THE LONELINESS OF THE LONG-DISTANCE READER

A review of the reading demand of a key stage 1 mathematics test

by

Anne Louise Woodman

A thesis submitted to The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Education
University of Birmingham
2007

ABSTRACT

This study describes how the reading demand of a typical statutory key stage 1 mathematics test was investigated for readability for its young audience. About 600 000 six- and seven-year old children in England take the test each year, many of whom are expected to read independently and to know when to ask for support with reading. In the belief that children themselves would be the best judges of readability, data were collected to establish how much help children requested during the key stage 1 2001 mathematics test and how well they could read test questions aloud. In addition, each written question was analysed for readability using word lists and sentence length as the main criteria. The conclusion is that the reading demand is inappropriate for many year 2 readers and that some children did not receive the reading support that they required to access the mathematics. This raises questions about test validity. Any unrecognised reading difficulties are a threat to validity since reading skills are not being assessed. The voices of the children tell a compelling story. Although too few in number to constitute a national sample, it is argued that the children who provided the data were and continue to be typical of the national cohort and comparable year 2 children can be found in schools across the country. This raises issues regarding the appropriateness of a statutory mathematics test for children who are still learning to read. Given the findings, the national policy of testing key stage 1 children in mathematics is queried.

To

my beloved parents, Evelyn and Walter

ACKNOWLEDGEMENTS

I would not have completed this study without the help and support of my family, friends and colleagues. I specifically wish to thank:

- children who told the story without which this thesis would never have been written;
- teaching staff who welcomed me into their schools to undertake field work;
- fellow members of The Birmingham Science Education Research Group (BSERG) who offered advice, constructive criticism and support;
- Gill Curry who gave freely of her time to read my thesis and to comment on it;
- Professor Leone Burton for her friendship, wisdom and constant encouragement;
- my husband and children for never doubting that I would succeed;
- my dogs, Mac, Jock and Meg who kept me company in the countless and unsociable hours I spent at my desk as my thesis took shape.

I also acknowledge gratefully the support of the Qualifications and Curriculum Authority who gave me permission to:

- collect data for my pilot study during a key stage 1 mathematics pre-test and to report on the findings;
- quote from the unpublished report, *The language and reading demands of mathematics tests* (Lupetti, Sainsbury & Schagen, 1995);
- collect data during the administration of the key stage 1 2001 mathematics test.

I acknowledge that QCA has ownership of data collected during the pilot study.

CONTENTS

CHAPTER 5 Hands up in the KS1 2001 mathematics test

CHAPTER 6 Look-listen in the KS1 2001 mathematics test

CHAPTER 7 Comparing Y2, Y3 and Y4 readers

CHAPTER 8 Analysis of test booklet

CHAPTER 9 The final chapter of my story

APPENDICES

1.1 Typical development of a KS1 mathematics test

3.1 Copy of correspondence for tracking sheet and items for interviews

3.2 KS1 mathematics pre-test results (School H)

3.3 Tracking sheets for children in focus groups at Schools A, B and H

3.4 Copy of Jack’s interview items with underlined text and observer’s comments

3.5 Transcribing conventions and Transcript 1

3.6 Transcribing conventions and Transcript 2

4.1 Copy of the KS1 2001 mathematics test booklet

4.2 Interview prompt sheet

5.1 Key to column headings and children’s results ordered by reading level within each school

5.2 Children sorted by reading level (combined data)

5.3 Categories used in transcripts for teachers at Schools K, N, P and Q

5.4 School N: administration of the KS1 2001 mathematics test

6.1 School U: administration of the KS1 2001 mathematics test

7.1 Transcript analysis of Year 4 children reading item 14

7.2 Transcript analysis of Year 3 children reading item 14

7.3 Transcript analysis of Year 2 children reading item 14

8.1 High frequency word lists

8.2 Analysis of items 20, 22, 27, 29, 30 and 31 for readability

8.3 Summative analysis of all words in KS1 2001 mathematics booklet (alphabetical order)

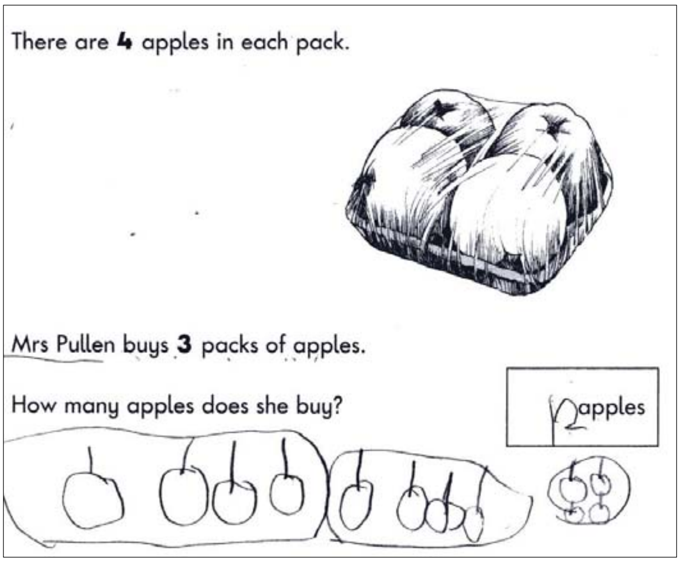
LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BERA	British Educational Research Association
DfEE	Department for Education and Employment
DfES	Department for Education and Science
EAL	English as an additional language
HMSO	Her Majesty's Stationery Office
KS1	Key stage 1
KS2	Key stage 2
KS3	Key stage 3
L1	Level 1
L2A	Level 2A
L2B	Level 2B
L2C	Level 2C
L3	Level 3
LEA	Local Education Authority
ME	Mathematical English
NAA	National Assessment Agency
NFER	National Foundation for Educational Research
OE	Ordinary English
QCA	Qualifications and Curriculum Authority
SCAA	School Curriculum and Assessment Authority
Y2	Year 2
Y3	Year 3
Y4	Year 4
Y5	Year 5
Y6	Year 6
YR	Reception year

CHAPTER 1



Marian was working at level 2C in mathematics.

CHAPTER 1

Beginning a story that needs to be told

Introducing myself

If my professional career had taken a more conventional route, this story would never have been written. It begins when I became a primary teacher in England in 1977. I quickly developed a particular interest in the teaching and learning of mathematics, an interest that has never left me. I moved on subsequently to be a teacher adviser for primary mathematics, organising and leading in-service courses for teachers but continuing to work in classrooms with teachers and children whenever the chance arose. I decided to become self-employed in the early 1990s, and, in 1995 was offered and accepted the part-time post of lead consultant for key stage 1 (KS1) mathematics with the School Curriculum and Assessment Authority (SCAA), later renamed the Qualifications and Curriculum Authority (QCA) then the National Assessment Agency (NAA) and this consultancy continued for nearly 10 years.

My responsibilities as lead consultant included helping to co-ordinate a small team of consultants who wrote and reviewed items for the statutory KS1 mathematics test. I was involved at all stages of the test development process and this gave me a good working knowledge of the development cycle of and processes within statutory test development, an experience that I believe is unique in professional terms.

In common with many other educationalists, I had reservations about statutory end of key stage tests for year 2 (Y2) children, most of whom were 7-years old with a minority being 6-years old. Insofar as the KS1 mathematics tests had to be developed and administered, I felt that, as a former primary teacher, I could at least empathise with the children and draw on my classroom experience to speak up on their behalf. This empathy led me to undertake this study, fuelled by concerns that had been building up in me over a period of years about the reading demand of the written test items for at least some Y2 children.

My initial concerns arose mainly through a ‘hunch’ based on my knowledge of Y2 children but this became more of a real issue through my marking of test booklets, observations of pre-tests and comments by teachers who had administered pre-tests. For example, when marking test booklets, the nature of the errors sometimes led me to suspect that children made errors because of failure to access the text accurately. An amusing but telling example is shown in Figure 1.1. This shows how two children misinterpreted the word *symmetry* as *cemetery* and *lines* as *lions*.

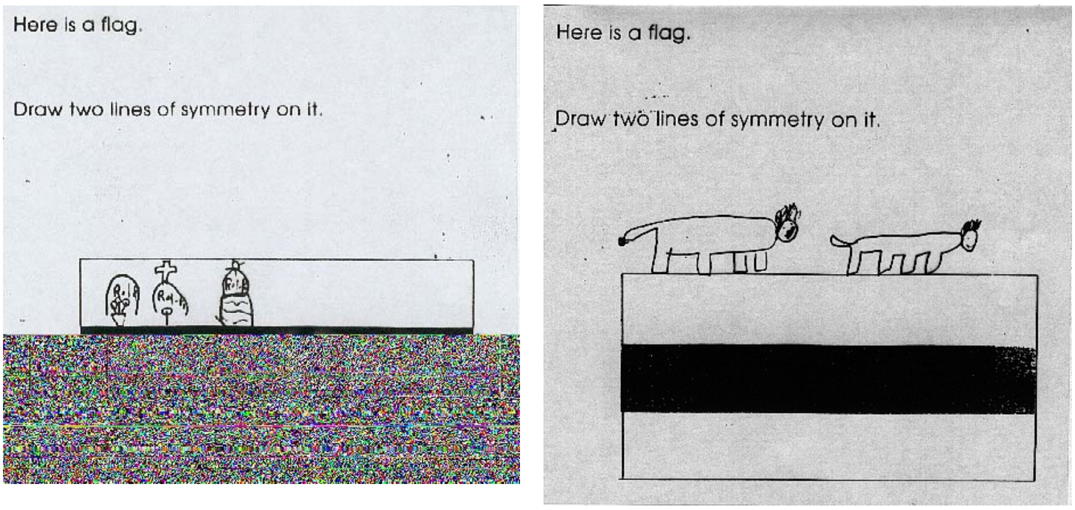


Figure 1.1: Children’s responses from item 29, KS1 1995 mathematics test

Part of my role as a consultant was also to observe pre-tests. These observations, undertaken each year of my consultancy, probably provided the greatest impetus towards my undertaking this study. During these observations I noted with concern that:

i. Some children spent a considerable amount of time asking for frequent help with reading.
ii. Some children who were identified as poor readers rarely asked for help.
iii. Some children appeared to be using phonics for words they found unfamiliar.

iv. Teaching staff were often faced with several children asking for help simultaneously and did not always know who had asked for help first.

v. Teaching staff sometimes had to assist with reading in a hurried manner in order to move on to the next child.

The first three of these observations suggested difficulties with reading; the latter two suggested administrative difficulties caused by requests for help with reading. Occasionally, when responding to children's requests for help with reading during pre-tests, I asked them to try to read text to me and sometimes found that they could not do so. This made me realise that some children were faced with reading demands that could prevent them from accessing the mathematics unless support was requested or offered. Over the years, teachers administering tests also expressed concern that the reading skills required by the booklet were excessive for some children and they were not always convinced that children asked for as much help as they needed.

These observations did not mean that readability was not taken into account in test development. My reservations were more about how the readability of the test was judged as I now explain.

Development of KS1 mathematics tests

Readers can find a detailed exposition of test development for key stages 1, 2 and 3 at http://www.qca.org.uk/12333.html. Nevertheless, I have included key points about the development of the KS1 mathematics test in Appendix 1.1 not only because I was about to embark on a critique of the reading demand but also because I wished to convince the reader that the development is a rigorous process, the intended outcome of which is that tests are appropriate for their intended audience. During the development cycle, for example, the reading demand of the items as well as sentence structure and layout are constantly under review through consultation with item writers, QCA staff, teachers administering pre-tests and the various advisory groups as all these aspects could affect accessibility.

The judgments of informed adults on the reading demand are, of course, an important element but, as I gained more experience of item writing, marking and observing children in pre-tests, I began to question increasingly whether the professionals involved at each stage of test development, including myself, were the best judges of the readability of the text. It seemed to me that what was missing from the development process was getting children of different reading competences to read aloud items from the pre-tests. Only children for whom a test was written could provide evidence of a test's actual rather than predicted readability. I was concerned that the children's 'voice' was missing and determined that children's involvement would be at the heart of my study. I hoped that I could somehow find a way to access what children actually did during a test and get their perspective on the reading demand.

Teachers' guides too are developed alongside test booklets. These acknowledge that some children are likely to have reading difficulties and suggest ways to support children. For example, the teacher's guide in 2001 (QCA, 2001d) stated that:

i. Teachers could give help with reading and should group children for the test to reflect their needs and their ability to work independently.

ii. Teachers should use their professional judgment 'to decide how best to make the tests accessible to all children …' (QCA, 2001d:5). For example, teachers could read an item to children with special educational needs, wait for children to write their answers and then continue with the next item.

iii. Before the test, teachers could display and read aloud in advance words that observers and teachers taking part in pre-tests had identified as causing children reading difficulties.

iv. Teachers should tell children that they could have as much help as they wanted with reading words and that they should put up their hands if they wanted help.

However, teachers are also instructed to tell children that, after a practice question, they will continue on their own so the expectation is that children will generally read independently and receive help with reading on request. The fact that reading support was available did not entirely reassure me. Since I was not aware of any research that identified which Y2 children were capable of reading a KS1 mathematics test independently, I was concerned that there might be individuals who were assumed to be capable of doing so but who did not ask for or receive appropriate support.

Content of KS1 mathematics tests

From 1997 to 2002, within the latter years of which most of my data collection was to take place, the structure and style of the KS1 mathematics test were not changed. Teachers read the first five items, using the script in the teacher's guide, and children wrote answers in the test booklet. Children then went on to answer about 30 written items, ramped in difficulty. Because the test items assessed number, handling data, shape, space and measures, the text varied from item to item so children had to cope with consistently changing text, a concern to which I return later in the section on children who take the test (see pages 6-10). Although guidance for administering written items remained largely unchanged post-2002, changes have been made to the test booklets and to the statutory arrangements; these are critiqued in my conclusions in CHAPTER 9. However, these changes did not affect the style or diversity of the items and I could see no reason why the reading demand would not be fairly comparable across all available tests.

Assessment arrangements

About 600 000 children in England take the KS1 mathematics test which teachers administer to children whom they judge to have attained level 2 or above in mathematics. Up until September 2004, KS1 children's levels of attainment in the national tests and teacher assessment levels had equal weighting even though the former was described as an end of key stage 'snapshot' (QCA, 1999:8) whilst continuous teacher assessment provided a much broader picture based on performance in a range of mathematical contexts throughout the key stage. However, it always seemed to me that, in the eyes of the public, national test results assumed more significance than teacher assessment, although, in theory, they were intended to complement each other. I did not accept that this was a satisfactory situation especially for such young children and the high status of national test results was another reason that nudged me towards undertaking this study. For example, after the tests are marked, teachers give each child a level and, for level 2, a sub-level. Most children are expected to attain level 2 with level 2B being 'regarded as the nationally expected level of achievement for most children at the end of key stage 1' (QCA, 1999:47). Only a few marks separate each of the sub-levels within level 2 and between levels 2 and 3 so any loss of marks because of unidentified reading difficulties could have a detrimental effect on levels of attainment. I was concerned that there might be children whose end of key stage mathematics level could be compromised because reading skills rather than mathematical skills were unintentionally being tested. I discuss this in more detail in the following section in which I consider some of the characteristics and experiences of Y2 children.

Children who take the KS1 mathematics test

In this section, I raise three further concerns:

i. how Y2 children might not receive the reading support they require;
ii. how the reading comprehension test and mathematics test differ;

iii. how normal mathematics lessons and the mathematics test differ.

Y2 children and reading support

Year 2 children taking statutory tests in May will range in age from about 6 years 9 months to 7 years 8 months with the average age being about 7 years 3 months. The youngest may have had as few as six and a half terms in school acquiring reading skills depending on when they entered reception year (YR) and early years experience prior to this. Others have the added challenge of learning English as an additional language. Consequently, Y2 children taking the test in May have widely varying reading competences, depending on aptitude and experience. The majority will be reading within level 2, but even within level 2, there are children who are reading above and below national expectations; a minority will be reading below level 2 or at level 3.

Since only one mainstream version of the mathematics test is developed and is not formally assessed for readability, I believed that it was unlikely that the reading demand would be suitable for all Y2 children, a fact that is acknowledged annually in the teachers' guides, when teachers are typically advised 'to be active in watching out for children who are having problems with reading …' (see, for example, QCA, 2001d:4). Consequently, when children are left to work through written items in the test independently, there is a dual responsibility. First, children need to acknowledge when they need help with reading and to be willing to ask for it; second, administering staff need to be on the lookout for children who need reading support and try to ensure that they receive it. Whether this responsibility should be given to Y2 children, especially those who are 'having problems with reading' (QCA, 2001d:4), is debatable because I doubt whether it is possible for a teacher to guarantee that all children have received optimum support. For example, children may incorrectly think that they are reading correctly, try to copy, resort to laborious decoding of words, guess what text says or be unwilling to show that they need help.

From my own teaching experience, I know that such children exist. It is also possible that teachers make assumptions about children's reading competence. Because some children show flair with reading from reading schemes and other texts that they enjoy does not necessarily mean that this flair is transferable to a mathematics test. Nor does it matter if occasional words are misread in blocks of running text, as found in stories, because illustrations, the context and repetition of words help to sustain meaning. However, items that make up the content of mathematics test booklets bear little resemblance to running text in the conventional sense. Much of the language is subject-specific, there is a range of contexts but little repetition of context words but if children fail to read items correctly, they could potentially lose marks.

Y2 children and the mathematics and reading comprehension tests

Most Y2 children reading at level 2 or above also take a statutory reading comprehension test, in which they work through a level 2 or level 3 booklet and answer questions on each page. Unlike the mathematics test, the readability of the text is carefully controlled to enable children to demonstrate attainment in reading at level 2 or level 3. This way of working is familiar to Y2 children for classroom work commonly entails reading comprehension in which children use knowledge of familiar words alongside strategies such as phonics. Indeed, to attain level 2, children are expected to use different reading strategies. For example, the attainment target for reading for level 2 states that children are expected to 'use more than one strategy, such as phonic, graphic, syntactic and contextual, in reading unfamiliar words and establishing meaning' (DfEE & QCA, 1999a:56). In other words, at the end of the key stage, the majority of Y2 children are not expected to be reading accurately and fluently since they are expected to find words unfamiliar on sight; they are still acquiring the tools of reading. Because it is the norm in classroom practice, children are likely to attempt decoding strategies such as word-building with phonics during the test rather than ask for help because this is what they are used to doing.

Similarly, they are often encouraged to try to read independently. Children may not recognise that doing so during a mathematics test is inappropriate. It is not until children are reading within level 3 that they are expected to 'read a range of texts accurately and fluently' (DfEE & QCA, 1999a:56) and independently. I was not convinced that all Y2 children reading the mathematics test independently would be able to read accurately and fluently since I expected the text to include words that were not in common usage as I explain next.

Y2 children and mathematics lessons

Unlike the reading comprehension tests, mathematics lessons in the style of KS1 mathematics tests are not part of familiar class routines. In a typical mathematics lesson, one theme, e.g. place value, weight or 3-D shapes, with its particular vocabulary, is the focus for teaching. Consequently, another concern I had was that written vocabulary for topics are unlikely to have been seen frequently enough to be recognised on sight during a test since:

i. teachers following the advice of the *Framework for Teaching Mathematics* (DfEE, 1999a) visit some topics for a few days only once or twice a term;
ii. teachers following the structure of the daily lesson plan recommended by the *Framework* (DfEE, 1999a) are unlikely to allocate much more than 20 minutes to written work;
iii. children may not have seen written forms of some mathematical vocabulary for several months before the test.

From my own experience and observations, classroom materials providing written work on a mathematics topic have a limited range of vocabulary. Not only that, the vocabulary is likely to have been rehearsed with the teacher by prior discussion, by using flashcards and by working through examples with children. Even when left to work independently, I have commonly observed children helping each other to read or reading aloud to help them gain meaning from the text. The mathematics test, by comparison, has a wide range of vocabulary covering many topics whose vocabulary cannot be rehearsed in advance through classroom discussion and the nature of statutory testing requires children to work without disturbing others or collaborating with peers. As a teacher, a strategy that I commonly observed young children using was ‘sounding out’ words to help them gain meaning, a strategy that most readers use at appropriate times even as adults. In a test situation, teachers discourage this practice because of possible disruption to others. The test, therefore, is far removed from normal classroom practice and such young children are unlikely to be coached in strategies that would make them ‘test-wise’.

I finish with a final concern. Most children are expected to have finished the test after about 45 minutes. Since most Y2 children are relatively inexperienced readers, I can only assume that a considerable part of that time will be taken up with reading and/or asking for help with reading which could affect their concentration on the mathematics. I felt that this could disadvantage less skilled readers in particular since they might spend a disproportionate amount of time on trying to access text.

Up to this point, I have drawn attention to various concerns I had surrounding the reading requirements of the KS1 mathematics tests. I now summarise the justification for my study.

Justification for the research

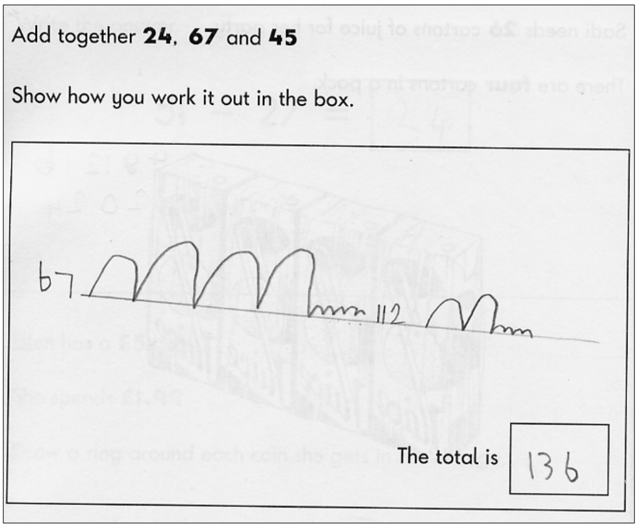
There has been a trend in recent years to hear the voice of children (see, for example, Griffiths & Davies, 1995; Lewis & Lindsay, 2000; McCallum, Hargreaves & Gipps, 2000; Wood, 2003). McCallum, Hargreaves & Gipps argued that ‘few studies have collected the views of young children on learning’ (2000:275); nor do there appear to be any large-scale studies that have sought the perspectives of KS1 children on statutory assessment. I agree with Wood (2003) that ‘[w]hile pupils are often considered the key stakeholders in education, rarely are their voices seriously taken into account … ‘ (p365). France, Bendelow & Williams (2000) argue that one reason for involving young children in research 'should be to work *for* children rather than *on* them and to describe their social worlds with a view of influencing social change' (p151). I could certainly commit myself to working *for* children; whether I could influence social change would remain an aspiration depending on what evidence I obtained.

My concerns about the reading requirements of the KS1 mathematics test came from two main sources, i.e. observations of pre-tests and my knowledge of Y2 children. Most concerns were not based on research evidence but on my recall of observations, anecdotal evidence and various assumptions. However, as Lloyd-Smith & Tarr (2000) stated, 'the reality experienced by children … in educational settings cannot be fully comprehended by inference and assumption' (p61). I felt the need to find out from Y2 children how readable they found the KS1 mathematics test. Not only that, I felt strongly that evidence should be gathered from a statutory test rather than a pre-test because of the impact of end of key stage results on children's lives. On behalf of Y2 children, I had become convinced that there was an urgent and overdue need to review:

i. the reading demand of the test; and
ii. teachers' arrangements for administering the test to take account of children's reading difficulties.

To address these two aspects would be the general aims underpinning my study. In the next chapter, I describe what I found in the relevant literature and how this helped me to refine my research project.

CHAPTER 2



Joe was working at level 3 in mathematics.

CHAPTER 2

Review of the literature

Finding a focus

In this chapter, I have summarised key findings from the research literature that I felt would guide me towards identifying research questions that would give a new slant to, or fill gaps in, the existing research literature. I anticipated that most of my research would involve two main but closely related foci: the first was the readability of the statutory KS1 mathematics test booklets, but, more importantly, the second was the Y2 children who take the tests. Consequently, the research literature I accessed focused on readability from two different perspectives, the test booklet itself and Y2 children as readers, as I explain next.

In CHAPTER 1, I expressed concerns, based mainly on anecdotal evidence and my own intuition as a former teacher of young children, that the reading demand of the KS1 mathematics test could disadvantage at least some children when it is left to them to ask administering staff for help with reading. To put this anecdotal evidence to the test, I needed to find research by others who had attempted to analyse the reading demand of text for this age of children and, of equal importance, others who had investigated the reading demand of mathematical text or tests. Hence I chose readability for my dominant theme, since, if the reading is difficult for at least some of the children, the factors that cause the difficulty need to be teased out. The test booklets, of course, include, other than text, features that form part of the assessment, e.g. the graphic elements of tables, sorting diagrams, graphs etc. It would have been tempting to widen my research to include how children coped with ‘reading’ these elements. However, since children cannot generally be assisted with the interpretation of these, other than by having any included words read to them, I decided to evaluate the ‘reading’ of these only where the children’s interaction with the text justified it. For the most part, the emphasis of my research had to be on how well or otherwise the children could read the text that provided access to the mathematical requirements of the test items. In the test, children reading independently need to read accurately and fluently so that they understand what the words are asking them to do even though the mathematical demand of an item might subsequently prevent a correct response. On the other hand, children who might be able to respond correctly may be prevented from doing so due to poor reading skills. As Clausen-May (2001) explained, in her consideration of test validity, it is possible for someone taking a test that is not assessing what it purports to assess to ‘have “failed” the test situation, not the test subject’ (p4). A similar point is made here:

> Validity was seen originally as a property of a test – a test was valid to the extent that it tested what it purported to test. However, a test may be more valid for one group than another – a particular mathematics test may involve complex language, and would tell us little of the mathematical abilities of poor readers, but might work quite well for fluent readers. (Wiliam, 2003:12)

There was no shortage of literature on readability. Indeed, Fry (1990) stated that ‘[t]he *Social Science Citation Index* shows that readability articles are among the most frequently cited research ….’ (p594). In the face of so much literature, I took heed of a warning by Verma & Mallick (1999) that some researchers fail to ‘stay within the topic limits’ (p142) and can read too widely but superficially. To focus my literature search, I had to keep going back to the concerns driven by my interests as outlined in CHAPTER 1 and to consider:

i. What landmark sources of literature exist that may help to shape my research?
ii. What are the strengths and limitations of the literature in relation to my own areas of interest?
iii. Where are the gaps in the research literature?

In the remainder of this chapter, I cite studies that I felt responded to the first question. I have tried to address the latter two questions by constructing my review of the literature in a critical voice. I conclude the chapter with the statement of my research questions and the justification of their originality.

An overview of readability

What is readability?

Later, in CHAPTER 5, for reasons that I substantiate during my consideration of this theme, I evaluate the reading demand of the text in a typical KS1 mathematics test. However, I also had to accept that there could be test items where the reading demand could not be evaluated in the absence of a consideration of other aspects. Mobley (1986:9), for example, stated that ‘readability refers to every aspect of a text which makes it either easy or difficult to read’ and identified these key aspects that could be relevant in evaluating readability:

i. legibility, e.g. the clarity and size of the type face, the length and spacing of the lines and the paper type;
ii. visual aspects, e.g. the quality and quantity of the illustrations;
iii. language aspects, e.g. the use of words and language structures which are familiar;
iv. content and conceptual difficulty, e.g. the amount of information introduced at one time;
v. clarity of meaning, e.g. absence of ambiguity;
vi. interest level, e.g. features which will motivate the reader;
vii. text simplification, e.g. keeping sentences short and the structure as simple as possible.

Consequently, I accepted that some of these aspects, and any others that might unexpectedly arise in context, might have to be considered in relation to particular test items or, indeed, to the whole booklet, where they appeared to affect the reading demand. Dale & Chall (1948) defined readability in more detail than Mobley (1986):

> In the broadest sense, readability is the sum total (including interactions) of all those elements within a piece of printed material that affects the success which a group of readers have with it. The success is the extent to which they understand it, read it at optimum speed, and find it interesting.
> (Dale & Chall, 1948 cited in Gilliland, 1972:12)

Both Mobley (1986) and Dale & Chall (1948) referred to the reader's interest in the text and Gilliland (1972) also considered that 'interest and motivation play a critical part in determining readability' (p13). I made the decision that researching attitudes such as 'interest' and 'motivation' in relation to the test would not be a major consideration since children have to take the test regardless of these. However, I had to recognise that a poor reader who was not supported appropriately during the test could have little motivation to succeed since 'a person with limited reading ability will soon become discouraged if he is given texts which are beyond his comprehension' (Gilliland, 1972:12). I empathised with the child in this situation as did Rye & Sjölander (1982) who stated that, when met with repeated failure in reading, 'the child is hardly likely to be highly motivated to try to understand the text or keen to develop any interaction' (p100).

To avoid the scope of my research extending too wide, I decided to restrict the definition of printed material to the text in the written items that administering staff were permitted to read to a child. I fully recognised that the test has other components that give it its unique characteristics, e.g. mathematical diagrams, graphs, symbols and abbreviations, all elements that have to be 'read' and decoded for a child to respond. However, apart from a small number of non-essential, decorative illustrations, these components form part of the mathematical assessment and, therefore, help cannot be given with interpreting them. On the other hand, the reading of text is not the focus of the assessment, at least not deliberately. The importance of success in reading text featured strongly in Dale & Chall's (1948) definition of readability. However, in a mathematics text or test, there are two distinct aspects. First, a child who reads fluently and accurately is likely to be successful in that aspect; second, a child who knows how to respond mathematically is likely to be successful in that aspect. Success in the first aspect does not guarantee success in the second aspect, but importantly for the KS1 test, success in the second might be precluded by lack of success in the first. I recognised that there was a fine dividing line between 'readability' and 'comprehensibility' that could complicate research into a subject-specific text. Consider the following statement:

> Readability is an attribute of text; comprehension is an attribute of readers. There is therefore a fundamental difference between the two concepts. Having made that distinction though, a moment's reflection makes it clear that the concepts are intimately related, in that very often when we use the term readability we mean in effect the comprehensibility of a text. (Harrison, 1980:33)

I envisaged that this close relationship could be difficult to unravel. Taylor, Graves & van den Broek (2000) believed that an essential feature for comprehension was for readers to 'construct a mental "picture" of the text: a representation in memory of the textual information and its interpretation' (p2). My research, therefore, had to focus on providing evidence that, where reading was unsupported, children were or were not reading with sufficient competence to comprehend what kind of response was required, i.e. through being able to 'hold the text' in their head, even if the mathematics were too difficult.

Speed of reading

I believed that the issue of 'optimum speed' (Gilliland, 1972:13) could be important. Technically, children whose reading lacks pace and fluency, who labour to read but who might persevere without requesting support, perhaps with few errors, may be assumed to be reading adequately. On the other hand, Kane et al (1974) pointed out that the nature of most mathematical text requires careful attention to detail. I interpreted this to mean that, in such circumstances, reading at ‘optimum speed’ might mean reading at a relatively slow speed to absorb every detail of the text. I anticipated that children in both of these categories would take part in the test, i.e. those who were slow because they were struggling to read and those who were deliberately reading slowly but accurately (or perhaps rereading). I believed that the latter children were more likely to obtain the ‘flow of meaning’ (Shuard & Rothery, 1984:66) from the text; the former were more likely to be struggling with the meaning of words or sentences. Klare (1963), in a study undertaken with associates, concluded through validity studies, using easy, medium and hard versions of text in which the same information was embedded, that reading speed was a function of readability. In the easy versions of the text, more words were read per second and the general results indicated clearly that ‘readability and reading speed are related’ (Klare, 1963:137). A study by Wilkinson (1980), to which I shall refer in more detail later in this chapter, also considered the speed of reading. For his research, samples of running prose were developed that contained a high percentage of words familiar to children. The readability of these samples was tested with children in grade 2 (7-year olds) to grade 6 (11-year olds). He concluded that there were three stages in the development of skilled reading, which he explained:

> In the first stage, the child reads accurately but slowly, with a loss of comprehension and memory. … At the second stage, corresponding to approximately the fourth grade, the child reads at a rate corresponding to the normal speaking rate. … A child in the third stage is able to read efficiently and rapidly …. (Wilkinson, 1980:568)

In the section on **Subject-specific text** that follows later in this chapter (see pages 22-23), Mobley (1986), although not commenting on the speed of reading, also referred to children who ‘go through the motions of reading’ (p31), which seemed to correlate well with Wilkinson’s first stage.

In the next section of this chapter, I refer to a study by QCA (2001a). A similar study was conducted a year later, again with Y2 children reading at level 2C (below-average) and level 2B (average) as they read books from the level 2 statutory reading task booklist. In this later study, there is a reference to the relationship between word recognition and speed of reading:

> For the purpose of this study, words were counted as recognised on sight if the normal pace of reading was maintained without evident pause for spoken or silent working. … Across the sample, children read 12.5% of words incorrectly. (QCA, 2002:24)

Having reflected on the references to the speed of reading, I identified four issues that could turn out to be relevant to my research. First, I felt that there were likely to be children taking the mathematics test who were at Wilkinson's first stage, or, indeed, who had not even arrived at this stage, i.e. were not reading accurately. At least some of these could fail to ask for, or receive, the amount of supported reading that would allow them to concentrate on the mathematics. Second, the children in Wilkinson's study (1980) were being tested on the reading of text that contained a large percentage of familiar words. The mathematics test is not written in running prose, nor is its subject-specific text likely to contain such a favourable ratio of familiar words. Third, although the samples contained a high percentage of familiar words, Wilkinson's (1980) evidence concluded that it was not until children were in the fourth grade, two years older than the children taking the KS1 mathematics test, that they were likely to read at normal speaking rate. This could imply that 7-year olds taking the test, even those whose reading skills are in line with national expectations, might struggle to read text that included unfamiliar mathematical vocabulary. As a result, they might be unlikely to read the text at optimum speed leading to possible lack of concentration and memory.

I would expect that such children would be disadvantaged without supported reading since they need to grasp the holistic meaning of the text before they can respond to the mathematics. Fourth, the QCA study (QCA, 2002) found that about one in five of the 161 children in its sample did not recognise some words generally expected to be familiar on sight, i.e. the normal pace of reading was not always maintained. This ratio would be likely to increase with the inclusion in the mathematics test of words that are not familiar sight words and with the amount of unrelated text that needs to be read to reach the end of the test.

These findings served to confirm my conviction stated in CHAPTER 1 that the only way that I could sensibly undertake research into the speed and accuracy at which children, of different aptitudes in reading, read mathematics test items would be to involve the children themselves, i.e. literally, to hear their voices. I could find nothing in the research literature in the public domain that suggested that similar research had been undertaken. The means of doing this was to become a major part of my methodology and research design and was influenced by the characteristics of Y2 readers as I explain next.

Characteristics of readers at the end of year 2

Because I intended to use Y2 children to find out how readable they found the text in a mathematics test, I recognised that I would also need to judge their reading skills. Consequently, it would be useful to have a checklist of characteristics of skilled and less skilled readers.

