

PAUL WILSON

<paul@wilson1123.freemove.co.uk>

## THE USE OF CAT SCREENING TESTS IN IDENTIFYING SPECIAL EDUCATIONAL NEEDS IN MATHEMATICS

### INTRODUCTION

Measurement of performance is a key component of current educational policy and practice. To help establish measures of “added value<sup>1</sup>”, LEAs and schools have adopted large-scale assessment procedures known as screening programmes. Many of the screening packages and tests used are claimed to have diagnostic potential for individual students. I will be investigating the test results for a large English 11-18 comprehensive school, whose LEA offers the NFER Cognitive Abilities Test (Thorndike & France, 1986) to screen all students early in year 7. In the administration manual accompanying the tests (Thorndike, Hagen & France, 1986), it is claimed that the profiles generated can be used to identify learning needs and appropriate teaching strategies for individual students. In the following account I will explore ways in which these tests (CATs) may be used to identify and support those students with special needs in mathematics. I will concentrate on two areas:

- The screening test profiles of those students identified (independently of the tests) as having special needs in mathematics
- How the screening test profiles may be used to identify students with special needs in mathematics and what those needs are.

While it may be reasonably asserted that all students have their own “special” (or particular) needs in learning mathematics, I will concentrate on those students at both ends of the ability/attainment continuum, those students for whom specialised support is most necessary. In the next section I will review current thinking relating to special needs in mathematics, intelligence, aptitude and

---

<sup>1</sup> The term “added value” is borrowed from commerce, where it refers to the increase in market value of a product resulting from the activities of a business. This term has been adopted in education to refer to gains in academic attainment over and above that consistent with typical intellectual development.

psychological testing. The following sections will look at the NFER CATs in detail, including their suggested interpretation before describing the research carried out to investigate this.

## SPECIAL EDUCATIONAL NEEDS IN MATHEMATICS

Although teachers may share some common elements in their understanding of the term “special educational needs”, the phrase can assume different shades of meaning in different contexts. At one extreme it may refer to a student’s legal entitlement to specific provision as a direct result of a Statement of Need (produced by a panel of professionals with responsibility for the student’s educational and personal welfare). It may refer to cognitive learning deficits of pupils (of varying degrees). At the other extreme it may refer to the need for an accelerated programme of learning for a particular student. Often, the terms “special needs” and “low attainment” are used interchangeably. This can cause confusion. Low attainment is not always the result of a reduced capacity to learn. It may result from a number of factors: illness, absence from school, change of schools, ineffective teaching, alienation from the learning environment etc. Clearly a specific physical impairment will have a significant affect on learning (e.g. blindness, deafness) and specific provision must be made to ensure that this does not handicap the learner. A special need of this nature can be relatively easy to identify, even at an early stage in the child’s education, thus allowing appropriate provision to be made for the student. However, some special needs giving rise to learning difficulties in mathematics are harder to identify at an early stage. In the 1993 Code of Practice on Special Educational Needs a “continuum of need” is described (DFE, 1993). This can range from problems causing the student to make little or no progress when taught with his or her peers to students having a reduced facility in some particular aspect of mathematics. There is also an increasing awareness in schools of the needs of very able students, who may require special approaches to learning that are also inappropriate for their peers. It is clear that a rigorous definition of special needs in mathematics is difficult to provide. However, for the purpose of this account I shall use the following definition.

*A student whose progress in learning mathematics is inhibited when taught within the limits of the normal curriculum using normal resources can be said to have a special need in mathematics.*

The use of the adjective “normal” clouds the issue somewhat, but it is a necessary qualification. If we accept that special needs form a continuum, then the severity of need is best described by comparison with the needs of the majority of students. It is important to recognise that some students have greater needs than others do and that, in the allocation of resources to cater for this, due regard must be given to the nature and magnitude of the problem.

Statutory recognition of special needs was made in the 1944 Education Act, in terms of eleven “clinical” categories of disability: (deaf, partially deaf, blind, partially sighted, physically handicapped, delicate, diabetic, epileptic educationally sub-normal, speech deficient and maladjusted). In addition, children with IQs less than 50 were considered to be outside the scope of education (Daniels & Anghileri, 1995). These “disabilities” were to be addressed directly, often in separate “special school” provision. Since the nineteen-forties the emphasis has shifted towards the psychological in interpretation of special needs. Subsequent educational reforms have been predicated on the desirability of integrating provision wherever possible and on the provision of a curriculum that will

*“allow teachers and schools to meet the particular needs of pupils with Special Educational Needs in ways which they judge to be relevant”* (Dearing, 1993: 53)

There has been a corresponding change in classification of special needs, from the 1944 categories to those arising from the 1981 education act, and in use today.

- Mild learning difficulty
- Moderate learning difficulty
- Emotional and behavioural difficulty
- Specific learning difficulty
- Physical disability
- Physical and neurological impairment
- Visual impairment
- Hearing impairment

(Daniels & Anghiler, 1995: 5)

The biggest innovation in special needs education has been the implementation of the Code of Practice on Special Educational Needs (part of the 1993 Education Act). This requires schools to make appropriate provision for students with special educational needs and to monitor five stages of assessment and intervention for these students. The first of these stages involves the class teacher consulting the Special Educational Needs Co-ordinator (SENCO). At the second stage the SENCO takes the lead in making appropriate provision for the child. Stage three involves support from outside the school. At stage four moves are made towards obtaining a statement of Special Educational Needs, which brings with it extra, statutory, provision. Stage five corresponds to the issue of the Statement.

This outlines the legal position but, as one may expect, students' special needs are not always clear-cut and easy to categorise. For instance, a student's needs in mathematics may be very different from his or her linguistic needs. Some needs may be seen to be a result of genetic make-up and may therefore be innate. Some may arise as the result of an accident (brain damage, loss of sight). Some may be a consequence of social and environmental factors, including low self-esteem and lack of confidence. In the following paragraphs I will outline how different needs will affect a students' mathematics education.

