7

Intelligence

- Measurement of intelligence, using the methods of PSYCHOMETRICS, is in principle similar to the measurement of any other physical or biological quantity.
- Intelligence tests should be both RELIABLE and VALID. Tests can show CRITERION VALIDITY, FACE VALIDITY, CONSTRUCT VALIDITY and INTERNAL CONSISTENCY.
- Individuals differ in their intelligence. Attribution of these differences to genetic or environmental factors can only be relative to a particular time and population. Probably 40–50% of variance is genetic and 50–60% environmental in western populations.
- Genetic influences on IQ are best demonstrated in the children of monozygotic twins, where the confounding effects of differential treatment of twins by their parents can be taken into account.
- Monozygotic twins are less intelligent than dizygotic twins because of their postnatal environment rather than their prenatal environment.
- Children from small families and those born early in their family are relatively more intelligent. These environmental effects can be explained by Zajonc's CONFLUENCE MODEL, which considers the total intellectual environment to which the child is exposed.

Previous chapters have implied that individuals are all rather similar; and indeed they do have similar basic processes of perception, learning, memory, cognition, language, etc, and have similar brain structures. However people also *differ* psychologically, in their intelligence, personality, and in how they view the world. How do we measure differences in intellectual ability (PSYCHOMETRICS) and where do they come from?

INTELLIGENCE can either mean COMPETENCE, the highest intellectual activity of which a person is *capable* (but might not ever perform), or PERFORMANCE, almost tautologically defined as 'that which is measured by intelligence tests'). When in 1905 Binet and Simon described the first intelligence test, their intention was totally pragmatic:

> 'Our aim is, when a child is put before us, to take the measurement of his intellectual powers, in order to establish whether he is normal or if he is retarded. . . . we shall not seek in any way to establish or

to prepare a prognosis and we leave the question of whether his backwardness is curable or not, capable of improvement or not, entirely unanswered'.

Doctors are mostly interested in intelligence as *performance*, when assessments are required for placing a backward child in an appropriate educational institution, in diagnosing reading or school difficulties, assessing head injury or brain disease in adults, or in research projects. Sometimes competence can be assessed: a child not doing well at school who has a high IQ is under-achieving, and hence competence is above performance.

Any form of measurement (be it of length, size, liver function or IQ) requires assumptions. For any form of measure, these assumptions can be assessed. To show the similarities I will compare a biochemical measurement (of liver function) with the psychological measurement of intelligence.

Measurement starts with an intuitive idea of how different items should be ordered. Liver function should be worse in an advanced cirrhotic than in mild hepatitis which in turn should be worse than for a normal individual. If a biochemical test did not put groups in that order it would be rejected. CRITERION VALIDITY is therefore the ability to predict a well-known and understood characteristic. Similarly an IQ test should, on average, put research scientists above bank clerks, and road sweepers above the congenitally idiot. This criterion is *necessary* for a measure, but it is *not sufficient*.

Patients with severe liver disease might also have impaired exercise tolerance, but it would not be a reasonable *measure* of liver failure because it has no obvious relation to liver metabolism *per se*. CONSTRUCT VALIDITY requires a plausible theoretical explanation for a test doing what it claims to do. Knowledge of Mozart's symphonies, or income, would probably show criterion validity as IQ tests, but would not show construct validity. Construct validity requires a definition, and for intelligence it is typically:

> 'The general ability to process abstract quantities, independent of cultural or education background, of memory, or of particular knowledge.'

IQ tests are therefore constructed to assess *only* intellectual ability, avoiding specific cultural knowledge, so that as far as possible they are CULTURE-FREE (although results from minority cultures must be interpreted carefully).

A patient will feel a blood test could reasonably assess liver function, but would be surprised if electrodes were put on their head. FACE VALIDITY means that those taking or use the test *feel* it is measuring what it should measure; and they have CONTENT VALIDITY when they measure *all* relevant aspects (so an intelligence test will assess verbal, spatial and other intellectual abilities, since all are included within the broad term 'intelligence').