In June 2000, a month after the administration of the statutory tests for that year, QCA commissioned a study to investigate how Y2 children used strategies in reading aloud from four books on the reading task booklist for level 2 (QCA, 2001a:18). In this study, coders used a running record and analysed the reading of 113 children who had already completed the statutory reading task with 59 having achieved level 2B (average) and 54 level 2C (below average). The analysis of each child included keeping a record of the use of phonic strategies, percentage of words read correctly, expression, fluency and understanding of what they read. The main findings were:

i. When applying phonic strategies, children reading at level 2B were more likely than not to be successful; for children reading at level 2C, the reverse was true because they lacked the facility to combine sounds to find the whole word.
ii. Whilst most of the weakest readers at level 2C were able to sound out single letters, they could generally only access short words with straightforward letter/sound correspondence (e.g. *dark*).
iii. Even children reading at level 2B were not consistent in being able to recognise clusters (e.g. *cl* in *clear*) and vowel phonemes (e.g. *ea* in *clear*).
iv. Children reading at level 2C were found to have difficulty in connecting an initial consonant or cluster to the following vowel.
v. When children reading at level 2C built up words letter by letter, they quickly gave up when building longer words in contrast with the children reading at level 2B who generally built up words in larger units (e.g. using clusters and phonemes) and were more successful.

QCA (2001a) concluded that the characteristics of children with lower accuracy rates in reading were that:

i. they analysed words into single letter sounds;
ii. they were unsuccessful at blending words;
iii. they made a high level of substitutions;
iv. they self-corrected infrequently.

The study also found that, overall, ‘children showed a high level of independence and were confident enough to attempt every word in most cases, but with widely varying success’ (QCA, 2001a:19). The study provided me with some insight into how children reading at these levels tried to make sense of text and what happened when they encountered difficult words. The same cohort of children also takes the mathematics test. It did not seem unreasonable to suppose that children, during the mathematics test, might choose to read independently, but possibly with ‘widely varying success’ (QCA, 2001a:19), particularly since the test included mathematical words that were less likely to be familiar on sight. Indeed, it is possible, because of mathematical terms, that above-average readers might also have some reading difficulties. Additionally, unbeknown to their teacher, the use of phonics by some children, correctly or incorrectly, would slow down the speed of reading, thus compromising fluency, the flow of meaning and progress through the test.

Any children who exhibit characteristics such as those mentioned above, and who do not recognise that they need reading support will be disadvantaged. If I were to find such children, I could argue more effectively that supported reading during the test should not be optional but essential, with few exceptions. At the other end of the reading spectrum, Gibson (1989) describes the characteristics of a fluent reader whom she states:

i. ‘builds on previous stages
ii. can process print details automatically
iii. can handle a variety of print forms independently
iv. can read at a rate appropriate to the print form’. (Gibson, 1989:110)

These characteristics seemed to describe best children reading within level 3, whom, as I explained in CHAPTER 1, are expected to ‘read a range of texts accurately and fluently’ (DfEE & QCA, 1999a:56) and independently. From the literature in this section, I now felt that I had some criteria against which I could judge the efficacy of children’s reading of the mathematics test.

Later in this chapter, at various points, I consider vocabulary that is likely to be familiar or unfamiliar in written form to Y2 children and the implications of this for my research. I start this consideration in the next section in which I focus on the unique features of subject-specific text in general and mathematical text in particular.

Mathematics and readability

Subject-specific text

Various writers have considered issues arising from the reading of subject-specific text (see, for example, Harrison, 1979; Shuard & Rothery, 1984; Mobley, 1986). Harrison's research (1979) included evaluating the readability of subject-specific text in four subjects, one of which was mathematical. His research involved children aged 10 to 15 years and so had only tenuous connections with the age-group of interest to me. However, his conclusions from discussions with secondary teachers in several LEAs 'indicated that many were concerned about the inability of children to cope with the reading demands of subject specific areas' and that 'children who were poor readers but were otherwise good in these subjects were unable to cope with the … examination questions which tested what had been learned …' (Harrison, 1979:75). It was whether this was also the case with KS1 children during the mathematics test that I felt needed to be determined.

Mobley (1986) made the point that subject-specific text books 'present vocabulary and language structures which are very different from the natural spoken language of the child' (p30). She also quoted research which showed that when children had to read complex sentences that they found too long or included unfamiliar vocabulary, they could 'go through the motions of reading, but absorb little of the meaning' (Mobley, 1986:31), a phenomenon to which she attributed the phrase 'non-reading'. A similar point was made in *A language for life*:

> If children are faced with texts containing more than a very small proportion of unfamiliar words they will spend far too much time struggling at frustration level and will derive neither meaning nor enjoyment.
> (Committee of Inquiry into reading and the use of English, 1975:103)

Although not referring specifically to subject-specific text, the term 'frustration level' was further explained in *A language for life* (Committee of Inquiry into reading and the use of English, 1975:253) in which a procedure, developed in America, for determining the appropriateness of a text for a particular child, was described. When reading a sample of 100 consecutive words, it was suggested that a child could be considered to be reading independently when 'able to read aloud in a natural and easy manner, without help from the teacher with 99 per cent accuracy in word recognition'. On the other hand, a child who made 10 or more errors in word recognition in the sample could be said to be operating at 'frustration level' and the material was too difficult, resulting in poor comprehension. Harrison (1980) described a similar procedure, a 'rule-of-thumb' that 'has proved its value in primary schools' (p29). Known as the '5% rule' (Harrison, 1980:29), it assumes that if a child makes more than one uncorrected error when reading aloud 20 continuous words in a book, the text is too difficult for the child. In independent reading during the mathematics test, I suspected that a significant number of children could be reading at 'frustration level', exacerbated by subject-specific vocabulary in the booklets as explained in the next section.

Mathematical English and Ordinary English

Kane (1967), Shuard & Rothery (1984), and Noonan (1990) considered the readability of mathematical text. Noonan (1990:58) stated that one reason that makes mathematical text more complex than ordinary text 'is that mathematics uses a technical vocabulary which overlaps with the vocabulary of ordinary English.' The term 'ordinary English' was not used by chance. Kane (1967) introduced the terms *Ordinary English* and *Mathematical English* to draw attention to the unique features of written mathematics:

> Mathematical english [sic] [ME] is a hybrid language. It is composed of ordinary english [sic] [OE] commingled with various brands of highly stylised formal symbol systems. (Kane, 1967:296)

Kane (1970) later remarked:

> Mathematical English and OE [sic] are sufficiently dissimilar that they require different skills and knowledge on the part of readers to achieve appropriate levels of reading comprehension. (Kane, 1970:579)

Yet another way in which mathematical text differed from ordinary prose was suggested by Shuard & Rothery (1984:24) and Noonan, (1990:58) when they explained how terms used in mathematical English fall into three main categories:

i. Ordinary English: words that have the same meaning in mathematical text and text that could be read, for example, in a story or comic, e.g. *house*, *child*, *when*, *hid*;
ii. mathematical English: words that are generally found only in mathematics and introduced through the medium of mathematics, e.g. *trapezium*, *cosine*;
iii. dual-meaning English: words that can arise in both Ordinary and Mathematical English but whose meaning can vary depending on the context, e.g. *product*.

Whilst this added layer of complexity would not affect readability, it was another dimension of language that could adversely affect children's comprehension. Shuard & Rothery (1984:1), in their consideration of mathematical text, used the word readability to 'convey the idea that in a readable text it is easy for the reader to *get the meaning from the page* [original italics]'. However, as Noonan (1990:59) pointed out, words in the ME category 'may cause reading difficulties because pupils do not meet them in any other context.' Furthermore, Shuard & Rothery (1984) claimed:

> Mathematical words are unlikely to be used at home or in the child's everyday speech, and so they cause reading difficulties because they are rare in the child's experience. (Shuard & Rothery, 1984:25)

The KS1 mathematics test does not include the 'highly stylised formal symbol systems' of the kind to which Kane (1967:296) referred so this aspect of Kane's study cannot be applied directly to my research. Indeed, for the present requirements at KS 1, children are required only to recognise numerals, most commonly with two or three digits, the four operation signs and the equals sign. Consequently, my view was that most children would be familiar with these symbols at the time of the test and would probably find them easy to read with meaning compared with text. The children were also likely to have had regular practice of doing calculations such as '9 + 8 = □' or '22 - □ = 18' whereas it was possible that word problems that could arise could be expressed in many forms, include a mixture of OE and ME, be drawn from many contexts and, therefore, could not be rehearsed in the same way in the classroom. For example, a subtraction involving money might be, A boy has 90p. He buys a pencil that costs 32p. How much money does he have left?, an example that contains OE words likely to be used in everyday speech and writing by 7-year olds. The children are also likely to be familiar with money contexts in their everyday lives through common usage. By contrast, the instruction, Draw the reflection of this pattern in the mirror line. contains words and word combinations such as 'mirror line' that are likely to be used exclusively in mathematics lessons and to be familiar in speech (and perhaps in occasional written form) only when children are learning about reflective symmetry, perhaps once a term. Although I would not expect 7-year olds to be confused by the dual-meaning of reflection to mean either a mirror image or a consideration, since the latter meaning would be unfamiliar to most 7-year olds, it does highlight how words with dual-meaning might further complicate issues related to reading.

The literature mentioned in this and the preceding section started to convince me that it would be the mathematical vocabulary in sentences, one aspect of ME, combined with unfamiliar OE words in the test, which children would be more likely to find 'hard reads'. I therefore decided to seek out literature specific to the readability of mathematics texts or tests.

Mathematics texts and readability

I could find no recent studies investigating the readability of mathematics texts (as opposed to tests) for primary school children. The most recent was that by Hubbard (née Jones) cited in Shuard & Rothery (1984). In her study, published in 1976, Hubbard explained how she applied a readability formula to samples of text from 45 different mathematics texts used by third-year juniors (rising 10-year olds), now renamed year 5 (Y5). Her analysis of the texts indicated that '[t]he reading ages obtained varied from 9 to 15 years, with most of them in the range from 12 years upwards' (Shuard & Rothery, 1984:83). To see if she would obtain similar results if she involved the children, she then made a survey of 310 third-year juniors, who were found to have a mean chronological age of 10.25 years but a mean reading age of 10.78 years. Hubbard provided each child with passages from texts they were using in mathematics lessons, adapted for cloze procedure, which I discuss in more detail in **Methods of measuring readability** later in this chapter. In brief, cloze procedure involves deleting individual words in text and replacing them by blank spaces. Readers try to predict the words that were deleted and the reader's degree of success gives one indication of how well the text is matched to the reader. Hubbard concluded that approximately three-quarters of the children had a reading age below that of the textbook they were using. Not only that, to use the words of Shuard & Rothery (1984), '[i]n all cases, the reading age of the book was above the mean reading age of the children for whom it was designed' (p84). The results from the use of the reading formula and the cloze tests shared the same conclusion, i.e. that the reading difficulty of the mathematics texts had been considerably under-estimated. To me, this implied that publishers, authors and teaching staff alike were unaware of the extra reading demand required by the subject-specific text. Although the results of the study concerned me, I recognised that, if the same phenomenon were to be true for mathematics tests, this would be a matter of even greater concern. In mathematics lessons, children often work collaboratively, by, for example, helping each other with reading or reading instructions together. In addition, the teacher often introduces any written material, usually covering one topic at a time, talks through examples, and can support reading during the lesson. These circumstances could, therefore, tend to mask potential reading difficulties. In most mathematics tests, including the national tests in England, collaborative work is not possible, various topics are covered, and reading can be unmediated. In a similar vein, the reading demand of the KS1 mathematics tests could be under-estimated by those involved in test development and administration. I consider this in the next section.

Mathematics tests and readability

No informed person would be likely to argue that the majority of text to be found in KS1 mathematics tests resembles the natural speech patterns of the typical 7-year old. Moreover, various studies have suggested that children are most likely to identify with text that they use in their everyday conversations. In its review of reading in the early years, *A language for life* (Committee of Inquiry into reading and the use of English,1975) acknowledged that 'a number of studies show that a printed text is easier to read the more closely its structures are related to those used by the reader in normal speech' (p72). Perera (1980) made a similar point in stating that there was ample evidence 'that children read more easily those sentence structures that they would themselves say or write than sentence patterns which occur predominantly in literary writing' (p156), and literary writing, in my opinion, is likely to be closer to the natural language of the child than the language of a mathematics test. The unique, and sometimes abstruse, language of mathematics tests makes it virtually impossible to make the text sound as if it had come from the mind of a Y2 child.

Whilst 7-year olds, in everyday speech, use little of the technical vocabulary of mathematics that they are expected to know and understand at the time of taking the test, reducing the text in the test only to words familiar in everyday speech would not be a realistic option. This is because the interpretation of Mathematical English (ME) words form part of the assessment. Shuard & Rothery drew attention to some shortcomings of such an approach:

> Omitting all technical words is a short-term policy which makes text easier to read but it may bring long-term disadvantage to the pupil. … Many technical terms have an essential place in mathematics; children cannot proceed without knowing them. … So a practice which may seem a kindness … may work against the pupil’s future comprehension of mathematics.
> (Shuard & Rothery, 1984:26)

Indeed, removing ME words would effectively reduce the majority of the test to abstract calculations at the expense of the breadth of coverage required by the programme of study. Consequently, I agreed with Shuard & Rothery (1984) that it would not be appropriate to remove the ME terms that assess the vocabulary taught during the key stage but it may not be appropriate for children to be expected to read these words. Another argument against an emphasis on abstract calculations might relate to one of the general aims of education, i.e. to help children and adults cope with the mathematics they need to manage their lives. In real life, mathematics generally arises in contexts where mathematics has to be applied. In the statutory mathematics tests, ‘real-life’ problems are included and these necessitate the use of words to expose the context and the problem to be solved.

Having accepted the necessary inclusion of ME terms and word problems in the test, I then searched for literature in which the reading of mathematical test items was under scrutiny. The most relevant literature came from studies by Newman (1977); Clements (1980), Lupetti, Sainsbury & Schagen (1995) and Shorrocks-Taylor & Hargreaves (1999, 2000). What particularly appealed to me in the first three of these studies was the direct involvement of children and how this was achieved. I review the first two of these studies in the next section since they focused on error analysis, part of which included identifying reading errors, unconnected to statutory testing. A review of the third and fourth studies follows in later sections of this chapter.

Reading errors in mathematics tests

In these two studies Newman (1977) and Clements (1980) conducted diagnostic interviews after children had taken a mathematics test to isolate the specific junctures within items at which children could fail, the first two of which were reading and comprehension. Newman (1977), a lecturer in reading education, expressed concern that ‘there have been very few attempts to analyze pupils’ difficulties by talking to them’ (p240). As a result, she developed a mathematical test of 40 single-step items that included number, measures, spatial and logic problems. The test was estimated to be suitable for children in the fourth-grade (nine-year olds) and ‘found to have a readability level of below grade four’ (Newman, 1977:242). The test was administered to 917 children in the sixth-grade in Melbourne, Australia, and, of these, 124 who were low achievers in the mathematics test were selected for diagnostic interviews. For the interviews, Newman (1977) established what she described as ‘a hierarchy of “performance strategies” that needed to be applied to solve written mathematical tasks successfully’ (p242). Each strategy was assumed to be an obstacle to overcome and failure to overcome one obstacle would prevent access to the next one. Five main stages or causes of errors, were identified, the first of which was ‘reading ability’ (Newman, 1977:242), i.e. establishing if the pupils could read the questions well enough to enable them to move on to the second stage. The second stage was ‘comprehension’ in which the children had to explain the questions in their own words or otherwise demonstrate their understanding of what they were being asked to do. The interviewers, using an agreed coding frame, identified the stage at which the children made initial errors in items that they had answered incorrectly or omitted. Newman (1977: 252) found that about 13% of the identified 3002 errors arose because the pupils were unable to read the questions accurately enough to move beyond this first stage. Only 1% of errors was due to faulty symbol recognition; the remaining 12% were due to faulty recognition of key words. The term ‘key words’ was also used by Fry (1990) who defined a key word ‘as a word necessary for understanding a passage’ and was of the opinion that ‘[u]sually the key words are the most difficult in the passage’ (p595). It was not unreasonable to suppose that some of the children who did not make errors in reading failed at the second stage, comprehension, because their reading lacked the fluency and pace that would give them efficient access to the meaning. Nor do we know from these studies if fluency and pace of reading were taken into account. What we do know is that although Newman’s test was constructed to be easy to read, there were individuals who did not find it so. Like Newman (1977), Clements (1980:7) expressed concern about analysing data obtained from children’s written work, claiming that ‘it is likely that such studies will never advance our understanding of why children make mistakes on written mathematical tasks’. In his study, Clements (1980) explained how Newman’s (1977) hierarchy of error causes was applied to a different sample of pupils from that of Newman (1977). Unlike Newman (1977) who concentrated on low-attainers, Clements’ study (1980:11) explained how teachers conducted diagnostic interviews with virtually all children in 21 classes from three different age groups (grades 5 – 7). The children represented the normal broad range of attainment in mathematics and took part in a common test, developed by Clements, but similar to that used by Newman (1977). The results showed that 8% of errors in grade 5 arose due to reading difficulties compared with only 5% in grade 6 and 2% in grade 7. Using a different sample of children but the same test and diagnostic assessment, Clements (1980:11) obtained another set of data. In this sample, 92 low and 92 average attainers in grade 7 were interviewed after taking the test and their errors were also classified using the Newman (1977) hierarchy. In this study, it was found that 8% of the low attainers made errors in reading compared with only 2% of the average attainers.

Both Newman's (1977) and Clements' (1980) studies showed that reading difficulties, even with carefully monitored text, could prevent children from comprehending what they were being asked to do. Newman's study in particular showed that, even in a mathematics test that she believed did not include any items that were 'hard reads' for pupils in the fourth-grade, the low attaining children in the sixth-grade made errors in word recognition in about three test items on average in the 40-item test. These errors then had a 'follow-through' negative effect on comprehension. I would expect, therefore, that for children who do not ask for supported reading during the KS1 mathematics test, the reading demand should be appropriate and not disadvantage them. Consequently, I recognised that part of my research had to be to investigate whether the reading demand is 'appropriate' or otherwise.

Clements' (1980) first set of data showed that pupils were more likely to make reading errors on identical test questions than those who were one or two years older and had more reading experience. Clements' (1980) second set of data showed that pupils of the same age with low attainment in the test made more errors in word recognition than those of average attainment. A reasonable conjecture to make, therefore, was that pupils who were the more experienced or competent readers may have been advantaged over those who made more errors in word recognition since more accurate and fluent reading would enable them to understand better what was being asked of them. For a sub-set of children in these studies, unsupported reading may have been responsible for them gaining lower test scores. Some children, therefore, who ended up as the low attainers in the test, may have done better had their reading difficulties been anticipated.

The style of test used by Newman (1977) and Clements (1980) was not dissimilar in content to the KS1 mathematics test although the former included only single-step items whereas the latter includes some multi-step items. However, the youngest pupils in their studies were in grade 5 (rising 10-year olds). For the children at the centre of my research, errors due to unsupported reading may be more significant since they are several years younger and still emergent readers. The long-term impact is also likely to be more significant since the children are taking part in a statutory test on which schools' performance is judged against national or local authority norms. The incidence and possible effect of undetected reading errors, therefore, had to provide one focus for my research.

What was not made clear from the data provided by Newman (1977) and Clements (1980) was the range of the pupils' errors in reading; the study reported only on the average number of reading errors for the pupils in the sample. It was likely that making a higher than average number of reading errors would have disadvantaged some children in particular. Nor was it made clear if there were any key words or questions that were a common source of reading difficulties. I felt that these issues were necessary to address because of my interest in the reading demand of the KS1 mathematics test.

Reading demands of KS1 mathematics tests

In the research literature, only one unpublished study (Lupetti, Sainsbury & Schagen, 1995) appeared to have considered the reading demand of the KS1 mathematics test. Since this research had the most in common with the issues that were of special interest to me, I evaluated it critically and in more detail than other studies. The authors (Lupetti, Sainsbury & Schagen, 1995) explained that SCAA, later renamed QCA, commissioned the NFER to conduct a survey to address concerns about the 'reading and language demands' (p1) made on children by the KS1 mathematics test. Three of the four parts of the research were of particular relevance to my own interests:

i. Structured interviews were carried out with children who had taken the test to ascertain its reading and language demands.
ii. The reading demand of the KS1 mathematics test was compared to that children were likely to meet in written mathematical tasks in class.
iii. A multi-level model analysis was conducted to ascertain whether any effect of reading attainment could be distinguished from mathematics attainment.

Structured interviews with Y2 children

About a month after taking the mathematics test in May 1995, 17 children from five schools, described as ‘able’ or ‘less able’ readers, were interviewed. The term ‘able’ was attributed to those who had read the mathematics test with minimal help and my interpretation led me to conclude that these children were above-average readers. In the interviews, the children reread each question in the booklet from the test. The main findings follow.

i. The analysis revealed that ‘certain words consistently cause reading problems even with the more able readers’ (Lupetti, Sainsbury & Schagen, 1995:9). These included such ME and OE words as ‘symmetry’ and ‘complete’.
ii. The eight poorer readers could not read the majority of words in the booklet unaided. It was the view of the teachers and researchers that these children would ‘not have been able to achieve Level 2 without most of the booklet being read to them’ (Lupetti, Sainsbury & Schagen, 1995:10). Indeed, for seven of these children, their teachers stated that they had administered the test individually so that the whole text could be read to them.
iii. The poorer readers looked for clues on the pages, e.g. illustrations that would cue them in or for items such as abstract calculations, which they felt they could do without reading text. One approach used ‘was to scan the page several times in order to work out what was required before attempting to read the text’ (Lupetti,

Sainsbury & Schagen, 1995:13). Where these children relied on decoding skills other than reading to find out what to do, they were likely to misinterpret what was being asked since most items were text-dependent. The study stated that ‘children made valiant attempts to read the text but did not have the fluency to cope with reading and comprehension at the same time’ (Lupetti, Sainsbury & Schagen, 1995:13).

iv. Interviews with children suggested that ‘they were not always able to show their mathematical ability through a written test because of their inability to process the reading and language involved’ (Lupetti, Sainsbury & Schagen, 1995:18).

The writers acknowledged that this was a small-scale study and, therefore, any conclusions would have to be confirmed or otherwise by further and more extensive research. Extending the scope of their study interested me. For example, although the researchers did note incorrect attempts at reading on interview schedules, these are not reproduced in the study so the nature and extent of these and the time taken to read (or misread) the items are not reported. I would also have been interested to know more about the nature and frequency of the reading difficulties of individual children compared with their reading competence. Not only that, the adjectives ‘able’ and ‘less able’ used to describe readers in this study were somewhat vague and I felt that there was more to investigate about the characteristics of children reading within different levels of the national curriculum.

Comparing the test with mathematics schemes

A further aim of the study by Lupetti, Sainsbury & Schagen (1995) was to compare the layout and ‘the language and reading demands’ (p19) of materials written for Y2 children in the five most popular mathematics schemes at that time to the KS1 1995 mathematics test booklet. Part of this involved producing vocabulary lists of ME words and phrases used in the schemes. They compared the vocabulary with the language used in the test ‘in order to gauge how familiar each was likely to be to the children’ and concluded that, of the 44 occurrences of ME vocabulary in the test, only about half of these could ‘be expected to be amongst those that children find familiar’ (Lupetti, Sainsbury & Schagen, 1995:31). The writers acknowledged that the frequency with which an ME word or phrase appeared in the schemes was not taken into account so a word that appeared infrequently would still be classed as ‘familiar’, a classification that might tempt the reader into thinking that the child should be able to read the word. However, my view was that an ME word or term that is familiar through limited topical usage in a classroom context may not be familiar in a test context or in a written form in a test.

At this moment of writing, about a decade later, most children are familiar with the layout as tests from earlier years are used for practice and publishers have not been slow to produce related materials for home and school use. Also, since the arrival of *The National Numeracy Strategy* (DfEE,1999a) and its related publication, *The National Numeracy Strategy: Mathematical Vocabulary* (DfEE, 1999b), the mathematical language used in tests, in mathematics lessons and, predictably, in the spate of new published materials since the arrival of the strategy, are very much in alignment. What I believe is unlikely to have changed however, is the reading demand of the tests. Indeed, it is possible that this might have increased due to the range and quantity of ME words that are now meant to be used and understood by Y2 children. The comparison of the language and layout used in popular mathematics schemes to that of the test was more relevant in 1995 than in the time frame of my research; the reading demand, however, remains a legitimate area for research, as topical now as in 1995.

The multi-level model analysis

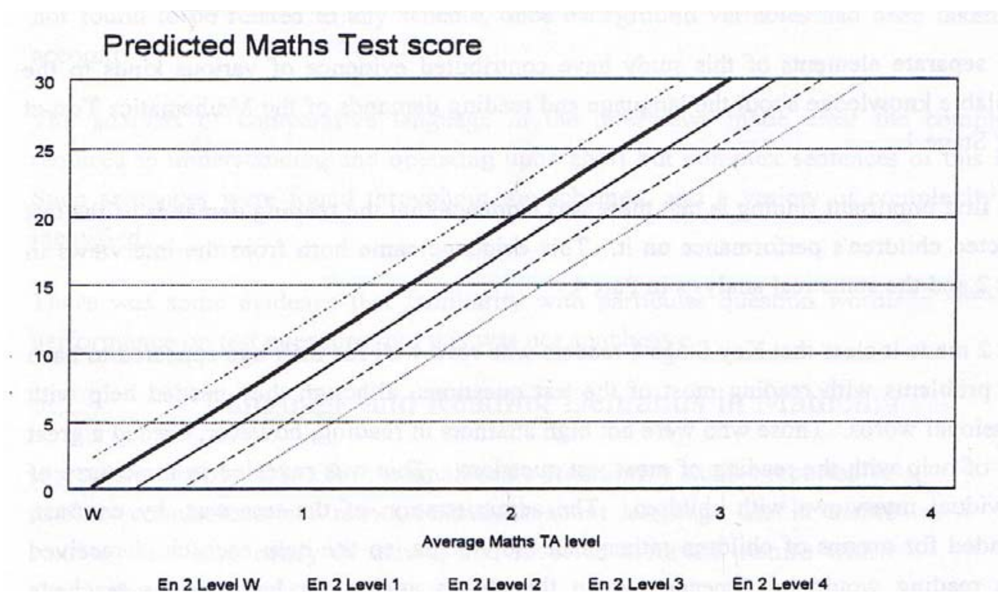
The multi-level model analysis was based on data from nearly 3200 children who had taken the mathematics test in 1995. The main purpose of the analysis was 'to explore the relationship between mathematics test score and reading ability, over and above a pupil's mathematical ability' (Lupetti, Sainsbury & Schagen, 1995:63).

Several models based on background variables, such as gender, ethnicity and English as an additional language, were explored but the final model included just the teacher assessment level for mathematics and the reading level from the statutory reading tasks/tests. Co-efficients were derived for these leading to a formula that was used to predict a child's mathematics test score. The conclusion of the study was that:

> When mathematics teacher assessment level [for mathematics] plus test level for En 2 [reading] were introduced, the results were highly significant. This implies that mathematics test score is not solely controlled by mathematics teacher assessment level, but is also influenced by reading level.
> (Lupetti, Sainsbury & Schagen, 1995:71)

For example, the application of the formula showed that a child who attained level 2 (average) in reading and who was teacher-assessed as working at level 2 in mathematics could be predicted to get a score of about 18 marks compared with a predicted score of about 20 marks for a child who attained level 3 (above average) in reading and who was teacher assessed at level 2 in mathematics. A one level increase in reading, therefore, was associated with about a two-mark increase in mathematics test score. Figure 2.1 (see page 37) reproduced from the study (Lupetti, Sainsbury & Schagen, 1995:71), shows the predicted effect.

It was not possible for me to replicate or adapt the multi-level analysis since the data used are not in the public domain. However, considered alongside the data from the interviews with the 17 children, the findings convinced me that further research was needed to determine whether children continue to have such reading difficulties since the style and content of the written questions are largely unchanged. Indeed, the analysis predicted that poorer readers would be positively disadvantaged. In the live test, if these children were to lose even one or two marks because of reading difficulties, their level of attainment in the test could be lowered.



Key: TA - teacher assessment   En 2 – reading   Level W - working towards level 1

Figure 2.1: Predicted mathematics test score as function of average teacher assessment level and reading level

The teachers interviewed in the study by Lupetti, Sainsbury & Schagen (1995) raised another issue that I felt I should address. It was reported that seven of the 15 children interviewed had every question read to them during the test due to reading difficulties. My own view was that children to whom all or most of the questions were read could be advantaged over those who attempted to read independently since those in the former category did not have to expend their time and energy on reading and could focus on the mathematics. Indeed, what was not known from this report was if any of the other eight children, who did not have every question read to them during the live test, were offered or requested help or made ‘valiant attempts to read’ (Lupetti, Sainsbury & Schagen, 1995:13) where it was not appropriate to do so. The study stated that even the more competent readers made at least one error such as substituting or omitting a word. It was not an aim of their study to determine if children received appropriate help with reading during the live test and I believed that more research was needed at that point in time.

Lupetti, Sainsbury & Schagen (1995) concluded that there was a need:

i. for teachers to give help with reading to allow access to the test items;
ii. to ensure that any complex language in the test was there because it is an essential part of what is being tested;
iii. for further research since ‘there is sufficient evidence … that language and reading demands have an effect on attainment in mathematics for this situation to warrant attention’ (Lupetti, Sainsbury & Schagen, 1995:73).

Since the outset of the KS1 mathematics tests, teachers have been allowed to give help with reading but it is possible that assumptions are made about children’s reading competence. As I explained in CHAPTER 1, an average Y2 reader who copes well with books from reading schemes, where words are constantly reinforced and become familiar through usage, may be considered capable of reading independently in the mathematics test where the language may not be familiar in written form. As previously discussed, ME words often fall into this category although understanding of ME vocabulary is part of what is being assessed; the reading is not. In CHAPTER 4, my methodology chapter, I explain how I planned to complement this research, i.e. establishing if the reading difficulties children had with the 1995 test would cross the boundaries of time to a test developed several years later. Not only that, I was also interested to extend it and to find out:

i. how teachers supported children with reading, why and to what extent;
ii. if it were possible to analyse the text in a typical test booklet for readability;
iii. if there were children who did not ask for reading support when they required it;

iv. if there would be a relationship between the frequency of help requested with reading and the reading competence of children who were assessed as comparable;
v. the nature and extent of the reading difficulties of children of different reading competences.

The writers of this study did not address these issues directly but it was my intention to do so. I would also have the advantage of being able to make my results public.

Reading demands of KS2 mathematics tests

In the studies of Lupetti, Sainsbury & Schagen (1995), Newman (1977) and Clements (1980), the key factor that linked and most influenced me was their use of diagnostic interviews resulting in evidence about the difficulties that some children had with reading items in a mathematics test. Although Shorrocks-Taylor & Hargreaves (1999, 2000) did not involve the children directly in the research described in their papers, they investigated why children might have difficulties in accessing the language in the statutory KS2 mathematics tests, taken by year 6 (Y6) children, and trialled methods by which reading difficulties might be measured.

In the earlier of their two papers, whose focus was a general review of language issues related to the test, Shorrocks-Taylor & Hargreaves (1999) considered factors similar to those identified by Mobley (1986), such as the format, the layout and the syntax of the test booklets. However, one of their stated concerns was that, in their review of the research literature, ‘there is very little research evidence about the precise issues surrounding the use of language in test contexts’ (Shorrocks-Taylor & Hargreaves, 1999:123). They argued that children taking the KS2 mathematics tests had more in common with their KS1 than their KS3 counterparts since ‘well-developed language skills cannot be assumed for much of the age cohort’ (Shorrocks-Taylor & Hargreaves, 1999:125). They considered skilled and less-skilled readers taking the tests, where their reading was not supported appropriately:

> There is ample research evidence that skilled reading involves rapid and efficient meaning construction, mostly by direct word recognition. Skilled readers can call upon back-up processes … but even skilled readers can become partially 'unskilled' when difficult or unusual text is encountered. (Shorrocks-Taylor & Hargreaves, 1999:127)

> Beginning readers have to move from being novice to being expert in this complex cognitive skill. … what is becoming increasingly clear is that there are many differences in strategy and approach among individual children and not all of them succeed. (Shorrocks-Taylor & Hargreaves, 1999:127)

Their views seemed to have much in common with those of Lupetti, Sainsbury & Schagen (1995) in drawing attention to the lack of research into this topic, the problems likely to be encountered by even skilled readers and the unsuccessful strategies that less skilled readers might apply.

In their later study, Shorrocks-Taylor & Hargreaves (2000) acknowledged that the teacher could read aloud an item at a pupil's request but also recognised that the pupil must have the confidence to ask or must know that they need help. They expressed concern that, since even pupils who cannot read competently generally work independently, it was possible that 'it is not the mathematical competence that is being assessed, but rather the ability to read and interpret information and instructions' (p40). Shorrocks-Taylor & Hargreaves (2000) were considering the KS2 mathematics test taken by 11-year olds, but I believed the following words to be relevant for KS1 children also:

> … although the focus of assessment is mathematics, the questions are presented through the medium of language, a factor that in itself may compromise the validity of the tests if the language in any way obscures or confuses the mathematical demand. (Shorrocks-Taylor & Hargreaves, 2000:40)

It was their view that, where pupils were expected to or chose to read independently, 'all the language should be as accessible as possible, employing agreed versions of mathematics vocabulary lists …' (Shorrocks-Taylor & Hargreaves, 1999:134). Agreed lists for each year group have since been introduced by the *National Numeracy Strategy* (DfEE, 1999b) but I was not convinced that such lists would make reading the mathematics test easier for the majority of Y2 children. By definition, mathematical vocabulary lists contain mainly ME words or phrases, many of which are likely to be 'hard reads' for reasons outlined previously (in **Mathematical English and Ordinary English***)*. Perera (1980) also argued that 'the presence of a word on a "familiar word list" will not necessarily mean that it is easy to read with comprehension' (p156). Indeed, I felt there could be a tension between test accessibility and such lists, which I consider in more detail in **Vocabulary and sentence length** later in this chapter. Although the strategy (DfEE, 1999a) does draw attention to some mathematical terms that Y2 children should learn to read, for the most part, it uses the phrase 'understand, use and begin to read' to preface most of the listed vocabulary. In a test of high status, expecting children to read independently seemed to me an unreasonable expectation when they were only expected to be beginning to read ME terms. Another factor that they considered could cause reading difficulties was that 'the *measured readability* of the text may be inappropriate' (Shorrocks-Taylor & Hargreaves, 1999:126), although they acknowledged that readability was not easy to capture in a measurable form. However, in the second of their studies, the aim was 'to investigate various methods for measuring the readability of the language of mathematics test questions' (Shorrocks-Taylor & Hargreaves, 2000:42) using the questions in a KS2 mathematics test. To do this, they compared the outcomes of applying several readability formulae to each question with the judgments of 15 experienced primary mathematics teachers. For each formula and for the teacher judgments, the questions were ranked in order of reading difficulty and the results compared. Later, in **Readability formulae,** I evaluate one of the formulae used by Shorrocks-Taylor & Hargreaves (2000) for its appropriateness for predicting the readability of texts for 7-year olds. The analysis of the teachers' judgments showed that they were influenced by criteria such as sentence length, supportive illustrations, mathematical vocabulary and unfamiliar words but 'the length of the questions may have been a major factor in the teacher decision-making processes' (Shorrocks-Taylor & Hargreaves, 2000:49). This came as no surprise since these criteria have been identified previously as potential causes of reading difficulties. However, the outcome of comparing teacher judgments and readability formulae was that, with the exception of one formula, 'teacher results were very discrepant with most of the results derived from the formal measures' (Shorrocks-Taylor & Hargreaves, 2000:59) even though there was a significant correlation between most of the reading formulae and between the judgments of most of the teachers. Consequently, it appeared that using teacher judgments in the way described would be less effective than using the children themselves to find out which questions they found most difficult to read, especially if at least some of these data were to be obtained during the live test. Data obtained in this way would remove the need for subjective judgments and establish actual rather than predicted reading difficulties. Harrison (1979) claimed that, although time-consuming, monitoring children's responses methodically was 'a perfectly valid method of assessing whether children can cope with a particular book' (p73). Nevertheless, I recognised that seeking Y2 teachers' views on the readability of the tests, and on how they supported the children with reading, would complement data obtained from children and from an analysis of the text in a test booklet for readability. I then searched existing literature to find out whether there were any systems or procedures that I could adopt or adapt for analysing the text.