Although the Piagetian model of intellectual development is no longer taken as definitive, it is still useful in describing the development of mathematical thinking (Piaget, 1952; Piaget & Inhelder, 1958). According to Piaget's theories, children pass through identifiable, fixed, hierarchies of cognitive development, where knowledge is assimilated and accommodated. This leads to cognitive reorganisation and, as a consequence, more sophisticated modes of reasoning. The Piagetian Stage Theory describes the forms of intellectual operation at each level of the hierarchy. The ages at which individuals reach each of these stages are variable, although all children pass through the same hierarchies in the same order. It is asserted that many children will not reach the final developmental stage during their years of schooling. Research carried out by CSMS (Hart et al, 1981) supports the usefulness of this model. In the light of this, it can be seen that many students with special needs in mathematics are making slow progress through the levels of the Piagetian hierarchy, rather than suffering from some cognitive dysfunction.

Slow progress can be the result of many factors:

- Lack of opportunity to develop concrete operations through disability or under-developed motor-skills.
- Emotional and/or behavioural problems
- Affective problems, low self-esteem, lack of confidence
- Poor reading skills
- Being taught in a language other than one's first
- Slow cognitive development

Neurological damage can lead to special needs in mathematics. In particular, I will describe dyslexia and dyscalculia. Dyslexia is a controversial issue in education. Although it was first described a hundred years ago, diagnosis of dyslexia has only become common in the last few decades. Commonly identified with mixing the orders of letters and words, educational psychologists recognise around twenty forms of dyslexia. Although there is a suspicion amongst some teachers that dyslexia is sometimes rather too freely diagnosed in order to obtain additional support for the student, there are recorded cases of dyslexia arising as the result of brain injury. With dyslexics' sequencing problems, mathematical processes involving place-value, arithmetic algorithms and mathematical symbolism can be problematic (Joffe, 1981; Miles 1983; Thomson, 1990). Dyscalculia is less commonly diagnosed, but again has been observed as a consequence of neurological damage where patients have lost a previous ability to perform calculations. Other associated problems are acalculia (sometimes referred to as acquired dyscalculia) and dysgraphia, which relates to graphs, tables and diagrams. However, many researchers are of the opinion that these syndromes are rare and that many purported cases of dyscalculia are likely to be examples of slow mathematical development.

## INTELLIGENCE, PSYCHOLOGICAL TESTING AND ABILITY

The term “intelligence” is widely used, often ambiguously. In common usage it can denote academic ability, but may refer to problem solving, communication or other intellectual facilities. In fact, many psychologists would claim that intelligence, as a single entity, does not exist. A compromise is to treat intelligence as an ability to do intelligence tests. This introduces a degree of reliability in the standard measures of intelligence and allows comparisons to be made. Most modern intelligence tests are composed of a number of separate sub-scales, off-setting the problem of intelligence as a single entity (but not if the results are merged to give a single result). However the issue of validity can be troublesome. There are different views on the development of intelligence. Some psychologists believe that intelligence is innate and immutable, in which case a well-designed test should be able to measure this with reasonable validity. Other psychologists believe that intelligence is influenced by cognitive experience and is therefore subject to change. This would mean that any measure of intelligence derived from testing would be, at best, temporary. Other theories suggest that intelligence is even less amenable to measurement. Horn (1971) proposed two forms of intelligence:

- crystallised – derived from experience and
- fluid – a general intelligence that can be applied to specific intellectual tasks.

Hann and Havinghurst (1957) and Gardner (1983) claim that there are many distinct components (multiple intelligences) which should be taken into account in quantifying intelligence.

The foundations of psychological testing were laid in the second half of the nineteenth century, as a way of discriminating between mental illness and low intellect. Methods were refined in the early years of the twentieth century when the notion of “intelligence tests” first arose. In 1905 Binet pioneered the use of “mental levels”, derived from a hierarchy of tests, cross-referenced to age from a standard sample. This gave rise to the use of “mental age”, though not by Binet himself (Anastasi, 1982). Subsequent developments led to the Stanford-Binet tests of 1916. These were the first to use the Intelligence Quotient, derived from the ratio of mental to chronological age (*ibid.*). (IQ is generally defined to be normally distributed with mean 100 and

standard deviation 15.) Around this time the distinction was drawn between intelligence tests and aptitude tests:

- Intelligence tests: single global score (such as IQ)
- Aptitude tests: measures of discrete “segments of ability” (ibid.)

There are two key measures that determine the value of any test, reliability and validity. Reliability is the property of yielding consistent results upon repetition and validity is the property of addressing those aspects under investigation rather than adjacent ones. Of these, validity is an area of contention. As illustrated above, notions of intelligence and aptitude are hard (or even impossible) to pin down, so tests purporting to measure intelligence and aptitude are open to critical evaluation.

Measurements of intelligence are useful if inferences can be made about an individual’s ability, in this case mathematical ability. Krutetskii (1976) identified nine components of mathematical ability, summarised by Orton (1987) as:

1. extracting formal mathematics from a problem and operating with it
2. generalising
3. using numbers and symbolism
4. spatial concepts
5. logical reasoning
6. shortening reasoning processes
7. flexibility in changing approach, avoiding fixations and reversing trains of thought
8. achieving clarity, simplicity, economy and rationality in argument and proof
9. a good memory for mathematical knowledge and ideas.

It is reasonable to assume that, although we may be born with predispositions to greater ability in some of these areas and less in others, all of these abilities can be developed through learning experiences. Too rigid a view of innate intelligence may lead to limited expectations and limiting educational experiences for some students. However, if psychological tests can indicate, for an

individual, those abilities that may be more advanced and those that may be less advanced, then this information will be of use in informing teaching and learning.

## THE COGNITIVE ABILITIES TEST

The Cognitive Abilities Test (CAT) is designed to measure a student's ability to operate with abstract and symbolic relationships. In its administration manual, emphasis is placed on relational thinking (Thorndike, Hagen & France, 1986). This is a multiple-choice test consisting of three batteries, comprising a total of ten subtests. Each of these subtests consists of a hierarchy of items arranged in six levels, to reflect the abilities of children from the ages of seven and a half (Level A) to nearly sixteen (Level F).

