Developmental tests and scales

Casual judgments about a child's development are often incorrect and can result in serious mismanagement and considerable parental distress. Accurate developmental diagnosis involves careful observation of intellectual, social, and motor abilities; the science of measuring these and other psychological functions is called psychometrics.¹ This article considers the relevance of psychometric principles to developmental paediatrics. The other components of developmental diagnosis, such as vision and hearing tests, have recently been reviewed elsewhere.²

Definitions in psychometrics

A standardised test is one in which the materials to be used and the method of administration are specified. Normative data for a test is the body of information telling us how other children perform on the same test and the extent to which the child being tested differs from an 'average' child.

The psychometric adequacy of a particular test is evaluated on the basis of its reliability and validity. Reliability is a measure of the extent to which the test will give the same result when given by different testers or on different occasions. Concurrent validity is concerned with questions such as 'Does this test measure what it claims to measure?' and 'How well do the results correlate with other measures of the same function?' Predictive validity is the ability of a test to predict future performance or achievement.

A developmental scale—for example, Sheridan's 'Developmental progress'³—consists of an inventory of abilities and milestones. It can be regarded as a simple form of psychometric test whose method of administration is not rigorously standardised and whose normative data consist only of approximate mean ages at which the various milestones are reached. A scale should be used primarily as a guide to *normal* development. Although a below average performance can be described in terms of age equivalence—that is, how many months behind—the significance of this result is difficult to evaluate because the psychometric data mentioned previously are not available.

The content of tests

Tests consist of many individual items in three distinct categories. Firstly, some make direct use of the parents' unique knowledge of their child's

current abilities.⁴ This is usually very reliable, though recall of past milestones is much less accurate.⁵ Tests such as the Vineland Adaptive Behaviour Scale,⁶ and some screening questionnaires,⁷ depend entirely on parental information.

Secondly, spontaneous behaviours such as sitting, walking, and social responsiveness can be observed. The third type of test item involves the use of special materials or activities to elicit particular skills—for example, bricks, form boards, and miniature toys are specified in many tests and scales. If the test battery uses information obtained by all three methods, the results are more likely to be valid, and this is particularly important if the parents' account is thought to be unreliable or the child is upset or excessively shy.

Test items are sometimes classified into functional groups, such as gross motor, fine motor, social, and language skills. This facilitates observation and recording, but the categories have limited diagnostic value in themselves. For example, inability to build a tower of bricks may be due to poor vision, ataxia, or mental handicap.

Interpretation

The interpretation of the results is as important as correct selection and administration of the test. It must be remembered that even the most sophisticated test can yield inadequate results, for instance if the child was unwell or anxious at the time of testing or if the test was not administered correctly.⁸

In the first two years of life low scores on developmental tests have limited predictive value for future progress unless they are in the very retarded range. Predictive value increases as the child grows older, but prediction is still subject to serious error because of individual fluctuations in development. Low scores on individual subtests or subsections may signify a specific disability, but sometimes they simply represent variations in the rate at which different abilities emerge.^{9 10}

Screening and assessment

There is an important distinction between screening and assessment. Screening is the process of checking the development of children whose parents believe them to be normal. Assessment is a problem solving exercise, which has several components. It is important to establish who is worried and the nature of the

problem. Information is collected by interviewing parents and others who know the child, followed by examination, appropriate psychological tests, hearing and vision checks, and medical investigations if indicated. Lastly, the results are interpreted and explained to the parents and a remedial programme is planned.

Tests for assessment

It is usual to begin with an informal interview and observation session. A developmental scale and various assessment kits such as the *Stycar* series developed by Sheridan are often useful. Formal psychometric testing, like other investigations, should be undertaken only when clear questions have been defined, and then the selection of appropriate tests is usually straightforward. In some cases, psychometric testing is irrelevant to the issues that concern the parents—for example, there may be conflict between parents and professionals over the interpretation of previous test results and the extent to which they should influence school placement.

Detailed testing is often needed when it needs to be known whether, and how much, a child is slow compared with his peers. For example, speech delay may be an isolated problem or may be associated with global retardation. Many tests are designed to answer this general question, such as the *Wechsler Intelligence Scale for Children*, the *Merrill Palmer*, the *British Ability Scales*, and the *Griffiths Developmental Scale*. The latter can be used by doctors who have received an approved training. The normative data for the *Griffiths* test recently underwent a much needed revision, but this had not yet been published.

