., 1972, 1974;
Hampson, 1982, chapter 2). In simple terms, people are **field dependent** if their thoughts and perceptions are strongly influenced by external factors (the 'field'), and they are **field independent** if their thoughts and perceptions depend mainly on internal cues. Many people are neither field dependent nor field independent but fall somewhere between these two extremes.

The most popular tests for measuring field dependence—Independence are the rod and frame test, the rotating room test and the embedded figures test. The **rod and frame test** is conducted in a completely darkened room. The subject's task is to adjust a luminous rod to the vertical position within a tilted rectangular frame. All the subject can see is the rod surrounded by the tilted frame. Field-dependent people are strongly influenced by the tilted frame and tend to align the rod with it. Field-independent people, on the other hand, ignore the frame and concentrate instead on internal gravitational cues in judging when the rod is in the upright position.

In the **rotating room test** the subject sits on a pivoted chair in a tilted room and tries to adjust the chair to the upright. This test also produces marked individual differences. Astonishingly, some field-dependent people think that they are sitting upright when the chair is, in reality, tilted as much as 30 degrees from the vertical.

The **embedded figures test** appears superficially to be quite different from the other two tests, but it also measures field dependence—Independence. In this test the subject tries to locate simple geometric figures embedded in a series of complex diagrams.

**SAQ 12**

Figure 7.5 presents a difficult item from the embedded figures test (Witkin et al., 1971). The simple figure on the right is embedded in the complex diagram on the left. Can you see it and trace it out? It took Witkin's student subjects an average of 2 minutes 24 seconds to solve this puzzle correctly.

![Figure 7.5](image)

Field-dependent people are distracted by the overall complexity of the diagrams and have difficulty finding the embedded figures. Most field-independent people find the figures popping out of the diagrams without effort on their part.

In these three tests, the field-independent cognitive style is associated with superior performance. Scores on all these tests correlate with one another: people who align the rod with the tilted frame also tend to align their bodies with the tilted room and have difficulty finding the embedded figures.
After decades of research into field-dependent and field-independent children and adults, it is clear that these cognitive style differences are associated with numerous other characteristics. Because they rely on their own internal cues, field-independent people are analytical and logical. They tend to gravitate towards occupations such as engineering, science teaching and experimental psychology, and they are often regarded by others as ambitious, inconsiderate and opportunistic. Field-dependent people, on the other hand, pay more attention to the environment and therefore excel at interpersonal relations. They are generally considered to be popular, friendly, warm and sensitive; they tend to prefer occupations such as social work, elementary school teaching and clinical psychology. Achievement in a particular occupation also relates to these cognitive style differences. In one study, for example, highly successful psychiatric nursing students were found to be much more field dependent than equally successful surgical nursing students, who tended to be field independent (Quinlan and Blatt, 1973).

Some researchers, notably the Canadian psychologist Philip Vernon (1972), have argued that field independence is merely an aspect of intelligence. Research has shown that field-independent people generally achieve higher, and field-dependent people lower, IQ scores on performance subtests of IQ tests such as the Wechsler scales. In other words, field independence and scores on these subtests are correlated. What is more, between the ages of about 5 and 17 years, people generally become increasingly field independent. This strengthens Vernon’s claim that field independence is merely an aspect of intelligence, because intelligence also increases with age up to about 17 years.

Nevertheless, there are good reasons for believing that field independence is not equivalent to general intelligence. Girls and women, for instance, are generally found to be more field dependent than boys and men, although there is no evidence of any difference in general intelligence between the sexes. It is, in fact, field dependence which is associated with superior functioning in various kinds of situations involving social interaction. Field-dependent people have been found, for example, to be better than field-independent people at interpreting information from the social environment. They are better at attending to, decoding and remembering facial expressions and verbal messages concerned with social relations and attitudes (Witkin, 1976; Hampson, 1982). The evidence suggests that field-dependent people may be more socially intelligent than field-independent people, and that IQ tests fail to measure this kind of intelligence.

5.2 Convergence–divergence

Convergent thinking is a cognitive style characterized by a tendency to focus on a single best solution to a problem. Convergent thinkers are at their best when dealing with problems that have unique solutions. At the opposite extreme is the divergent thinking style, characterized by the fluent production of a variety of novel ideas relevant to the problem in hand. Divergent thinkers prefer, and perform better at, open-ended problems which do not have unique solutions. These cognitive styles, first identified and named by the American
psychologist J.P. Guilford in the 1950s, have been investigated by Getzels and Jackson (1962), Wallach and Kogan (1965), Hudson (1966, 1968) and many others.