The verbal battery reflects the importance of verbal symbolism in education and is designed to test flexibility in the use of concepts. It consists of vocabulary, sentence construction, verbal classification and verbal analogies. The vocabulary items involve identifying synonyms, ranging from "leap" at Level A to "compassion" at Level F. Sentence construction requires the student to insert a word to make sense of a sentence. Verbal classification involves finding the word that belongs to the same set as three or four others: from "cat, dog, cow" at Level A to "concealed, hidden, potential, dormant" at Level F. In the verbal analogies items, the student is given two related words as an example. Another word is given and the student is required to select a word from the list which has the same relationship to it. Examples range from "funny → laugh : sad → " (Level A) to "acute → severe : weak → " (Level F).

The quantitative battery is designed to "require almost no reading of verbal symbols" (ibid: 1). It consists of quantitative relationships, number series and equation building. The quantitative relationship items require the student to state whether the first of a pair of numbers is greater than, equal to or less than the other. These items vary in difficulty from  $3 \times 2$  and  $2+2+2$  at Level A to  $(a^2+a)(a^2-a)$  and  $a^4-a^2$  (assuming  $a=2$ ) at Level F. The student has to find the next term in a sequence in the number series items. Examples range from 6, 7, 8, 9, 10, 11, ? (Level A), to  $8/5$ ,



7/6, 6/7, 5/8, ? (Level D) to 1, 3, 7, 15, 31, ? (Level F). The equation building items require the students to arrange a series of numbers and operations to produce one of the listed results. The items at level D and above involve brackets and those at Level F involve square roots.

Verbal reasoning and quantitative reasoning together are related to academic ability. The items in the non-verbal battery use neither words nor numbers and have little relationship with the formal school curriculum and it is claimed that this battery provides a measure of fluid intelligence. The non-verbal battery consists of, figure classification, figure analogies and figure synthesis. In the classification items, the student has to select a figure belonging to the same set as three others. Shading, closure, intersections, linearity and many other properties can classify the figures. In the figure analogies items, a figure and its image are given. The image corresponding to a new object is required. In the figure synthesis items several “component” shapes are shown and the student is required to specify which of a selection of shapes can be formed by from these components.

The administration of these tests requires the examiner to read a lot of instructions and “demands ...a certain amount of listening ability on the part of the test taker” (Buros 1972, 640). Some of the instructions are very wordy, particularly in the non-verbal battery. This does minimise the need to read instructions; “Influence of reading competence is eliminated by the use of paced, oral, item-by-item directions.” (ibid, 638). One drawback of the tests is the dependency on visual discrimination between multiple choice answers, many of which appear as small pictures (ibid).

Six different levels of test (A to F) are provided, with indications as to appropriate target age. The six levels of test are printed in one booklet and students answer on separate optical mark reader sheets, which are processed by a computer. The responses can be processed centrally or by the user to arrive at a variety of profiles, most expressed in terms of Standard Age Scores (SAS). As with intelligence tests, these scores are normally distributed with mean 100 and standard deviation 15. Scores are given for each of the three batteries, along with a mean score. It is possible to produce profiles by stanine (the allocation to one of nine ability bands) for each of the ten sub-sets identified above.

The results for these tests were restandardised in 1982, using a representative sample of approximately 2000 UK schoolchildren at each level. Reliability values of 0.8, 0.75 and 0.7 were obtained for the verbal, quantitative and non-verbal batteries respectively over a period of 2 to 3 years, with reasonably good correlations (~0.8) between the secondary school age tests and GCSE grades. NFER claims that the test is very reliable in the lower attainment ranges, but caution should be exercised in interpreting the scores of students in the top 10% in the quantitative and non-verbal batteries. NFER points out that the results indicate *current* developmental levels and warns users to exercise caution in three respects:

- developed abilities will affect the tests
- the scores should be looked at in the context of other knowledge of the individual
- all tests are subject to statistical error.(ibid: 47-49)

The CAT has proved useful in providing information about a cohort of students. NFER claims that it also provides useful information at an individual level (ibid: 53-55), and can support teachers in:

- choice of curriculum materials
- organisation of small groupings of students
- maintaining an appropriate pace of learning
- choice of teaching approaches
- target-setting.
- identifying learning difficulties
- advising students

In discussing the choice of curriculum materials, it is suggested that students with high verbal and quantitative abilities should be able to work at a higher level than the majority of their peers, in particular students with high verbal abilities should be particularly suited to independent learning. These students need to be challenged. Students with low reasoning skills in all three batteries need learning materials that are broken down into more accessible units, involving sufficient practice. Students with low verbal scores are unlikely to benefit from independent learning (with its reliance on text) and need specific teaching. Students with an uneven profile may have

different needs. A pupil with high quantitative and non-verbal but low verbal reasoning skills may have a good facility for mathematics (as long as problems with text can be minimised), but not in other areas. Particular attention should be given to students whose non-verbal skills are significantly higher than their verbal and quantitative skills. This is indicative of fluid intelligence. Curriculum materials should be carefully matched to the needs of these students.

It is pointed out that it is unwise to group children with very low scores in all three batteries together with those with very high scores. It is suggested that low-scoring children will benefit most from being grouped with children whose reasoning skills are a little more developed. Teachers may group pupils to accommodate learning needs. For example, a student with a high verbal reasoning score, but experiencing difficulty in mathematics, needs a different teaching approach from a child with low scores in all three batteries.

In identifying learning difficulties, three distinct categories are identified. Students with high quantitative and non-verbal, but low verbal scores should receive intensive help to develop reading and language development. A pupil with high non-verbal, but low verbal and quantitative scores has good fluid reasoning ability, but is showing a lack of ability in those areas most closely related to the school curriculum. A pupil with three very high scores, but whose academic work is average, may be under-achieving.

Teachers are recommended to use CAT scores in determining the pace of teaching and learning, the methods of instruction, advising students about curriculum choice and in setting targets. NFER provides tables mapping mean CAT scores onto projected KS3 and GCSE outcomes, for all subjects except English, in which the verbal score is shown to be a better indicator. (It is interesting to note that the mean score is a better indicator of performance in mathematics examinations than the quantitative score.)