The measurement of readability

Methods of measuring readability

Since involving children directly appealed to me, I started off by evaluating the cloze procedure, a process that 'involves deleting words from samples of the text at regular intervals, and requires the reader to fill in the blanks' (Mobley, 1986:7). By choosing appropriate words to replace the deletions, the reader can demonstrate comprehension of the text. In **Mathematics text and readability**, earlier in this chapter, I explained how Hubbard (cited in Shuard & Rothery, 1984) had successfully used cloze procedure to adapt passages of mathematical text. This procedure initially, but briefly, appealed to me because the child has to interact with the text, the process is familiar through classroom practice and it can be adapted for oral reading. I quickly realised, however, in relation to my area of interest, the limitations of such an approach far outweighed its benefits. For me, its greatest limitation was that all the words in the first two or three sentences in the text under scrutiny need to be left in place, or, as Harrison (1980) recommended for his study, that the deletions should begin 'at or beyond the fiftieth word' (p89). In this way, the child can build up a sense of the context prior to predicting the words that should replace the deletions. Since most items in the test booklet contained fewer than three sentences or, indeed 50 words, the process was entirely inappropriate. In a prolonged critique of cloze procedure, I could identify other limitations that would lead me to the same conclusion, but these became inconsequential in the absence of appropriate text.

Of the various methods used to assess reading difficulty, two are the most common but do not involve children directly. The first of these is a process of seeking the subjective opinions of those with relevant professional experience and pooling their judgments to obtain an estimate of the reading level of a text. However, in the previous section, I argued that asking Y2 children to read test items would produce more robust data on the readability of items than seeking the subjective judgments of their teachers. Consequently, I rejected the notion of pooled judgments in favour of asking teachers their general views on the reading demand of the tests and using children themselves to read test items and be the 'judges'. I then considered the second method, the application of readability formulae, in the hope that it would guide me towards a valid and reliable means of analysing the reading demand of the test booklet. If successful, I could then compare results with the children's judgments.

Readability formulae

Readability formulae are 'devices which provide a quantitative estimate of readability' (Klare, 1963:2). Their application attempts to measure the reading difficulty of several samples of text, often randomly chosen, leading to a prediction of the probable reading level of the entire text. This is usually expressed as an American school grade or a chronological age. The application of such formulae does not cover the range of aspects of readability I mentioned earlier in this chapter in **What is readability?** since elements such as, for example, graphic elements, a common feature in mathematics texts, are ignored. However, whilst I could not ignore such elements in the test booklet, the main focus of my research was to determine if it was reasonable to expect children to read the text without support. I needed to decide, therefore, if it would be appropriate to apply a formula to the test, and, if so, which one. Indeed, I found no shortage of research literature that considered the strengths and limitations of readability formulae (see, for example, Klare, 1963; Gilliland, 1972; Harrison, 1979; Harrison, 1980; Perera, 1980; Shuard & Rothery, 1984; and Mobley, 1986). For example, Perera (1980) pointed out that there was evidence to show that 'formulae are broadly reliable when used to find the average reading level of a number of books' but 'do not give a dependable reading age for any one book' (p152). Similarly, Wiest (2003) considered that they were best used 'for crude prediction, such as ordering the relative difficulty of texts' (p2). Of course, I did not wish to compare books for readability; nor did I wish to obtain only a 'crude prediction' (Wiest, 2003:2) of the reading age required for the test. Not only that, one concern of Harrison (1979) was that there is 'no causal relationship between difficulty level as predicted by a formula and the actual difficulty a reader is likely to encounter' (p103). To me, this suggested that, even if the text was found to be an appropriate ‘read’ for Y2 children in general, there would still be individuals who would find the reading demand inappropriate. In spite of growing reservations, I decided that further deliberation was needed before I finally made my decision.

Selection of sample text for analysis

I quickly realised from the literature that choosing samples to determine the readability of short bursts of text, sometimes single sentences, as found in the KS1 mathematics test, would be problematic. Most formulae measure at least one sample of running prose, commonly a passage of 100 consecutive words, a feature that is not present in the test. Furthermore, Harrison (1980) explained that ‘[r]eadability formulae were derived from an analysis of normal expository or narrative prose, …’ (p44) and that ‘it is only valid to use a formula on narrative or expository prose, and any deviation from this can be dangerous’ (p117). This view was upheld in another study:

> When test developers consider readability, they treat the entire test as one continuous prose unit and apply a traditional readability formula to the entire unit. The result is an average readability for the whole test. When test items are grouped together in this way … individual test items may be above or below the intended reading level of the total test…. (Homan, Hewitt & Linder, 1994:349)

Perera (1980:152) expressed similar concerns, stating ‘that the reliability of formulae decreases sharply on some types of written material’ and cited research by Bormuth (1966) who pointed out that ‘they cannot be validly applied to small samples of language, such as questions on workcards and examination papers ….’. She also made the point that even if 100 words of continuous text were chosen:

> It would be possible for two passages to achieve the same readability score when one had, perhaps, three very difficult sentences and four very easy ones while the other had seven relatively simple straightforward sentences … young or inexpert readers are not helped over a difficult section by easy sentences further on in the text. (Perera, 1980:153)

In the KS1 mathematics test, if 100 consecutive words were to be grouped and a formula applied, these would span several questions that covered different topics and the text would be unlikely to be consistently easy or difficult to read across the sample. If only one of the test items in the sample was a 'hard read' and the rest 'easy reads' for a particular child, the average readability for the whole sample could appear to be appropriate; this clearly would not be the case. Not only that, the test includes words that are not part of sentences, e.g. in table headings and on the labels of sorting diagrams, so samples would not constitute 'continuous text'. Continuous text is generally associated with a theme in which clues to unfamiliar words are likely to be prompted by the context. In a KS1 mathematics test, each item is a miniature context with its unique language, presented in a format that, in few instances, bears little resemblance to the style of text used in reading and story books. Even if 100 words were grouped together and treated as a sample of text, this would include several contexts. A readability formula could not compensate for the lack of a sustained context and could perhaps underestimate the readability of the sample.

Shuard & Rothery (1984) drew attention to a further weakness of reading formulae. They found that, when they applied different formulae to the same sample of text, they got 'very different estimates of the readability of the same passage' (p80), with two estimates differing by as much as six years. Thus they concluded that 'the formulae are not very helpful in assessing whether a particular reader will be able to read the text with meaning' (Shuard & Rothery, 1984:80). Harrison (1979) and Rye (1985) also reported differences in readability scores when a range of formulae was applied to identical samples of text. These findings further reduced my confidence in the use of a reading formula for the test since there would be no way of knowing how accurately it would predict the age level of the samples, one aspect of formulae reliability known as 'age level accuracy' (Harrison, 1980:53).

By this stage, I had recognised that applying a readability formula designed for continuous prose would not be appropriate for the mathematics test. However, in the research literature, I found three reading formulae whose use I consider in the next section since they were designed to predict the readability of shorter or mathematical samples of text.

Specialised readability formulae

Homan, Hewitt & Linder (1994) expressed concern that the readabilities of test items ‘are rarely taken into consideration by test developers even though standardized and other test scores are sometimes used as the basis for decisions which seriously affect the lives of those being tested’ (p350). To start to address this concern, they developed a formula for assessing the readability of single-sentence, multiple-choice test items, related to the social studies curriculum for the purpose of their research. Several aspects of this formula made me realise that it could not be applied with any accuracy to the KS1 mathematics test. The most important of these were that:

i. the items were not written to test mathematics and so did not take the unique features of ME text into account;
ii. the formula used, as one variable, a familiar word list developed in America around 1980 and was, therefore, likely to be unreliable to apply to text read by 7-year olds in England over 20 years later;
iii. the KS1 test does not comprise solely single-sentence or multiple-choice items. Indeed, the items that have more than one sentence, often contextual problems, could turn out to be the most problematic for a non-fluent reader.

I did, however, share their concerns about the possible impact of reading difficulties on children since no formal or semi-formal instruments have been used to predict the readability of the KS1 mathematics test, and ‘when readability is not controlled it might impact the validity of test scores’ (Homan, Hewitt & Linder, 1994:349). The conclusion reached by Lupetti, Sainsbury & Schagen (1995), discussed previously on pages 37-38, showed that their concerns were justified.

I then considered the readability formula for short passages developed by Fry (1990). Using this formula, it is possible to calculate a predicted reading grade-level for a sample of text containing three sentences with at least 40 words. Fry (1990) considered the formula suitable for ‘short but important passages such as those in … maths textbooks’ (p594). I rejected this formula too for the first two reasons mentioned in the previous paragraph but also because it could only predict whether a passage was likely to be suitable for an average reader in grade-levels 4 to 12, thus excluding its appropriateness for the majority of children taking the KS1 mathematics test whose average reading age would be just over 7 years (approximately equivalent to grade-level 2). At best, it could only report a readability score of 4 or below for a passage as suitable for ‘4th grade or below’. This would be of no practical use for my intended research.

Two readability formulae for mathematics text were designed by Kane et al (1974). Their second formula, known as Kane’s Formula II, considered to be the more reliable, took into account ‘the number of familiar OE and ME words, the number of familiar mathematical symbols and the number of changes from a word to a mathematical symbol and back again, within a piece of text’ (Shorrocks-Taylor & Hargreaves, 2000:42) and the number of question marks. For my research, I had to eliminate this formula also, the main reasons being:

i. individual items in the KS1 mathematics test contain far fewer than the required 400 ‘tokens’, i.e. occurrences of OE and ME words, symbols etc. for the formula to be applied;
ii. the formula does not provide a predicted reading age but allows mathematical texts to be compared and ranked by reading difficulty, an outcome that would not contribute to determining the readability of the KS1 mathematics test;
iii. the OE and ME words were taken from written sources, including mathematical texts, used by American children in the 7th and 8th grades in the early 1970s so

use of the word lists would be inappropriate for the next generation of children in England who were at least five years younger.

In this respect, it suffers from the same problem as the first two formulae mentioned in this section.

Shorrocks-Taylor & Hargreaves (2000) applied several formulae, including Fry's (1990) formula for shorter passages and Kane's Formula II (Kane et al, 1974) to items from one of the KS2 mathematics tests used in 1996. They had to adapt the latter formula to allow for fewer than 400 tokens in an item and acknowledged that all but the Kane's Formula II (Kane et al, 1974) were 'deemed inappropriate for measuring the difficulty of mathematical text' (Shorrocks-Taylor & Hargreaves, 2000:50). Nevertheless, they found that when the different formulae in the test 'are applied to the same text, they correlate at significant levels in most cases' (Shorrocks-Taylor & Hargreaves, 2000:59). However, when compared with teacher judgments on the same items, these 'were very discrepant with most of the results derived from the formal measures, …' (Shorrocks-Taylor & Hargreaves, 2000:59). This discrepancy further confirmed my resolve to use the children themselves as 'judges', particularly if this could be done at the time of a live test. Not only that, I had also recognised that it would be entirely inappropriate to apply a readability formulae for reasons I summarise below.

The rejection of readability formulae for the KS1 mathematics test

In this and the preceding section, I limited my argument against applying a reading formula to the main factors which informed my decision that it would not be valid or reliable to do so. Other factors that could affect the reliability and validity of formulae and further influenced my decision have been well documented in the prolific literature on readability. The factors I considered led me to concur with the view that '[u]se of readability formulae is questionable for most types of text and is especially debatable for mathematical text' (Wiest, 2003). More recently, a study commissioned by QCA concluded that ‘[t]here are significant elements …which would cast doubt on the effectiveness of using readability formulae on exam questions’ (Allan, McGhee & van Krieken, 2005:11). This certainly appeared to be the case for ME text written for my target readership. However, Harrison (1980) suggested that text difficulty could be evaluated at sentence level using ‘a word-frequency list’ (p55), by examining the syntax and by monitoring the materials as they are used by children. I have made no secret of my intention to include children in monitoring the test’s readability. Having concluded that it would be inappropriate to apply readability formulae to the KS1 mathematics test, I recognised that the variables that were the basis of most formulae, including word-frequency lists as suggested by Harrison (1980), might be useful tools to predict the readability of the test items. I consider these next.

Variables used in readability formulae

Vocabulary and sentence length

A reading formula is derived when at least two variables ‘which correlate best with reading difficulty are combined’ (Shuard & Rothery, 1984:77) and a calculation performed to produce a score that indicates the predicted reading difficulty of the text. The three most common variables are:

i. word length, i.e. the number of letters or syllables per word;
ii. word frequency, i.e. how often the word arises in everyday use; and
iii. sentence length, i.e. the number of words in a sentence.

I considered these with a view to how they might contribute individually to my intended analysis of the text in the test.

Word length

Harrison (1980:18) claimed that ‘research studies consistently find vocabulary to be the surest single predictor of text difficulty’. Indeed, Klare (1963:164) stated that ‘the most important unit in the study of readability is the “word” ’. However, in their assessment of word difficulty, Shuard & Rothery (1984:30) and Perera (1984:154) pointed out that long words, often multi-syllabic, are commonly equated with reading difficulty, with the result that, in many readability formulae, short words are generally regarded as easier to read. However, I knew that the KS1 mathematics test can include short ME words that are mono-syllabic but not phonetically simple, e.g. *length*, *height*, *graph*, as well as multi-syllabic ME words such as *reflection*. In such circumstances, I felt that using word length as the key tool for evaluation could lead to an under-estimate of the reading difficulty since standard readability formulae were not designed for analysis of ME text. Generally, word length as a variable in reading formulae is designed to be applied to OE text and the research literature so far has indicated that ME words are the most likely to be difficult to read. As previously established in **Mathematical English and Ordinary English,** such words are also unlikely to arise in everyday use, least of all in written form. This led me to consider lists of words that occur most frequently in writing since such lists provide the other principal means of assessing the reading difficulty of vocabulary.

Word frequency

Some readability formulae use the two variables of sentence length and ‘the number of difficult or unfamiliar words per hundred as determined by an established word list’ (Paul, Nibbelink & Hoover, 1986:165). Klare (1974:97) believed that word lists were slightly better at predicting readability than measuring word length. In referring to such lists, Harrison stated that ‘[t]he idea is that the greater the proportion of infrequently used words, the more difficult the reader is likely to find the passage’ (1979:78). Such lists generally comprise words that occur most frequently in writing or ‘in ordinary usage’ (Harrison, 1980:20). However, various criticisms of word lists were raised by Perera (1980) and Stuart et al (2003). These included that word lists quickly become out of date and are commonly derived in the United States of America where speech patterns can be different from those in other English-speaking countries. As explained previously, the latter criticism was one of the factors that led me to reject readability formulae. I did, nevertheless, identify strongly with Mobley's view (1986) that unfamiliar words 'are difficult to predict' and that '[t]he reader may have to pause to decode them, or to work out their meaning' (p27), a situation that would clearly be undesirable in a test that was assessing mathematical but not reading competence.

From my stance, a third criticism could have been that word lists are generally derived from analysing the frequency of words in narrative and expository passages of OE text. However, for the intended analysis of the KS1 mathematics test for readability, this could be to my advantage if I could make use of recent and relevant word lists whose validity would be less likely to be questioned.

Recent and relevant word lists

In **Mathematics and readability**, I referred to various sources in which explanations were given as to why ME words are often hard to read. I believed, therefore, that if words were used in test items that did not feature on lists appropriate for 7-year olds, I could conclude that they were likely to be unfamiliar in written form at least. Some of these were likely to be ME words. It would then make sense for me to compare a word-list analysis of the test with what actually happened when children working at different reading levels read items aloud to me. The latter analysis would be important, since, as Perera (2000) stated, '[i]f words on a 'familiar word list' are not really familiar, there is no guarantee that their presence in a text will contribute to reading ease' (p155). Word lists, however recent or relevant, cannot take account of the reading strengths or weaknesses of the individual. Nevertheless, I considered that the following four word lists were recent and relevant enough to contribute to my analysis:

List 1: High frequency words

The *National Literacy Strategy* (DfEE, 1998) lists about 200 essential 'high frequency words to be taught as "sight recognition" words through YR [reception] to Y2 [year 2]; these are words that 'pupils will need even to tackle very simple texts' (p60). It was explained that:

> These words usually play an important part in holding together the general coherence of texts and … will help pupils get pace and accuracy into their reading at an early stage. … By the end of Y2, pupils should be able to read all these words easily, in and out of context. (DfEE, 1998:60)

Additionally, the strategy summarises phonics to be covered by the end of year 2. This too could be useful for any analysis of the text.

List 2: Children's early reading vocabulary – word lists

Stuart et al (2003) analysed '[t]exts from 685 books from reading schemes and story books read by 5-7 year-old children' (p585) in England to create a word frequency list. I immediately recognised that this list, created in the period 1994-1996, satisfied my criterion of being up to date and relevant to Y2 children in English schools. Not only that, it had the added bonus of being downloadable from a website either ordered alphabetically or by word frequency. Both formats would have their uses if I were to decide to compare words in the test booklet with those that children in year 2 were most likely to find familiar through their reading and story books.

Of further interest to me was the conclusion to one of Stuart et al's (2003) research questions, '[H]ow many times does a child have to experience a word in print before they can name it accurately out of context?' (p586). Stuart et al (2003) found that just over half of the 9748 words entered onto their database appeared only once or twice and believed that this large number of low frequency words presented a problem since 'children may well not see these words repeated often enough … to commit them to memory at all' (p588). They backed up this statement by referring to an earlier study

(Stuart, Masterson & Dixon, 2000) in which they concluded that, for 5-year olds, '36 experiences of reading a word in text proved insufficient to guarantee that the words became stored in the children's memories to subserve subsequent recognition' (Stuart et al, 2003:588). Only 2.3% of words on their list appeared 36 times or more. By the time of taking the mathematics test, common sense and teaching experience led me to believe that children in year 2 are unlikely to have heard and certainly not to have read all but a few mathematical words with anywhere approaching that frequency. Even when words are introduced in written form in a mathematical topic, that topic may be the focus for a few days only each term. Not only that, the emphasis in year 2 is on practical work and discussion so exposure to written words will be infrequent, short in duration and, consequently, 'insufficient to guarantee that the words became stored in the children's memories' (Stuart et al., 2003:588 referring to Stuart, Masterson & Dixon, 2000). In topics that are not taught regularly, the words will not be revisited until weeks or months later, at which point I considered that the written forms were likely to remain unfamiliar to all but the most competent readers. Nor can it be assumed that teachers' exposition and reinforcement of even spoken forms of vocabulary related to a topic are effective. Raiker (2002:55) collected data from six groups of teachers and children across different year groups, including year 2, on the use of spoken language in mathematics lessons based on the *National Numeracy Strategy* (DfEE, 1999a) and stated:

> [T]he teachers in the study, even though they were using vocabulary recommended by *Mathematical Vocabulary*, were not aware of the importance of this vocabulary and did not plan for its introduction, explanation of meaning and repetition. … Understanding of this key vocabulary was not included in the assessment of achievement of learning objectives. (Raiker, 2002:59)

If Raiker's (2002) conclusions were to be representative of teachers nationally, this would add an extra layer of difficulty for children in the mathematics test. Not only could they be struggling to read unfamiliar words but also to unravel the meaning of ME words that they encounter infrequently in spoken and written forms and that may have been explained inadequately.

List 3: Mathematical vocabulary

The *National Numeracy Strategy: Mathematical vocabulary* (DfEE, 1999b) provides a checklist of vocabulary that teachers of children in each year group from YR to Y6 should plan to introduce during mathematics lessons. Teachers are advised to introduce new words in context, to ‘explain their meanings carefully and to rehearse them several times’ (DfEE, 1999b:4). The checklist also includes words and terms ‘commonly used when giving instructions … in national tests and in published resources’ (p2). Used in conjunction with the *National Numeracy Strategy* (DfEE, 1999a), it is possible to determine which words children are expected to read by the end of year 2. There are supplements of examples in the strategy, described as a selection of ‘what children should know and be able to do’ (DfEE, 1999a:38) by the end of each year. In the supplement for place value and ordering for year 2, (DfEE, 1999a:9 in the Y123 [sic] examples) it uses two different phrases preceding lists of words: ‘Read these words:’ and ‘Understand, use and begin to read:’ However, the former phrase is rarely used for the Y2 examples. Most examples suggest that Y2 children should be beginning to read the mathematical words by the end of the year. By contrast, where the phrase ‘Understand, use and begin to read:’ precedes most lists of words in the Y2 examples, the phrase ‘Use, read and begin to write:’ precedes the same lists in the year 3 examples. To me this implied that:

i. the reading of most words on the vocabulary list was not expected until year 3, at which time children have had an extra year of experience in mathematics and reading;

ii. all but the most competent readers in year 2, whose reading age was at least one year ahead of their chronological age, were likely to experience reading difficulties during the test in the absence of supported reading;
iii. unsupported reading could threaten accessibility to and the validity of the test.

I anticipated two further problems for less competent readers. First, they might not be able to read words expected of their more competent peers. Second, some of the words that children in year 2 were expected to read by the end of the academic year, according to the *National Numeracy Strategy* (DfEE, 1999a), were not on the high frequency word list in the *National Literacy Strategy* (DfEE, 1998:61), as discussed at the start of this section. I knew that this did not necessarily mean that children could not read them but that they might revert to using phonics or hesitate, thus compromising fluency and concentration on the mathematics.

List 4: Words used by 7-year olds in their writing

In this study (Reid, 1989), 979 scripts from 7-year olds in British schools, that demonstrated ‘a variety of purposes of writing – imaginative, factual, descriptive, etc.’ (p3), were analysed. The outcome of the analysis was a list of words used by the children ordered by frequency of usage. This list too could be useful since it is likely that words that children commonly use in their writing are familiar to them in reading materials also.

Selected word lists

I was confident that the four word lists mentioned in the preceding section would make different but complementary contributions to an analysis of a mathematics test booklet and help me to decide how appropriate the reading demand was for its intended readership. Through their availability, I would be able to determine for Y2 readers:

i. words they are expected to recognise on sight;

ii. the likelihood that words are familiar in written form through reading, story books and their own writing;

iii. ME words that they are expected to recognise on sight and distinguish these from ME words that they should only be starting to read.

Sentence length

The average length of sentences, usually calculated from 100-word samples of continuous text, is used in most popular formulae as one variable to predict reading difficulty using the principle that 'the longer a sentence, the harder it is to read' (Shuard & Rothery, 1984:32). However, a criticism common to both word and sentence length is that shorter does not necessarily mean easier. Gilliland (1972) Harrison (1979), Perera (1980), Shuard & Rothery (1984) and Fry (1990) gave various explanations and examples why shorter sentences were not always associated with increased readability. Gilliland (1972) stated that '[a]t the sentence level, a short sentence of unusual structure may be more difficult to read than a longer, more familiar structure' (p96). Indeed, on page 27, I quoted from Perera (1980) who found evidence that 'children read more easily those sentence structures that they would themselves say or write' (p156). In the mathematics test, some of the sentences are relatively short but not necessarily familiar in structure or content, if likened to what children normally say, write, or, indeed, meet in reading schemes and story books. As Harrison (1979) pointed out, 'many of the sentences analysed in mathematics texts were brief instructions rather than expository text', and could, therefore, because of their short sentence length, 'tend to lower the reading-level score' (p84). Although the comments related to secondary texts, I felt that these comments were also relevant to the KS1 mathematics test, as I explain next.

At first glance, most sentences in various KS1 mathematics test booklets seem relatively short and the syntax relatively simple. I knew from my involvement with QCA that great efforts are made to monitor sentence length and syntax. However, the requirement to include items on many different topics and contexts means that the test overall is language-laden and, as is often the nature of subject-specific text according to Mobley (1986), ‘packed tight with meaning’ (p30). Children who do not ask for, or who are not offered support, have to read at least one sentence in each item, and, in some cases, several. Additionally, the children often have to relate the sentences to other graphic content such as tables, which adds another element of complexity and even more words to read. I recognised that, in a test item, even if introduced by a sentence comprising a brief instruction, the children would need to concentrate on the holistic meaning of the sentence, rather than on individual words or phrases to respond to the mathematics. I anticipated that many children in year 2 would find such an approach difficult, particularly if ME or unfamiliar everyday words were included. In such situations, sentence length alone as a predictor of reading difficulty could lead to an under-estimate.

Nevertheless, I could not ignore sentence length entirely, since, as Gilliland (1972) explained, sentence length can ‘be a reflection of memory span, since the longer a sentence is, the more difficult it will be to remember the parts and so the more difficult it will be to understand’ (p91). I believed that the effect on memory span, caused by hesitant reading, would also apply to short sentences that included unfamiliar sight vocabulary, as could be the case in the KS1 mathematics test. Irrespective of my predicted effect of unfamiliar words even in short sentences, I was curious to find out what length of sentence it was reasonable for an average child in year 2 to read. On page 48, I wrote a short critique of Fry’s (1990) readability formula that he considered suitable for ‘short but important passages’ (p594) such as mathematics text. In this, Fry (1990) included sentence length as one variable by averaging the number of words in each sentence in the sample to obtain an ‘average sentence difficulty’ (p596), expressed as a grade level from grade 1 onwards. For example, he estimated that if the average sentence difficulty was between 6.7 and 8.6 words, it was appropriate for children of grade level 2 (7-year olds). Although Fry's (1990) formula was for application to continuous text at least three sentences long, I believed that sentences approximating this length would not be unreasonable in a KS1 mathematics test, but *only* if all words were familiar to the reader, which now seemed unlikely. Sentence length alone as a measure of readability would fade into insignificance for a child reading a sentence that included unfamiliar words. Indeed, it was Harrison's (1979) opinion that worksheets or instructions for a task 'ought to be in simpler prose than anything else a child reads' (p84). More specifically, Mobley's (1986) view was that '[i]f the reading is to be unsupported, it may need to be two years *below* [sic] the pupil's reading level' (p49). If true, this could mean that the language in the mathematics test should be at the reading level of a 5-year old child, which would clearly be nonsensical! In the preceding paragraph, I referred to Gilliland's (1972) concern about the possible negative effect of longer sentences on 'memory span' (p91), i.e. there could be too much concise information to 'hold in the head'. I also had concerns about the total number of sentences in the test, even if considered of appropriate length, that independent readers could be expected to read before the cumulative effect of so much reading took its toll. For example, to reach the end of the 2001 test (QCA, 2001c), a typical example, a child who asked for no support in reading would have been expected to read more than 50 sentences (not including words that were not part of sentences, e.g. axes on graphs) to respond to 30 written items. This was the most extreme scenario, because, in reality, I knew from classroom observations that children asked for varying amounts of help. Nevertheless, some asked for little or none. For these children, the 'concept load' (see section on 'Assessing readability of text' in Wiest, 2003), arising jointly from the reading and mathematical demand, was likely to result in children becoming increasingly mentally fatigued as the test progressed and having cumulative difficulties in remembering what they had read. To add to this potential burden of reading, some sentences would also include ME words, generally categorised as difficult to read.

Most of all, I anticipated that the amount of text would be excessive for less competent readers whose reading was unmediated. I could find no studies relevant to my research that had investigated the correlation between the amount of text to be read and the effect on a child's memory, concentration, and, indeed motivation, since the text is imposed and not chosen. I also recognised that the readability of a sentence, regardless of length, is affected by its syntax. Shuard & Rothery (1984) stated that 'the importance of syntax is as great as that of vocabulary' (p34). For example, both Shuard & Rothery (1984) and Clausen-May (2001) recommended the avoidance of conditional phrases and the passive tense in instructions or questions in mathematics texts. However, I knew from my involvement with QCA that item writers for KS1 mathematics were instructed to avoid these and I did not expect unfamiliar syntax to exacerbate reading difficulties. However, any findings to the contrary would have to be acknowledged.

Having considered the three main components of reading formulae, word length, word lists and sentence length, I had to decide which of these would be the most effective in analysing the reading difficulty of the test. I concluded that:

i. word length was likely to be the least effective since some ME words used in the tests can be short or mono-syllabic but phonetically complex, e.g. length;
ii. recent words lists were likely to be the most useful since they would identify words that were likely to be familiar in written form to readers in year 2;
iii. sentence length would also be useful but only on an item by item basis since component words could affect readability more than length.

For my test booklet analysis, therefore, I would have to consider whether it was possible to combine a qualitative approach using 'commonsense, intuition and experience' (Perera, 1980:152) with a quantitative approach using recent and relevant word lists and sentence length as predictors of readability. Of course, the latter approach would have an important limitation, because reading formulae, and by implication, their component parts, are at their 'weakest when being used to link individuals with texts' (Rye, 1985:111). Since children, even those assessed as being of comparable competence, have individual strengths and weaknesses in reading, any analysis would complement but be of secondary importance to hearing Y2 children read test items to determine the extent and nature of individuals' difficulties. For example, I could identify children who were reading at 'frustration level' as described in *A language for Life* (Committee of Inquiry into reading and the use of English, 1975:254) or exhibiting the characteristics of poor readers as described on pages 19-21. Only then could I find out if the reading demand 'in any way obscures or confuses the mathematical demand' (Shorrocks & Hargreaves, 2000:40). Nevertheless, an analysis of the booklet would enable a prediction to be made about whether the reading demand was appropriate for the general readership of Y2 children. For example, Klare (1974) believed that:

> … counts of the 2 [sic] simple variables of word length and sentence length are sufficient to make relatively good predictions of readability. No argument that they cause ease or difficulty is intended; they are merely good indices of difficulty. (Klare, 1974:97)

Whilst I could never combine these two variables in a formula, I could at least consider words and sentences as independent variables if I were to find items that children had difficulty reading. I now agreed with Wiest's (2003) comment that '[the] limitations of determining the readability of mathematical text are particularly great'. However, the reasons that make this so, some of which have been considered at various points in this chapter, had not persuaded me that evaluating the readability of the KS1 mathematics test was an impossible task, especially if children were to assist me.

Towards a theoretical framework

Learning from the literature

I learned much from the various studies read during the evolution of this chapter and had no doubt that they would guide the direction of and improve the quality of my own research. Literature on readability in general was abundant. However, recent and relevant literature on the readability of statutory mathematics tests was scarce. Considering the high status that has been given to such tests, I was surprised that so little research had been undertaken and even more surprised about the lack of studies that allowed the voices of KS1 children to be heard.

Of the few studies that had investigated the readability of mathematics tests those by Newman (1977), Clements (1980), Lupetti, Sainsbury & Schagen (1995) and Shorrocks-Taylor & Hargreaves, (1999; 2000) convinced me that my intended study was worth pursuing since the conclusion from each of their studies was that the reading demand of the tests was not always appropriate for its audience. One disappointment on reading these studies, however, was that they were based upon empirical data collections and did not appear to attach themselves to any particular theoretical stance. The result was that I could not use this literature as the basis for a 'theory-testing' (De Vaus, 2001:6) or 'theory verification' (Robson, 2002:62) approach but would have to construct for myself a context or personal theory in which to position my research. Nevertheless, I was particularly influenced by the three earlier studies (Newman, 1977; Clements, 1980; Lupetti, Sainsbury & Schagen, 1995), since their epistemological stance was more in tune with how I wished to develop my study. This was because the authors used the voices of children rather than teacher judgments or measures of readability to establish the nature and extent of the reading difficulties that children of various ages and reading aptitudes had. In the relevant sections in this chapter, I have indicated possibilities for enhancing their contribution to knowledge.

Various studies also recommended that further research was needed on the readability of mathematical text. For example, a general recommendation from Shuard & Rothery (1984) was:

> Despite the problems of using readability formulae with ME, it remains important to assess whether a particular piece of ME writing might be 'easy' or 'difficult' or 'about right' for a particular child. Making such an assessment is not a simple matter. To arrive at such a judgment, it is necessary to look closely at the styles of writing used in ME, and the ways in which children respond to what they read. (Shuard & Rothery, 1984:2)

When considering statutory tests, Shorrocks-Taylor & Hargreaves (2000) recommended that:

> …further research is urgently needed, especially since the tests in the UK contain a wide range of question types, many of which are rich in information and necessarily language-heavy. (Shorrocks-Taylor & Hargreaves, 2000:59)

More specifically, Shorrocks-Taylor & Hargreaves (2000) acknowledged that further research 'will need to involve pupils themselves in judging the difficulty and comprehensibility of what they are being asked to read in mathematics, especially under the stress of test conditions' (p59), a recommendation that I intended to pursue. The result of my review of the literature was that the issues I wished to investigate had been identified as areas needing research but had not yet been studied. To the best of my knowledge, no data had ever been made public about:

i. how Y2 children coped with reading the KS1 mathematics test during an administration;
ii. the amount of help requested with reading during administrations of the test;
iii. the efficacy of teachers' support with reading during the test.

This was what I wished to research. I could find no evidence of a similar undertaking since the unpublished research of Lupetti, Sainsbury & Schagen (1995). Mindful of the literature mentioned in this section, I needed to create a way to proceed so that I could interpret those aspects of the children's world in which my interests lay. Consequently, I now consider my study's significance through its theoretical framework.

My theoretical framework

Linden (2002) claimed that a researcher's theoretical perspective 'provides a "mind-set" for exploring, exposing, noticing and interpreting' (p70) the phenomenon of interest. My theoretical framework arose out of the conjunction of two personal experiences. The first was my experience as a primary teacher; the second as a consultant writing items for the KS1 mathematics test.

As a primary teacher, I knew that even the most competent rising 7-year old children were still acquiring reading skills. When reading aloud, mainly from structured reading schemes such as those studied by Stuart et al (2003), children had reading difficulties to varying degrees. For example, they could fail to correct errors, misread words that could change the meaning of the text, use phonics inappropriately or read with little intonation or phrasing. For such children, direct input from the teacher is essential, and, for texts read regularly in class, teachers are aware of children's strengths and weaknesses through continuous teacher assessment. In a one-off mathematics test, the teacher is not in a position to know how well or otherwise children of different reading competences can cope with text that bears little resemblance to standard reading materials, as discussed earlier in this chapter. Furthermore, opportunities to rehearse the text with the children are not possible so children's competence to read text independently is an unknown. Unless children are heard reading from the test, it is not possible to know whether they will demonstrate the characteristics of competent or poor readers (see for example, QCA, 2001a and Gibson,1989). Part of my theoretical context is the degree to which children's reading affects their access to the mathematics.