A dozen different tests of liver function should all show broadly similar results; they should be correlated, or show INTERNAL CONSISTENCY. If the items of an IQ test are ordered from most difficult to easiest then a person answering any item should also answer correctly all easier items. A consequence is that if two scores are calculated, from the odd and even numbered items, then the scores should correlate highly (SPLIT-HALF RELIABILITY). Similarly if the same person is tested again after a few weeks both scores should be similar (TEST-RETEST RELIABILITY). Good tests should be reliable *and* valid (although tests may be highly reliable and yet have little validity).

Returning to liver failure it might be objected that there can be no such thing as 'liver failure' in the abstract. Perhaps there should be two separate measures, one of parenchymal liver failure, due to hepatocyte damage, and a second due to extrinsic damage of the biliary system, and each could have its own specific tests. And indeed abnormal levels of SGOT, SGPT or glutamine transaminase imply parenchymal damage, and abnormal alkaline phosphatase or 5-NT levels assess biliary tract damage. There then will be greater correlations between the items within one group than between items in different groups. The same applies to intelligence; items assessing verbal ability correlate more closely with other verbal items than with those measuring spatial ability. On this basis, Thurstone and Guilford suggested that there is no single, unitary intelligence, but only specific intelligences, each independent in some sense. Returning to liver failure it can be argued that whilst liver failure may have different causes, all that matters to the patient is the overall degree of failure, particularly as each may cause the other anyway. Therefore, it may be better to measure overall failure, and also to create a second measure of 'relatively parenchymal' versus 'relatively biliary'. A similar argument applies to intelligence. Since verbal, spatial and arithmetic sub-tests all intercorrelate, Burt, Spearman and others argued for measures of GENERAL INTELLIGENCE (G) and of COGNITIVE STYLE ('relatively verbal' versus 'relatively spatial'). The important thing about such controversies is that they are pseudo-controversies. Intelligence and liver failure are multidimensional concepts; the best single number to describe them is, in a real sense, arbitrary, depending on the purpose of our measure. Similarly, a rectangle can be described by its height and breadth, or by its area and perimeter; both are equivalent and contain equal information; area is better if you are a surveyor, and height and width if you are a carpenter. With liver failure, an overall measure is required to assess overall fitness for work, whereas deciding whether or not to operate requires a specific measure of the extent of extrinsic damage. With intelligence, different measures are needed for asking if a person is suitable for a sheltered workshop or if they have a specific reading defect.

Given that we can measure intellectual performance, how and why do individuals differ? Intelligence levels are normally distributed, with a mean of 100 and standard deviation of 15. This is hardly interesting as the tests are constructed to have such properties. Five per cent of the population therefore have an IQ above 125, and 5% below 75; and 0.1% above 143 and 0.1% below 57. To be more practical, an IQ of about 100 is needed to gain GCSE grades A or B, 110 to obtain A-levels and 120 to enter university; medical students have an average IQ of 125. Arguably an IQ of 70–75 is sufficient for independent existence in a technological society.

Many intelligence tests are in common use. Some, like RAVEN'S PROGRESSIVE MATRICES, cover the entire range of ability and are completely non-verbal. Tests in clinical use are the WAIS (WECHSLER ADULT INTELLIGENCE SCALE), the WISC (WECHSLER INTELLIGENCE SCALE FOR CHILDREN) and the WPPSI (WECHSLER PRE-SCHOOL AND PRIMARY SCALE OF INTELLIGENCE), which measure specific abilities as well as general intelligence. A similar, more recently introduced test is the BRITISH ABILITY SCALES. Tests such as the AH4, AH5 and AH6 are designed to distinguish between university students, and most items are too difficult for the average person. Tests like the AH6 give separate measures of verbal, numerical and spatial intelligence, and discrepancies between the scores can be important.

The causes of differences in intelligence are controversial. Environmental theorists argue that IQ is mainly dependent upon differences in environment (diet, education, culture, etc.) whereas genetic theorists emphasize differences in genetic inheritance. Crudely, it is obvious that low IQ associated with cerebral palsy due to anoxia during delivery is environmental, whereas low IQ in untreated phenylketonuria is genetic. But the latter case shows the problems of attributing causality, for a phenylketonuric child in an environment with no dietary phenylalanine develops a normal IQ. There are no *absolute* answers to such questions, we can merely say genes or environment are *relatively* more important in a particular population at a particular time. Consider the less controversial topic of height. Half a millennium ago adult height in Britain was mainly a function of diet; now with no major nutritional deficits, differences in height are principally genetic. Ironically, it should be realized that the imposition of a uniform social and intellectual environment will result in all remaining differences being genetic; and in genetically identical clones, all differences must be environmental in origin.