## RATIONALE

The aim of this research was to identify how NFER CAT screening tests can be used to identify students with special needs in mathematics and what those needs are. In the large comprehensive school in which this research was carried out, the deputy principal with responsibility for monitoring student performance circulated papers to all members of the teaching staff to support them in their use of CAT scores. Several of the assertions in these papers are worth investigating and form a basis for research. One note claimed: "15 or more difference between the tests could identify specific learning difficulty." This paper is largely derived from the CAT administration manual (Thorndike, Hagen & France, 1986) although it contains several new suggestions as to how the data may be used. The same deputy principal and other members of the senior management team at the midlands comprehensive school where he had worked previously produced the first draft of this paper. The final page provides specific suggestions as to how school departments should use CAT scores. These suggestions extend the guidance given in the manual and were used as a framework for collecting data concerning the learning styles of students in mathematics lessons. Because these suggestions have a direct bearing on my research, I have listed them here, in full.

### *Verbal Scores: High*

*These students have an excellent grasp of written material and can be set more challenging work by departments. They need to be stretched and given assurances that they have considerable potential. Differentiated work, linked to challenging resources could be identified. These students could also be chosen to explain more difficult concepts to weaker peers.*

### *Verbal Scores; Low*

*These students are finding that the low scores in this area are seriously affecting their progress. Departments can use differentiated resources to interest and raise achievement. Units of work should be short and easy to follow, but should challenge and have specific targets attached to them. It's important to have high expectations of these students and to keep on encouraging and supporting them.*

*Numeracy Scores: High*

*These students can easily retain and remember mathematical facts and they have the skills and knowledge to tackle more challenging questions. They need tasks and questions which will allow them to “think through” the best way to achieve an answer by identifying the knowledge needed and the skills to work out the answer. They will also have the ability to check the validity of their answers. These students could be used to lead small groups, to motivate others and to work with weaker students.*

*Numeracy Scores: Low*

*These students will probably find basic concepts difficult to understand and will only have a limited knowledge, they will not easily retain the relevant facts. Their competence in number handling will be low and they will require much support whenever basic maths is required.*

*Non-verbal Scores: High*

*These students have good reasoning skills. They will tend to be able to explain concepts clearly. They need to be given more challenging problem-solving activities.*

*Non-verbal scores: Low*

*These students find problem-solving activities hard and may give up easily. They need a gradual build-up to the activity and more specialised resources.*

*If the Non-verbal Score is Significantly Higher than the Verbal Score*

*These students have a good intake (sic) ability and are good at problem solving. They can express themselves well but their reading and language is poor. (Appendix 2: 3)*

## METHODS

The data used were obtained from an 11-18 comprehensive school in the south-west of England, with just over 2000 students on roll. The CAT data were obtained from the test at level D (targeted at Y7) early in the autumn of 1998 and were available for 388 students (almost the entire

cohort). The students were taught in mixed-ability groupings (by tutor-group) until January, when they were set (by attainment in tests based on the mathematics syllabus) in four populations with four hierarchical sets in three of the populations and three sets in the other population. A further group of 27 students, identified by the school as having special needs in mathematics, were extracted from all mathematics lessons from the middle of the Autumn Term (this group is referred to as the “Resource Base”). The data concerning mathematical attainment was collated from mathematics department records at the end of the Summer Term 1999. The data relating to characteristics of the students’ learning was collected at the same time from their mathematics teachers (including the Resource Base teachers).

The research related CAT score and student performance in two ways:

1. The CAT scores of students known to have special needs in mathematics
2. How CAT scores might indicate special needs in mathematics

CAT scores were produced for each of the top four sets, the bottom four sets and the Resource Base as these should contain all of the students with special needs at either end of the ability continuum. In addition, teachers of these classes were asked to evaluate their students in relation to a set of criteria derived from the Guidance On Using CAT Results (listed above). They were provided with a set of characteristics and corresponding descriptions (appendix 3). These are listed below.

High attainers

Low attainers

Good grasp of written material

Has trouble maintaining sustained work

Willingness to take on challenging tasks

Has trouble following written tasks

Ability to help other students

Very low competence in basic arithmetic

Ability to pursue original lines of thought

Poor retention of facts

Self-checking and correcting

Needs constant support in problem-solving

Explains concepts clearly

For each student in these groups, teachers were asked to indicate either the absence of the characteristic, the presence of the characteristic or the presence of the characteristic in abundance.

This “light touch” individual assessment approach was chosen, as anything more rigorous would have made excessive demands on the time of the teachers concerned. Although the results are partly subjective, based on teachers’ interpretations of the descriptors, and none fully comparable from one class to the next, they should be reliable within individual classes and the professional experience of the teachers concerned should ensure a reasonable degree of validity. These results allow us to test the assertions detailed on pages 12 and 13.

The three CAT scores and the mean were considered independently, to identify indications of high and low mathematical ability. (Some caution must be exercised in using the mean, as its standard deviation can be shown theoretically to be 8.7.) In addition, the differences between the three CAT scores were calculated for each individual and, where these were pronounced, they were related to characteristics of the students’ abilities in mathematics.

## RESULTS

Before outlining the results for the specific areas of enquiry, it is worth outlining some general results for the whole sample. For 376 of the students, most of them excluding the Resource Base, complete sets of CAT scores and the aggregate of two common mathematical assessments were available. Correlation coefficients were calculated between the aggregate scores and the verbal, quantitative, non-verbal and mean scores. They were:

Verbal	Quantitative	Non-Quantitative	Mean SAT score
0.74	0.73	0.69	0.81

This is consistent with the NFER findings that the *mean* of the three SAT scores correlates best with KS3 and GCSE results. This reflects the fact that mathematical attainment, even as it is measured by the examinations system, consists of more than an ability to perform numerical work, other abilities have a significant part to play.

### 1) CAT Scores Obtained by Students with Special Needs in Mathematics

We shall look at all of the students in the Resource Base and extend the population to consider other students in the bottom sets for mathematics. The Resource-Base for mathematics consists of those students considered to have the greatest special needs in mathematics (at the lower end of the ability continuum). The table below shows a summary of the CAT data for these 26 students. The data were unavailable for two students, (a significantly higher proportion of the group than for the rest of the school).