In CHAPTER 1, referring to my role of test consultant, I explained that observations of pre-tests in particular had led me to believe that the reading demand of the KS1 mathematics test could be excessive for at least some Y2 children and that they might not be receiving the necessary support from teaching staff under 'the stress of test conditions' (Shorrocks-Taylor & Hargreaves, 2000:59). Any assessment is influenced by the context in which the assessment takes place. In my observations of pre-tests, children were treated differently by teaching staff and behaved differently. For example, some teachers kept in the background and only intervened when a child asked for help with reading; others were pro-active and offered help automatically. Children too could range from keeping a low profile and avoiding teacher interaction to requesting considerable amounts of help. Common sense also told me that children who were above-average readers were likely to have an advantage over those who were less skilled when reading independently because they would have easier access to the text. The behaviour and attitudes of teachers and children during the test, with regard to reading support, could affect children's test score. I had a hunch or 'personal theory' (Robson, 2002:62) that this could be the case but needed to formulate research questions that would then lead me to find an appropriate methodology and research design to test my emerging theory.

As someone who had worked in the classroom as a teacher of young children and as a consultant in test development, I could position myself in both contexts and still have the same concerns. Consequently, my theoretical position would be as a spokesperson for Y2 children since I was openly ideological in the sense that I believed that a mathematics test should not covertly assess reading skills. In a mathematics test, children reading independently need to do so fluently and accurately so that they know with what mathematics they are being asked to engage, regardless of how difficult they might find the content. For those who lack reading skills to work independently, support should be available. However, when children are left to read independently, it is impossible to tell whether their reading skills are adequate or whether they receive the support that they require and that is their entitlement. Children could be losing marks because of unidentified reading difficulties and this could adversely affect their final score or the end of key stage level awarded to them. For any child hindered by reading difficulties, the validity of a mathematics test would be compromised. In this respect, I would also be contributing to assessment theory.

As I explained in CHAPTER 1, I was aware of the growing literature that looked at the world through children's eyes and listened to children's voices. I wanted to contribute to that in such a way that I could enable young children's voices to be heard. Consequently, I wished to position my study in the context of a KS1 mathematics test where children could not speak up for themselves and over which they had no control. I now consider theoretical issues raised by Marshall & Rossman (1995):

> In examining a specific setting or set of individuals, the writer should show how she is studying a case of a larger phenomenon. By linking the specific research questions to larger theoretical constructs or to national policy issues, the writer shows that the particulars of the study serve to illuminate larger issues and, therefore, are of significance. (Marshall & Rossman, 1995:7)

I could never be a spokesperson for the national cohort of nearly 600 000 children who take the KS1 mathematics test; nor could I involve enough children to generalise for the cohort. However, by involving sufficient typical Y2 children in the test setting, my rationale was that children with similar reading skills could be found in most schools in England. In this respect, I would be studying a '*case* of a larger phenomenon' (Marshall & Rossman, 1995:7). I would also be contributing to socio-cultural theory because the responses of children to the reading would be affected by who they were, e.g. 6- and 7-year olds who were learning to read and who were not 'test-wise', and the unique setting in which they found themselves. Although statutory end of key stage tests were part of government policy, investigating 'national policy issues' (Marshall & Rossman, 1995:7) was never the driving force of my study. However, any data I collected had to show if there was a problem caused by unanticipated reading difficulties and the effect of that assessment policy on KS1 children. I also identified with Wellington (2000) who stated that '[a] theory may be … a framework for understanding or making sense of things which happen in education' (p27). I hoped that I could bring about a better understanding of how readable children found the test and how effectively teachers supported children's reading, which were the main aims of my study as stated in CHAPTER 1.

A theoretical framework provides a rationale and justification for a research proposal. In my framework, I have argued that *what* I want to study is important, most of all to the Y2 children. Next, I had to try to link *what* I wanted to study with *how* I intended to conduct my enquiry so that it fulfilled the important criterion of 'do-ability'. Miles & Huberman (1994) write about 'focusing and bounding the collection of qualitative data in the field' (p18), one aspect of which was to derive a theoretical framework. They also advised that 'conceptual frameworks and research questions are the best defence against overload' (p55). My research questions appear in the following section. I considered these to be the minimum required to 'provide explanations of the phenomenon under analysis' (Denzin, 1970:56), one function of socio-cultural theory.

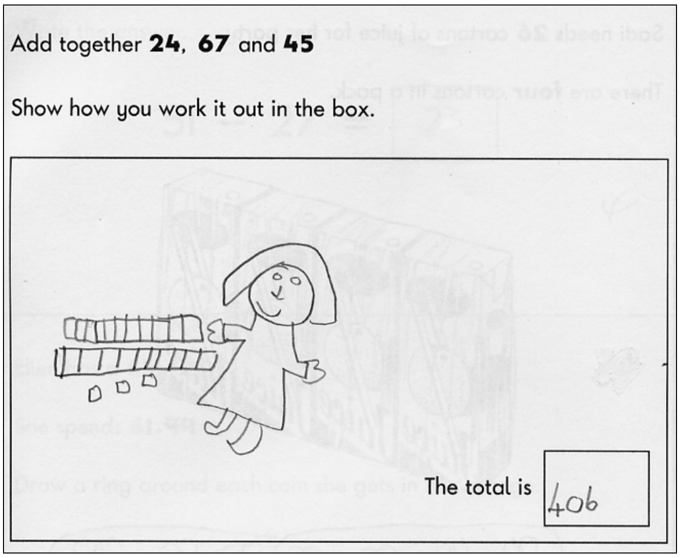
Research questions

I believed that I could fulfil the aims of my study by attempting to answer these research questions:

i. *Is the reading demand of the text in a typical KS1 mathematics test appropriate for the target children in year 2?*
ii. *What is the extent of the help needed by children of different reading competences during the test?*
iii. *What are the views of year 2 teachers on the reading demand of the test?*
iv. *Is teacher support with reading available to children during the test whether they ask for it or not?*
v. *What is the reading age required by a Y2 child in order to read the text in a typical KS1 mathematics test fluently and accurately?*

Although I limited my search of the literature to focus on readability, a theme that has been explored in many studies, I believed that I was about to undertake research on a phenomenon linked to KS1 statutory testing that had theoretical links with readability and about which little was known. I expected my contribution to be original and was optimistic that my research questions would lead me into a world of enquiry in which the boundaries were well-defined and that had young children at its centre. Since I wished to represent the children's voices, I intended to do this as effectively as I could. In CHAPTER 4, after my pilot study in CHAPTER 3, I explain how I developed a methodology and research design that was in sympathy with my theoretical position and took account of my research questions.

CHAPTER 3



Laura was working at level 2A in mathematics.

CHAPTER 3

Trial and improve - my pilot study

Organising the pilot study

Justifying a pilot study

In November 2000, the first pre-test of items, from which items for statutory KS1 mathematics tests in later years would be chosen, took place. The first pre-test was administered to about 1400 children in year 3 rather than in year 2. This was because Y2 children who took the tests each May (until revised arrangements started in September 2004) had not covered enough of the mathematics that forms the basis of the test at that point in the academic year. I chose this event to collect data for my pilot study. I had done some background reading about methodology and would have preferred to conduct the pilot study after I had written at least a draft of my methodology chapter. In so doing, I could have thought through in depth and justified the proposed evolution of my study. However, a delay was not advisable because the next opportunity for pre-test observations would not have been for at least another seven months and would overlap with the time-frame when I intended to collect data during and after the KS1 2001 mathematics test. This was to be the main source of data collection of my study and a new contribution to research. The live test was to be administered during May 2001 and failure to collect appropriate data would mean that I would have to wait until May 2002, the next administration of the statutory test. I could also see potential advantages in conducting a pilot study prior to writing the methodology chapter since I would establish ‘the need and right to determine the precise focus of the research *after* these first days in the field, after that experience begins to clarify relevant themes and patterns' (Marshall & Rossman, 1995:43). I believed it would provide a background of initial experience where I could make mistakes and learn from them. In turn, this would assist the writing of my methodology and help me to justify my approach and acceptance or rejection of qualitative methods.

Pilot study design

Nine consultants including myself, chosen mainly from the KS1 mathematics writing team, and approved by QCA, each, as observers, visited a school administering the first pre-test. The schools were chosen to be in different geographical locations across England for representativeness. The letters A to I will be used as identifiers for these schools. With the permission of QCA, I organised the study to be carried out during the pre-test with the intention of trialling and evaluating methods:

i. for measuring the time children of different reading competences took to complete the pre-test;
ii. for finding out the extent to which children asked for help with reading during the administration of the pre-test;
iii. for assessing how well or otherwise children of different reading competences could read selected test items.

In particular, I wished to trial and evaluate methods to obtain data on the first two criteria in a pilot study so that I could be reasonably confident that, when I wished to gather data during the administration of the statutory test in May 2001 with Y2 children, my methods had been carefully planned so as not to disrupt or worry them. This was because I believed that the best way to find out if the reading demand of the test was appropriate, the main aim of my study, was to obtain data during the live test. I also believed that there were data that could be gathered during the live test that would not be possible to obtain at any other time, e.g. the number of requests for help with reading compared with reading level.

I was influenced in particular by studies undertaken by Lupetti, Sainsbury & Schagen (1995) and Newman (1977) who developed observation sheets for tracking children's difficulties as they worked through mathematics test items. Both studies identified reading difficulties as a source of concern, as explained in my review of the literature. As I also explained, the former study concluded that the reading demand of the KS1 1995 mathematics test was too difficult for the Y2 children but the study was unpublished. I wished to find out if the same difficulties were also present in the 2001 test but also to build on the contributions made by these studies, e.g. by:

i. collecting data on requests for help with reading during the live test; and
ii. making audio-recordings of children reading items from the test so that detailed coding could be undertaken after rather than during the reading as was done in these two studies;
iii. interviewing Y2 teachers about the reading demand and on how they supported children with reading.

Another aspect of interest was the speed of reading as discussed in my review of the literature. In this pilot study, I also wished to trial approaches:

i. to find out whether children were reading at the 'optimum speed' (Gilliland, 1972:13) that supports understanding of text or 'at a rate appropriate to the print form' (Gibson, 1989:110); and
ii. to record evidence of how accurately or otherwise children were reading.

I recognised that children might be reading at an appropriate speed but not reading accurately so the two aspects were closely interlinked. These data too would make a contribution to existing knowledge.

Applying an evaluative approach

During my pilot study and thereafter, I saw myself as participating in evaluative research, seeking to ‘describe, interpret or explain’ (Bassey, 1995:6) as well as make a judgment about:

i. the appropriateness of the reading requirements of a KS1 statutory mathematics test for its intended audience;
ii. the degree to which the children that were its audience received the amount of reading support that they needed.

I discuss this approach in more detail in my methodological stance in CHAPTER 4.

To set up the data collection process, I telephoned the headteachers of schools who had agreed to a visit from an observer to get permission to carry out the research. I followed the initial telephone contact with a confirmatory letter and a letter to parents asking for approval for children to be interviewed. I also planned in advance an observer’s pack for each consultant to help them prepare for the visit and to try to standardise what they should do during the visit. Clear guidelines were essential because there could be no retakes of the pre-test and the visits were ‘one-off’. Consequently, all data had to be collected on the day. The consultants were asked to liaise with the headteachers of schools taking part in the pre-test to confirm approval:

i. to observe the administration of the test;
ii. to underline any text for which the children asked help with reading at the time of the request;
iii. to track the speed at which a focus group of three children of different reading competences (identified by teachers as below-average, average and above-average) worked through the test booklet;
iv. to interview children in the focus group with the permission of parents or guardians and, on the day, children themselves.

(See Appendix 3.1 for a copy of correspondence relating to the tracking sheet and test items used for the interviews.) The rest of this chapter presents findings from the pilot, an evaluation of the methods and materials used with a view to informing my methodological choices.

Focus groups

Research methods: factors that may skew findings

The end of this chapter includes recommendations for improving some of the research methods and materials used in the pilot study. However, some factors that may have skewed the findings need to be drawn to the reader's attention. First, only five of the nine schools nominated three children who included a below-average, average and above-average reader. Of the 27 children taking part, almost one-half were above-average readers, mainly reading at level 3 rather than level 2A, with about one-quarter being average and one-quarter being below-average readers, according to their statutory reading results from May 2000. My intention had been to have one third in each category. This meant that any averaging of the data obtained would have to reflect the high proportion of above-average readers, which could skew data in an already small sample. Second, although not known in time to make other arrangements, schools C and E (italicised in Table 3.1 on page 75) were private independent schools with a selective intake. School E could not provide children reading at levels 2C or 2B but the observer did not know this until arrival at the school. Third, one observer, who visited schools C and G, forgot to underline text from the pre-test booklets for which help was requested with reading. This was particularly disappointing in relation to school G since a high proportion of children had English as an additional language, the only school to which this criterion applied, and a factor that could be significant. However, in spite of these disappointments, the strengths and weaknesses of the resultant data made the amount of work involved in the pilot study worthwhile as a steep learning curve in the journey towards writing up and justifying my methodological choices.

Focus group results

After marking the pre-test booklets for the children in the focus groups, I prepared Table 3.1 to give an overview of results for the children whose progress through the test was tracked and who were interviewed. At the time of the pre-test, the children were about six months older than when they took the statutory tests but the reading and mathematics levels from these were the latest available. The children were likely to have progressed in these two subjects in the intervening period, and several teachers made this point but I believed that children who were, for example, below-average readers in May 2000 were likely to maintain that status six months later.
I return to the tracking sheets and interview data later in this chapter. This table includes the time taken for children to complete the pre-test and the number of requests for help with reading, aspects in which I had a special interest.

Table 3.1: Data for children in the focus groups sorted by school

		May 2000 statutory task/test level		November 2000 mathematics pre-test marks			
School	Name of child	Reading	Maths	Oral items (max 6)	Written items (max 30)	Time taken on written items (minutes)	No. of items where help requested
A	Luke	2C	2C	5	3	30	0
	Rosie	2B	2B	6	17	35	0
	Nisha	3	3	6	24	50	0
B	Katy	2C	2A	2	18	58	13
	Miles	2B	3	5	20	45	7
	Izzy	3	2A	6	22	39	1
C	*Kate*	*2C*	*3*	*6*	*18*	*32*	*n/k*
	Suzie	*2B*	*3*	*6*	*15*	*52*	*n/k*
	Alice	*3*	*3*	*6*	*17*	*25*	*n/k*
D	Bianca	2C	2C	3	7	38	0
	Aaron	2B	3	6	26	33	0
	Alexandra	3	3	6	28	37	0
E	*Rhaki*	*3*	*2A*	*4*	*20*	*52*	*0*
	Sarah	*3*	*2A*	*5*	*22*	*50*	*0*
	Stephanie	*3*	*3*	*6*	*25*	*34*	*0*
F	Bobby	2C	3	3	13	30	0
	Charlotte	2B	2A	6	16	25	0
	Kerry	3	3	5	29	25	0
G	Osman	2C	2A	4	19	35	*n/k*
	Manisha	2A	2A	6	20	24	*n/k*
	Leigh	3	2A	6	23	22	*n/k*
H	Jack	2C	2B	5	10	50	9
	Matthew	2A	3	6	24	30	0
	Roseanne	3	3	6	21	35	0
I	Ben	2B	2B	5	14	20	0
	Krystal	2A	2A	5	11	37	0
	Ashlyn	3	2A	5	21	40	0

Shaded rows: results for children attaining L2C in reading (below-average) in the May 2000 reading task/test
Italics: private independent schools with selective intake
n/k: not known since observer forgot to organise underlining in the booklets.

Extent to which children asked for help with reading: focus groups

My first reaction was surprise at how few requests for help were made by children at all reading levels, but particularly by those who were average or below-average readers. Of the 21 children for whom data are available, as shown in the final column of Table 3.1, 17 children requested no help at all. Of the four who did, three were at School B, and only two, Katy and Jack, were below-average readers. Even allowing for the fact that these children had had six months to improve their reading skills since obtaining their end of key stage 1 reading level, I was not convinced that all children had requested or were offered appropriate reading support. For example, compare Luke's and Katy's results. Both were awarded level 2C in reading in May 2000. Luke coped better than Katy in the oral items, gaining 5 marks out of 6 compared to Katy's 2 marks out of 6 showing that he responded well to teacher-read items. In contrast, Luke, one of the youngest children in the class, asked for no help with reading and gained only 3 marks on the written items compared with Katy who requested help with reading 13 items and gained 18 marks. This behaviour may have been replicated in the KS1 mathematics test six months earlier since Katy gained level 2A compared with Luke's level 2C. Children's timings for completing the tests were interesting but, without complementary evidence, provided data of little value. QCA's advice is that children should be able to demonstrate what they can do in about 45 minutes. The time children took to complete the written items ranged from 20 minutes to 58 minutes, with no obvious link to reading competence, or indeed, mathematical competence. I calculated from Table 3.1 that each child took an average of 36 minutes to complete the written questions compared with 39 minutes for below-average readers, a less than anticipated time-difference of three minutes. However, the three children who asked for most help with reading, Katy, Jack and Miles, were among the seven children who took 45 minutes or more to complete the test. I was concerned about these children because I knew through my own teaching experience that Y3 children would generally find it difficult to concentrate for that length of time. At the other extreme, I was also concerned about children like Luke and Bobby who were below-average readers, asked for no help with reading but completed the test in 30 minutes. I return to the time children worked on the written items later in this chapter in my analysis of the tracking sheets.

In spite of the missing data and the high proportion of above-average readers, the underlining of text revealed that four children did not have the confidence to read the whole booklet independently, and, of these, Katy, Jack and Miles requested frequent help. This provided limited evidence that, for some children, the reading demand was likely to be excessive.

Extent to which children asked for help with reading: focus class

Because teachers or observers had underlined text for which all children taking the test had requested help with reading, not only the focus groups, I decided to produce Table 3.2, in a layout similar to Table 3.1 but for a whole year group. This was because I expected to be collecting data from whole Y2 year groups during the live test rather than from small groups. I chose School H because it had more children taking the pre-test than the other focus schools and included children reading from level 1 to level 3. Timings for completing the test were available only for the three children in the focus group so these data are not included.

Table 3.2: Data for Y3 children in School H sorted by booklet number

	May 2000 statutory test level		November 2000 mathematics pre-test marks			
Child's booklet number	Reading	Maths	Oral	Written	No. of items where help requested	Other details
1	2B	2C	5	14	1	
2	2C	2A	5	18	3	
3	2C	2C	5	16	6	
4	2C	2B	4	10	10	SEN stage 1
5	2C	2C	5	7	3	
6	2C	2B	5	10	9	Jack: see Table 3.1. Omitted from Q24
7	2A	2C	6	11	7	Omitted from Q22
8	2A	2C	3	4	5	
9	2A	2C	4	9	10	
10	2A	2C	5	9	6	
11	3	2B	5	15	5	Omitted from Q31
12	3	2A	5	17	1	Single word read
13	2A	3	2	25	0	
14	3	2B	5	18	0	
15	3	3	6	28	0	
16	3	3	6	21	0	Roseanne: see Table 3.1
17	3	3	6	24	0	
18	2A	2A	4	14	11	
19	2B	2B	4	14	0	
20	2C	3	5	27	0	
21	2C	2A	5	21	0	SEN stage 1
22	2C	3	6	27	0	
23	2C	2B	5	13	0	
24	2C	3	5	24	0	
25	2A	2B	5	11	0	Omitted from Q24
26	2B	2C	6	15	0	
27	2B	2C	6	16	2	
28	2B	2C	3	11	0	
29	1	2C	3	7	16	Omitted from Q28
30	2C	2C	3	10	1	SEN stage 1
31	2B	2A	5	15	0	
32	2A	2A	4	18	0	
33	2A	2A	4	24	0	
34	2A	2A	5	21	0	
35	2A	3	6	24	0	Matthew: see Table 3.1
36	2A	3	5	18	0	
37	2C	2A	5	19	13	Omitted from Q29
38	2A	3	6	26	0	
39	2A	3	6	29	0	
40	2A	3	6	29	0	
41	2A	2A	5	25	0	
42	2A	2A	6	26	0	

Key
Oral: number of oral items correct (maximum 6 marks)
Written: number of written items correct (maximum 30 marks)
Shaded rows: children who attained L1 or L2C in the statutory KS1 reading task/test in May 2000 (below-average expectation)

Table 3.2 gave a richer source of data than Table 3.1 because it enabled me to compare results across a year group within one school. I chose not to spend much time analysing the results since my intention was to pilot a way of recording data that would be manageable and relevant in an analysis of the reading demand. However, I did note that the frequency of requests for help with reading for below-average readers ranged from 0-16 showing that some children were given considerably more reading support than others. I wondered if this range would increase with Y2 children who would be six months younger on average than these children in the KS1 mathematics test in May 2001. At that later time, I would also analyse results from average and above-average readers for comparison.

After designing this table, I felt that an electronic spreadsheet would have been a more effective means of managing the data, especially if I were to analyse subsequently data from a greater number of children. For example, I could sort the data in different ways, e.g. ordered by frequency of requests for help with reading, by reading level etc. Each way of manipulating data would provide a different perspective on children's patterns of behaviour during the test compared with their peers. Not only could I see the potential to compare results within schools but across schools, especially if these were from different catchment areas. Consequently, I adapted the data on Table 3.2 and produced the spreadsheet in Appendix 3.2 whose design was to remain largely unchanged when used at a later date to manage data collected during a statutory test. This had the advantage of tracking in a concise way the frequency of requests for help by individuals and in which items.

By now, I was confident that the strategy of underlining text for which children asked help with reading as the teacher read it was effective, and, more importantly for the live test, not distracting for the children. Observers commented that the strategy was manageable and caused no logistical difficulties. I too found this to be so in my own observations.

Comparing the speed of working

In this section, I describe my approach to measuring the time children of different reading competences took to complete the pre-test. There are various factors that affect the speed of working in a written test, some of which have been discussed in CHAPTER 2. However, where a child is reading text in a test of about 30 written items, particularly where little or no reading assistance is given, it seems reasonable to predict that completion time may be protracted by lack of reading proficiency. To see what data might emerge, the observers recorded the item on which the 27 children in the focus groups were working at five-minute intervals using the tracking sheets or ‘chronologs’ (Cohen, Manion & Morrison, 2000:312) provided for them. See Appendix 3.3 for samples of annotated tracking sheets that show children’s progress throughout the test and other information considered relevant by observers. These samples were chosen because they show data for Luke (School A), Katy (School B) and Jack (School H) who feature at various points in the remainder of this chapter. On analysing the sheets, I was surprised to find that children in the focus groups took from about 20 minutes to almost an hour to complete the written items, with no obvious link to reading or, indeed, mathematical competence, a fact I commented on in page 76. On reflection, I should not have been surprised because there was no way observers could tell how time was distributed between reading text and engaging in the mathematics. Another complication was that some below-average readers attained above-average results in the statutory mathematics test, which could imply that their main effort was on the reading rather than the mathematics, or that they had been supported with reading, but the data could not make this distinction. Nevertheless, from the tracking sheets, I obtained data that I had not predicted, for example, patterns of behaviour that children exhibited during the test, as well as some that I had predicted, i.e. clues that suggested difficulties with reading.

In general, the tracking sheets showed that average and above-average readers worked through the test in a systematic way, with about half of them using time after going through the test once to retry items they had omitted or, more commonly, to check work. Some observers noted that these children paced themselves and concentrated well. The notes for the below-average readers, however, suggested some difficulties that might have been overcome by teacher intervention. For example, the observer noted that Luke (School A) 'took one look at some questions and immediately decided not to attempt them …. He finished his first run through questions after 16 minutes and went back to do those he had skimmed over'. Since he gained five out of six marks for the oral items, his difficulties with the written items may have been more to do with reading than mathematics. Indeed, I noted that the three items he got correct were abstract calculations that were not text-dependent. Other factors for Luke's poor performance could not be ignored, e.g. lack of motivation or laziness, but it is likely that he would have gained more marks had more help with reading been requested or offered. Bobby (School F) asked for no help with reading in the pre-test and gained 13 marks out of 30. In the statutory test in May 2000, he had every item read to him because he was a below-average reader and gained level 3 in mathematics, showing that he had gained at least 20 marks in the written items. The drop in marks in the pre-test could be attributed mainly to inappropriate independent reading and lack of teacher intervention. Although it is not known if Osman (School G) requested help with reading, the observer noted that he lost concentration and interest after about 20 minutes and spent most of the remaining time flipping pages to and fro. Katy (School B) and Jack (School H) took longer than other below-average readers to complete the test. Katy asked for help with reading 13 items and the wait for help with reading for each request would have protracted the completion time, however promptly given. Indeed, Katy had just completed the test when the booklets were taken in. However, the annotations on her time-log showed that she worked systematically to the end and gained 18 marks for the written questions, with thirteen of which she had been given help with reading. Jack asked for help with reading nine items but the observer noted that the teacher was busy dealing with other children needing help and could not always help him promptly. His time-log suggested that he spent up to five minutes on some items and was only two-thirds of the way through the test booklet after 50 minutes at which point the test was stopped. Jack gained five out of six marks in the oral questions but only 10 out of 30 marks in the written items suggesting that more effective reading support may have improved his results.

Evaluation of the tracking sheets

As explained at the outset of this chapter, I hoped to measure the time children of different reading competences took to complete the pre-test. In this respect, the tracking was successful but I learned little that was directly relevant to the role that the reading demand played. The time-logs are interesting in their own right, however, since they showed differences in behaviour between children of different reading competences, with above-average readers in particular working more systematically and checking answers. This was not the case with below-average readers like Luke and Osman who showed some evidence of being under stress. Indeed, I was not convinced that all children received adequate reading support but much more evidence is needed to support this claim. As Gibson (1989) explains, fluent readers can revert to being less than fluent when faced with ‘unfamiliar text types’ (p110), in this case a mathematics test, a phenomenon whose significance teachers could underestimate.

The tracking sheets gave me a general insight into children’s progress through a mathematics test and I am unaware of any other study that has undertaken such a data-collection exercise. However, I decided that I would not use this instrument in the live test. Logistically, in the absence of generous staffing, it would be impossible to track all children in a class. Instead, I would focus on how well or otherwise Y2 children could read individual items, taking into account accuracy, fluency and speed of reading and analyse the results in detail, an approach that I justify in my methodology.

Interviews with children in focus groups

The observers interviewed each child in the focus groups individually and made audio-tapes of the interviews as well as underlining text that the child misread or asked help with reading. At the start of the interview, each child was provided with five written items taken from previous mathematics test booklets (see Appendix 3.1). This was a compromise, because it would have been more sensible to select items from the pre-test booklet but these were confidential and could not be reproduced. The items were chosen because they were text-dependent, i.e. children could not guess what to do from diagrams etc., covered different aspects of the programme of study and were graded in difficulty. The observer had to ask the children:

i. to read the item (with help with reading being given only if requested);
ii. to explain in their own words what the item was asking them to do;
iii. to explain how they would work out the answer;
iv. to record their answer.

The observers' protocol was influenced by the study undertaken by Newman (1997) as explained in detail in ***Reading errors in mathematics tests*** in CHAPTER 2 (see pages 29-32).

I listened to the audiotapes with no small amount of disappointment. About half of the interviews were inaudible to some extent because of hissing, the interviewee talking quietly or background noise in the classroom. Poor quality recordings included children who were below-average readers who had asked for no help with reading and I was particularly interested in hearing them read. My least concern was for children reading at level 3, since, apart from Izzy (see last entry on Table 3.3 on page 91), they read the chosen items accurately and fluently. However, Bianca (School D) requested no help with reading during the pre-test but made reading errors in every item during her interview. Her recording was inaudible but see Table 3.3 (page 91) for the vocabulary that she could not read. From recordings that were audible, I fully transcribed interviews with three children, Luke (School A), Jack (School H) and Katy (School B) as part of their profile but also to practise transcribing, interpreting and comparing interview data. These children were chosen because they gained level 2C in reading in the KS1 2000 reading test/task but gained different levels in the KS1 2000 mathematics test, levels 2C, 2B, and 2A. Of the three profiles, I found Jack's the most interesting for reasons I give within my commentary which starts below. In the transcript, numbers in the left-hand column are from the audio-tape counter. The equivalent data for Luke and Katy are available if required. See Appendix 3.4 for Jack's interview items with underlined text and observer's comments.

Jack's data (School H)

Background information:

May 2000 statutory mathematics test level:	level 2B
May 2000 statutory reading test/task level:	level 2C
English as an additional language:	no
Ethnic background:	white
Special educational needs:	0
Date of birth:	3. 10. 92
School:	Inner city primary in disadvantaged area

Pre-test information

Results from pre-test:	oral items:	5 out of 6 marks
	written items:	10 out of 30 marks
Help with reading:	requested for 9 items (last request on item 23 [17th written item])	
Underlined text:	13 words; 1 phrase; 6 sentences	
Other comments:	Omitted items 24 to 35 [18th to 29th written items]	

Comments from observer: *The year group was divided into two ‘ability’ groups with about 20 children in each. The upper group could manage their reading with the aid of one teacher. Children in the lower group (Jack’s group) had one teacher plus one ancillary plus me buzzing in and out, and we could barely keep up with the reading needs with three of us at it.*

Time taken to complete written items: 50 minutes (finishing on item 23)

Jack’s transcript

Item 1

009	**JV**	Let’s have a look at this (interview booklet) and see if you can read the words for me. Read this for me (item 1). Can you read the numbers?
012	**Jack**	Two, six, three, four, five.
013	**JV**	That’s right. Will you read these words here for me now?
014	**Jack**	You can use two cards to make a … number – number …
017	**JV**	Good boy.
017	**Jack**	… less than thirty.
018	**JV**	So here is a number less than 30 being made (indicating 24 on sheet). What number is this?
020	**Jack**	Twenty-four.
020	**JV**	Well done. Right. Come on, read these words for me now.
021	**Jack**	Use two of the cards to make a number between 30 and 40.
024	**JV**	OK. Can you do that? Can you write in here (indicating answer box) what it asks you to do? ……That was quick. Brilliant. Well done. Here’s the next one, OK? Read the words first of all.

Item 2

030	**Jack**	(silence of about 10 seconds)
032	**JV**	Don’t know? Can you say the first sound? What does ‘c’ and ‘h’ say together?
034	**Jack**	‘Ch’… choose …
035	**JV**	Good boy.
036	**Jack**	… a word …from … the box … to … finish … each … sentence. (obvious hesitations between words)
038	**JV**	Good. Well done. Can you read these words (indicating words in box)?
039	**Jack**	(sounds of attempt to read words)
040	**JV**	Do the first bit.
041	**Jack**	(sounds of attempt to sound out ‘kilograms’)
042	**JV**	This says ‘kilograms’. That’s a hard one, isn’t it? What about this one (indicating ‘litres’)?
044	**Jack**	Metres.
045	**JV**	That says ‘litres’. Good guess. Good guess. What about this one (indicating metres)?
046	**Jack**	Metres.
047	**JV**	And this one (indicating ‘hours’)? Quite a hard word that. Can you read that bit (pointing to ‘ours’ by masking the ‘h’)?
049	**Jack**	(sound of attempt to read ‘ours’ read as ‘oors’)
049	**JV**	Good guess. ‘Hours’ that says.
050	**Jack**	Hours.
051	**JV**	Right, read this sentence here for me (item 2).
052	**Jack**	I can … I can (pause of about ten seconds) –
053	**JV**	What sound does it start with?
054	**Jack**	‘M’.
054	**JV**	Measure. I can measure –
055	**Jack**	I can measure the…length…of the ……….. class –

057 **JV** Class?

058 **Jack** - classroom in –

059 **JV** Now, what you have got to do here –

060 **Jack** --metres (stating his answer)

061 **JV** Well done. That's right. Good boy. So which one of these says 'metres'. (Jack indicates.) That one isn't it, because you look at the first sound. That's easy to do. (Jack records answer.) Read this one for me now (item 3).

Item 3

067 **Jack** How is one triangle –

067 **JV** Look at the first word.

068 **Jack** He - r (sounded out separately)

069 **JV** Here –

070 **Jack** Here is one triangle.

071 **JV** Well done.

072 **Jack** Here are four solid triangles.

073 **JV** What does this word say (pointing to small)?

074 **Jack** Small triangles. They make a big triangle – bigger triangle.

075 **JV** OK. Do this bit now (indicating instruction).

076 **Jack** Use nine of the small triangles to make a bigger triangle (read hesitantly).

078 **JV** Do you want to have a go at that? Do you think you can do it? Have a go then.

081 **Jack** (Jack quickly draws an incorrect shape bearing no relationship to the answer.)

085 **JV** OK. Yes. This is the last one now (items 4 and 5). You are doing really well.
See if you can read these words then.

Item 4

087 **Jack** This ………(pause as he thinks about pronunciation of 'graph')

089 **JV** Do the first bit.

090 **Jack** Gr … grap, I think.

090 **JV** Graph. This graph –

091 **Jack** This graph shows the time it …took Tim to ……tell …--

096 **JV** -- travel – to travel --

099 **Jack** --to travel to school in one week.

100 **JV** Right. OK. So this is the graph. It shows how long it took Tim to travel to school in one week. OK. So read this item then (item 4).

Item 4

103 **Jack** How long did it take Tim … on Monday?

105 **JV** OK, then. What does this say here (indicating ‘minutes’ on the graph)?

106 **Jack** Thirty.

107 **JV** What does this say (still pointing to ‘minutes’)?

108 **Jack** Metres.

108 **JV** Not metres.

109 **Jack** Minutes.

109 **JV** OK then. So how long did it take him on Monday then?

110 **Jack** Thirty.

110 **JV** Write thirty in there for me then. OK. And read this now (item 5).

Item 5

112 **Jack** How much …long –

114 **JV** How much – what does this word say (pointing to ‘longer’)? You’re nearly there.

116 **Jack** --longer --

116 **JV** Good boy.

117 **Jack** How much longer did Tim take on Monday than on Friday (all read hesitantly)?
…. There’s Monday. Twenty.

120 **JV** OK. Write it in for me. OK. That’s you – I’ll just check – yes – you’ve done everything.

Reflections on Jack

Jack had just had his eighth birthday when he took the pre-test and was seven months older than Luke. Jack was in the lower ‘ability’ set due to his reading difficulties. Like Luke, Jack gained five out of six marks in the oral items, showing a good response to teacher-read items. Unlike Luke, Jack was willing to ask for reading support and gained 10 marks in the written items compared with three marks gained by Luke. However, Jack last requested help on item 23, the 17th written item and omitted the remaining items because time allocated for the test then ran out due to the start of dinner time. His time-log and comment from the observer, about how much pressure was teaching staff under to manage requests for help, suggest that his progress was slowed by waits for help with reading. For example, if I assume that he worked through the test items in order, which is the indication on his time-log, he took 25 minutes to work from item 16 to 23 and was still on item 23 five minutes later. I can only conclude that although Jack recognised the need to ask for help, this was not always available when he wanted it or given promptly enough for him to progress through the test at a reasonable pace. It is possible that he would have gained more marks later in the pre-test if he had had the opportunity to attempt the items he omitted. After reaching only two-thirds of the way through the written items after 50 minutes, he may have been pleased that no more time was available.