A common method for assessing contributions of genes and environment involves monozygotic (MZ) twins. If a character is genetic, then since MZ twins have identical genes they should be more similar one to another than dizygotic (DZ) twins, who share only half their genes. However, MZ twins not only share genes, but also have more similar environments, being treated more similarly by their parents. In addition they sometimes develop PRIVATE LANGUAGES, which interfere with normal intellectual development. MZ twins have an average IQ of 95, compared with 100 for DZ twins and singletons. This discrepancy is not due to prenatal influences since if one member of the pair dies during childbirth, the IQ of the remaining twin averages 100. It is postnatal treatment that makes MZ twins have a lower IQ.

MZ twins separated at birth and reared in different families have identical genes but different environments, and hence any similarities ought to be genetic in origin. However, such twins are very rare (perhaps 300 cases in the world's literature), and their circumstances so special, often illegitimate or evacuated in war-time, or with retarded parents, that they cannot resolve the gene–environment controversy.

In the last two decades, a new method has resolved most of the problems by looking not at twins themselves, but at their children, who are mainly singletons. Think of two female MZ twins, each with children. The children of one twin are as closely related to their aunt as to their mother; and while in a normal family a child shares one eighth of a cousin's genes (half of the mother's genes, who has half the aunt's genes, who has half the cousin's genes), in MZ twin families a child shares a quarter of its cousin's genes. Cousins are raised in different families, and hence different environments, so it is possible, by using sophisticated statistical models to compare twin families with ordinary families and thereby find if intelligence has genetic or environmental components. Probably most researchers would now agree that about 40–50% of IQ variance has genetic origins, with the remaining 50–60% of environmental origin.

What then are the environmental influences upon IQ? Many factors contribute: for instance the FAMILY LEARNING ENVIRONMENT, parental provision of books, conversation and intellectual support, has an influence over and beyond the fact that parents of higher intelligence are more likely to provide a good environment. The differences in IQ between birth cohorts (see Chapter 17) also suggest an important role for environmental effects. Particularly interesting work has examined the influence of birth order and family size. Since genes cannot know the order in which they are born, or how many siblings they have, then effects of birth order and family size must be environmental. Later born children are of lower average IQ than firstborns, and children in smaller families have lower IQ, both effects being independent of social class (Fig. 7.1). Zajonc's CONFLUENCE MODEL accounts for findings of this sort. A first child initially has the undivided attention of its parents, and hence is exposed to undiluted adult intelligence. A second child necessarily receives less parental attention than its older sibling, since the parents' time is shared with the older sibling (although it does receive the partial compensation of being exposed to the rather lower intellect of its older sibling). The intellectual