No data were available to give any SAS values less than 70. Those students with scores less than 70 form a significant subset of the students with special needs in mathematics. Any sample excluding these students would be so biased as to be useless. It was decided that it was important to include these students in the analysis. A score of 69 was the maximum possible for these students, lower scores would produce more dramatic results. In order to avoid exaggerated inferences, *69 was adopted to represent these students*. Inevitably, this will affect any correlation coefficients involving these students, especially for sub-groups of the entire population which contain a significant proportion of low-attainers. The correlation coefficients in this account are used primarily to illustrate the data, rather than as specific descriptors. In this case the “feel” of the data is as important as the fine detail. The issue of using 69 to indicate scores less than 70 should have little affect on the results for the whole population, but caution must be exercised in using descriptive statistics to describe the lowest attainers. As a consequence of this estimation, the results below will have artificially high means and reduced standard deviations. The number of students scoring less than 70 is also indicated.

	Verbal	Quant	Non-verb	Mean
mean	76.2	75.1	79.3	76.8
s dev	9.5	7.2	7.8	6.4
largest	108	98	98	98
smallest	69	69	69	69
no. <70	10	8	2	2



One student has scores of 108, 98, 88 and a mean score of 98. Only on the non-verbal battery is this student's score less than three standard deviations above the mean for the Resource Base. At the very least, this student's scores are not typical of the group as a whole. If we were to exclude this student's results (because of their unrepresentative nature), we would obtain these results.

	Verbal	Quant	Non-verb	Mean
mean	75	74.1	78.9	76
s dev	7.2	5.6	7.8	4.8
largest	98	90	98	86
smallest	69	69	69	69
no. <70	10	8	2	2

It is not surprising to find that the quantitative score is lowest. One would expect this to be the best indicator of difficulty in mathematics. It should be noted, however, that there isn't much difference between any of the mean scores. The standard deviation of these results is included merely to indicate the spread of the data, rather than for any analytic purpose.

The correlation coefficients are low, although just significant for association (using product-moment correlation coefficients, one tailed test, 5% significance level).

V to Q: 0.394          V to N: 0.385          Q to N: 0.475

Thus there is no evidence to suggest that, for this group of students, low scores in two particular batteries constitute strong evidence of special needs in mathematics.

In the verbal battery there were only 3 students with scores of 85 or above (85, 98, 108\*). In the quantitative battery there were also 3 scores of 85 or above (85, 90, 98\*). In the non-verbal battery there were seven such scores (85, 86, 88\*, 88, 88, 94, 98). One student indicated by \* (and mentioned above), appears three times in this list, three more occur twice: (85, 76, 98), (83, 85, 85), (69, 90, 86).

According to the theoretical distribution of this SAS (normally distributed, mean 100, standard deviation 15), the percentage of all students nationally with scores less than 85 is ~16%. From the whole sample of Year 7 students from this school, those students with quantitative scores of less

than 85 were considered, to discover whether this is an indicator of special needs in mathematics. Of the 388 students tested, 55 had quantitative scores of less than 85, this constitutes 14% of the population, 22 of these students were in the Resource Base. That leaves 33 students outside of the Resource Base, all but one of whom were in the bottom sets for mathematics (there are four populations, therefore four bottom sets). The one student unaccounted for (scores 97, 84, 81) was in the third set of four in her population. In setting year 7, decisions were made on the basis of departmental assessments, teacher recommendations and KS2 results. The CAT scores were only referred to in borderline cases to indicate possible anomalies in setting and had no direct bearing on the actual setting decisions. Thus after a term at the school, the mathematics department had assessed all but one of the students with a quantitative score of less than 85 as being low attainers. Two terms after the initial setting, departmental assessments based entirely on items from the Year 7 curriculum were finalised for Year 7. Of the 33 students scoring less than 85 in the quantitative battery and in bottom sets, 25 were amongst the lowest attainers in their sets. Six of this group of students were median attainers and one (scores 80, 82, 90) was one of the highest attainers in his bottom set. The student in the third set of four was one of the lowest attainers in her set. From its successful record of assessment against national standards, the department has every reason to believe that its assessments produce reliable rankings. From the data, it appears that students scoring less than 85 on the quantitative battery are amongst the lowest attainers for this school. *For this sample, a quantitative score of less than 85 is a good indicator that a student from this population may have special needs in mathematics.*

## 2 Learning Characteristics of Students with Special Needs in Mathematics

Data were collected to investigate the claims made in the document concerning interpretation of CAT scores. As the data were collected towards the end of the school year, they had to be collected quickly. To ensure teachers' cooperation they also had to be easily derived. An easy assessment procedure was necessary. Teachers were given set-lists with headed columns and asked to evaluate their students in terms of each of these categories:

### High attainers

A Good grasp of written material

B Willing to take on challenges

### Low attainers

Has trouble maintaining sustained work

Has trouble following written tasks

- C Ability to help other students      Very low competence in basic arithmetic
- D Original lines of thought      Poor retention of facts
- E Self-checking and correcting      Needs constant support in problem-solving
- F Explains concepts clearly

Teachers were asked to indicate the absence, presence, or presence in abundance of these characteristics. A blank, 1 or 2 indicate these respectively. The data derived were not cross-moderated so there are problems of reliability of comparison between groups. However, for specific classes, all of the data were generated by one teacher and should be reasonably reliable (although issues such as teachers' stereotyping of students' abilities could present problems). To demonstrate ways of analysing these data I will look at the results for the Resource-Base group.