Interview information

Jack’s reading throughout was hesitant and lacked accuracy. He used phonic strategies to word-build on occasions but with intermittent success. He also had to be helped to read several words by the observer, e.g. after reading *minutes* as *metres*. He omitted the endings of comparative words, e.g. reading *big* for *bigger* and *long* for *longer*. Occasionally, he self-corrected words after the interviewer queried what he had read, e.g. correcting *solid* to *small* when describing triangles. Jack lacked the reading skills to access the mathematics required by these items. His reading skills had more in common with those of an ‘early reader’ than a ‘fluent reader’ as defined by Gibson (1989:110).

Words underlined in Jack’s interview items: choose, kilograms, litres, hours, measure, classroom, here, small, graph, travel, minutes, longer

Conclusions

- In the pre-test, Jack correctly recognised that he needed reading support but there may have been items that he thought he was reading correctly where this was not the case. In the interview script, he misread words without sensing that he was doing so. There is no reason to assume that this would not be the case during the pre-test.
- Although reading support was available, Jack appears to have been kept waiting to receive this and did not have time to finish the last third of the pre-test.
- Jack should not have been given the responsibility to know when to ask for reading support; this should have been given automatically and promptly so that he could work through the pre-test at a pace to suit his requirements.
- Most of the words that Jack could not read were ME terms that would be seen rarely in ordinary texts.

In spite of the disappointment of inaudible recordings, I was able to salvage enough material for transcripts as evidence that below-average readers at least struggled to read items of comparable demand and similar style to those in the pre-test booklet. Jack has certainly helped to convince me of this. Even in the absence of recordings, the underlining and comments on the interview items provided supplementary evidence of reading difficulties. A benefit of the annotated items was that I could analyse afterwards which children had reading difficulties and with which items. For example, all seven children who were below-average readers made reading errors as well as a few who were average or above-average readers, as shown in Table 3.3.

Table 3.3: Focus groups: reading errors identified in interview items

Name	May 2000 statutory reading level	No. of requests for help in pre-test	Words misread	No. of items where words misread (max. 5)
Luke	L2C	0	metres, triangles	2
Katy	L2C	13	between, cards, finish, sentence, measure, length, hours, graph, took, travel	5
Kate	L2C	n/k	measure, metres, graph, travel	2
Bianca	L2C	0	these, use, between, word, finish, sentence, kilograms, litres, measure, length, triangle, graph, shows, Tim, travel	5
Bobby	L2C	0	measure, metres, hours, travel	2
Osman	L2C	n/k	measure, travel	2
Jack	L2C	9	choose, kilograms, litres, hours, measure, classroom, here, small, graph, travel, minutes, longer	4
Miles	L2B	7	measure	1
Suzie	L2B	n/k	minutes	1
Izzy	L3	1	measure, hours, graph	1

The column showing misread words was of particular interest since it lists mainly ME words that are phonetically irregular and likely to appear only in ME texts, such as worksheets given to reinforce teaching, when a topic is being taught. In year 3, children were unlikely have seen them regularly enough in written form for them to become sight vocabulary. Here was some evidence that terms included in KS1 mathematics tests could exceed the reading skills of children several months older than those for whom the test is written. Even with limited data, Table 3.3 shows that words that were essential to access the mathematics caused several children reading difficulties. For example, five out of 10 children could not read *graph*; eight out of 10 children could not read *measure*. At least three of the below-average readers, Luke, Bianca and Bobby, asked for no reading assistance during the pre-test but misread key words in the interviews. I could only conclude that such children would have benefited from reading assistance. The difficulties associated with reading ME text were discussed at length in my review of the literature, e.g.

> Mathematical English and OE [sic] are sufficiently dissimilar that they require different skills and knowledge on the part of readers to achieve appropriate levels of reading comprehension.
> (Kane, 1970:579)

These data were a reminder that, during my study, I must take into account the impact that ME text could have on the reading demand of the 2001 test. In the interviews, there were only five items and children made reading errors; in the statutory test, there are about 30 written items covering a much broader range of vocabulary so reading errors may increase proportionally.

Evaluation of children's interview data and transcripts

I considered three aspects in my evaluation. First, I reviewed the interviews. Because of the interaction with the observers, I felt that the voices of the children did not come through as strongly as I had hoped. For future interviews, I would wish to hear children reading items from start to finish with minimal interruption by the interviewer so that children's reading skills were the main focus of the interview. In some instances, I felt that interviewers intervened too quickly to help children read and, consequently, the extent of children's reading difficulties could be understated. Hearing children read parts of items interspersed with interruptions from the observers made it difficult to imagine how that child would have read the item independently in a test. That was what I needed to investigate. Consequently, I needed to rethink how future interviews would be conducted. Fortunately, the underlining of text that children could not read worked as well in the interviews as in the pre-test. All observers felt that it was a simple yet effective means of identifying problem vocabulary that did not seem to bother the children. This gave me confidence that I could adopt a similar procedure with Y2 children in the statutory test in May 2001 but with a modified interview structure. What I was pleased to obtain through the interviews was emerging evidence of reading difficulties through children's voices which gave me confidence that my research questions were worth pursuing.

Second, I considered the audio-recordings. I commented previously on the poor quality of some audio-recordings which seemed to pick up every possible sound in the classroom. Since I wished to make transcripts of children reading items from the KS1 2001 mathematics test, I needed to investigate conditions and equipment that would give good quality recordings.

Third, I reviewed the transcripts. Although I had been warned by colleagues about how long it took to transcribe interviews, it still came as a shock to find out how many hours it took me to transcribe the interviews with Luke, Katy and Jack. Much of the content of these transcripts are the words of interviewers so that what children read did not stand out in the pages of text. I needed to use time more effectively to transcribe in detail what children read, the fluency and accuracy of their reading and to measure the speed of reading in individual items. To do this, I recognised that I would need to design transcripts that would truly represent the voices of the children so that their reading skills were evident for all to see.

Group interview with teachers of children in focus groups

Two of my research questions required me to seek Y2 teachers' views on the reading demand of the test and on how they provided reading support. I took the opportunity to make audio-tapes of some teachers I encountered during this pilot study with a view to transcribing and analysing the data as a learning experience in preparation for interviews with Y2 teachers after the 2001 test. The first of the two interviews was with four teachers, who had taken part in the pre-test and was spontaneous, taking an opportunity that unexpectedly presented itself. See Appendix 3.5 for a copy of the transcript on which I show my first attempts at assigning categories to the data, mindful of advice from Thomas (1998) that ‘classification categories must be defined precisely enough to enable any reasonably-informed judge to place data in their correct locations’ (p184). I also took Thomas’s advice that informing judges by giving them an example or explanation that typifies each code would make it more likely that they would categorise data correctly. This advice gave a focus to my categorisation also. I now explain the codes:

Reading demand: teachers’ views on the reading demand of the test

Administration: how children were organised for the test, e.g. a whole class setting or in groups; also availability of staff etc.

Support: how staff supported children with reading during the test

For the convenience of readers, I colour-coded the most relevant text yellow (reading demand), green (administration) and blue (support).

Interpretation of interview 1

Teacher A expressed concern and surprise that some children ‘were not reading the questions properly’ even when expected to, and ‘were guessing some of the answers’ instead of asking for help with reading’. She also felt that the total amount of text in test booklets was ‘overwhelming’. Like Teacher B, she had observed children using phonics during the pre-test, an indication that words were unfamiliar and therefore ‘difficult to predict’ (Mobley 1986:27), especially if phonetically irregular. During the pre-test, no classroom helpers were present and the 14 children had to request reading support. Agreeing with Teacher C, she felt that one advantage of reading support was that important words could be emphasised. As she explained, ‘It’s like how you read a question, your voice, your emphasis on the words, I think that helps so much’.

Teacher B did not comment specifically on the reading demand other than on children using phonics, an indication in itself that the reading was too difficult. On one child using phonics, she commented, '[h]is problem is language and it is a barrier to his actually completing a piece of maths work unsupported, definitely'. Because of lack of staff, Teacher B had no extra support for Y3 mathematics tests and said, 'There's only me and the thirty children'. As with Teachers A and D, she expected children to request help with reading, but, of the four teachers interviewed, she had the most children to manage.

Teacher C had been the catalyst for the initial discussion about the reading because the others were surprised that she had 'three other invigilating staff in the class' as well as herself to oversee the pre-test. She read each question with the class, a technique called look-listen, and the others circulated to make sure that children were 'starting to read at the correct place'. She believed that children of different reading competences benefited from this approach and, like Teacher A, felt that she could emphasise important words.

Teacher D believed that children's reading had improved since taking the KS1 2000 mathematics test six months previously but that children 'feel far more comfortable in that first part of the test, the oral part, because you are reading, everything is clear to them'. However, she also had noted some children using phonics during the pre-test 'because when I went round the ones that had put their hands up to ask for help, and I put my finger to get them to read it and they were sounding it out … I took over because I thought, *This is going to take ages'*. Like Teacher B, Teacher D agreed that the text acted as a barrier when describing the unwillingness of the 'bright ones' to read text carefully. This teacher did not comment on her mode of administration but mentioned no classroom support. However, her comments showed that children had help with reading on request only.

Some teachers commented on arrangements for administering the Y2 statutory mathematics test in their schools and for reading support. Teacher A, who had a mixed class of Y2 and Y3 children, said that she did 'all my year 2s at the same time but there are only usually 16 or 17 and the help (extra staff) does go to the rest of my class'. The Y2 children in Teacher B's and Teacher D's school were taken out in 'ability' groups and children had to request help with reading. The Y2 class in Teacher C's school were also taken out in 'ability' groups 'so that one particular child isn't slowing down the pace' and had every item read to them.

From the interviews, my main findings were:

i. All four teachers referred to children needing help with reading, thus indicating that, even in year 3, some children had difficulties reading the text in test items designed for Y2 children.

ii. Two teachers who provided reading support on request believed that hearing the text read by them with appropriate emphases on words helped the children. Children reading independently only heard this emphasis when they requested reading support.

iii. In three of the four classes, children had to request reading support. In contrast, children in Teacher C's class heard all items read aloud.

iv. The administrative arrangements for the pre-test varied from children in small groups having all items read to them to children taking the test as a class and having to request help with reading. There was also considerable variation from school to school for the administration of the Y2 mathematics test.

I was surprised at the differences in how teachers administered the pre-test and provided reading support. The arrangements also seemed to vary considerably for the Y2 statutory mathematics test. Whilst the tests were standard, I felt that such variations must advantage some children over others. This was a matter that I needed to investigate further in May 2001.

Evaluation of interview 1 and transcript 1

Although the spontaneous interview provided some relevant data, there were aspects that could have been improved. First, the built-in microphone on the recorder did not pick up clearly the voices of the two teachers furthest from it, a fault that could have been avoided with a trial recording. Not only that, because the teachers' voices were unfamiliar, I found it hard to distinguish between them when transcribing. This could have been overcome by asking them to give their name each time before speaking. Second, my lack of experience in conducting an unstructured interview was obvious, as I learned later from the transcripts. For example, although I tried not to intervene excessively, I sometimes asked leading questions or prompted too much, e.g. *Why did you …? Was it because …?*, a fault that I must try to avoid. Nor was I aware that the discussion focused too much on one topic, i.e. 'look-listen' which, at that time, was not expected to be a major consideration in this study. I also noticed that Teacher C had dominated the discussion whereas Teacher D contributed little.

Even at the transcript stage, I found the data complicated to code and analyse, perhaps because of the unfamiliar voices and the unstructured format of the interview. This led me to design a prompt sheet of for my second interview, the final version of which is filed in Appendix 4.2. In view of Teacher C's use of look-listen, I included questions to seek views about this mode of administration.

Even with the small focus group, two teachers described very different scenarios for the pre-test. I was concerned that children in Teacher C's class had reading support given without request for the whole pre-test whereas children in Teacher B's class had to read independently and ask for reading support. Not only that, staffing arrangements were more generous in Teacher C's class. In the live test, I recognised it would be useful to compare administrative and staffing arrangements and I would need to collect data on these during teacher interviews. The ratio of staff to children could have an impact on how well children were supported with reading and advantage some children over others.

Individual interview with a Y2 teacher

I conducted Interview 2 with an individual teacher since I wished to compare a group with an individual interview to see which I preferred and why (see Appendix 3.6 for the coded transcript). The interviewee was an observer of the pre-test but also a practising Y2 teacher of many years experience.

Interpretation of interview 2

When asked about the reading demand of the pre-test, JV commented that ‘it was quite demanding with only level 3 readers able to cope without help, and this was the opinion of the teachers that I spoke to in the schools where I did the pre-test observations as well as what I think from my own experience’. In her opinion, some questions ‘seemed to have a lot of language or instructions before you could address the mathematics’. When referring to her experience of administering the KS1 mathematics test to Y2 children, she pointed out that when too many children in the ‘lower ability group’ were taking the test at the same time, the teacher ‘is really stretched to get round them all’ because of too many requests for help with reading. She thought that look-listen would probably advantage:

i. poorer readers who need all text reading to them individually and thereby ‘save the teachers rushing around and also the children getting stressed having to wait’;
ii. children who believed they were reading competently but were not;
iii. children who were reluctant to ask for help e.g. not wishing to be identified as poor readers.

She identified three possible disadvantages with look-listen, namely:

i. copying could be a problem since all children would be working on the same item at the same time;
ii. preventing children moving on whilst waiting for the slowest to answer a question;
iii. competent readers might be bored waiting for slower children to catch up.

Most of JV's views on look-listen were raised by teachers in Interview 1 so there was a consensus of opinion. However, the key point arising from this interview is that JV considered the reading in the pre-test to be 'demanding' and only suitable for children reading at level 3. If correct, this means the majority of children in year 2 need reading support as most are reading at level 2.

Evaluation of interview 2 and transcript 2

The transcript was straightforward although I decided to introduce a fourth code for look-listen since this was a particular type of support. In this interview, the extra code was useful since it distinguished between JV's actual experience of administering the test from her views on look-listen, a form of administration that she had never used. This interview solved most of the problems identified with Interview 1, e.g. the interviewee could sit close to the microphone for good audio reproduction and I had more control of the interview because of the prompt-sheet. I decided not to revise this since it provided a good basis for a focused interview. I also felt that with experience, I could use it as a guide but improvise more depending on teachers' comments. The one element that I missed compared with the group interview was the interaction between the interviewees as they shared their different perspectives and experiences. However, on balance, I preferred the individual interview but felt that I could obtain and manage comparable data with up to two interviewees if such a situation arose. There is a possible argument that transcribing one group interview with four teachers would be more time-efficient than transcribing one interview with one teacher. I would counter that by arguing that, in the group interview, the data were time-consuming to extract and analyse whereas the data in the individual interview were easier to manage.

Codes

Defining the three categories for Interview 1 was useful to help keep the categories discrete. There was no difficulty in summarising comments on the reading demand of the test. I was also able to code comments on administration and support separately but, in writing up my findings, it was difficult to separate these because administrative arrangements were often made with a view to supporting children. However, I still believed it useful to make the distinction between the two categories since support is a specific sub-set of administration that could advantage some children over others. Some teachers expressed views on look-listen even though they had not administered the test in this way. Consequently, for the transcript for Interview 2, I added an extra category to distinguish views on look-listen from experience of administering look-listen. There were also a few comments that were relevant but could not be placed in the chosen categories. For any future transcripts, I would include a category of 'Other' to make it possible to include relevant comments that lay outside the chosen categories but which did not justify a category of their own.

Moving towards a methodology

At the outset of this chapter, I stated my intention to trial and evaluate methods to answer my research questions. I summarise briefly my views on the efficacy of these and the decisions I made.

i. My strategy of underlining text in test booklets for collecting data on requests for help with reading was simple but effective. This method of annotation would carry forward to the live test and provide data about which items children were most likely to ask for help with reading and the frequency of individual children's requests. The design of Table 3.2 and Appendix 3.2 give an idea of how such data could be presented.

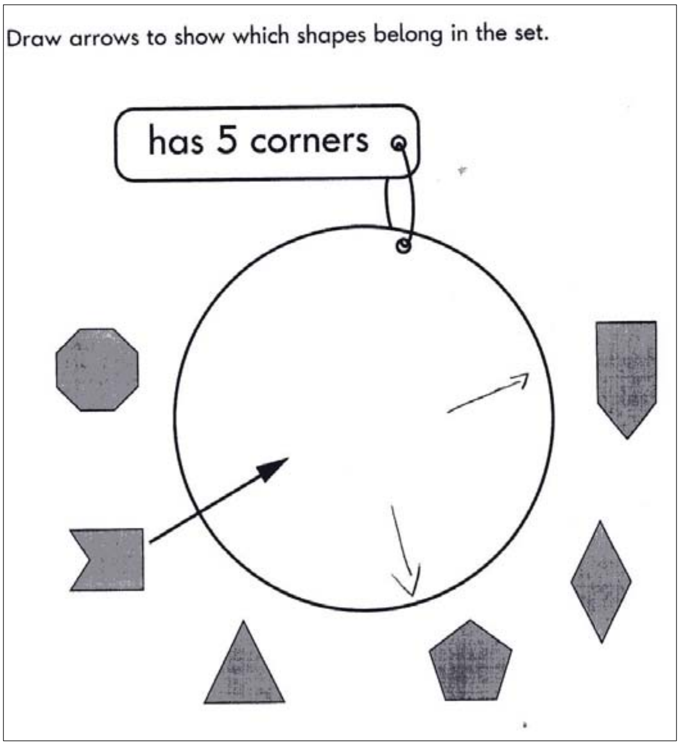
ii. I would discontinue the use of the time-logs to track the times taken by children to progress through the test. These provided interesting data about children's behaviour during the pre-test but did not provide data that would have been useful, i.e. whether children were reading 'at a rate appropriate to the print form' (Gibson, 1989:110). Instead, I would time and compare how long children reading at different reading levels took to read individual items as one means of deciding reading competence.

iii. Because excessive interventions by interviewers made children's reading of items hard to analyse, I would undertake all future interviews with children personally in their schools. I would continue to make audio-recordings so that I could analyse what they read aloud in detail, and decide the best format for transcripts and coding. The transcripts would provide evidence of how accurately or otherwise children were reading. I would also find out how well children reading at different levels could read common items. The voice of the children would be the dominant data.

iv. Having found the recorded individual interview with a teacher much easier to analyse than the focus group interview with teachers, I would continue to interview teachers individually or possibly in pairs for manageability. I would use a prompt-sheet to ensure that I gathered data I needed to compare teachers' views and administrative arrangements. I would code comments in transcripts that complemented data from the children.

In spite of some disappointments, most methods had successful elements on which I would build for the methodology of my study. I felt more confident that I could structure and justify methods to help me find answers to my research questions and make a new contribution to existing knowledge.

CHAPTER 4



Monica was working at level 2B in mathematics.

CHAPTER 4

Methodology

First thoughts on methodology

Myself as a researcher

As I typed the first tentative words for this chapter, it would be less than honest not to admit that I had vague impressions but no clear vision of how it would unfold. I had favourite books on developing qualitative research and was surprised to find that, in some of these, the term ‘methodology’ was absent from the index whereas ‘methods’ was generally included. The relationship between the two terms started to make more sense when I read a chapter by Burton (2002) in which she makes explicit the relationship between methods and methodology and her views on the role of the researcher. I considered the latter first since I identified with Burton’s (2002) view when she stated that she ‘did not believe that there is ever a case where the researcher’s beliefs, attitudes, and values have not influenced a study’ (p3). These personal attributes must surely compel the researcher from the moment he or she chooses to undertake a study of a particular phenomenon until the last word is written. However, it was not until I tried to rationalise what my methodological stance would be that I recognised the need to make explicit the beliefs, attitudes and values that were the catalyst for my study.

Beliefs, attitudes and values

As the sole owner of this study and ‘the key ‘instrument’’ (Wellington, 2000:41), I did not need to be told that I had special responsibilities since, as I explain later in this chapter, I would be interacting with and intruding on the lives of children, mainly 6- and 7-year olds, whose voices I wished to represent. There has been a trend in recent years towards studies that listen to pupils' voices. For example, Crozier & Tracey (2000) believed that ''there has been increasing and overdue interest in hearing the voices of young people themselves about their educational experiences' (p174) and that 'pupils' perspectives are of particular value' (p174). I was happy to contribute to this trend for two main reasons:

i. I believed that it would only be by involving the children themselves that I would be able to tell the 'story' of how well or otherwise they could read the text in a KS1 mathematics test;
ii. I found few large-scale studies that had KS1 statutory testing as their focus.

Because I undertook this study on the 'hunch' that the reading could be too demanding for some Y2 children, I had to accept this bias but also to be willing to proceed honestly and critically, even if my 'hunch' eventually proved groundless. Not only that, there was limited but convincing evidence from my pilot study that led me to conclude that reading difficulties would be likely. Wolcott (1995) did not consider a researcher's acknowledged bias a threat to integrity and I empathised with the following belief:

> I regard bias as entry-level theorising, a thought-about position from which the researcher as inquirer feels drawn to an issue or problem and seeks to construct a firmer basis in both knowledge and understanding.
> (Wolcott, 1995:185)

I felt the need to pursue my 'hunch' or 'entry-level theorising' (Wolcott, 1995:185) in the belief that the KS1 mathematics test 'should assess mathematical, not linguistic skills and abilities' (Fisher-Hoch & Hughes, 1996:4) i.e. assess what it is claimed to be assessed with no threat to validity, for example by undetected or excessive reading difficulties. My insider knowledge of test development also affected my role as a researcher. During the development of a test, a rigorous procedure as outlined in CHAPTER 1, professionals are involved in evaluating the test items and have a chance to comment on any aspect of the items including wording and readability. However, I was convinced that:

i. the best judges of the readability of the items would be the children for whom the test was intended and their direct input in this matter had not been built into the formal reviewing procedures;
ii. Y2 children could provide me with the most convincing data to establish if assumptions were being made about their reading competence.

These convictions led me to consider my ethical stance, not only related to my proposed involvement with children, who were to become the leading participants in my data collection, but as a wider concern.

Ethical stance

Through reading pertinent literature (particularly Bassey, 1995; BERA, 2000; Cohen, Manion & Morrison, 2000; Lewis & Lindsay, 2000; Clough & Nutbrown, 2002 and Robson, 2002), I was mindful of the ethical issues that confront any researcher. Bassey (1995) synthesised the BERA (2000) guidelines that I intended to use as a checklist to underpin my ethical stance, to three main issues:

i. respect for persons;
ii. respect for truth;
iii. respect for democratic values, i.e. my personal freedom to undertake research provided that I was respectful of persons and the truth.

Later, at appropriate times, I draw attention to the ethical considerations and, in some cases, dilemmas for each method of enquiry I chose. I also had to reflect carefully on my long-standing relationship of nearly 10 years as a mathematics consultant with QCA. Much of my role in test development involved access to confidential material. I determined not to abuse the position of trust or professional privilege arising out of my consultancy. At the same time, I wished to use the professional experience I had gained in test development that gave me the confidence to undertake this study. I was only too aware that I would be immersed in the politics of educational research. Consider the following statement:

> Research that changes nothing – not even the researcher - is not research at all. And since all social research takes place in policy contexts of one form or another research itself must therefore be seen as inevitably political.
> (Clough & Nutbrown, 2002:12)

Since its outset, statutory testing at the end of key stage 1 in particular has been politically sensitive yet there appear to be few large-scale studies that have that focus. My review of the literature revealed only one unpublished study (Lupetti, Sainsbury & Schagen, 1995) on the reading demand of the KS1 statutory mathematics tests and, I believed, further research was urgently needed. Mindful of BERA's (2000) guidance that 'honesty and openness should characterize the relationship between researchers, participants and institutional representatives' (p2), I wished the research to be overt and consensual, with teaching staff, children and parents/guardians aware of its purpose, how I intended to pursue it and why. I also made the decision to inform QCA of my field of study and to share relevant data with them. I had no wish to take an adversarial stance, but I also recognised that a possible ethical dilemma could arise if my findings were unwelcome or embarrassing to QCA. By adopting a policy of openness, I was of the opinion that I was not abusing the trust of my colleagues and would be in a better position to influence policy in some way if I subsequently concluded that children's reading difficulties were a threat to test validity. I lived in the idealistic hope that if my 'hunch' were accurate, I could make a difference by representing the largely unheard 'voice' of children.

The decision to place young children at the heart of my research and to use my voice to tell their story, of course, would bring with it particular ethical responsibilities. Data that I collected from young children, particularly where no other observer was present, could be misrepresented by me to support my 'hunch' since 6- and 7-year olds would not have the maturity to evaluate my presentation or analysis of their contribution. Potentially, I could 'exploit the powerless' (Cohen, Manion & Morrison, 2000:121). Consequently, it was my ethical responsibility to show respect for the participating children by reporting truthfully on their reading competence in a politically-sensitive test. I also had to be able to convince any reader of the integrity of my study and this consideration also shaped my methodological stance.

Moving towards a personal methodology

I now return to my unravelling of the distinction between 'methods' and 'methodology'. Hart defined methodology as:

> [a] system of methods and rules to facilitate the collection and analysis of data. It provides the starting point for choosing an approach made up of theories, ideas, concepts and definitions of the topic; therefore the basis of a critical activity consisting of making choices about the nature and character of the social worlds (assumptions). (Hart, 1998:28)

Burton (2002) expressed concern that all too often the emphasis in studies has been on explaining *how* research was undertaken but not *why*, namely 'what influenced the researcher to choose to do the research in the manner described' (p1). The importance of the *why* was also stressed by Wellington (2000) and Clough & Nutbrown (2002) who argued that 'one of the tasks for a methodology is to explain and justify the particular methods used in a given study' (p27). I recognised that it would only be by developing and rationalising my own methodological stance that I could subsequently and genuinely make the claim to a unique contribution to knowledge.

Having read about the key features of different types of research, e.g. descriptive and predictive, I concluded that evaluative research best described the approach that I wished to pursue. Evaluative research has become increasingly popular in recent years (Verma & Mallick 1999; Cohen, Manion & Morrison, 2000) because of its emphasis on making 'value judgments' (Burgess, 1995:6) about how effective are established policies or their implementation. I identified with the following definition for its relevance to my study:

> The term ‘evaluative research’ is often used to refer to the systematic methods used in the collection and analysis of data regarding the effectiveness of an educational experience. (Verma & Mallick, 1999:46)

Similarly, Robson (2002) stated that ‘most evaluations are concerned with the effectiveness and appropriateness of a … programme in a specific setting’ (p205). The government policy of requiring children to sit a statutory written mathematics test, a specific setting, at the end of key stage 1 and for their teachers to implement it, led me indirectly to undertake this study. However, the children who take the test were the more direct and driving influence because I wished to ‘capture, interpret and explain’ (Robinson, 1993:88) how Y2 children tried to make sense of the text in the written items of a statutory mathematics test and to compare how teachers supported their reading. Whatever the greater influence, I recognised, as stated in the previous section, that my research would be regarded as ‘inevitably political’ (Clough & Nutbrown, 2002:12); evaluative research rarely stands free of political underpinning. Such research is commonly funded by sponsors or self-funded by employers but I did not wish to receive funding because it could ‘enable others to set the research agenda’ (Cohen, Manion & Morrison, 2000:38). I initiated the idea for this study and, subsequent to the pilot study, was an independent and sole researcher, albeit with the knowledge and consent of QCA. My professional independence was essential so that I would have ownership of data that I collected, freedom to analyse them and to report on my findings. I wished to be a spokesperson for the children rather than for a sponsor.

I had concerns that some of the typical features of evaluative research were not such a ‘good fit’ for my intended approach. For example, Cohen, Manion & Morrison (2000) suggested that such research was not only less likely to be exploratory and but also less likely to contribute ‘something original to the substantive field’ or to extend ‘the frontiers of knowledge’ (p38). However, my intention was to explore the context in which a statutory test took place and its effect on its target audience. In so doing, I expected to make an original contribution to knowledge and to extend the boundaries of what is known about children's behaviour and attitudes in the context of a statutory test. However, some literature does not make such a distinction, and Cohen, Manion & Morrison (2000) also acknowledged that the boundaries between research and evaluation are not always clearly defined. For example, Robson (2002) suggests that:

> [E]valuation research is essentially indistinguishable from other research in terms of design, data collection techniques and methods of analysis.
> (Robson, 2002:204).

Fortunately, evaluative research offers the same range of methodological choices as other research disciplines. My interpretation of the literature led me to conclude that what makes this type of research different is that judgments have to be made about the 'effectiveness and appropriateness' (Robson, 2002:205) of the phenomenon under investigation. The purpose of the written test is to 'provide a standard "snap-shot" of attainment at the end of the key stage' (QCA, 1999:8). I wished to make judgments about whether children's attainment was likely to be compromised by the reading requirements; any unrecognised reading difficulties would affect the validity of the test, and consequently, its worth. As someone who wished to represent the voice of Y2 children, this described exactly my intentions.

In my review of the literature, I read much that provided background information on and improved my knowledge of issues related to readability. However, I read little that gave detailed or practical suggestions as to how I might answer my research questions. As a result, all my methods were not chosen at the outset and I had to develop aspects of my study as I went along. Writers on research methods (see, for example, Marshal & Rossman, 1995; Robson, 2002; and Charmaz, 2006) stress the importance of having a flexible research design; I had to be prepared to work flexibly to allow for unexpected phenomena or unsatisfactory methodological decisions.

Methods chosen to answer research questions

To answer my research questions, as enunciated in my literature review, I recognised that interviewing and observation would be the most effective to suit the purposes of my study, complemented by documentary analysis as I now explain in Table 4.1 through consideration of my research questions.

Table 4.1: Methods chosen to answer research questions

Research question	Method
i. Is the reading demand of the text in a typical KS1 mathematics test appropriate for the target children in year 2?	Documentary annotation Diagnostic interviews with Y2 children Documentary analysis
ii. What is the extent of the help needed by children of different reading competences during the test?	Documentary annotation
iii. What are the views of year 2 teachers on the reading demand of the test?	Interviews with year 2 teachers
iv. Is teacher support with reading available to children during the test whether they ask for it or not?	Diagnostic interviews with Y2 children Documentary annotation Interviews with Y2 teachers
v. What is the reading age required by a Y2 child in order to read the text in a typical KS1 mathematics test fluently and accurately?	Comparative diagnostic interviews with Y2, Y3 and Y4 children

I provide details of how data acquired from each method were treated in the relevant chapters. I include the broad details of my methodological choices in the following sections; the finer details are included in the relevant chapters where they are better positioned to set the scene for the reader.

Although not planned at the outset of this study, the data collection and analysis fell into four chronological phases as shown in Table 4.2.

Table 4.2: Four phases of study

Timescale	Timescale	Main focus of activity
Phase 1	November 2000 – April 2001	KS1 mathematics pre-test (pilot study): collect data from Y3 children during pre-test; analyse data
Phase 2	June 2001 – May 2002	KS1 2001 mathematics test: collect data from Y2 children and Y2 teachers during and after test; analyse data
Phase 3	June 2002 – March 2003	KS1 2001 mathematics test: collect data from Y2, Y3 and Y4 children; analyse data
Phase 4	August 2005 – March 2006	KS1 2001 mathematics test: analyse booklet for readability

This information is given here since references are made to phases as I explain my methodological stance. The detailed timetable for data management during these phases is given in Table 4.3.

Research question 1

Is the reading demand of the text in a typical KS1 mathematics test appropriate for the target children in year 2?

As I explained in CHAPTER 2, I could find no published literature that had investigated the reading demand of a typical test. Not only did I wish to carry out such an investigation but I wished to collect data during a statutory test, i.e. capture how children responded to reading text in a test that was part of their cultural practice.

Documentary annotation

In Phase 1 of my research, my pilot study carried out in November 2000 (see CHAPTER 3), I explained how underlining text for which children requested help with reading, at the time of the request and as the text was read aloud, worked successfully during a pre-test and whilst children read test items to interviewers. To generate data to launch Phase 2 of my study, I decided that I would repeat this procedure, a method that I have described as documentary annotation in Table 4.1, in the KS1 2001 mathematics test six months later. This method was chosen since it had not disturbed the children and was the least obtrusive measure of which I could think to collect the data I sought; I wished to 'affect the setting as little as possible' (Gitlin, Siegel & Boru, 1993:196) during a statutory test. My concern that the reading might be too difficult for some children was of a general nature but I chose the KS1 2001 mathematics test for particular analysis since it was the first to follow my pilot study. However, I considered that the 2001 test was typical and that, from any emergent data, I would be able to make some generalisations about the reading demand of mathematics tests in other years. Since the test booklet (QCA, 2001c) is referred to frequently in this and subsequent chapters, I include a reference copy for readers in Appendix 4.1.

The underlining technique would have the advantage that other staff administering the test could annotate on my behalf when I was busy with another child or when I could not be present at a particular administration since all data had to be collected during May 2001, the time-frame for administering the test. However, the main advantage of this approach was that I would be able to:

i. analyse and compare the annotated booklets after the administration;
ii. experiment with the best format to organise the data for analysis;
iii. gain an overview of how much or little help individual children requested, relative to their aptitude in reading;
iv. identify items where requests for help were most frequent so that they could be a particular focus for analysis;
v. compare results within and between schools.

I applied to QCA for permission to annotate booklets in this way since a statutory test would be involved and am grateful that this was given. Having obtained this, I found six headteachers in different catchment areas across three LEAs who would be willing to participate in the underlining procedure with all Y2 children taking the test in their schools and to allow me to interview selected children in subsequent weeks. I believed it would be important to have access to children in different catchment areas, including children who had English as an additional language. I would need to convince readers that, whilst I could never generalise about the national cohort of about 600 000 Y2 children, these children were likely to be typical of their peers nationally. I believed that collecting data from at least 150 booklets would provide sufficient data; many more than that might make data analysis unmanageable. In the event, the six schools provided 164 booklets with annotations for analysis. Because I was asking questions of the *How many …?* type, I would obtain quantitative data. However, using quantitative data to be a catalyst for and complement qualitative data is not uncommon. I agreed with Wellington (2000) that:

> Background statistics … can set the scene for an in-depth qualitative study. When it comes to data collection, most methods in educational research will yield both qualitative and quantitative data. (Wellington, 2000:17)

Indeed, qualitative analysis of background statistics was to feature throughout this study because I considered that electronic tables and spreadsheets would be the most sensible means of organising and interrogating large amounts of data in preparation for qualitative analysis. The interrogation of these data would provide direction for my subsequent methodological approach. For example, after the analysis, I would be in a better position to choose which children I would like to interview and why, as I explain later in this chapter.

Documentary analysis

The documentary analysis of a test booklet for readability, Phase 4 of my study almost five years later, also resulted in the collection of quantitative data. In my review of the literature, I concluded that there were no appropriate measures of readability to apply to the test. QCA did seek the views of teachers and other professionals on readability of the text, as I explained in CHAPTER 1, but no formal attempts were made to predict readability. I felt that I should try to do so because of its importance to and impact on the children.