More specific problems may need specialised test—for example, speech and language delay may be assessed using the *Reynell Developmental Language Scale*. The intelligence of a child who is deaf, dysphasic, or blind is easily misjudged if one uses a test whose norms were based on a sample of children without these disabilities. Appropriate tests are available for these and many other complex problems; details will be found in standard texts.⁸

The tests described above provide normative data that allow a child's development to be compared with that of his peers, but this may not be very helpful if the child is already known to be handicapped. It is more important to find out how the child has progressed over a period and what intermediate steps are needed to achieve the next milestone. Some tests have been designed specifically for this purpose, and these usually form part of

a teaching package—for example, the *Portage* checklist and the *PIP* (parent involvement project).

Screening tests

There is little value in a programme capable only of detecting obvious major defects, because these are most often recognised at the neonatal or six week examination or are actively sought after a neurological illness. The remainder are suspected first by the parents.² Screening might be worth while if subtle treatable disorders previously overlooked by the parents could be detected—for example, language deficits or potential learning difficulties. The rather low predictive value of developmental tests means that they can only indicate how much a child deviates from the norm. Subjective clinical judgment is needed to decide whether there is a problem.

For this reason it can be argued that developmental screening requires an experienced doctor who can treat each examination as a 'mini-assessment'. Whether this is an economic use of medical manpower is a separate and contentious issue. The alternative approach is the identification of high risk groups¹¹ by paramedical personnel, using a test with arbitrary pass/fail criteria. The best known example of such a test is the *Denver Developmental Screening Test*, which has satisfactory normative data and reasonable, though inevitably imperfect, validity.¹²

Much further work is needed before either type of developmental screening programme can be said to fulfil the criteria for screening that were laid down by Wilson and Jungner.¹³

Conclusion

The sale of most psychometric tests is restricted to suitably qualified persons. Developmental paediatricians can acquire the necessary training if they so wish, but many prefer to consult their colleagues in psychology for detailed psychometric testing and for advice on interpretation.

In some cases, assessment may be performed by only one person, while in others a large multidisciplinary team may be involved. Whoever carries out the actual testing, it is the paediatrician's responsibility to ensure that parents are told the results and the interpretation of the entire assessment procedure, preferably in writing as well as verbally.¹⁴ If a problem has been identified, however tentatively, the parents will need information on its severity, its cause if known, and the potential for improvement. Lastly, they should be encouraged to participate in the planning and execution of a remedial pro-

gramme that is designed to give the child as normal a life as possible.

We are grateful to Dr M Berger and Mr R Smith for helpful discussion and advice.

References

- ¹ Berger M. Towards an educated use of IQ tests: a reappraisal of intelligence testing. In: Lakey B, Kazdin A, eds. *Advances in clinical child psychology*, Vol. 9. New York: Plenum, 1986.
- ² Baird G, Hall DMB. Developmental paediatrics in primary care: what should we teach? *Br Med J* 1985;**291**:583-6.
- ³ Sheridan M. *Children's developmental progress: from birth to 5 years*. Windsor: NFER, 1975.
- ⁴ Newson E. Parents as a resource in diagnosis and assessment. In: *Early management of handicapping disorders*. IRMMH review of research and practice. Vol. 19. Associated Scientific Publishers, 1976:105-17.
- ⁵ Hart H, Bax M, Jenkins S. The value of a developmental history. *Dev Med Child Neurol* 1978;**20**:442-52.
- ⁶ Sparrow S, Balla D, Cicchetti D. *Vineland adaptive behaviour scales*. Minnesota: American Guidance Service, 1984.
- ⁷ Frankenburg W, van Doorninck W, Lidell T, Dick N. The Denver pre-screening developmental questionnaire (PDQ). *Pediatrics* 1976;**57**:744-53.
- ⁸ Berger M, Yule W. Psychometric approaches. In: Hogg J, Raynes N, eds. *Which assessment? A guide to tests, batteries and checklists in mental handicap*. Beckenham: Croom Helm, 1986.
- ⁹ Pennington B, Smith S. Genetic influences on learning disabilities and speech and language disorders. *Child Dev* 1983;**54**:369-87.
- ¹⁰ Robson P. Shuffling, hitching, scooting or sliding: observations in 30 otherwise normal children. *Dev Med Child Neurol* 1970;**12**:608-17.
- ¹¹ Silva P, McGee R, Williams S. Developmental language delay from 3 to 7 years and its significance for low intelligence and reading difficulties at age 7. *Dev Med Child Neurol* 1983;**25**:783-93.
- ¹² Frankenburg W, Camp B, van Natta P. Validity of the Denver developmental screening test. *Child Dev* 1971;**42**:475-85.
- ¹³ Wilson J, Jungner G. *Principles and practice of screening for disease*. Public Health Papers no 34. Geneva: WHO, 1968.
- ¹⁴ Hall D. *The child with a handicap*. Oxford: Blackwell Scientific Publications, 1984:134.

D M B HALL
St George's Hospital Medical School,
London SW17, and
G BAIRD
Guy's Hospital,
London SE1