#### RESOURCE-BASE

PUPIL	A	B	C	D	E	V	Q	N	Mean
1	2	2	2	2	2	76	69	73	73
2	2				1	85	76	98	86
4		1			1	72	80	84	79
5	1	1	1	1	1	83	85	85	84
6	2	1		2	2	79	78	72	76
7	2	2	1	2	2	69	76	69	71
8	1		1	1	2	81	69	69	73
9	1	1		1		82	78	88	83
10	1	1		1	1	82	69	82	78
11			2	1	1	80	69	88	79
12	1		2		1	98	69	79	82
13	2	2	2	2	2	69	74	73	72
14	1	2	1	1	1	69	90	86	82
15		2			1	69	73	72	71
16	2	2	2	2	2	69	69	70	69
17	2	2	2+	2	2	70	72	74	72
18	1	1	1		1	69	76	77	74
19	1	1	1	1	1	71		76	74
20			1		1	76	77	80	78
21	2	2	2	2	2				
22	1	2			1	69	71	84	75
23		2				69	71	80	73
24	2	1		2	2	73	80	94	82
25	1	2+	2+	2	2	69	69	70	69
26						108	98	88	98
27	1		2	2	2				
28	1	2+	1	1	1	69	70	77	72
29	2	2	2	2	2	76	69	73	73

Without resorting to any sophisticated statistical analysis, it can be observed the majority of these students exhibit each of the problems identified in the document, especially the need for constant support (which is consistent with having a special need). According to the document:

- A should relate to V and N
- B should relate to V
- C should relate to Q
- D should relate to Q
- E should relate to N.

These relationships can be shown in one table.

VERBAL			QUANTITATIVE			NON-VERBAL		
V	A	B	Q	C	D	N	A	E
69	2	2	69	2	2	69	2	2
69	2	2	69	1	1	69	1	2
69	1	2	69		1	70	2	2
69		2	69	2	1	70	1	2
69	2	2	69	2		72	2	2
69	1	1	69	2	2	72		1
69	1	2	69	2+	2	73	2	2
69		2	69	2	2	73	2	2
69	1	2+	70	1	1	73	2	2
69	1	2+	71			74	2	2
70	2	2	71			76	1	1
71	1	1	72	2+	2	77	1	1
72		1	73			77	1	1
73	2	1	74	2	2	79	1	1
76	2	2	76			80		1
76			76	1	2	80		
76	2	2	76	1		82	1	1
79	2	1	77	1		84		1
80			78		2	84	1	1
81	1		78		1	85	1	1
82	1	1	80			86	1	1
82	1	1	80		2	88	1	
83	1	1	85	1	1	88		1
85	2		90	1	1	88		
98	1		98			94	2	2
108						98	2	1

By inspection it can be seen that the strongest relations are VB, NA, NE and to a lesser extent QC.

That these relationships should exist is not surprising. One would expect students with low verbal scores to experience difficulty in written mathematical tasks. Similarly it can be inferred from theories of fluid intelligence that students with low non-verbal scores will have difficulty

sustaining work and will need help in problem solving. Although it is not surprising that students with low quantitative scores should have difficulty with basic arithmetic, it is worth noting that nine of twenty-two students with quantitative scores less than 85 were not identified as having problems with basic arithmetic.

This analysis was repeated for each of the four bottom and the four top sets. In addition, correlation coefficients were calculated between the battery scores and the categories for each class. These results are summarised in the table below. Significance tests using these correlation coefficients are not applicable in this case since the data are not bivariate normal. However the correlation coefficients can be used to indicate relationships. To support this, squares have been shaded to indicate those coefficients which would (under normal conditions) have been significant (5% one-tail).

Correlations - Low Attainers

Verbal

	A	B	C	D	E
Res Base	-0.2	-0.72	-0.14	-0.32	-0.38
Class 1	-0.19	-0.04	-0.08	-0.12	-0.07
Class 2	-0.04	-0.02	-0.19	-0.24	-0.23
Class 3	-0.1	-0.17	0.064	-0.19	-0.1
Class 4	-0.34	-0.14	-0.24	-0.23	-0.31

Quantitative

	A	B	C	D	E
Res Base	-0.22	-0.24	-0.42	-0.23	-0.4
Class 1	0.031	-0.32	-0.32	-0.43	-0.2
Class 2	0.197	-0.08	-0.21	-0.12	-0.47
Class 3	-0.35	-0.34	0.155	-0.29	-0.17
Class 4	-0.08	-0.18	-0.23	-0.38	-0.33

Non-verbal

	A	B	C	D	E
Res Base	-0.22	-0.47	-0.5	-0.44	-0.56
Class 1	-0.11	-0.21	-0.26	-0.12	-0.21
Class 2	-0.06	0.114	0.15	0.035	-0.18
Class 3	0.005	-0.2	0.084	-0.24	-0.05
Class 4	-0.01	-0.04	0.017	-0.06	-0.02

Correlations - High Attainers

Verbal

	A	B	C	D	E	F
Class 5	0.576	0.468	0.482	0.454	0.335	0.495
Class 6	0.394	0.152	0.256	0.219	0.317	0.223
Class 7	0.578	0.523	0.535	0.614	0.555	0.496
Class 8	*	0.439	0.363	0.404	0.343	0.432

Quantitative

	A	B	C	D	E	F
Class 5	0.485	0.481	0.453	0.303	0.48	0.375
Class 6	-0.22	-0.31	-0.37	-0.43	-0.01	-0.16
Class 7	0.697	0.496	0.554	0.46	0.584	0.526
Class 8	*	0.512	0.499	0.339	0.319	0.264

Non-verbal

	A	B	C	D	E	F
Class 5	0.479	0.371	0.23	0.297	0.185	0.265
Class 6	-0.21	-0.38	-0.27	-0.41	-0.39	-0.37
Class 7	0.618	0.601	0.445	0.547	0.547	0.58
Class 8	*	0.312	0.237	0.372	0.359	0.332

\* - not available (all entries the same)

These data may suggest some evidence of correlation; however, there is no evidence at all of strong correlations. For low attainers, the guidance document suggests that the strongest relationships should be between VA, VB, QC, QD, NA & NE. If we consider the Resource Base

first, the best indications of correlation are for VB, NE, NC, NB, ND, NC, NE & VE in that order (only VB, NE & QC from the list). Excluding the Resource Base students, there is little indication of any correlation at all for the lower groups. It is clear that the relationships between the learning characteristics and the SAS scores are more pronounced for the higher ability students. Most students in the top sets have at least some of the attributes suggested. The one anomalous set of results is for class 6, where a much higher proportion of the students (nearly all) were considered to manifest the learning characteristics in some form. The outcomes for the other three classes are strikingly similar. Although statistical inference can not be drawn from this particular set of data, the data suggest that the characteristics listed for high ability students are general for the top sets at the school. However, these do not appear to be special characteristics of the most able. Those very able students with special needs in mathematics only constitute a small subset of these top four sets. To allow for this we shall compare CAT scores with the results in the Junior Mathematics Challenge.