Convergent thinking is measured by conventional IQ test items which have unique correct answers (though you may think that the answer to SAQ 10, which involved the anagrams of boys' and girls' names, required a heap of 'divergent' imagination). Tests of divergent thinking tap more creative types of thinking. A typical item from a test to measure divergent thinking is the following: 'How many uses can you think of for a brick?' In his study of English schoolboys, Hudson (1966) found that many boys could think of only three or four answers in about 3 minutes, but most of those whom he identified as divergent thinkers gave ten or more answers. One of Hudson's divergent thinkers, for example, gave the following set of answers: 'To break windows for robbery, to determine the depth of wells, to use as ammunition, as pendulum, to practise carving, wall building, to demonstrate Archimedes' Principle, as part of abstract sculpture, cosh, ballast, weight for dropping things in river, etc., as a hammer, keep door open, footwiper, use as rubble for path filling, chock, weight on scale, to prop up wobbly table, paperweight, as firehearth, to block up rabbit hole'. Some other items from tests of divergent thinking are given in Box D.

**BOX D Some tests of divergent thinking**

1. How many words can you think of ending in '-tion'?
2. How many words can you think of that mean roughly the same as 'hard'?
3. How many objects can you think of that are round and edible?
4. Suppose that people no longer needed or wanted to sleep. How many consequences of this can you think of?

(Adapted from Guilford, 1959)

Several interesting differences between convergent and divergent thinkers have been discovered. Hudson (1966, 1968) found that, even while they are still at school, convergers usually specialize in physical sciences, mathematics, or classics, hold conventional attitudes and opinions, pursue technical or mechanical interests in their spare time, and tend to be emotionally inhibited. Divergers, on the other hand, usually specialize in the arts or biology, hold unconventional attitudes and opinions, have spare-time interests involving interaction with other people, and tend to be emotionally uninhibited.

Some psychologists believe that divergent thinking is equivalent to creativity. It is certainly true that conventional IQ tests measure only convergent thinking and that, beyond an IQ of about 140, intelligence plays a negligible role in explaining the original contributions of creative people in various fields. It may also be true, though this is more controversial, that divergent thinking is a necessary component of creative ability. But scores on divergent thinking tests correlate only modestly with independent criteria of creative
ability, such as critical acclaim among professional artists. There is more to creativity than divergent thinking. Personality characteristics such as a willingness to take risks, self-confidence, perseverance and curiosity no doubt play a part in creative achievement.

If divergent thinking is an essential prerequisite of exceptional intellectual performance, it is ironic that candidates for higher education are selected largely on the basis of their ability to do well in school examinations which, for the most part, tap only convergent thinking. It is worth remembering that many creative geniuses, including scientists like Albert Einstein, had undistinguished school records.

5.3 Concluding comments

Psychologists have identified numerous cognitive styles apart from the ones discussed in this section. Field dependence-independence has been investigated more thoroughly than any other cognitive style, but research into convergence-divergence has also contributed significantly to our understanding of individual differences, especially individual differences in thinking.

The psychometric approach to the study of intelligence, discussed in earlier sections, concentrates on individual differences in global intellectual ability as measured by IQ tests. The object of research into cognitive styles, on the other hand, is to describe and understand individual differences in people's manner of thinking rather than in their ability to think. Research into cognitive styles highlights the fact that there are important cognitive differences between people, apart from overall differences in intelligence.

Some recent critics have argued that field independence in particular, and cognitive styles in general, are simply specialized mental abilities (Morrison, 1988). The implication is that cognitive styles are really aspects of intelligence. It is certainly true that field-independent people are better than field-dependent people at ignoring external distractions when judging whether luminous rods or their own bodies are upright and when locating a simple figure embedded in a complex diagram. Field-independent people therefore perform better on these tests than field-dependent people. But there are situations requiring social skills and creative thinking in which the field-independent cognitive style could be a handicap rather than an advantage (Witkin, 1976; Hampson, 1982).

Although tests of cognitive styles inevitably appear to favour one particular cognitive style rather than its opposite, researchers have sometimes managed to devise alternative tests of the same cognitive styles in which the advantage is reversed. Whether or not this could be done with all cognitive styles is an open question.