This phase was to entail an in-depth rather than a superficial analysis of the readability of the test booklet as had been my original intention, since, by that time, children participating in my study had already provided me with convincing evidence that the reading demand was excessive. However, out of interest and to see if it would be possible, I wished to obtain further evidence as to *why* this was so. Wellington (2000) stated that:

> …documents can be of value at different stages of research and can be 'brought in' to the research process for different purposes: to open up and explore a field; to complement other research approaches and methods; and to conclude or consolidate research …. (Wellington, 2000:114)

A detailed analysis would indeed complement other methods I had used and consolidate research to date. From my review of the literature, I concluded that, in the absence of a suitable readability formula, words likely to be familiar in written form, considered alongside sentence length, would be the best predictors of readability of the test. Unfortunately, in the early years of my study, I could not find enough recent and relevant lists that identified words likely to be familiar or unfamiliar to Y2 children. Without these, a detailed analysis would not have been possible but I changed my mind when appropriate word lists were published during my study (see Stuart et al, 2003), as discussed in my review of the literature. I then made the decision that, however tedious and time-consuming it would be, I would code each word in the booklet into one of three categories according to how familiar they were likely to be to Y2 children as sight vocabulary, i.e. read accurately without pausing to decode. Taken in consideration with sentence length and other relevant features, this would give me the basis for deciding how readable each item was likely to be. I could then establish which words, and what percentage of these, were likely to be the least familiar, and by implication, most difficult to read, for an item-by-item and an overall analysis of the test booklet. If my data were to provide evidence that such words existed, they would enable me to make a stronger argument that it is not reasonable to expect children to read such vocabulary without assistance in a high stakes test.

Interviews with Y2 children

I believed that my analysis of the booklet would be an apt conclusion to the research phase of my study since it would contribute to answering research question 1 and be used to triangulate data collected earlier in the study. Indeed, the test booklet was to underpin the study but I wished to involve its intended audience directly. Who better than Y2 children who had taken the test could demonstrate if the reading demand was appropriate? Even though I had already obtained evidence from the underlining technique that children needed considerable help with reading, proof in itself that the reading was too difficult, hearing children read test items would help me discover *why* they had needed help. Underlining indicated which text children found difficult but not the nature of the difficulties; only children could provide those data. Consequently, I decided that I would interview Y2 children in the months following the test.

Having learned from the pilot study that my interview structure did not capture children's reading skills precisely enough, I decided that I would conduct interviews personally for consistency and minimise spoken interventions from me so that the focus would be on the child and not the interviewer. I explained in CHAPTER 2 that Lupetti, Sainsbury & Schagen (1995) and Newman (1977) developed observation sheets that included codes for recording children's reading difficulties as they worked through mathematics test items. I decided that I would not code or take notes during the interviews. This decision was taken partly because I lacked the confidence to code spontaneously but also because I could not predict in advance the codes that might be needed. From the children's perspective, I agreed with Costley (2000) that '[i]t would have been difficult, and possibly intrusive or disruptive' (p171) to do so. Instead, I chose to record interviews on audio-tape. In so doing, I would be able to decide afterwards the best way of presenting data in transcript form. In this respect, I used the flexibility of a grounded theory approach, for as Charmaz (2006) explains, '[t]he logic of grounded theory coding differs from quantitative logic that applies *preconceived* categories or codes to the data' (p46). Added benefits of audio-recordings over coding sheets were that:

i. I could provide more substantive and revealing evidence. For example, rather than record that a child had misread a word, I chose to write verbatim what the child said in transcripts that left little to the reader's imagination.
ii. I could time and compare the speeds at which children of different reading competences read items, an approach that could be intrusive during interviews.
iii. I could provide proof through audio-recordings that the transcripts were a true record of what children read and that I had not succumbed to 'observer bias' (Fraenkel & Wallen, 2006:452) in my interpretation of them.

Initially, I used the term 'interviewing' to describe my interaction with children with some reluctance because a more enlightened term eluded me and because I did not wish to seek their opinions, judgments, attitudes or beliefs as is often the case in interviews. Nor did I wish to engage in a chatty 'two-way conversation' (Cohen, Manion & Morrison, 2000:269) since the conversational style used by interviewers in my pilot study only served to mask children's reading difficulties. Mindful of Newman (1977) and Lupetti, Sainsbury & Schagen (1995), as discussed earlier in this chapter, I identified one feature that their studies shared, i.e. their interviews with children were diagnostic. Their approach corresponded with one purpose of interviewing identified by Cohen, Manion & Morrison (2000) which was 'to evaluate or assess a person in some respect' (p268). This described my intention well; i.e. to conduct interviews that were diagnostic to assess reading competence. To arrive at a judgment about readability, Shuard & Rothery (1984) believed that 'it is necessary to look closely at the styles of writing used in ME, and the ways in which children respond to what they read' (p2). The former judgment would be made from an analysis of the booklet in Phase 4 of my study; the latter, and more important judgment from the children.

I recognised that my interview schedule would have to be structured but adaptable for each interviewee. This was because I wished to find out if children could read fluently text that they had read independently during the test. Consequently, I would need to identify, in advance of any interviews, items in a child’s test booklet where text had underlining. I would then ask each child chosen for interview to read aloud several items where only part or no text for particular items was underlined. I would have to keep an open mind about how many interviews I would transcribe and in how much detail. My guiding principle was that I would transcribe as much of each interview and as many interviews as needed until some trend or pattern in the responses emerged, at which point I would decide how to organise and analyse them. I knew that I would need to develop a coding system for the transcripts to illuminate any reading difficulties but I wished to defer a decision about what this would be until after the interviews, i.e. the data would determine the codes. As Charmaz (2006) stated:

> Coding is the pivotal link between collecting data and developing an emergent theory to explain these data. Through coding, you *define* what is happening in the data and begin to grapple with what it means. (Charmaz, 2006:46)

My original intention had been to ask children with different competences in reading to read the chosen items to me and to explain what the items were asking them to do because I felt that the explanation would reveal the level of comprehension of the text. However, as I explain in later chapters, Y2 children could rarely give an explanation even if they read fluently or had given the correct response. This made the quality of my transcripts all the more important because I would have to restrict my judgment about how well children were likely to understand the text to the analysis and interpretation of transcripts only.

An added benefit of transcribing items would be the opportunity to time and compare the speed at which different children read items as close to the time of the live-test as possible, when their reading skills would have little time to improve, and to analyse why differences existed. My review of the literature revealed that reading at ‘optimum speed’ (Gilliland, 1972:13) or ‘at a rate appropriate to the print form’ (Gibson, 1989:110) were indicators that might give clues as to how well readers coped with text. In my pilot study, I collected data on the time taken by children to complete the written items in a similar mathematics test. I also explained at the end of that chapter that I abandoned this approach because I could not make the distinction between time spent on reading and mathematics. However, through timing children as they read individual questions, I could focus exclusively on the speed of reading.

To some extent, I addressed research questions 2 and 4 in my commentary on this question, *Is the reading demand of the text in a typical KS1 mathematics test appropriate for the target children in year 2?* since they are inter-related. However, since the questions require different answers, I summarise briefly in the appropriate sections the shift in emphasis in my methodology.

Research question 2

What is the extent of the help needed by children of different reading competences during the test?

I explained in my commentary on research question 1 that I decided to enter on an electronic database each occurrence of when a child requested reading support for an item, indicated by underlining of a word at least. From these data, I would be able to calculate the mean number of requests for help and the ranges relative to reading level and for the overall sample of children. Answering this question would be less about helping me to choose suitable children for interview as in the previous question but about analysing and interpreting the data quantitatively then qualitatively. My intention here would be to gain an overview of children’s responses to reading the text and to look for trends or results that concerned me, e.g. a high proportion of children requesting reading support for particular items or children who asked for little support relative to their reading level. Items where requests for help were most frequent would then become a particular focus for my analysis of the readability of the test booklet as I explain in CHAPTER 8.

Research question 3

What are the views of Y2 teachers on the reading demand of the test?

Since I wished to obtain answers to questions relating to the reading demand of the test as well as how teachers supported children with reading (see research question 4), I decided that I would prepare a ‘pre-structured’ (Robson, 2002:274) list of open questions that would allow teachers to express opinions and share their knowledge and experience with me. A copy of the questions in the preferred sequence, but flexible enough to be changed depending on responses from interviewees, is included in Appendix 4.2.

Originally, my intention had been to interview only the seven teachers who had given me access to their Y2 classes for the underlining procedure. I would then transcribe the interviews and code comments that contributed to answering my research questions as I had done in my pilot study. The coding had worked well then and I saw no reason to change it. However, I subsequently reviewed my research design and decided that I would also interview teachers who used the mode of administration commonly called look-listen where the teacher intentionally reads each item in turn with the children. At the time of contacting schools to take part in my research, I had originally contacted 12 headteachers. Of these, six agreed to take part in the underlining procedure; the other six said that the procedure was inappropriate because the Y2 teachers in their schools used look-listen. Indeed, it turned out that more than half of the 17 teachers whom I contacted to provide data from the test had changed to the look-listen mode of administration. I believed that the latter teachers’ perspective on the reading demand would also contribute to answering this research question and arranged to interview them. For ease of comparison, I would code transcripts of their interview data in the same way as those taking part in the underlining procedure.

Research question 4

Is teacher support with reading available to children during the test whether they ask for it or not?

As explained earlier, the underlining procedure would enable me to identify whether a child had or had not asked for support with reading a particular item during the live test. By comparing the frequency of requests with children's reading level, I expected to find that children would fall into two categories in particular that would make them relevant candidates for interview:

i. those who asked for little or no help compared to others assessed to be reading at a comparable reading level or above;
ii. below-average readers who did not ask for support with reading particular items compared with the frequency of requests by others in the overall sample.

If I could find evidence that children could not read items for which they had not requested reading support, I would then be able to say that children had not been given the degree of support with reading that they needed and was their entitlement. I needed to find out if children would be likely to lose marks because of undetected reading difficulties since children were awarded an end of key stage level in mathematics according to test scores.

I would also ask the seven teachers who took part in the underlining procedure how they supported children with reading. I could then compare strategies that they used. Because teachers would be in the best position to know the strengths and weaknesses of children's reading skills, especially towards the end of the academic year, I wished to be reassured that their administrative arrangements included appropriate support so that every child had equal access to the text.

The higher than expected incidence of look-listen also led me to ask teachers using this approach how they supported children during the test and which children received this support. Seeking this information was important because it introduced yet another variation in delivery of the test that I sensed would benefit less competent readers and probably provide a better option than children having to request help. This mode of administration was also raised spontaneously in the interview with Y3 teachers in the pilot study so is likely to be uncommon (see Appendix 3.5). Responses from teachers using look-listen would also contribute to answering this research question about whether help with reading was available even when not requested and for whom. Indeed, having been made aware of the prevalence of the look-listen mode of administration, I decided that I would extend my list of questions to seek views on look-listen from some teachers that were involved with the underlining of text because I was interested in their views on this mode of testing.

Research question 5

What is the reading age required by a Y2 child in order to read the text in a typical KS1 mathematics test fluently and accurately?

I had hoped that I would answer this question through interpreting interviews with Y2 children after they took the test in 2001 but, in the event, I lacked the evidence that would suggest a reading age. This was because my interviews had focused on children reading across the range of level 2 and they demonstrated reading difficulties. I needed more data on children reading at level 3. Once again, I had to review my research design and incorporate this ‘add-on’ which was to become Phase 3 of my study.

The obvious choice would have been to interview more Y2 children who were reading at level 3. Unfortunately, from the personal data I had available, there was no way of telling whether Y2 children awarded level 3 in the statutory reading arrangements had just moved over the border from level 2 to level 3 or were reading securely within level 3. However, I knew that:

i. most average Y3 children were likely to be reading at the bottom end of level 3 at the end of year 3; and
ii. most average Y4 children were likely to be reading securely within level 3 at the end of year 4.

Developing from my original plan, I therefore decided that I would use the same procedures adopted in 2001 to interview children in years 2, 3 and 4 who had been identified as average readers by their teachers. I followed the same ethical guidelines and procedures as for the interviews in 2001. I chose June 2002 to carry out my plan since Y2 children had just taken the KS1 2002 mathematics test, and were, therefore, of the same age and at the same point in the key stage as the 2001 cohort one year before. This meant that I would then be able to compare results across the three year groups for similarities and differences in reading competence. A bonus was that I would also be able to compare Y2 children in 2002 with those in 2001 to convince myself and any reader that the latter were not atypical. I felt that I would have to hear older children read to establish at what age children could read the text fluently and accurately, and by implication, independently. For example, if average readers in year 3 could read test items with no apparent difficulty, their reading skills would equate broadly to those of above-average readers in year 2, reading at the threshold of level 3. I could then argue that year 2 children needed to read as well as an average 8-year old at the time of the test if reading was unmediated.

At this point, I had to decide which items I would ask children to read. I chose four common items from the 2001 test for each child. These were chosen because many Y2 children at that time could not read them independently. For manageability, I restricted my interviews to 24 children, i.e. six children, two in each year group, in four schools in different catchment areas. I made the decision that I would:

i. transcribe all interview data in full so that the extent of any reading difficulties would be illuminated; and

ii. code in more detail than in the transcripts of children who took the test in 2001 so that I could compare the frequency of misreads, use of phonics, hesitations, speed of reading etc.

I felt that this amount of detail would be necessary to make an informed judgment. Only by an in-depth comparison of children across different year groups, reading the same test items, could I hope to find an answer.

Timetable for data management

Having attempted to justify the approach and methods used to find answers to my research questions, I summarise in Table 4.3 how I transformed my methodological aspirations into reality, i.e. 'the practicalities of the research' (Cohen, Manion & Morrison, 2000:75) or 'logistics' (Hart, 1998:49).

Table 4.3: Timetable for data management

Timescale	Activity	Treatment
Phase 1 November 2000 – April 2001	Pilot study: i. Documentary annotations made in 27 pre-test booklets (three children each at nine schools) to determine amount of reading support given. ii. Track speed at which the 27 children work through pre-test using observation sheet. iii. Conduct interviews with 27 Y3 children whose booklets were annotated. iv. Conduct interviews with four pre-test teachers and one observer. Number of visits: Nine QCA observers spent one half-day at a selected pre-test school	i. Documentary annotations: quantitative analysis of booklets leading to qualitative analysis of amount of help with reading requested. ii. Compare children's progress through pre-test taking into account reading level. iii. Transcribe in full three interviews of children reading at levels 2C, 2B and 2A; quantitative analysis to compare reading skills. iv. Transcribe in full group interview with four pre-test teachers; transcribe in full interview with pre-test observer; qualitative analysis of views on reading demand of pre-test.
Phase 2 June 2001 – May 2002	KS1 2001 mathematics test: i. Documentary annotations in 164 test booklets to determine amount of reading support given. ii. Conduct interviews with 30 Y2 children who took test. iii. Conduct interviews with 17 teachers who had administered test. Number of visits: 12 personal visits of about one half-day to the six schools involved in data collection during live test to interview children and teachers. Seven personal visits of about one hour each to interview teachers using look-listen.	i. Documentary annotations: quantitative analysis of booklets leading to qualitative analysis of amount of help with reading requested and items where help was most frequently requested; ii. Transcribe in full three interviews of children reading at levels 2C, 2B and 2A; partial transcriptions of seven children reading items for which they did not request help with reading; code transcripts to identify reading weaknesses; quantitative analysis to compare reading skills. iii. Transcribe in full interviews with five teachers who had assisted with documentary annotations; transcribe in full interviews with 10 teachers who read all questions to children; code transcripts to undertake comparative qualitative analysis of views on reading demand of test and on how children were supported with reading.

Table 4.3: Timetable for data management (continued)

Phase 3 June 2002 – March 2003	Years 2, 3 and 4 comparative study: i. Interview eight average readers in each year group near time of statutory testing. Number of visits: One personal visit of about one-half day to each of the four schools to interview children.	i. Transcribe in full interviews with 24 children at four schools; children to read same four items found difficult to read in 2001 test. ii. Code reading errors in detail to identify use of phonics, hesitations etc. iii. Compare most and least competent readers for each item within each year group and across year groups.
Phase 4 August 2005 – March 2006	Documentary analysis: i. Analysis of test booklet for readability for Y2 children	i. Code each word in test booklet according to how likely it was to be familiar to Y2 children and by sentence length. ii. Conduct quantitative analysis of data leading to qualitative interpretation. iii. Make detailed analysis of 8 items found easiest and hardest to read. iv. Make semi-detailed analysis of next six items found hardest to read. v. Comparison with data from interviews with Y2 children.

The starting point for Phases 1 to 3 was pre-determined by the fixed dates in which the tests could be administered. Because so much data were generated, I ensured that there would be gaps of several months between each phase to allow me time to organise and interpret data from the field before the next phase. The intervals between phases would also give me time to reflect on the efficacy of my methods and on data that would be useful to collect subsequently, to find answers to outstanding or partially-addressed research questions. At first glance, the chronological progression in Table 4.3 appears linear. In reality, the evolution of the study was cyclical in the manner described by Bassey (1995), i.e. my focus was redefined after each stage of data collection and analysis and, in some instances, required 'a new round of investigations carried out in an attempt to get closer to whatever truth is being sought' (p59). Similarly, Charmaz (2006) said, '[s]ome of our best ideas may occur to us late in the process and may lure us back to the field to gain a deeper view' (p10). For example, in Phase 2, the interviews with teachers using look-listen to administer the test were included because their approach was in stark contrast to teachers who expected children to request help with reading. Phase 3 was also unanticipated in the development of my original plan but designed to investigate an unanswered research question. Phase 4, for reasons explained earlier, became an in-depth rather than a cursory analysis of the readability of the text. This was because of happenstance, i.e. data from a new study made it possible for me to devise a coding system to predict the readability of individual test items and the test as a whole. However, these revisions, in spite of increasing data management and the time-scale of the study substantively, enabled me to answer my research questions more effectively so were entirely justified.

Triangulation

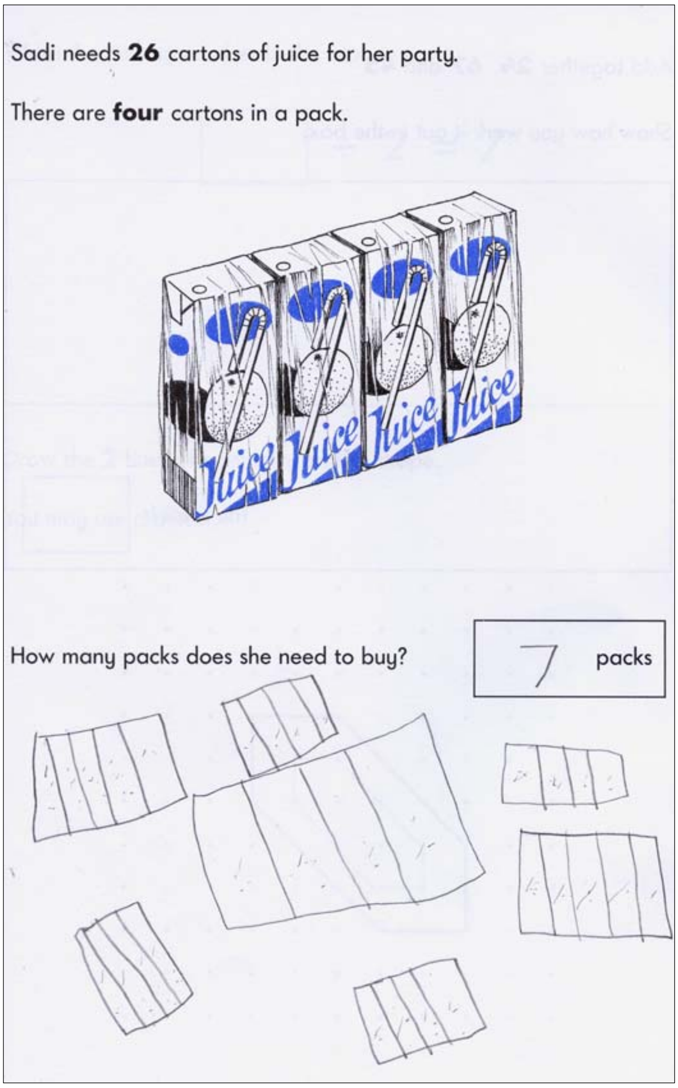
Of the various methods chosen to answer my research questions, I considered that the diagnostic interviews, in which Y2 children read test items from the KS1 2001 mathematics test, would provide the most compelling of all the evidence I obtained. Hearing and reporting on children's voices were at the heart of my study, but, as I explained in my ethical stance, children could not comment on whether my interpretation, through coded transcripts, reflected accurately what they had read to me. I could not, therefore, use 'respondent validation' (Cohen, Manion & Morrison, 2000:116) to confirm the trust-worthiness or credibility of my findings. Nor could I use 'investigator triangulation' (Cohen, Manion & Morrison, 2000:113) where more than one person investigates and cross-checks data from the same phenomena, since, after the pilot study, I was the sole researcher. Consequently, I used a multi-method approach or 'methodological triangulation' (Denzin, 1970:307) to support and corroborate the evidence from the children.

Underpinning all research questions was a concern that the reading skills required by the test could be too demanding for at least some Y2 readers. Each method chosen was designed to build up cumulative evidence from different perspectives about whether my concern was justified. If the evidence from the underlining of text during the test and the analysis of the booklet for readability corroborated the evidence from the interviews with children, I would be more confident about claiming that my concern was not without foundation. I did not view the interviews with teachers as part of the triangulation of my data since teachers could only provide opinions and anecdotal evidence about the reading demand. However, I was further reassured by the consistency of their views since these supported the interpretation of the evidence I had obtained. In CHAPTER 8, at the summation of all my evidence, I reflect again on the triangulation of my data when I reconsider my multi-method approach for its significance at that point in my study.

In this chapter, I have explained how I sought answers to my research questions and given a rationale for my preferred methods. In subsequent chapters, I explain how I applied my methodology to obtain the evidence I sought. The relevant chapters with their main foci are:

CHAPTER 5	Documentary annotations of test booklets
	Interviews with Y2 teachers
	Diagnostic interviews with Y2 children
CHAPTER 6	Interviews with Y2 teachers who use look-listen
CHAPTER 7	Comparing Y2, Y3 and Y4 children reading test items
CHAPTER 8	Analysis of KS1 2001 mathematics test booklet for readability

CHAPTER 5


Sadi needs **26** cartons of juice for her party.

There are **four** cartons in a pack.

How many packs does she need to buy? 7 packs

Andrew was working at level 2A in mathematics.

CHAPTER 5

Hands up in the KS1 2001 mathematics test

An overview

Main foci

In my methodology chapter, I justified how and why I intended to obtain data about the frequency of help with reading requested by children during the administration of the KS1 mathematics test in May 2001. Two main justifications were that I could:

i. gain an overview of how much or little help individual children requested, relative to their aptitude in reading;
ii. identify items where requests for help were more frequent so that they could be a particular focus for analysis.

This chapter includes the analysis of these data and complementary data collected during and in the months following the administration of the test. The overall analysis has three main sources that were discussed in my methodological choices and research design:

i. annotations in 164 test booklets indicating where children asked for help with reading;
ii. views of teachers about the reading demand of the test and how they supported children with reading;
iii. transcripts of diagnostic interviews with some Y2 children.

In the next section, I seek to obtain data to answer, at least partially, my research questions, *Is teacher support with reading available to children during the test whether* *they ask for it or not?* and *What is the extent of help needed by children of different reading competences during the test?*

Analysis of annotations in test booklets

Manipulation of the data: research design

Appendix 5.1 page 1 explains the headings used in the analysis sheets, some of which are also filed as appendices and considered later in the chapter. They summarise data collected by examining children's test booklets from six schools, later identified by the letters K, N, O, P, Q and R, whose staff kindly agreed to let me analyse the booklets in detail. Text underlined in any of the 30 written items in a booklet indicated that a child had requested help with reading that text during the live test. Although I did not code for such occurrences, I made notes to the effect that, in the majority of cases, children were inclined to ask for help with reading parts of items only, often individual words. Possible reasons and implications for this are discussed later in this chapter. Data obtained from the annotations informed my choice of children to interview, as I described in my methodology.

Since its layout was compact and concise, I used the spreadsheet design from my pilot study (Appendix 3.2), adapted slightly to improve clarity, to tabulate the data. I manipulated the data in five ways since each presented the frequency of the annotations from different perspectives. These were:

i. a school by school analysis in which I ranked children by reading level;
ii. filtering the 164 children's results by reading level, e.g. all children reading at the same level;
iii. combining all 164 children's results in ascending order by reading level;
iv. combining all 164 children's results in ascending order by frequency of requests for help with reading;
v. combining all 164 children's results by mathematics test score.

I now consider each of these analyses in turn.

School by school analysis: matters of surprise and concern

Appendix 5.1 pages 2-7 shows the results for each of the six schools taking part with Y2 children taking the test ranked in ascending order by reading level (Re L). My first reaction on examining the data on the worksheets was surprise and concern in equal measure. This was because of the high frequency of requests for help with reading across all six schools, particularly for children reading at level 2B or below. This gave some credence to my conjecture, outlined in my theoretical framework, that the reading demand was too difficult for many of the children.

Although the data showed that, in general, less competent readers were more likely to ask for help than more competent readers, there were some noticeable variations between children who attained the same reading level within and across schools. It was only for above-average readers (levels 2A or 3) that the variations became less pronounced as shown by Table 5.1.

Table 5.1: A cross-school comparison of frequency of requests for help with reading individual items (maximum 30)

	All readers		Readers at level 2C or below		Readers at level 2B		Readers at levels 2A or 3	
School	mean	range	mean	range	mean	range	mean	range
K	7.0	0 – 18	10.1	5 – 17	6.8	0 – 18	4.2	0 - 13
N	4.5	0 – 19	8.7	0 – 19	5.6	0 – 16	2.4	0 - 16
O	5.5	0 – 17	13.0	9 – 17	8.7	4 – 13	0.9	0 - 5
P	3.3	0 – 13	8.8	3 – 13	1.0	0 – 2	3.3	0 – 4
Q	7.6	0 – 23	14.3	9 – 23	10.0	6 – 15	2.8	0 – 9
R	13.7	0 – 29	21.3	18 – 29	11.6	8 – 17	2.6	0 – 6
All schools	6.3	0 – 29	12.6	0 – 29	6.6	0 – 18	2.3	0 – 16

(All arithmetic means rounded to nearest tenth)

Whilst I could claim that differences in the frequencies of requests for help in children of comparable reading levels were possibly due to personality traits, differences between schools were harder to rationalise. However, as I report later in this chapter, interviews with teachers and children provided some evidence that administrative arrangements could have been at least part of the reason for the observed effects.

The results for school R were of particular interest since the mean number of requests was from about two to four times higher than in the other schools. In School R, for example, children who were average or below-average readers asked for help about four times as frequently as those in School P despite their similarity of intake. I was also surprised by differences in results between Schools R and K. In School K, children were almost half as likely to ask for reading assistance in spite of the fact that six of the 10 children reading at level 2C and below and four of the nine children reading at level 2B had English as an additional language (EAL). Although EAL was not my particular research interest or area of expertise, I was not convinced that children in School K would require less reading assistance than those in School R.

My disquiet for children whose reading was not above-average came from two opposing positions. At one extreme, children who asked for little help may have guessed what to do or spent excessive time and mental energy trying to read text that could have been too difficult. Consequently, they may not have known or misread what was being asked of them. At the other extreme, children who asked for frequent help had the advantage of hearing at least part of the text read by an experienced reader but the disruption and wait for each request to be met would have extended the overall time children spent on the test, probably at the expense of concentration and mental energy. Table 5.1 shows that for readers at level 2C or below, both extremes were evident with requests for help ranging from 0–29 times. The range at 0–18 was narrower for readers at level 2B but I saw this also as a cause for concern for the same reasons.

All 164 children sorted by reading level

In my literature review, I discussed in some detail the characteristics of readers at levels 2C and 2B based on research by QCA (QCA, 2001a). The study found that

i. children reading at levels 2C and 2B used phonics to decode unfamiliar words, with this strategy being less successful with readers at level 2C;

ii. most of the weakest readers at level 2C decoded unfamiliar words by combining single letter sounds and could generally only sight-read short words with straightforward letter/sound correspondence;

iii. children showed 'a high level of independence and were confident enough to attempt every word in most cases, but with widely varying success' (QCA, 2001a:19).

In their study of the reading demand of the KS1 mathematics test, Lupetti, Sainsbury & Schagen (1995), also found evidence that 'children made valiant attempts to read the text but did not have the fluency to cope with reading and comprehension at the same time' (p13).

Appendix 5.2 shows an example of entries for the 164 children made by merging the data from the six schools into one database, and, in this case sorted by reading level (see Column AK). Simply by looking at page 1 of Appendix 5.2, that shows data for below-average readers only, the number of requests ranged from 0 – 29. Whilst I had not expected children reading at the same level to ask for exactly the same amount of help, I did not expect the ranges to be so extreme. Even allowing for slight differences in aptitude within the same reading level, I could not explain why average or below-average readers in particular should request either so little or so much help. Children reading at level 1 were not part of the QCA (2001a) study, nor had I expected to find them in this dataset of children who were left to read independently since, at level 1, children are only expected to read 'familiar words in simple texts … use knowledge of letters and sound symbol relationships .... In these activities, they sometimes require support' (DfEE & QCA, 1999a:57). Of equal concern was the number of children reading at level 2C, particularly in view of the weaknesses in their reading skills identified in the QCA (2001a) research.

All 164 children sorted by frequency of requests for help (1)

I then created another spreadsheet that showed data on children reading at the same level, but sub-sorted by frequency of requests for help with reading. Children working at level 1 and level 2C, all below-average readers, represented just over one quarter of participants in the dataset. On average, these children requested help with reading just less than half the total number of items even though the range was 0 – 29 requests. Knowing the limitations of their reading skills, I found it hard to imagine how children who requested little help coped with the text and how they could have been put in a situation where they were expected to try, even with encouragement to request help. The pattern was not dissimilar but less extreme for children reading at level 2B who made up about one quarter of participants and who would be regarded nationally as average readers. The frequency of requests for help ranged from 0 – 18 with an average number of requests of 6.6. It should be remembered, however, that some of these children would have gained enough marks to attain level 2B in reading but would be just over the level 2C boundary so their reading skills would be only marginally better.

Later analyses of interviews with children and the test booklet will show that the reading demand of the booklet was too challenging for most children. I could only conclude, therefore, that some children were ignoring text, guessing what to do, copying, using diagrams to try to find answers (see Lupetti, Sainsbury & Schagen, 1995:13), using phonics or reading text independently where it was not appropriate. The DfES acknowledged that '[y]ounger and poorer readers often do not recognise when they have not understood a text' (DfES, 2005:5) so it is likely that some children thought, incorrectly, that they did understand.

This dataset also showed that readers at levels 2A and 3 generally read more independently that those at lower reading levels with the mean number of requests for help dropping to 2.6 for level 2A readers and to 1.6 for level 3 readers. These results suggested more confidence in independent reading, and, by implication, greater fluency based on levels attained in the reading test. Nevertheless, a minority of these readers requested frequent help. Taking into account all 164 children, the mean number of requests for help was 6.3 but about 12% of children reading at level 2A and 4% reading at level 3 exceeded this number. In the level 2A reading band, the range was 0-16 with the outliers at the top end of the range possibly reading on the borders of levels 2B and 2A. Children reading at level 3 fared better with the range of requests for help narrowing to 0 – 7 times, and with only one child asking for help seven times. For the latter group, requests for help also seemed to be clustered towards the end of the test where children may have been getting tired after the combined intensity of the reading and mathematics. However, it is not unreasonable to assume that above-average readers would have the advantage over the remaining children by being able to concentrate on the maths in the absence of laboured reading or constant interruptions for help.

All 164 children sorted by frequency of requests for help (2)

Although I did not expect to gain much further information, I manipulated the data again to show all 164 children sorted by frequency of requests for help, irrespective of reading level, to see if further trends or patterns emerged. I found this layout particularly revealing. For example, it showed that:

i. of the 37 children who asked for no help with reading, most were reading at levels 2A and 3 but the group included readers at levels 2C and 2B also;

ii. of the children who asked for help six or fewer times, i.e. less than the mean frequency of 6.3, representing more than half the participants, their reading levels ranged from levels 2C to level 3;

iii. of the 28 children who asked for help with reading more than 12 times, i.e. more than the mean frequency for level 2C readers, most were reading at levels 1 and 2C but the group included readers at levels 2B and 2A.

All 164 children grouped by mathematics test score

My final dataset showed all children ranked by the number of marks they were awarded out of 36. Only a small number of marks separate the levels or grades, i.e. sub-levels within level 2 and between levels 2 and 3, that can be awarded as shown in Table 5.2.

Table 5.2: KS1 2001 mathematics - relationship between test scores and levels

Number of marks	0-4	5-7	8-13	14-19	20-25	26-36
Level awarded	None (W)	Level 1 awarded (L1)	Level 2C awarded (L2C)	Level 2B awarded (L2B)	Level 2A awarded (L2A)	Level 3 awarded (L3)

Just over 20% of children in the dataset missed being awarded the next highest level by one or two marks. The scores for these children have been shaded in column AI. Most children are expected to be working within level 2 at the end of key stage 1. Two children missed level 2C, the lowest grade at the target level, by two marks. Indeed, teachers generally enter for the test only children who are expected to attain levels 2 or 3 in mathematics and the fact that nine children failed to attain level 2 could be attributed to the fact that their reading levels were either levels 1 or 2C. Only one of these children asked for help with reading more than half the written items. A child who attains level 2B is regarded as average nationally and the data showed that 11 children attaining level 2C missed attaining level 2B by one or two marks. For these children, the range of requests for help with reading was 1–22, with most at the lower end of the range. Even for those who did ask for frequent help, it was not always the case that the teacher read all the text and children may not have read the remaining text accurately. A similar number missed attaining level 2A, the highest grade at level 2, scoring 18 or 19 instead of the required 20 marks. This group included mainly above-average readers as did the few who just missed attaining level 3. However, both of these groups included a minority of children reading at levels 2C and 2B who asked for little help with reading. It was not unreasonable for me to make the conjecture that children who were average or below-average readers may have lost one or two marks, possibly even more, because they did not receive appropriate help with reading. The analyses in These five analyses, exemplified in Appendices 5.1 – 5.2 provided me with convincing evidence that this was a likely outcome. This predicted outcome was more extreme than the conclusions reached by Lupetti, Sainsbury & Schagen (1995). I agreed with the prediction, resulting from their multi-level analysis, that mathematics test scores were influenced by reading level but felt that the prediction that '[a] one level increase in reading is associated with about a two mark increase in mathematics test score' (Lupetti, Sainsbury & Schagen, 1995:71) was likely to be an under-estimate of the effect of reading level. As discussed in CHAPTER 2, Newman (1977) also found that in a mathematics test taken by low-achieving 11-year olds, with carefully controlled vocabulary, about 12% of errors were made because children could not read key words accurately. The vocabulary in the KS1 mathematics test is not rigorously controlled for readability, a matter to which I return in my analysis of the test booklet in CHAPTER 8. The children's voices so far suggested that average as well as below-average readers could not read with sufficient skill to be expected to read independently. However, at this point in my study, the picture for above-average readers was, as yet, less clear.