The Junior Mathematics Challenge is a national mathematics contest for the most able students in Years 7, 8 and 9. The test items are designed to be amenable to students from all three years. Since the aspects of the mathematics curriculum covered by such a range of pupils is extremely varied, the curriculum knowledge required by the tests is appropriate for an able Year 7 student. The questions themselves are challenging, requiring originality of thought, an ability to explore and develop unfamiliar areas of mathematics and well developed problem-solving capacities.

JMC	V	Q	N
75	117	118	122
59	130	118	119
53	122	114	119
51	105	115	121
50	131	115	122
50	113	126	129
46	121	116	115
43	131	115	114
42	113	125	130
42	125	123	119
38	119	115	115
36	110	114	114
34	125	120	126
33	125	118	112
31	115	114	115
31	103	118	115
27	114	105	117
24	116	120	115
2	111	109	105
Corrs	0.30	0.36	0.58

NFER standardise the CATs and claim in their literature that each of the scores is normally distributed. Accepting this, it is reasonable to assume bivariate normality, hence correlation coefficients may be treated with some confidence. It is worth noting that the correlation between these scores and the quantitative battery is not significant. The only significant correlation coefficient is between the JMC score and the non-verbal score. This is not surprising, as success in the JMC requires mental agility, the ability to explore mathematical concepts, make connections and deduce solutions (all characteristics of fluid intelligence). Although the student's learned skills and knowledge have to be reasonably good to begin with; mere repetition of learned techniques will not lead to success in this items. The higher correlation between the JMC score and the non-verbal score supports the assertion that the non-verbal score is a measure of fluid intelligence

Summarising, it would appear that there is evidence to suggest that, *for this sample*;

1. Low scores on the verbal battery indicate problems with written tasks.
2. Low scores on the non-verbal battery indicate general learning problems, in particular difficulty in problem solving.
3. Low scores on the quantitative battery indicate problems with basic and numerical mathematics.

4. More able mathematics students exhibit many of the characteristics listed, a good grasp of written material, a willingness to take on challenges, the ability to help others, the pursuit of original lines of thought, self-checking and correcting and an ability to explain concepts clearly. However these characteristics *are not confined to the exceptionally able* and do not necessarily indicate a special need in mathematics.
5. The most able students in mathematics generally have high non-verbal reasoning scores (and, almost certainly, high verbal and quantitative reasoning scores).

### 3) Differences Between CAT Scores

According to the guidance document (appendix 2), a difference of 15 between two CAT scores may indicate a specific learning difficulty. This is questionable. Theoretically, SAS values are supposed to be distributed normally with mean 100 and standard deviation 15. Thus the difference between two scores should be distributed normally with mean zero and standard deviation  $15\sqrt{2} \cong 21.2$ . Through normal statistical variability, the probability of the difference between two scores with the same underlying mean being >15 points apart is approximately 0.48. In other words, the probability of two students, with identical abilities, having a difference greater than 15 between any of their scores is just under a half. (Issues of independence make a similar analysis impossible for the difference between one student's scores.) It is possible that the adoption of a CAT difference of greater than 15 as an indicator of a specific learning difficulty may be based on the mistaken assumption that the standard deviation of the difference between two scores is 15. (This would lead to the incorrect conclusion that the probability of two students of identical ability having scores more than 15 apart is about .32, i.e. a very high significance level of 32%). From the sample, 124 of the 388 students (32%) had a difference of 15 or more between their scores. The probabilities of two similar students having differences between their scores of 20 or more and 25 or more are approximately 0.34 and 0.24 respectively. From the sample, 55 students (14%) had a difference of 20 or more between their scores and 22 (6%) had a difference of 25 or more between their scores

Each of those students with a difference of 20 or more between two of their CAT scores was looked at in terms of their scores on common mathematics tests, and their membership of teaching



groups. Correlation coefficients between the sets of three differences in CAT scores and the common test results were found: Q-V: -0.0306 Q-N: 0.0820 V-N: 0.1012. None of these is significant. In particular, those year seven students thus far identified by the school as having specific learning difficulties are listed in the table below. This table shows position on the special needs register, CAT scores and the greatest difference between them.

Register	V	Q	N	Mean	Maxdiff
1	92	95	98	95	6
3	69	70	77	72	8
3	80	82	90	84	10
5	69	71	80	73	11
3	92	80	81	84	12
3	71	84	78	78	13
3	90	98	110	99	20
3	69	90	86	82	21
1	69	90	87	82	21
3	74	94	96	88	22
5	83	87	106	92	23
3	83	104	107	98	24
5	76	83	107	89	31

There is no discernible evidence of any connection between the greatest difference between two CAT scores and the position on the scale of those students with specific learning difficulties. Six out of thirteen of these students have a greatest difference less than or equal to 15, and only one has a greatest difference greater than 25. The three stated students have greatest differences of 11, 23 and 31. If a student with specific learning difficulties has generally very low CAT scores then the differences between them will be minimal. If students with all three scores less than 85 are removed from the list, then we are left with ten students, seven of whom have greatest differences between CAT scores of twenty or more. If we look at all the students in the school with greatest test differences of at least 20 and exclude those for whom all three scores are less than 85 we have a possible test for specific learning difficulty. This is not to say that *all* of these students have a specific difficulty in mathematics, but students with specific learning difficulties in mathematics may well be amongst this group. This process would not help identify specific learning difficulties in low attainers however, nor would it necessarily identify difficulties in mathematics. From the school sample, there are 55 students whose greatest difference between CAT scores was 20 or more; six of these are identified as having specific learning difficulties. It is possible that some of the 55 students identified above may have specific learning difficulties in

mathematics, which have not yet been identified. Unfortunately insufficient data were available to explore this further. This may provide scope for further investigation.