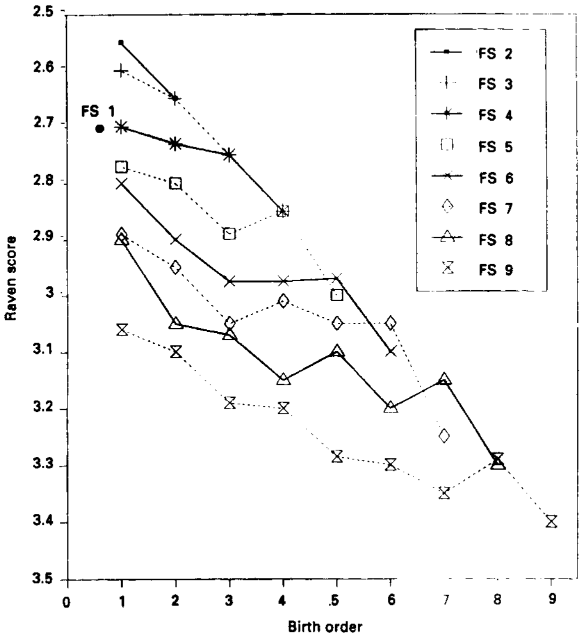

Fig. 7.1 Intelligence as measured by Raven's Progressive Matrices in 386 114 nineteen-year-old men who had been tested on entry into Dutch military conscription, in relation to the size of their family (FS) and their order in that sequence (birth order). Note that Family Size 1 is slightly laterally displaced to avoid confusion with the first born of family size 4. Because of the method of scoring used, low scores (e.g. 2.5) indicate *higher* intelligence than higher scores (e.g. 3.5). Adapted from Belmont L and Marolla F A (1973), Birth order, family size and intelligence, *Science*, **182**, 1096–1101.

environment of a second child is therefore less rich than the first child, and the first child's environment is also diluted; and so on for each additional child. The precise mathematical model predicts that first born children and children from smaller families should be more intelligent, as is seen in Figure 7.1. In the extended, non-nuclear families typical of developing countries, the last born in large families has a relatively increased IQ (see Fig. 7.2). The model copes well with this since although parental attention is further diluted, older siblings are now adult and continue to live in the parental home, where they can enrich the youngest children's intellectual environment.

The basic model needs one further refinement, since only children have a lower IQ than the first-born of families of two (see Fig. 7.1). This is ascribed to the beneficial effects to an older child of teaching their younger siblings: for teaching benefits teachers as much as pupils.

IQ is often criticized as a practical measure in that those with

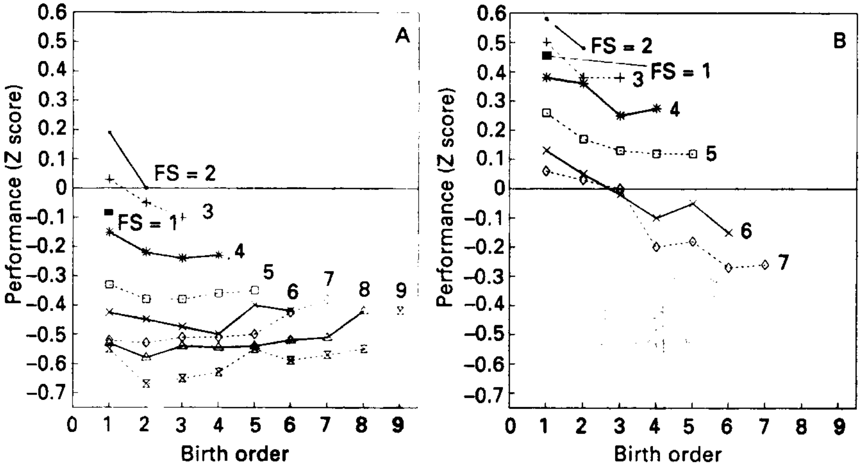

Fig. 7.2 Intelligence test scores (expressed as standard normal deviates, with mean of zero and standard deviation of one) of 191 993 Israeli school children, in relation to family size and birth order (singletons being indicated by a solid square). *a*) Students of predominantly Middle Eastern or North African origin (N = 109 304) and hence predominantly from extended families; *b*) Students of European, American and Australasian origin (N = 82 689) and predominantly from nuclear families. Adapted from Davis D J, Cahan S and Bashi J (1977), Birth order and intellectual development: the confluence model in the light of cross-cultural evidence, *Science*, **196**, 1470–2.

highest intelligence do not usually become leaders of society, politicians, captains of industry, etc. This is not due to intelligence being invalid as a measure, but rather reflects the complexity of the social psychology of leadership. Although high intelligence means greater reasoning power, and hence greater ability, this is of diminishing benefit as IQ increases; it matters little to be more intellectually able than 99% of the population rather than 98%. More importantly, if the arguments of those of high intelligence are not understood by those of lesser intelligence, then they will have no influence. Individuals typically only influence by argument those down to about ten or fifteen IQ points below themselves, and hence the highest intellects influence very few. Intelligence alone is not the only requirement of social success and influence, and hence the most successful are often of only moderately high intelligence (to the chagrin of the highly intelligent).