The analysis of children's requests for help: key points

At the outset of this chapter, I announced my intention to find out how much or little help individual children requested relative to their reading level. The data I collected did this successfully but the analysis left me feeling uneasy for several reasons.

i. The amount of help requested by some children was excessive and suggested that they did not have the confidence or reading skills to cope with the text. In contrast, too many children reading at levels 2B and below in particular showed more independence in reading than would seem reasonable to expect, given the weaknesses in their reading skills as identified by QCA (2001a) and the DfES (2005). For some of these children, the number of requests was minimal or even non-existent relative to their reading level, and, as my later analysis of the booklet will show, at least two-thirds of the written items, covering a wide range of contexts and vocabulary, are text-dependent and have to be read accurately to obtain a correct answer.

ii. I was incredulous that readers at levels 1 and 2C were put in a situation where they were expected to try to read independently and to have to request support. All children in the levels 1 and 2C reading bands have the most basic of reading skills and certainly insufficient to cope with the textual content of the test.

iii. About one in five children, most of whom did not ask for the amount of help with reading that I would have expected, missed attaining a higher level (or sub-level) by one or two marks.

iv. The differences in frequency of requests for help with reading within children of comparable reading competence and within schools of comparable intake point to a variable in a standard test that could work to the advantage or disadvantage of particular children. In a test designed to assess mathematics, it appears that some children's lack of reading skills would have had a detrimental effect; these children would not have had equal access to the mathematics.

v. Because no such analysis had been undertaken previously, administering staff were probably unaware that children working at comparable reading levels varied so much in the frequency of help they requested. This phenomenon could have gone unnoticed in the 'business' of managing the test. I was satisfied that my analysis had made a small but convincing start in highlighting the disparities.

The analysis also enabled me to identify items where requests for help with reading were most and least likely. I deal with these data in the next section.

Items where requests for help were most frequent

The red numbers in column 171 of Appendix 5.2 show the number of times the 164 children in the dataset asked for help with reading particular items. To make comparisons easier, I arranged the items by descending frequency of requests for help as shown in Table 5.3, expecting that the items that would be easiest to read would come mainly towards the front of the booklet. However, Table 5.3 shows that this was not the case. Rows with shading indicate items whose mathematical demand was meant to be accessible to the majority of children taking the test, i.e. items towards the beginning and middle of the test, targeted at children working at levels 2C and 2B.

Table 5.3: Written items in descending frequency of requests for help with reading

Qu. no.	Frequency (max. 164)	Target level	% (children)
14	76	2C	46%
18	62	2B	38%
11	57	2C	35%
17	56	2B	34%
29	55	3	34%
20	48	2B	29%
31	43	3	26%
16	42	2B	26%
30	42	3	26%
15	41	2B	25%
22	41	2A	25%
27	41	2A	25%
24	37	2A	23%
25	37	2A	23%
26	37	3	23%
19	34	2B	21%
35	33	3	20%
33	31	3	19%
21	30	2A	18%
12	29	2B	18%
10	28	2C	17%
28*	21	3	13%
13*	20	2B	12%
8*	19	2C	12%
34*	19	3	12%
23*	15	2A	9%
7	9	2C	5%
9*	9	2C	5%
32*	9	3	5%
6	7	2C	4%
Total	**1028**		

* indicates an abstract calculation
(Percentages rounded to nearest whole number)

The mean number of requests for help with each of the 30 written items was 34.3. Put another way, an average of just over 20% of children requested help with reading every written item. The table shows that seven out of 10 items that were meant to be accessible to children working at levels 2B and 2C generated the most requests for help with reading and exceeded the mean. The likely effect was that children who were not competent readers but did not ask for help would not have accessed the mathematics. In contrast, all children showed more independence in coping with the 10 items at the lower end of the table, which included items spanning the range of mathematical demand. Most of these were abstract calculations where, even if the introductory line of text was ignored, children appeared to find the presentation of equations familiar. One limitation of this dataset was that, whilst it does show the frequency of requests for help, it cannot reveal whether children could read accurately parts of or whole items that were not read for them so the data may understate the amount of help children needed but did not request.

The high frequency of requests for help with some items was evidence that many children found the reading demand was excessive but did not provide evidence of why this was so. Such evidence would come only from hearing children read and from an analysis of the test booklet for readability. I return to how I obtained this evidence later. There were also poor readers who did not ask for help with items that came near the top of Table 5.3. This too I found puzzling but some possible reasons are given in interviews with teachers reported later in this chapter, e.g. peer-group pressure.

Interpretation of the analysis of the booklets

i. The evidence from the children showed convincingly that the reading demand of some items, meant to be accessible to children working at levels 2C and 2B in mathematics, was under-estimated. Reasons for this will be considered in CHAPTERS 7 and 8.

ii. In schools, where the ratio of teaching staff to children during test administration was not so generous, I did not find it hard to imagine how hard-pressed staff would be to deal promptly with requests for help and to know the order in which children put up their hands. This could result in rushed or inadequate reading support or undue delays for children.

The following section deals with administrative arrangements, including the ratio of staff to children, and leads to the identification of another variable that could affect children's attainment in the mathematics test.

Interviews with teachers participating in the analysis

Data collection: research design

In this section, through the voices of teachers rather than children, I return to my research questions, *Is teacher support with reading available to children during the test whether they ask for it or not? What are the views of year 2 teachers on the reading demand of the test?* Implicit in the questions is the reasonable assumption that teachers know the strengths and weaknesses of the reading skills of their pupils, how much support they are likely to need and how to provide it.

I interviewed seven Y2 teachers who had administered the KS1 2001 mathematics test in the six schools participating in the analysis previously discussed in this chapter. Due to a faulty tape, I lost interviews with teachers at two of the schools and had to rely on notes taken as I listened to the tape for the first and only time it played correctly. To provide some structure to the interviews, I prepared questions whose answers would complement the data from the analysis. I was interested to know:

i. if the mode of administration of the test or other factors might have influenced how much or how little help children requested with reading;
ii. how standard their support was and whether some teachers might provide more advantageous test conditions than others.

Some interview questions related to their views on the look-listen mode of administration that these teachers had not used. I shall return to their opinions in the next chapter to compare their views with those of teachers who administered the test using look-listen. The three main aspects of the interviews I report here are:

i. teachers' views on the reading demand of the mathematics test;

ii. how they organised children to administer the test;

iii. how they supported children whom they considered could not read without assistance.

Appendix 5.3 explains how I coded the transcripts of teachers from Schools K, N, P and Q. The transcript of the two teachers in School N in Appendix 5.4 shows how I applied the coding; the other transcripts are available if required.

Teachers views on the reading demand

All teachers used terms like ‘difficult’ and ‘demanding’ to describe the reading in the test. In School K, which had a high percentage of EAL children and non-EAL children who were poor readers, the teacher felt that the reading put ‘an extra demand on them which doesn’t measure their maths. Teacher MA at School N described the language as ‘quite demanding’ and, consequently, said that ‘[t]here was a lot of rushing around from child to child to try and read some of the language for them’. In some items, teacher WK, also at School N, felt that ‘the actual reading was more difficult than the maths’ and there was ‘way too much reading for the amount of [mathematical] work that was required’. In school P, the teacher considered the reading demand ‘quite high’ particularly for slow learners whom, she felt, could do some of the mathematics but couldn’t ‘access the information’ from the written text. The reading was described as ‘far too demanding’ for readers at levels 2C and at 2B by the teacher at School Q. She had some minor concerns about readers at level 2A and felt that it was only when children could read at level 3 that she was confident that they could cope. This was a view shared by JV, the Y2 teacher interviewed in my pilot study.

Modes of administration

Each school had its own procedure, the key features of which I now describe.

School K: the teacher divided the 29 children entered for the test into two groups of 14 and 15 children. The teacher was supported by one classroom assistant. Children who had more difficulty with English or had general reading difficulties were in one administration and those more confident at reading English in the second administration.

School N: The same mode of administration was used by both teachers who had entered 24 and 23 children respectively for the test. They each administered the test to half of the class whilst the other half was supervised elsewhere.

School O: All 18 children took the test at the same time with the teacher supported by a nursery nurse.

School P: The teacher administered the test three times to approximately one third of the class of 28 children each time. She explained that children were grouped by ‘ability’ with ‘brighter children in one group, the middle ones in another and the slow ones in another, so I could take more time with the slower group to give them a chance – plenty of time to do the actual items’.

School Q: The 24 children entered for the test all took the test simultaneously. The teacher was supported by a classroom assistant.

School R: The teacher organised the 18 children into small groups, usually of three children, and administered the test to each group in a small room kept for group work. The remaining children were supervised by other staff.

These descriptions built up a picture of different administrative arrangements, key features of which I summarise in Table 5.4. Table 5.1 (see page 130) showed variations in requests for help within and between schools. Table 5.4 complements Table 5.1 by showing another variable, i.e. differences in the ratio of staff to children.

Table 5.4: Ratio of staff to children at each administration of test

School	Number of children entered for test	Mean number of children at each administration	Number of administrations	Number of administering staff	Ratio of staff to children
K	29	14.5	2	2	1:7.3*
N (1)	24	12	2	1	1:12
N (2)	23	11.5	2	1	1:11.5
O	18	18	1	2	1:9
P	28	9.3*	3	1	1:9.3
Q	24	24	1	2	1:12
R	18	3	6	1	1:3

*rounded Mean ratio for 6 schools (7 classes) = 1:9.2 (to nearest tenth)

To make it easier to compare how the frequency of requests for help compared with the ratio of staff to children, I constructed Table 5.5.

Table 5.5: Ratio of staff to children in ascending frequency

School	Ratio of staff to children	Mean number of requests for help with reading
R	1:3	13.7
K	1:7.3	7.0
O	1:9	5.5
P	1:9.3	3.3
N (2)	1:11.5	4.5
N (1)	1:12	4.5
Q	1:12	7.6

I had expected to find that, where the ratio of staff to children was lower, children would be more inclined to ask for help and teachers would be in a better position to monitor and give prompt help. Indeed, this seemed to be the case for School R in which children took the test in groups of three, and, on average, asked for reading assistance for 13.7 out of 30 items. However, in School P, where children asked for the least help with an average of 3.3 requests, the ratio of staff to children was 1:9.3, approximately equal to the mean ratio of 1:9.2. Even more surprising was that School Q, with the equal highest ratio of staff to children at 1:12, came second to School R in the frequency of requests for help with a mean of 7.6. However, this was about half the frequency of School R. One result that possibly pointed to a higher staff to children ratio discouraging children from requesting help came from School N. In the two classes, the ratio of staff to children was 1:11.5 and 1:12 and the mean number of requests by each child was low at 4.5. However, these children were taking a standard test that was intended to assess mathematics ‘in a fair and comparable way’ (QCA, 2004:23), i.e. inclusive to everyone. However, it was already clear that variables such as the amount of help with reading or the ratio of staff to children during test administration made for large differences between the experiences of children that made me question how ‘standard’ such tests were. Nevertheless, any conclusions from these ratios have to be treated with caution since the ratio of staff to children may have been less important than the quality and quantity of support the children received as discussed in the following section.

Support for children with reading difficulties

School K: For the test, the 14 children regarded as less competent readers were grouped together mainly because of lack of familiarity with English. Indeed, six children who were below-average readers also had English as an additional language but were left to read independently. The teacher ‘positively encouraged them to put their hand up to ask us to read the items so that it was the norm rather than something exceptional’. She believed that the poorest readers ‘were happy to put their hands up for every item’ although data from the children (see Appendix 5.1, page 2) show that this belief was not upheld. The two administering staff also ‘monitored the children who would need most help’. The teacher also said that the ‘two of us were kept busy the whole of the session, particularly with the children who had more difficulty with the reading’. This could suggest that children were not always helped as promptly or monitored as carefully as the staff would have wished. Those regarded as more competent readers took the test in a second administration. The teacher felt that they could read confidently but might ask for help with the ‘odd word or two’.

School N: Teachers MA and WK, who each supervised groups of about 12 children, expressed concern about children whom they thought were not asking for appropriate help in spite of encouragement to do so. Teacher MA positively intervened with children whom she considered reluctant to ask for help, by asking, for example, *Do you need any help?.* Indeed, the two teachers identified shy children as being unwilling to ask for appropriate support and the size of the group may have contributed to this unwillingness.

Both teachers read only the text for which children requested help although about half the children were reading at level 2B or below. This surprised me because they must have assumed that children could read the remaining text without assistance. Alternatively, with about a dozen children to supervise, they may have had more than one child requesting help at the same time and had to compromise to respond promptly. Teacher WK felt obliged to restrict help to reading only what was requested in case she exceeded the permitted assistance. (This was a misunderstanding since any text could be read.)

School O: The teacher and nursery nurse supervised 18 children taking the test. The two teaching staff shared the role of helping children with reading but the teacher concentrated on children reading at level 2C and at the lower end of level 2B. The teacher read all the text in an item to these children even though they may have requested help with part of it only. The nursery nurse, concentrating on the more competent readers, read only what was requested. This could range from a single word to the complete text for an item. Indeed, the teacher admitted that she had never thought to standardise the approach for reading assistance with the nursery nurse and had assumed that she adopted the same practice.

School P: Teacher SH administered the test to one third of her class at a time. One of these administrations was for the third of the class described as those ‘who have difficulties’ to give them more time. For all administrations, SH read all the text in items when children requested help. Before the test, she wrote words that she thought children might find difficult to read, as permitted by the teacher’s guide, and read them through with the children but she thought that this was probably too much information for them to take in and to remember during the test. This view was shared by JV, the Y2 teacher interviewed during my pilot study. None of the other schools suggested that they had used this strategy to support reading.

School Q: In common with School O, the whole class took the test simultaneously, with two staff supervising the 24 children. The teacher and classroom assistant both worked in the same way, i.e. if a child asked for help, they read the whole item. They also monitored the poorest readers, and ‘[i]f we saw any of them … staring into space, we would go over and read the item’.

School R: With only three children taking the test at one time in a small room off the classroom, the teacher felt that the advantages were being able to read all the text when help was requested, to respond quickly and to monitor each child closely. She also offered to reread items that contained a lot of information so that children had the chance to hear these items read aloud twice. She intervened occasionally where she felt that a child was finding the reading difficult, asking, for example, ‘Would you like me to read that?’. Children in this school, possibly because of a comfortable setting and a small group situation, may have been more relaxed than in a larger group. They might also have had fewer inhibitions about asking for help which was likely to have been given promptly.

Interpretation of support for children with reading difficulties

There was a general impression from the transcripts that teachers were not concerned about above-average readers since support for these children was rarely mentioned. Teachers' responses describing how they supported children with reading difficulties revealed yet another variable that I felt could have affected children's performance in the test for better or worse. Apart from differences in administrative arrangements and the ratio of staff to children, there were also differences in teachers' approaches to requests for help from children. For example:

i. In School R, where the ratio of staff to children was 1 to 3, children had all the text of an item read to them, at least once, even when they had asked for help with reading only part of it. In School N, where the ratio of staff to children was about four times as high, teachers read only text that children indicated.

ii. Schools O and Q were the only two schools where all the children entered for the test took it as a class, with two staff supervising each class. In School Q, the two staff had agreed how to support children with reading difficulties but this was not the case in School O where a common approach had not been discussed.

Other issues of interest

What staff at these six schools had in common was expecting children to know and indicate when they wanted reading support. This put a lot of responsibility on 6- and 7-year olds and I was interested to know if the teachers felt that children were asking for appropriate help. The teacher at School K was the only teacher, of the five who commented, who was convinced that children were asking for such help. The reason given was 'because we made more of a point of it, that we were very happy for them to put their hands up for every item if they felt that they needed it'. However, as shown by the transcripts of interviews with children from School K later in this chapter, her conviction was overly optimistic. However, one teacher at School N felt that timid children ‘weren’t keen to ask’ and ‘would sit and just quietly try and get on’. Her colleague shared this view, believing that some children were ‘quiet and tentative’ and looked ‘as though they are doing something’ but she was not confident that they always requested appropriate help. At School Q, the teacher said that ‘not all children want to put their hand up … which is another problem…and they are spending time trying to work out what it actually says when you know … that they are capable of doing that particular item’. A similar point was made by the teacher in School P who said that ‘there are some children who are shy about putting their hands up because they don’t like to … mark themselves out ….’ She considered that such children might guess what to do or use illustrations that could mislead. My interviews with children later in this chapter show that the view of the majority was correct. Another concern raised by teachers at School N was the length of the test. One of them said that ‘the length of the test is too long, way too long. The children run out of steam, definitely, even the able ones run out of steam, and obviously the amount of reading in that, that’s a factor ….’

Interpretation of other issues of interest

The majority opinion of teachers was uncertainty about whether all children asked for the amount of help they needed with reading. Not only that, strategies discussed in the section ***Support for children with reading difficulties*** (see page 145) showed that only about half of the teachers positively intervened when they suspected that children were reluctant to ask for help. Lack of intervention by the remaining teachers may have deprived some children of necessary reading support. In the next chapter, I compare the ‘hands-up for help with reading’ approach to the ‘look-listen’ approach where assumptions are not made about children having appropriate reading skills or recognising the need to ask for help.

Conclusions from interviews with teachers

Across the six schools, administrative arrangements varied considerably. Nevertheless, no teacher exceeded the permitted assistance given in the teacher's guide but, for a test that was meant to be 'standard' nationally, I considered that I was building up a picture of non-standard management of the test involving more than one 'intervening variable' (Silverman, 2005:101). I was convinced that these variables could affect some children's attainment in the test and may have advantaged some children over others. I had no doubt that all teachers genuinely believed that they were making the best possible arrangements for administering the test, taking into account available staff and knowledge of the children. Teachers are judged in part by pupils' results in national tests so it is in their best interests to do so.

Returning to children's voices

Shuard & Rothery (1984) recommended that one way of assessing the readability of mathematical text was to 'look closely at …the ways in which children respond to what they read' (p2). Following this recommendation, I now provide transcripts of some children reading items from the test. From these, I hoped to determine whether the reading demand was appropriate and whether children did receive appropriate help.

Interviews with children

Time frame

Interviewing children after the analysis of their booklets was not without problems. Because of in-school and LEA auditing procedures, I was unable to obtain access to the booklets until late June or early July 2001 and, generally, had to return them by the end of term. By the time I had analysed the booklets, there was little time to carry out interviews before the end of term with the result that most children were interviewed the following September and October, i.e. up to five months after the administration of the test. This was not ideal since some teachers said that children had improved their reading level in the intervening period.

Choosing children to participate

I interviewed at least four children from six schools, concentrating on children reading at level 2 since this is the level most children are expected to attain by the end of key stage 1. I made sure that I included a good balance between boys and girls and some children who had English as an additional language. I interviewed more children reading at levels 2C and 2B than 2A, and, at that point in time, chose not to interview those reading at level 3 who were least likely to need extensive reading support. This was because, in my pilot study, I found that children in year 3 assessed as reading at level 3 generally coped well with reading items from previous KS1 mathematics tests. I lost interviews with children at two schools due to the faulty tape, mentioned previously. However, I interviewed 24 children at the remaining four schools. In spite of lost data, I collected sufficient to contribute towards answering the research questions, *Is the reading demand of the text in a typical KS1 mathematics test appropriate for the target children in year 2?* and *Is teacher support with reading available to children during the test whether they ask for it or not?*

Having found out from my pilot study how long it takes to transcribe interviews with teachers and children, I chose not to transcribe all 24 interviews but stopped when I felt that my selection typified the standard of reading at each of the sub-levels 2C, 2B and 2A. This decision followed the advice to keep 'the number of interviews to a minimum for adequate representativeness' (Gillham, 2000:12) but also for effective time management.

I interviewed children individually using an audio-tape recorder since I needed to refer to the fine detail of their responses for my analysis. I wanted the interviews to feel informal so that the children would be relaxed. I showed the children the test booklet and explained that I wanted them to read some items to me so that I could find out what words they found easiest to read. I also explained that they did not have to take part, could stop at any time, and did not have to find or write answers unless they wanted to. However, it turned out that some children offered answers without prompting, and, in some instances, the answers gave me more information about their competence at using reading to engage with the mathematics. Where appropriate, I refer to the responses in the analysis.

By letting the children say their name and answer informal questions into the tape recorder and playing their responses back to them, I put them at ease with the machine. To further increase the sense of informality, I asked the children to explain what each item was asking them to do so that my interaction with them was more conversational and child-friendly. However, it turned out that few children in year 2 could do this even if they went on to give the correct answer without prompting; others read so badly that there was no point in asking if they knew what to do. Not only that, as Harrison (1980) said:

> Some poor readers are unreliable reporters of whether or not they are in difficulty with a text, but most of us know whether or not we are understanding what we read. (Harrison, 1980:140)

Whilst I agreed with the first viewpoint, I was not convinced that all year 2 children would know if they were reading ME or unfamiliar sight words with understanding. My emphasis had to be on using evidence from the transcripts and professional experience to judge whether the children were reading with sufficient competence rather than rely on explanations of the response required. I agreed with the following view:

> Requiring the reader to summarise or précis the text is a frequently used technique for assessing the level of understanding. However, it is an inadequate tool for use in readability studies since the translation and production of a response include many skills which bear little or no relationship to those required in comprehension. (Gilliland, 1972:88)

Although tempted, I chose not to take field notes during the interviews because, as stated in my methodology, I agreed with Costley (2000) that '[i]t would have been difficult, and possibly intrusive or disruptive' (p171) to do so.

Sample of test items to be analysed

Unlike the research method used by Lupetti, Sainsbury & Schagen (1995), I decided not to ask the children to read every test item, as they did, but to confine each interview to a 10-minute limit during which I would concentrate on a few items only. In so doing, I felt that the children would have minimal disruption to their normal classroom routine and were unlikely to lose concentration or interest.

So that I could transcribe faithfully and review transcripts as often as required, I made audio-tapes of each child reading several items from the test, mainly those for which they had not asked help with reading. As with all interviews with children in this study, I asked their permission to do this and told them that they could discontinue the interview at any time.

I organised the interview data in two ways to provide different perspectives. These were:

i. transcribing recordings of individuals to find out how well they coped with reading several items;
ii. transcribing the same item read by several children reading at different levels to compare their reading skills.

Transcription method

I considered carefully how much detail I could cope with in the transcriptions, because 'transcriptions concerned with linguistic features have … often been subjected to very detailed and intricate transcription' (Powney & Watts, 1987:146). Another decision I had to make was whether to code the text exactly as written and set out in the test booklet, mark the codes above or between the relevant words and then comment on the codes. This would have been the easier option, but, after listening to the tapes, I felt that it would be a less effective system for showing the extent of some children's struggle with the reading. Powney & Watts (1987:146) also pointed out that '[d]iffering kinds of interviews require different kinds of transcription'. By transcribing verbatim what the children said, the extra time was well spent because the transcripts became a true and literal record of children's voices. I also made a note of pauses in their reading by timing pauses within or between sentences and the time taken to read the item from start to finish. I believed that these data collectively would provide adequate evidence about children's reading competence since they placed an emphasis on accuracy, fluency and pace. Fluency and pace are closely connected since the former 'is the extent to which a person can read a given text at optimum speed' (Gilliland, 1972:13). The codes used, adapted from Heritage (1984, cited in Silverman, 1993:118) were as shown in Table 5.6.

Table 5.6: Symbols used for transcription analysis

Symbol	Explanation of symbol
()	Closed parentheses indicate a brief silence of up to 1 second in duration.
(4)	Numbers in parentheses indicate a silence measured in seconds. Normal expected pauses between sentences were not coded except where there was a silence of an exceptional duration. The rare spoken interventions by me were excluded from the timings.
(Comments)	Italicised comments in brackets provide supporting information by me, recorded immediately after the interview and/or during the transcription.

The transcriptions follow, starting with responses from individuals reading several items.

Individual readers

Table 5.7: Y2 boy (L2C reader/EAL) from School K: Priya – July 2001

Item number	Text read (verbatim)	Time taken (secs)	Comments
11	Draw arowone to () sorted the () numbers.	8	Priya did not ask for help with reading this. He appears to have recognised the diagram as one that was used for sorting and guessed incorrectly what he was being asked to do. His lack of reading accuracy was likely to have prevented him from knowing what the item was asking him to do. Priya omitted this item and may not have attempted to read it in the live test.
17	Ann (5) m-m-mipiger (2) the (2) spoiler that two () do () in () bocks. Bocks min- mind () bocks t-taller in the lard (2) d-doll	43	Reading stilted throughout. Priya struggled to decode the words and could make no sense of the sentences. His intonation was monotone with no apparent awareness of meaning. Priya did ask for help with reading this item and the transcript shows the extent to which he needed help. He gave the correct answer in the live test.
18	*(Omitted text above diagram.)* Sar () sor the () nu six () and twenty-seven. The () the (5) *(I encourage him to carry on.)* this in the cret bo-box () of () the (1) number.	37	Priya had all text for this item read to him during the test and responded correctly. Read numbers confidently but otherwise struggled with the text.
19	*(Omits upper two lines of text and tries to read text positioned alongside the number 19.)* *(Reading indecipherable – see notes that follow.)*	n/a	Did not ask for help with reading this item. Gave an incorrect response.

Priya had English as an additional language and was regarded as becoming confident as a user of English. His teacher explained that he was of average attainment in mathematics but had below-average reading skills. He asked help with reading 10 items with all but two of these requests in the first half of the written questions. I attempted to transcribe his reading of item 19, for which he did not request help, but I gave up since his attempt bore so little relationship to the text, consisting mainly of what I could only describe as grunting noises. When asked to read item 24, he said that he couldn't read it, at which point I asked if he recognised any of the words. He identified 'the'. Indeed, the transcript shows that he could read simple everyday words such as 'this', 'in' 'and' etc. His reading of item 27, like item 19, was not intelligible for transcription. His sight vocabulary was totally inadequate for independent reading of the test as shown by misreads, numerous pauses between words and the overall time taken to read items. Without prompting, he told me that a friend told him answers to some items that he could not read. After item 19, Priya asked the teacher for help with reading only two items and omitted many later items.

I was shocked at the extent of Priya's reading difficulties. Although he was positively encouraged to ask for help, the responsibility to do so was put on him and, consequently, he did not receive the help that he obviously needed although his teacher believed that children did request appropriate help. Without assistance, he could generally read only simple words that occur frequently in ordinary texts. After my attempt at transcribing his reading of items 11, 17 and 18, I believed that the child's reading was so poor that he should never have been expected to attempt any of the reading without assistance. Priya missed being awarded L2B by two marks, although it was not clear whether some marks had been gained through his friend's support. Nevertheless, the lack of professional reading support may have affected his final score in the test since his teacher assessed him as working at level 2B in mathematics.

Table 5.8: Y2 girl (L2B reader/EAL) from School K: Stephanie – July 2001

Item number	Text read (verbatim)	Time taken (secs)	Comments
14	The number in the (4) sh-shadow squares makes a square s-s-squares. Con (2) can (3) con (1) con (1) tin (*She looks to me for help. I ask her to try again.*) contin the squares by [shading] (*pronounced with soft 'a'*) more squares.	38	Seemed to have some awareness that what she read did not make sense but could not self-correct. Words 'shaded', 'sequence' and 'continue' caused particular problems. Gained the mark for this item so the diagram probably provided sufficient prompts for her to know what to do.
16	Write the (2) carrest (1) car-carrest (2) seen () sing in the (*self-corrects*) each box.	17	Had problems with words correct and sign. Response incorrect.
17	*(She starts at 2nd line of text.)* How many blocks t-taller (1) taller taller *(read three times with 'a' as in tally*) is the longer (2) doll? *(I ask her to read first line.)* Ann m-m-mestured mestured the sheent height () of () these () two () dolls in blocks. How many blocks taller *('a' as in tally)* in in () the (1) longer doll?	41	Missed first line of text completely. Had problems with key word taller. Could not read key word 'measured'. 'Longer' incorrect but sensible substitution, probably prompted by diagram. Gave answer of 17 blocks i.e. height of taller doll.
18	Start .. *(I ask her to start at top of page.)* Here is a (1) sort()ing diagram. The number thirty-four has been (1) start. St () Stuck () suct () st () sart () sart () start the numbers six and twenty-seven. Write them in the correct box on the diagram. *(I ask her to read labels on diagram.)* more than 10, less than 10 even number, odd number even number, odd number	46	Read slowly but with reasonable accuracy apart from the words 'sorted' and 'sort' although she had less difficulty reading 'sorting'. She gained the mark for this item, perhaps because she found it possible to interpret the diagram correctly without having to decode all the explanatory text. She was fortunate that her misread of 'number' for 'numbers' in the arrowed labels of the diagram did not affect its interpretation.

Stephanie had English as an additional language but was considered to be fluent in most social contexts. In the statutory tests, she gained level 2B in reading and level 2C in mathematics. She requested help with reading nine items, slightly more than the class mean, but not for any of the items above that came towards the middle of the test. The transcript provides evidence that she did not obtain help with reading items for which she lacked reading competence. In items 14 and 18, she responded correctly. In these, she is likely to have interpreted what to do from the diagrams alone since her reading lacked fluency and pace even when accurate in parts. In some cases she misread key words, for example, *measured* in item 17. She missed the award of level 2B in mathematics by two marks, which could have resulted from lack of help with reading. She also had reading difficulties with other items later in the booklet. These are on audio-tape but have not been transcribed since the selected extracts were sufficient to convince me that she had unidentified reading difficulties.

Table 5.9: Y2 boy (L2B reader) from School Q: Marius – September 2001

Item number	Text read (verbatim)	Time taken (secs)	Comments
14	The number in () the () sned shed shed sheded question (1) swe (2) se () seconds. Count the () seconds by seding () s- s-sedding more (*next word unclear*)	41	Did not ask for help with reading. Shaded, squares, continue, sequence and shading caused particular reading difficulties. Marius tried hard to make sense of the words but they were too difficult for him to decode.
15	*(After being asked to read, he waits about 10 seconds just looking at the text before he starts. He then starts to read text below graphic and I indicate for him to start on top line.)* There are () four () apples () in () each pack. Miss Pull () buys () three () packs () of apples. Now () men () many () apples () does () she () buy?	29	The child's reading throughout was stilted with momentary pauses between most words. He substituted 'Miss Pull' for 'Mrs Pullen', which, in this instance, did not affect the overall meaning. Marius gave an answer of 4 for this item, the number of apples shown in the artwork.
16	With the con conte (2) [cards] signe sing (2) singing in each box.	17	Marius could not read 'correct or 'sign'. Since the item comprised abstract equations for which he had to write in missing signs, he is likely to have thought that he knew what to do without reading the text. Unfortunately, he inserted numbers instead of signs in the boxes.

Marius requested help with reading 7 items. He asked for help only three times before item 19, the last item for which he gained a mark. From thereon, he gave incorrect answers or omitted items. He read items 14, 15 and 16 without assistance, all items that children could possibly interpret by ignoring text and focussing on graphic clues and he was awarded the mark for item 14. He said that he put his hand up for help with reading only when he thought that he did not know what the words were. Marius also told me that he would have liked to have the teacher read all the items to him without putting up his hand all the time. This child was awarded the minimum number of marks to attain level 2B (average) but he gained four of these marks from the maximum of five available in the oral items, showing that he responded well to teacher–read items. The evidence suggests that his reading competence at level 2B was not sufficient for independent reading of the mathematics test.

Table 5.10: Y2 girl (L2B reader) from School P: Jade – October 2001

Item number	Text read (verbatim)	Time taken (secs)	Comments
14	The numbers in the shade-(1)ed squares make a sentence. (3) *[She looks at me with puzzled expression.]* Counting the questions by shading more squares.	23	Substituted 'sentence' for 'sequence', 'counting' for 'continue' and 'sequence' for 'questions' (2nd occurrence). Asked for help with reading 'sequence' only in live test. Puzzled expression suggested that she realised that first sentence did not make sense but made no attempt to self-correct. Read 2nd sentence fluently but made a substitution. Jade asked for reading support for two questions only and missed the award of L2B in mathematics by three marks. She may have gained more marks had maximum reading support been available.
15	There are four apples in each pack. (2) Mrs (1) Pullen buys three packs of apples. How many apples does she buy?	15	Read well apart from short delay before 'Mrs' and 'Pullen'.
17	Ann () measures () the () height of these () two () dolls in () blocks. How many blocks () taller () is the large doll?	15	Asked for help with reading 'measured' in live test. Read mainly accurately but not fluently due to brief hesitations between words.

Jade asked for help twice only for items 14 and 17 and gained L2C in maths. During the test, Jade's teacher read the first occurrence of *sequence* only and Jade read the remaining text without assistance. Even five months later, she made errors, although she seemed to sense that this was the case. She missed being awarded L2B in mathematics by three marks. She was awarded level 2B in reading in May 2001 but her teacher believed that her reading had improved since then. Indeed, her reading was more fluent and accurate than that of Marius, also reading at level 2B. Nevertheless, she continued to make errors in spite of the improvement in reading and she would have benefited from more reading support during the test.

Table 5.11: Y2 boy (L2A reader) from School N: Daniel – October 2001

14	The numbers in the () shaded () squares make () a () squ- () se-() se-()quence. *(pronounced as two words.)* Continue to sequence by shading more squares.	16	Asked for help with reading sequence in test. Stumbled with first sentence but coped well with second sentence apart from substituting ‘to’ for ‘the’.
17	Ann mea-()sured *(pronounced as two words)* the height () of () these two () dolls in block()-s. How many blocks taller is the large doll?	15	Read this independently in test. Coped well apart from slight hesitations between words.
18	Here is a sorting (2) diagram. The number 34 has been () sorted. Sort the numbers 6 and 27. Write them in the correct boxes on the (1) diagram.	20	Read this independently in test. Read fluently and accurately apart from hesitation before ‘diagram’.
29	Draw the two lines of (1) sys-(1)tem-try *(pronounced as three words)* on this shape. You may use a mirror.	12	Asked for help with reading ‘symmetry’ only in live test. Read fluently apart from misread of ME word ‘symmetry’.

Daniel asked for help with reading items 14 and 29 only and gained the minimum number of marks to be awarded L2A in mathematics. He read with good intonation and, in parts, with fluency and accuracy but, in some sentences, his reading was punctuated by momentary hesitations between words and reading errors with ME words *sequence*, *measured* and *symmetry*. Once he had decoded a word in a sentence, he could generally read this more fluently when that word, or a derivation of it, occurred later in the text. For example, he read *sequence* fluently on his second attempt in item 14. It is possible that Daniel read items with sufficient skill to decode most of the text but, taking into account the intermittent lack of fluency and accuracy, I was not convinced that he would not have benefited from more reading support.

The transcriptions that follow in Tables 5.12 and 5.13 record how several children reading at levels 2C, 2B and 2A read items 14 and 17. An adult reader takes about 8 seconds to read the text of either of these with understanding, as I found out using several volunteer teachers. (Because of the difficulties of getting convenient page breaks in the two tables, each table is continuous.)

Individual items

Table 5.12: Transcript of Y2 children reading item 14

Text from item 14: The numbers in the shaded squares make a sequence. Continue the sequence by shading more squares. (See Appendix 4.1 for artwork.)