## CONCLUSIONS

### Methods

The aim of the project was to investigate the use of NFER CATs as screening tests for special needs in mathematics. In the school concerned, written guidance had been provided as to how the data might be used to identify and make provision for individual learning needs. Data were collected to evaluate this guidance and to explore the use of CATs generally in diagnosing special needs in mathematics. Several issues should be borne in mind when considering the findings.

- This was a short research project and the data were collected quickly, right at the end of a term.
- CAT scores less than 70 are not identified. As a consequence the data on those students with the greatest needs are the least informative and some of the statistical analyses are obtained using a score of 69 in these cases (erring on the cautious side).
- The data would have been better had specific diagnostic tests been used and compared with the CAT data.
- The method for collecting data on learning characteristics was subjective and could have been better developed.
- Specific measures could be found for each of the categories increasing reliability and validity.
- The data concerning specific learning difficulties in mathematics amongst students other than the least able was inadequate to support any conclusions.
- Many of the samples were small and reflect the school population. As a consequence, it is unwise to put too much emphasis on implications for mathematics teaching

## Results

Several conclusions can be made about the use of CATs in identifying special needs in mathematics for this particular school population.

1. CAT scores correlate positively to attainment in mathematics. The mean gives the best correlation. One would expect there to be a correlation between mathematical attainment and the quantitative and non-verbal battery scores. From the analysis of the subtests, it can be seen that between them they assess at least three of the components of mathematical ability cited by Orton (1987) and mentioned earlier. The ability to extract formal mathematics from a problem and to explain clearly what is happening is contingent upon linguistic capability. This would suggest that verbal ability has some bearing on mathematical attainment. In explaining why the best correlation is with the mean, it is important to remember that sample means will show less variation than individual, "point-samples". This accepted, the evidence from this sample nevertheless supports the notion that mathematical ability draws on abilities other than quantitative.
2. A quantitative score of less than 85 may indicate the possibility of special educational needs in mathematics. If CATs measure cognitive ability, as claimed in the Administration Manual (Thorndike, Hagen & France, 1986), then students with special needs in mathematics should have low CAT scores, particularly in the quantitative battery. For this sample, the group of students with quantitative scores of less than 85 contained nearly all of the lowest attainers in mathematics and only one student whose attainment was not low. Although it would not be reasonable to say that all students with quantitative scores of less than 85 have special needs, it is highly likely that a student with special needs in mathematics will be in this group.
3. Low scores on the verbal battery indicate problems with written tasks. As was mentioned above, verbal ability will have a bearing on a student's mathematical attainment. Mathematics teachers may find it useful to take their students' verbal ability into account when planning mathematical activities, as suggested in the manual (ibid).
4. The non-verbal battery is designed to measure fluid intelligence (ibid: 2). Together with verbal and quantitative abilities, in the mean CAT score, this has been shown to correlate with

mathematical attainment. Low scores in this battery indicate general problems in applying knowledge and skills in new contexts, in particular difficulty in problem solving.

5. As described above, the quantitative battery deals with number relationships, number patterns and equations. As might be expected, there is evidence from this project is that low scores on this battery indicate problems with basic numerical work.
6. The most able students in mathematics generally have high non-verbal reasoning scores (and, usually, high verbal and quantitative reasoning scores). Students with special needs in mathematics arising from a high level of ability will have high CAT scores. However not all students with high CAT scores will have pronounced needs in mathematics.
7. A difference between CAT scores of 20 points or more *may* help to identify students with specific learning difficulties, though differences of this magnitude do not automatically imply that this is the case. Because the CATs return a lowest SAS of <70, differences between scores will not be pronounced for students with low scores in each battery.
8. The absence of SAS scores below 70 means that it is impossible to identify exceptionally needy students from this test. According to the IQ model of intelligence, there is a huge difference in ability between a student with a score of 70 and one with a score of 50. Yet according to the 1944 Education Act, only those children with IQs of less than 50 fall outside the scope of normal education (Daniels & Anghileri, 1995). The CAT scores provide no information to help discriminate between these most needy children.

## BIBLIOGRAPHY

- Anastasi, A. (1982) *Psychological Testing: Fifth Edition*, New York: Macmillan
- Buros, O. K. (Ed) (1972) *The Seventh Mental Measurements Yearbook*, New Jersey: The Gryphon Press
- Daniels, H., Anghileri, J. (1995) *Secondary Mathematics and Special Educational Needs*, London: Cassell
- Dearing, R. (1993) *The National Curriculum and Its Assessment: Final Report*, London: SCAA

- Department for Education (1993) *Education Act 1993: Draft Code of Practice on the Identification and Assessment of Special Educational Needs*, London: HMSO
- Gardner, H. (1983) *Frames of Mind; the theory of multiple intelligences*, New York: Basic Books
- Hart, K. M. (Ed) (1981) *Children's Understanding of Mathematics: 11-16*, London: John Murray
- Hann, R., De, Havinghurst, R. (1957) *Educating Gifted Children*, Chicago: The University of Chicago Press
- Horn, J.L. (1971) *Intelligence – Why it Grows, why it Declines*: Ripple (1971) 82-94
- Joffe, L. (1981) *School mathematics and Dyslexia*, Unpublished PhD thesis, university of Aston, Birmingham
- Kruteskii, V. A. (1976) *The Psychology of Mathematical Abilities in Schoolchildren* (Translated by J. Teller), Chicago: University of Chicago Press
- Miles, T. R. (1983) *Dyslexia: The Pattern of Difficulties*, London: Granada
- Orton, A. (1992) *Learning Mathematics* (2<sup>nd</sup> edition), London: Cassell
- Piaget, J. (1952) *How Children Form Mathematical Concepts*: Scientific American (November)
- Piaget, J., Inhelder, B. (1958) *The Growth of Logical Thinking: From Childhood to Adolescence*, New York: Basic Books
- Thomson, M. E. (1990) *Developmental Dyslexia*, London: Cole and Whurr
- Thorndike, R. L. and Hagen, E. (1986) *Cognitive Abilities Test Levels A to F: Second Edition*, Windsor: NFER-Nelson
- Thorndike, R. L., Hagen, E. and France, N. (1986) *Cognitive Abilities Test Levels A to F: Second Edition: Administration Manual*, Windsor: NFER-Nelson