Most of this chapter has focused on the psychometric approach, which assumes that IQ tests measure all important aspects of intelligence. IQ tests certainly seem to tap intellectual skills such as logical thinking and the ability to acquire information and to use it effectively. But there are arguments to
suggest that a broader approach to intelligence might be useful. A recent theorist, Sternberg (1984), has emphasized the importance of analysing the underlying processes and skills that are required to do specific tasks. In his view, intelligence includes the ability to transfer skills to novel tasks and to execute mental operations rapidly, smoothly and effortlessly. Analysed in this way, the concept of intelligence is broadly applicable to performance on practical problems in everyday life as well as formal intelligence tests. In general, research into individuals' intelligence and cognitive styles has drawn attention to the variety of intellectual abilities and ways of thinking.

**Summary of Section 5**

- Cognitive styles refer to differences in people's manner of thinking rather than their overall level of intellectual ability.
- Field independence is characterized by the ability to make judgements based on internal cues. Field dependence involves more reliance on external factors.
- Convergent thinkers are best at logical thinking and solving problems with unique solutions. Divergers tend to produce a wide variety of novel ideas, and tend thus to perform better at problems which do not have unique solutions.
- Cognitive styles have advantages and disadvantages. People who score highly on field independence and convergence tend to do well on IQ tests. Divergers tend to be more intuitive and creative and, if they are also field dependent, are more likely to display social intelligence.

**Further reading**

The following two books contain excellent accounts of the history of intelligence testing. The second also explains the basic statistical ideas very clearly. They are both very well written:


For an orthodox account of psychometric approaches to intelligence, with masses of factual information clearly presented, you might refer to:


A summary of the arguments and evidence on both sides of the nature-nurture debate is presented in the following book. The authors are leading figures in the debate representing opposite points of view:

The main controversies surrounding intelligence and a few other issues in psychology are dealt with in:


The best general introductions to the study of cognitive styles are found in the following two books. The first contains contributions from leading researchers into cognitive styles. The second is an introductory textbook.


References


London: Macmillan.
CHAPTER 7 ASPECTS OF INTELLIGENCE


Answers to SAQs

SAQ 1
(a) Moderately positive;
(b) highly positive;
(c) negative.

SAQ 2
About 10 minutes past 11.

SAQ 3
(a) They do not correlate with one another or with students' examination marks.
(b) Older children perform better on his tests than do younger children.
(c) Galton’s tests measure sensory and motor functions; Binet’s tests measure the ability to deal with problems intelligently.

SAQ 4
(a) Elizabeth: \(\frac{5}{6} \times 100 = 80\); Andrew: \(\frac{5}{6} \times 100 = 100\); William: \(\frac{5}{6} \times 100 = 12\)
(b) Mark: \(2\frac{3}{5} \times 100 = 100\); Philip: \(2\frac{3}{4} \times 100 = 50\). This is absurd because there is no reason to believe that Philip's intelligence is below average. We would not think an adult of 40 mentally retarded just because he obtained the same score on an IQ test as a 20 year old.

SAQ 5
(a) \(9\); 
(b) \(2\), i.e. scores of 8 and 9; 
(c) \(2\); 
(d) \(0 + 2 + (2 \times 3) + (3 \times 4) + (5 \times 5) + (4 \times 6) + (2 \times 7) + 8 + 9 = 100, + 20 = 5\).
SAQ 6
(a) 2.14%;
(b) 2.14%;
(c) 2.14% + 13.59% = 15.73%.

SAQ 7
(a) 80 = 3 standard deviations above (+ 3 SD) the mean, i.e. $3 \times 15 = 45$ IQ points above 100; $100 + 45 = 145$ IQ points.
(b) 30 = −2 SD = $2 \times 15 = 30$ points below 100; 100 − 30 = 70 IQ points.
(c) 70 = + 2 SD = $2 \times 15 = 30$ points above 100; 100 + 30 = 130 IQ points.

SAQ 8
(a) 0.81;
(b) 0.74 (you will have had to look down the Vocabulary column to find the correlation with Comprehension. The blank half of the matrix is a mirror image of the other half);
(c) 0.68;
(d) 0.63.

SAQ 9
Diagram number 6.

SAQ 10
(a) (ii) (there are two letters before C and after X in the alphabet, and there are five letters before F and after U).
(b) (ii) (the anagrams Olga, Rose, Ruth and Dawn are female names; Neil is a male name). We congratulate you if you got this one right!
(c) The missing number is 32 (the differences between adjacent numbers are 2, 6, 10, 14—that is, every other even number).

SAQ 11
Galton’s critics pointed out that people with eminent relatives may be successful because of superior nurture, in the form of wealth, education and personal connections, rather than superior nature, in the form of hereditary factors.

SAQ 12

Figure 7.6 A difficult item from the embedded figures test with the embedded figure picked out
(Source: Witkin, 1950)