Key: K-G-L2C means school K, girl, reading at level 2C
(EAL) means having English as an additional language

Child	Text read (verbatim)	Time (secs)	Comments
Q-G-L2C Charlene	The () numbers () in () the () shade (7) squares make a s- (10) (*I intervene and read sequence.*) Count () the () space() by () using () more (6) shades.	53	(September 2001) Did not request help with reading this. Words 'shaded', 'sequence' and 'squares' caused most problems. Reading stilted throughout with poor intonation. Charlene asked for help only nine time during the test. If the reading of this item is typical, she is likely to have struggled with reading other items where help was not requested. The effort of reading may also have used up her energy at the expense of the maths. She gained 5 marks, failing to gain L2C by three marks
P-B-L2C Andre	The numbers is the s-s-e-senti-(3) sentiment () s-e-q-d (2) *(He says he can't read that word so I said to carry on and we would come back to it.)* make a s-s-s-s-s-song. Could () could the sequ () e () be s-h-a-ding - s-h-a-ding make more sentiments of five.	60	(October 2001) Did not request help with reading this. Could read only common sight words such as make, the, more etc. and used phonics ineffectively for unfamiliar words. After reading text, went on to explain correctly what to do by indicating the diagram. He said that he knew what to do because of the diagram.

			Asked for help three times only during test but his reading of this item shows lack of fluency, accuracy and dependence on phonics. He gained L2C but, with three marks more, he would have gained L2B, a possibility with maximum reading support.
P-B-L2B Lee	Continue the (*I ask him to start at the top line of text.)* The numbers on the shading squares making a () sink (10) (*I ask* *him to try reading the word*.) [sinquint] *(He looks to me for help* *and I read sequence.)* C-con-contining () the s-s-sequ (4) (*He looks to me for help and I read* *sequence again.)* by shading more squares.	34	(October 2001) Did not request help with reading any items. Started reading at line of text by item number when text started above this. 'Continue' and 'sequence' found hard to read. However, he could indicate on diagram what numbers had to be shaded. Lee gained four marks out of five for the oral items showing that he responded well to teacher-read instructions. The remaining 18 marks were from the written items suggesting that he gained enough meaning from the text / diagrams to obtain an above-average score. However, this does not mean that he would not have benefited from further reading support as his reading of item 14 suggests.
K-G-L2B (EAL) Stephanie	The number in the (4) sh-shadow squares makes a square s-s squares. Con (2) can (3) con (1) con (1) tin (*She looks to me for help. I ask her* *to try again.*) contin the squares by shading (*pronounced with soft 'a'*) more squares.	38	(July 2001) Did not request help with reading this. Words 'shaded', 'sequence' and 'continue' caused particular problems. Stephanie requested help with reading nine items. It would appear from her reading of this item that she would have benefited from further assistance. She was awarded L2C in maths and missed attaining L2B by two marks. Lack of reading support may have contributed to her failure to do so.

Q-B-L2B Marius	The number in () the () sned shed shed sheded question (1) swe (2) se () seconds Count the () seconds by seding () s- s-sedding more (*next word unclear*)	41	(September 2001) Did not ask for help with reading. 'Shaded', 'squares', 'continue', 'sequence' and 'shading' caused particular reading difficulties. Marius tried hard to make sense of the words but they were too difficult for him to decode. Marius asked for help seven times in total and gained the minimum number of marks to obtain L2B in mathematics. However, he gained four of his 14 marks from the five oral items showing that he responded well to this mode of testing.
P-G-L2B Jade	The numbers in the (1) shad-ed squares make a sentence (5). *(Looks at me with puzzled* *expression. I ask her to carry on.)* Counting the questions by shading more squares.	20	(October 2001) Asked for help with reading 'sequence' only in live test. Seemed to realise that first sentence did not make sense but made no attempt to self-correct. Read 2nd sentence fluently but misread 'continue' and 'sequence'. Jade asked for reading support for two items only and missed the award of L2B in mathematics by three marks. She may have gained more marks had maximum reading support been available.
N-G-L2B Sharon	The number in the (4) sheded square make a (3) squince. Contin the sqince by sheding more squares.	19	(October 2001) Asked for help with reading. 'Shaded', 'shading', 'continue' and 'sequence' found hard to read. Explained that she had to shade squares but did not know which ones. Could not, however, read 'shaded' or 'shading' within the text. Sharon requested help five times only during the test, the last being for item 17. She gained 8 marks in total, the lowest mark to attain level 2C.

N-B-L2A Jaspreet	The number in the sh-shedded –sh- (1) shaded () question (4) [Says he doesn't know that word.] make a s- s-(1) se-*sequence (Mispronounces slightly then say he doesn't know that word.)* Sequ-sequ- (7) sequence by shading *(pronounced with soft 'a')* shading *(now pronounced correctly)* more squares. *(He omits 'Continue the'.)*	58	(October 2001) Asked for help with reading 'sequence' and 'squares' only. 'Shaded', 'shading', 'continue' and 'sequence' found hard to read. Jaspreet requested help four times only during the test and missed being awarded level 3 in maths by two marks.
P-G-L2A Amber	The numbers in the shaded squares make a (2) sentence. Continue the sentence by shading more squares.	11	(October 2001) Did not ask for help with reading but gave the correct answer, probably because diagram alone gives a sense of the response required. Although she misread 'sequence' as 'sentence', she read confidently and fluently. However, she may have made similar but more crucial errors in other items since she asked for help only twice during the test. She gained 13 marks giving her a level 2C, regarded as below-average. Had she gained 14 marks, she would have been awarded L2B, average.

Interpretation of item 14

Almost half the children in the dataset asked for help with reading this item (see Table 5.3 on page 139) and the transcripts above show that readers across the range of level 2 had difficulties. Of the five readers at level 2B, Jade appeared to be the most fluent but still made reading errors. Children reading at level 2A also varied in fluency but had difficulties. Predictably, children reading at level 2C had most difficulty. The words *shaded*, *sequence*, *continue* and *shading* caused most problems, with *squares* also not recognised by some children. Two children misread *sequence* as *sentence*. Both words have a similar structure but *sentence* is likely to be more familiar through written work in English. I could not tell whether children who did not request reading assistance during the test guessed what to do or struggled to read the text. In the former circumstance, they might have guessed correctly since the diagram offers prompts. In the latter circumstance, the evidence shows that children reading at this level were unlikely to gain holistic meaning from the text and to waste time and mental energy to no avail. Further examples of children reading this item are available on audio-tape but this selection typifies the hesitations and miscues that were made by children reading within level 2.

Table 5.13: Transcript of Y2 children reading item 17

Text from item 17: Ann measured the height of these two dolls in blocks. (See Appendix 4.1 for artwork.) How many blocks **taller** is the large doll?

Child	Text read (verbatim)	Time (secs)	Comments
K-B-L2C (EAL) Priya	Ann (5) m-m-mipiger (2) the (2) spoiler that two ()do () in () bocks. Bocks min- mind () bocks t-taller in the lard (2) d-doll.	43	(July 2001) Reading stilted throughout. Priya struggled to decode the words and could make no sense of the sentences. His intonation was monotone with no apparent awareness of meaning. Priya did ask for help with reading this question and the transcript shows the extent to which he needed it. Gained the mark for the correct answer.

N-B-L2C Scott	Ann (1) m-(14)-(*Says he doesn't know next word; I read it.)* [measured] the height of these two dolls in blocks. How many blocks (1) taller is the large doll?	33	(October 2001) The speed of reading was considerably extended by the ME word, 'measured' that would be seen rarely in standard texts read by Y2 children. Scott asked for no help with reading during the test, which is surprising since he was a below-average reader. He took the test with about 13 or 14 other children of different reading competences and his failure to ask for help may have gone unnoticed in such a large group. He was awarded L2C in the maths test
K-G-L2B (EAL) Stephanie	*(She starts at 2nd line of text.)* How many blocks t-taller (1) taller taller *(read three times with 'a' as in tally*) is the longer (2) doll? *(I ask her to read first line.)* Ann m-m-mestured mestured the sheet () height () of () these () two () dolls in blocks. How many blocks taller *('a' as in tally)* in () the (1) longer doll?	41	(July 2001) Missed first line of text completely. Had problems with key word 'taller'. Could not read key word 'measured' 'Longer' incorrect but sensible substitution, probably prompted by diagram. Lacking throughout in appropriate intonation. Gave answer of 17 blocks i.e. height of taller doll. Stephanie requested help with nine questions. It would appear from her reading of this item that she would have benefited from further assistance. She was awarded L2C in maths and missed attaining L2B by two marks. Lack of reading support may have contributed to her failure to do so.
N-B-L2B Joshua	Ann *(He mouths the next word silently.)* (7) mish *(Looks to me to read 'measured' ,which I do.)* [measured] the (1) head of () the two dolls in blocks. How many blocks () tall is () the lady doll?	34	(October 2001) Asked help with reading 'measured' only. Could not read 'measured' or 'height', both key mathematical terms. 'Lady' a sensible substitution in this instance. Read hesitantly throughout. Gave answer of 17 indicating that he had counted how tall (rather than taller) the larger doll was.

			Error probably due to misread. Joshua asked for help with reading eight times. Of the 15 marks he gained, four of these were from the five marks for the oral questions always read by the teacher.
P-B-L2B Jade	Ann () measures () the () height of these () two () dolls in () blocks. How many blocks () taller () is the large doll?	15	(October 2001) Asked for help with reading this. Afterwards, Jade's teacher told me that she now regarded Jade as reading at level 2A. Read mainly accurately but not fluently due to brief hesitations between words. Jade asked for help twice only during the test five months earlier and gained L2C in maths. She missed being awarded L2B by three marks. Her reading had improved since taking the test so it is not unreasonable to assume that she would have benefited from more reading support in May.
N-B-L2A Daniel	Ann meas-(1)ured () the height () of these () two () dolls in block-s. How many blocks taller () is () the large doll?	14	(October 2001) Daniel did not ask for help with reading this. He read accurately but his fluency was punctuated by short pauses. This was the case in all the items he read to me and he made occasional errors, e.g. he could not read 'sequence' in item 14. Daniel asked for help twice during the test but scored the minimum number of marks to be awarded L2A.

Interpretation of item 17

About one third of children in the data set asked for help with reading this item (see Table 5.3 on page 139). Like item 14, some children may have been tempted into responding by referring to the diagram only. However, the item was text-dependent.

The words most likely to cause reading difficulties were *measured* and *height*, both ME words that would be seen rarely in ordinary reading materials. As for item 14, children reading at levels 2C had most difficulty but children reading at levels 2B also failed to read accurately or fluently. Daniel and Jade, whom her teacher assessed as reading at level 2A at the time of her interview, showed more accuracy but were still hesitant. Joshua made the unfortunate mistake of reading *tall* instead of *taller* and therefore did not recognise that a comparison was needed. Most of the children in this selection were unlikely to have gained meaning from the text in unmediated reading.

In both items 14 and 17, even the poorest readers could generally read simple words met commonly in texts, e.g. *in*, *that*, *the*, *doll* etc. but the skills of children reading at level 2 were generally not sufficient to cope with some of the vocabulary, some of which was mathematical English e.g. *sequence* and some ordinary English, e.g. *continue*. As the transcripts of children's reading show, struggling to read one unfamiliar word independently could take many seconds, was rarely successful, and, presumably, had an adverse effect on the overall meaning of the text. An item is only as easy to read as its hardest part, even if this is a single word. I empathised with the belief that:

> In some subjects, skipping a word here or misreading a word there will not have a damaging effect on overall comprehension, but in mathematics it may alter the meaning drastically. (Harrison, 1980:124)

An overview of children's reading of the items

Like Lupetti, Sainsbury & Schagen (1995), I found that the children made 'valiant attempts to read' (p13) but few were fluent readers. In my review of the literature in CHAPTER 2, various studies (for example, see Committee of Inquiry into reading and the use of English, 1975 and Harrison, 1980) used terms such as 'frustration level' to describe the struggle children had to read texts that included too many unfamiliar words. The majority of Y2 interviewees reading at levels 2B and below and some reading at level 2A, found at least some of the text too difficult and were frustrated in their attempts to read. If the ‘5% rule’ (Harrison, 1980:29) were applied, i.e. making more than one uncorrected error when reading 20 continuous words, an indication that text is too difficult, most children showed evidence that they came into this category. Not only that, the ‘true difficulty of reading a mathematics text is greater than that of merely comprehending the prose’ (Shuard & Rothery, 1984:85) since children have to respond mathematically, not just retrieve information from the text. In the review of the literature, I also sought studies on the speed of reading (for example, see Dale & Chall, 1948; Klare, 1963; Wilkinson, 1980 and QCA, 2001a in the section ***Speed of reading*** (see page 16). Few children whom I interviewed showed evidence of reading at the ‘optimum speed’ for comprehension (Gilliland, 1972:13) or of maintaining ‘the normal pace of reading … without evident pause for spoken or silent working’ (QCA, 2002:24), the criterion used to identify words that children recognised on sight. Gilliland (1972) observed that ‘a poor reader may have such difficulty with word recognition that factors involved at the sentence level are irrelevant’ (p98). It would appear that most of the children reading at level 2B or below were so intent on decoding words that they would rarely if ever be able to consider sentences as units of meaning.

I archived the audio-tapes of children whose reading is not transcribed here. These include children reading at levels 2C, 2B and 2A and focus on items for which children did not request help with reading. Having listened to all of them, I have included a selection representative of most of the children but not including the least competent. In a small number of instances, I found it impossible to transcribe scripts of some children reading at level 2C since their efforts consisted mainly of grunts and sounding out individual letters, in some cases with no breaks between words. Any tapes in the archive are available for readers who require further evidence of children’s reading difficulties.

Conclusions

Answers to research questions

Collectively, the three main sources of data in this chapter, i.e. the analysis of the booklets, diagnostic interviews with children and interviews with teachers contributed to answering the following research questions:

Is the reading demand of the text in a typical KS1 mathematics test appropriate for the target children in year 2?

Is teacher support with reading available to children during the test whether they ask for it or not?

What is the extent of the help needed by children of different reading competences during the test?

The analysis provided enough evidence to convince me that the reading demand of the 2001 mathematics test would have exceeded the reading skills of many Y2 children. Evidence that it was not appropriate was shown by the number of requests for help by individual children and for individual questions during the administration of the test, particularly by readers across the range of level 2. I also felt that I could claim that the analysis was an effective instrument in providing evidence of the extent of the help needed by children of different competences during the test. In this respect, what surprised and concerned me most were the variations in the frequency of requests for help with reading by children of comparable reading competence within and between schools. There was limited evidence from interviews with teachers that differences in administrative arrangements may have contributed to these variations but I was concerned that teachers did not always feel that children asked for appropriate support. This suggested that not all children were getting the support to which they were entitled and it could be inferred that appropriate support was not always available whether it was requested or not. Indeed, the level of reading support appeared to have been minimal for some children, even when requested, because teachers sometimes only read words that children indicated and left children to read the remainder. Consequently, I accepted that the analysis may have understated children's reading difficulties or lack of support because it was too complex to code for whether whole or part questions were read to children. The most compelling evidence, however, came from the children themselves. The samples of their reading in Tables 5.7 – 5.13 reveal the scale of their reading difficulties and the extent to which their reading should have been supported. Homan, Hewitt & Linder (1994:350) refer to research by Drum, Calfee & Cook (1981) and paraphrase the view of the latter that 'when a child is struggling to recognize words, there is diminished attention to the content …'. However, children did struggle to decode words, often chose not to ask for help and may have lost marks, or, worse, have gained a lower mathematics level because most items were text-dependent. For such children, there will always be doubts about whether lack of reading skills or teacher intervention had a detrimental effect on their test score. Not only that, an absence of underlining did not necessarily mean that children could read the text independently. Some of the interview transcripts with children earlier in this chapter show that children could not always read items for which reading support was not given.

I recognised that, compared to the national cohort of about 600 000 KS1 children who take the mathematics test, my study touched the lives of very few. Nevertheless, I have tried to represent their voices for the first time through their involvement. I believe that the children who kindly assisted me were typical of their peers nationally. I make this claim for several reasons:

i. all headteachers in maintained schools nationally are required to implement the national curriculum and the statutory assessment arrangements;
ii. all children in my study attended maintained schools, had followed the KS1 programmes of study for mathematics and English (of which reading is one

component), and been subject to the KS1 statutory assessment arrangements for mathematics and English;

iii. the children came from schools in various catchment areas.

Having acknowledged that the results are not generalisable for the national cohort of Y2 children, I still hope that I have convinced the reader that my concerns about the reading demand being excessive were justified. As the data evolved in this chapter, my concerns grew proportionally. I felt a sense of unfairness:

i. that children of comparable reading levels within and between schools received such varying amounts of help;
ii. that children were given the responsibility of knowing when to ask for help;
iii. that literacy skills may unintentionally have been assessed to the possible detriment of mathematical skills;
iv. that children who were competent readers were probably advantaged since their reading skills would have allowed them easier and quicker access to the mathematics.

In a mathematics test, I did not think it unreasonable to expect that there should be a 'level playing field' to make the test fair for all participating children, i.e. that all of them should have an equal opportunity to know what items were asking them to do. Only in this circumstance could the test be fair or valid. To be valid, '[t]he test tests what it claims to test' (Clausen-May, 2001:4). and '[i]f the sentence structure and layout of questions used in a test make it difficult for some pupils to understand what is being asked, then it is not a valid measure of what it purports to measure' (Clausen-May, 2001:35). Where children could not read items and were not offered reading support, the assessment of their mathematical competence was, therefore, not valid. Not only that, there is a statutory inclusion policy for all children participating in the National Curriculum, e.g. teachers must 'make provision, where necessary, to support individuals and groups of pupils to participate effectively in …assessment activities' as quoted in Section B of *Providing effective learning opportunities for all pupils* (QCA, 2004). Some teachers appeared to have been unwittingly naïve in this respect and made assumptions that children could read independently or would ask for adequate reading support. Teachers' judgments about children's competences to read independently would almost certainly be linked to performance in reading schemes but in *A language for life* (Committee of Inquiry into reading and the use of English, 1975) the authors found evidence of children 'who had made good progress through a scheme and were now struggling at frustration level in other kinds of reading …' (p113). My evaluative judgment now was that the reading age required to read the test independently and accurately exceeded the reading competence of most children who were 'rising' 7-year olds.

To the best of my knowledge, no research had been undertaken previously that has gathered data about administrative arrangements that could affect the quality of reading support. Nor have data been gathered about the frequency of requests for help with reading during a KS1 mathematics test, or about the reading competences of the participants. The latter were 'naturally occurring data' (Silverman, 1993:208) whose occurrences had never been drawn to the attention of the test developers. Indeed, in the rigorous test development procedure, as explained in CHAPTER 1, views of professionals, but not of the children, are sought on the appropriateness of the reading. The children themselves must surely be the best judges of the accessibility of the text. The evidence from their transcripts showed that professionals, including myself, made errors of judgment in not recognising potential reading difficulties in a context where children were expected to know when to ask for help or chose to read independently.

Evaluation of the data

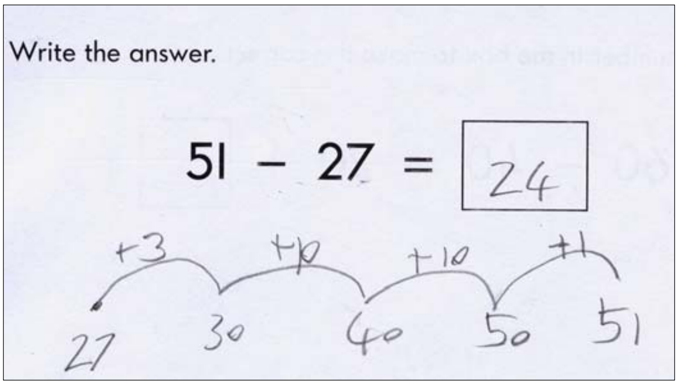
In general, I was satisfied that the data from the test contributed to answering some of my research questions. However, I had some concerns that I wished to address in later chapters:

i. Interviewing children five months after the statutory tests, which were taken in May 2001, may not have given a true picture of children's reading competences at the time of the test. Therefore, I decided to interview more Y2 children reading from the same test in June 2002, a more appropriate time to evaluate how well they could read test items.

ii. I chose to concentrate on interviewing children reading at level 2 for this phase of the study. Since I wished to establish the reading age at which children could read the text fluently and accurately, I needed to include some children reading at level 3 to compare them with those reading at level 2 since the evidence so far suggested that children reading at level 2A were still having some difficulties with the text.

In CHAPTER 7, therefore, I address these concerns through further data collection and analysis. However, first, in CHAPTER 6, I consider another form of administration that is permissible but removes the responsibility for reading from the children.

CHAPTER 6



Shanice was working at level 2B in mathematics.

CHAPTER 6

Look-listen in the KS1 2001 mathematics test

Other modes of administration of the test

Look-listen

In the search for teachers willing to participate in my research, I found that, out of the 17 teachers I contacted, several were intending to alter their mode of administration of the 2001 test to the look-listen mode for at least some of the children. In look-listen mode, the teacher reads all the text for each question, with the child generally following the text with the teacher, thus removing from the child any need to request help with reading. Unlike the hands-up mode, the child is not expected to read independently or to know when to ask for help. Some of my contacts had already made this change pre-2001. Indeed, at least half the teachers I contacted stated their intention to use look-listen to some extent. Perhaps I should not have been surprised since, as I stated in my literature review, Lupetti, Schagen & Sainsbury (1995), in their study of the reading demand of the KS1 1995 mathematics test, found that seven of the eight children regarded as poor readers had had the whole text read to them on an individual basis. Teachers and researchers alike believed that these children would 'not have been able to achieve Level 2 without most of the booklet being read to them' (Lupetti, Schagen & Sainsbury, 1995:10). To see if 'look-listen' teachers had similar views, I interviewed 10 staff in seven schools and sought the same information as I had from the 'hands-up' teachers. The 'look-listen' staff shared one common viewpoint, i.e. that the change to look-listen had evolved because of concerns about the reading demand of the test.

Interviews with teachers using look-listen

Data collection

I interviewed and coded transcripts for teachers from Schools L, S, T, U, X, Y and Z respectively. I used the same codes on the teachers' transcripts as in the preceding chapter (see Appendix 5.3) to inform my commentary below. I include a sample transcript from School U in Appendix 6.1. Each teacher had a slightly different story to tell and a summary of each story follows in the next section.

Teachers' stories of look-listen

Staff in School L used the look-listen method for the first time in the 2001 mathematics test. Twenty-one of the 36 children in year 2, consisting of children reading at levels 2A and 3, took the test in the traditional mode and had to request help with reading. The 15 remaining children, reading mainly at levels 2C and 2B, were split into two groups and had each question read to them as a group. The teacher had changed administrative arrangements because of her belief 'that the majority of children just don't bother' to ask for help, are 'too embarrassed' or 'feel that they can read it all' even though 'there are a lot of demands with the reading'. Indeed, she felt that even children reading at level 3, who read independently, may not have asked for adequate help and that their scores could have been higher if they had done so. In contrast, she felt that the look-listen mode had benefited the less competent readers and that it made 'a difference between a level with a lot of children because I read it all for them …'. She was influenced by other teachers who had used look-listen and decided to 'use it with some and not with others and see what the differences were'. For future years, she stated her intention to read all questions to all children in small groups, regardless of reading level, 'because I think they should all be given the same opportunity'.

In School X, the teacher gave similar reasons for moving to look-listen. She believed that:

> The reading demand is very high. Not only are a lot of words too difficult for them to read, but there's too many words. The children don't even bother trying to read it. They just guess what they have to do.

In the 2000 test, she had used look-listen for poor readers, as did the teacher at School L in 2001, but 'the rest just got on with it on their own'. In the 2001 test, she used look-listen exclusively and administered the test to groups of up to seven children at a time. She grouped the least confident children into fours and 'tended to read the question several times to them … because their concentration isn't quite so good'. The children were also given a break 'after 50 minutes …because their concentration was waning' before doing the final few items. Like the teacher in School L, she believed that results had improved and 'were better than expected' compared with results from trial tests administered in the preceding few weeks using the hands-up method. Having learned from the experience of look-listen in the 2000 and 2001 tests, she thought that her only modification for the 2002 test might be to put children that 'were notoriously slow together' so that others were not disadvantaged by lack of pace. She was committed to using look-listen in future years because she felt that, since independent reading was not involved, the children 'had a maths result which was a true reflection of their mathematical ability'.

School S was the only school in my study where there were two Y2 classes with children set by 'ability'. In the class with the lower set, children's reading levels ranged from W to 2B with most children reading at levels 1 and 2C. In the class with the upper set, reading levels ranged from 2C to 3, with most children reading at levels 2B and 2A. Each teacher administered the test to groups of four children at a time but used a different approach. In the lower set, teacher LA used look-listen with all children since she believed that the reading was 'demanding'. She explained how she started off by reading the same questions simultaneously to the four children in a group, but, because of the uneven pace of working, she ended up reading questions to suit the pace of individuals. She felt that individualised reading in this way would be ‘unworkable’ with a larger group. Although admitting that the administration was difficult, she was adamant that ‘she did not want any child to be penalised for their lack of reading when it’s a mathematics test’. In the upper set, children had to request help with reading even though some were reading at the same level as those in the lower set. Furthermore, their teacher, UA, felt that ‘some were looking at the questions rather than reading what was asked of them and they went ahead and did what they thought they needed to do’. However, she ‘knew which groups would need more help with reading’ and supported them accordingly. I found myself wondering why, with support staff available to free both teachers to work with small groups, both had not adopted the look-listen approach, particularly in light of the reservations of teacher UA. An opportunity had been missed particularly since the upper set included average and below-average readers, many of whom had English as an additional language.

In School T, the teacher considered that the reading in the test was ‘demanding’ because of the ‘new words that come up in the stories’, i.e. contextual problems and ‘the actual mathematical reading’. Originally, she administered the test to the whole class and children requested help with reading. However, she believed that some children did not ask for the help they needed and that others ‘were all asking for help at the same time for the same question’, resulting in the teacher ‘running round like a chicken; it was just inefficient’. A similar point was made by the teacher in School N when discussing the hands-up approach, who said ‘it’s a nightmare because you see about six children with their hands up and who do you go to first’? In 2001, VM at School T administered the tests to two groups of about fifteen children but minimised the need for children to ask for help. One group comprised the more fluent readers. She used different strategies to support these children, e.g. by going through the test in advance and reading the more difficult text or by reading a double-page spread at a time. For the second administration with the less fluent readers, she either read a double-page spread at a time to all children or ‘with the very slow ones, we read literally every question'. She did not make assumptions about children's reading competence, believing that 'most of them are disadvantaged by the reading demand'. In common with other look-listen teachers, she wished to distinguish between children's mathematical and reading competences, stating that 'you don't want them to get it wrong because they haven't grasped the reading'.

In School U also, the teacher felt that most children 'can't understand what they are reading' in the mathematics test. She made different administrative arrangements for groups of children depending on their reading competence but she grouped children whom she thought would work at about the same pace. Through using KS1 mathematics tests from previous years as practice, she identified a group of seven children out of the 33 involved whom she felt could cope confidently with the reading and the mathematics. The first administration was with this group. She read the first few questions with them then left them to continue since some children asked to carry on at their own pace. For the others, comprising groups of four to five children for each administration, she read all the questions by holding up a test booklet 'so they don't have all the words in front of them.   I have got my finger on the actual sum they are doing and they follow the words with me … they feel more confident'. She could also reread questions, if necessary. Each administration was carried out in a small room with the relevant group 'and it is very cosy for them'. She started using look-listen to administer the KS1 test after finding that she had to use this mode of delivery with children taking the KS2 test because of reading difficulties.

The administrative arrangements in School Y had some similarities to those in school U. For example, the teacher administered the test to her 'brightest group' of mathematicians first and supported their reading where required. In doing so, she became familiar with the content and text of the test. She then used her knowledge of the test to administer it to small groups 'so that they have got my attention …and for some children it was a case of reading the questions with them, every one, because they just could not read any of it'. By that time, she had identified the questions that she 'felt they would be confident in answering first' and she 'guided the children to those questions and read them out'. She then asked them to go back to attempt the remaining questions whilst she supported reading. She justified this way of working by explaining that there were questions towards the end of the test that she thought children could attempt but would never have reached without her direction because they were so far into the test. She did not feel that she gave them an unfair advantage because she knew what mathematics had been covered in class work, directed the children to questions whose content was likely to be familiar and read the text to them. She considered that she provided an exam technique that children could not be expected to acquire until they were older.

Teacher CB in School Z felt that the reading demand of the 2001 test was 'quite high and, consequently, the teachers 'had to read it to the majority of the children'. There were two Y2 classes, but, unlike the classes at School S, they were not set. However, children were set for the test. For those who were '[b]etter at reading and mathematics – the top half of both classes', the teachers started off reading the questions but then left the children to get on and to request help with reading. A similar approach was used with the middle set. With the 'bottom group', children were further divided into 'very small groups' and one question at a time was read to them. After each block of two or three questions, 'they'd have a little break …and we did that with the very, very poor ones'. As Teacher JL explained, '[S]ome of them have got such limited concentration. That's the only way you can get through them (the questions) because they are used to doing small chunks of work, not working through great big long booklets'. The 'middle' and 'lower' groups were also given a break during the test.

Teachers' views on the reading demand

Like the teachers who contributed to the analysis of the test booklets, as described in the previous chapter, the look-listen teachers also felt that the reading demand was too high. However, they reacted to this concern in a different way by reading all questions to at least some of the children. Although the fine details of management of the test varied from school to school, there was general agreement that look-listen was suitable for the least competent or least confident readers but teachers in all but one school felt that their most competent readers did not need that level of support. Indeed, where attempts to use look-listen with more confident children had been made, e.g. by teachers reading the first few questions with them, children often wanted to move at a pace faster than that dictated by the teacher. Nevertheless, the most common administration in these schools was to use look-listen for children who were often described in terms such as 'least able' and for the most competent readers to request help with reading.

Modes of administration

As with teachers using the hands-up approach, each teacher using look-listen had slightly different administrative arrangements. The majority chose to let children judged to be competent readers work independently but to use look-listen with the others. Only in two of the seven schools did all children have the items read to them. For look-listen, teachers most commonly had groups of from four to seven children with the poorest readers usually being grouped together for maximum support. School T had the largest grouping with 15 children at one administration.

None of the teachers expressed any reason why they would wish to discontinue using look-listen. Teachers at Schools L and X believed that test results had improved due to look-listen but a particular advantage, raised by teachers in Schools L, S, T and X was the issue of equality of opportunity for the children. They believed that it was only by using look-listen that results in the tests would be, as the teacher in School X said, ‘a true reflection of their mathematical ability’.

Support for children

Hearing all items read was the main form of support for children who were considered to lack the reading skills to work independently. Teachers suggested various other strategies that supported children. These included:

i. giving children short breaks during the test (Schools S and Z);
ii. reading questions more than once for the least competent readers (School X);
iii. directing children to and reading items that were considered most accessible before allowing children to attempt the remaining items (School Y);
iv. allowing children to work at their own pace but reading each item to individuals (School S).

Interpretation of the look-listen data

I thought it reasonable to conclude that look-listen solved many of the problems caused by reading difficulties that I discussed in detail in the preceding chapter. It did however raise further concerns about non-standard forms of administration to add to those I raised at that time. Even in this group of teachers, there were differences in administrative arrangements, though their common link was look-listen. For example, the size of the groups and the way questions were read, e.g. a double-page spread at a time or by the teacher pointing to the words on her copy of the test booklet, varied from school to school. In some schools, children were also given at least one break during the administration. Even with this mode of administration, it was possible that some children had more advantageous test conditions than others. I did wonder also about whether children who were excluded from look-listen and had to ask for help with reading may have been disadvantaged for reasons discussed in the preceding chapter, i.e. that children who had above-average reading skills did not always ask for appropriate help.

Views of 'hands-up' teachers on look-listen

Advantages and disadvantages

Since I was aware of the shift towards the look-listen mode of administration through my contact schools, I also sought the views on look-listen of the teachers in Schools K, N, P and Q who asked all children to request help with reading. These teachers featured in CHAPTER 5. I categorised their views into advantages and disadvantages. All teachers agreed that look-listen would advantage children who had reading difficulties. In School N, these were considered to be children reading at level 2C and at least some at level 2B. The teacher in School Q felt that another advantage would be children's improved concentration 'if they are all listening to what you are reading out'.

Only one of the teachers felt that look-listen would advantage all children. The others felt that the main disadvantage would be a delivery that lacked pace for the most competent readers. A typical comment came from a teacher at School N who said, '[T]he more able children get frustrated because they can't work at their own speed. …So I'm not convinced it is the best way to do it for them'. The only other disadvantage raised was the question of staffing. Teachers at Schools P and Q felt that staffing would have to be generous because look-listen lent itself to small group administration. Indeed, the teacher at School Q felt that administering look-listen, even though she would prefer it to the hands-up approach, would be 'problematic' because, as she explained:

> I haven't got the support and I'd have to probably administer the test four or five times to five different groupings and that would have implications. I would have to have different teachers taking the rest and we just have not got the staffing or the money.

Two teachers at Schools N and Q raised the issue of fairness. In the words of Teacher WK in School N:

> It's not a fair test, is it, due to it being administered in so many different ways? It should perhaps be more clearly defined how we should administer all tests, really.

I contacted these four schools again a year later to review how they administered the KS1 2002 mathematics test. School K had introduced look-listen for all children. School N had introduced look-listen for less competent readers, approximately half of the Y2 cohort. Teachers in Schools P and Q had kept the same mode of administration. However, the latter school continued to have staffing difficulties that prevented small group administration that the teacher felt was essential for look-listen.

In CHAPTER 5, I provided evidence that

i. the reading demand was excessive for many children;
ii. teachers were not always confident that children asked for appropriate help;
iii. teacher support was not always available since some children could not read essential text for which they had not requested help.

In this chapter, I have reported on teachers' views on the look-listen approach that they adopted to ensure that children could concentrate on the mathematics. In CHAPTER 2, where I reviewed appropriate literature, I did not think to read up on the effect of look-listen in test administration since, at that point, I was unaware that this approach was in such common usage for the KS1 mathematics test. Since I considered its effect to be relevant to answering my research questions, I retrospectively sought literature that had considered this mode of testing.

Literature on look-listen

Review of the literature

Literature on different approaches to test administration was scarce considering the high profile of test results such as those at the end of key stages. Of the research undertaken on national assessment in the public domain, none appeared to have considered oral delivery of the written mathematics test items to KS1 children or its impact on children's results. This was all the more surprising since most KS1 children are unlikely to have developed fluency in reading text with specialist or unusual terms, an issue that was discussed in detail in CHAPTER 2 and CHAPTER 5. This is further complicated by the fact that, where a child gets an incorrect answer, it is unclear whether it is the reading or the mathematics that is too difficult. Indeed, Newman (1977:252), whose study was discussed in CHAPTER 2, found that about 12% of errors made by eleven-year olds, in a mathematics test designed to be easy to read by nine-year olds, arose because children made errors reading key words. The text in the KS1 mathematics test is not rigorously controlled, e.g. by applying measures of readability, and the evidence I obtained in diagnostic interviews with children, discussed in CHAPTER 5 and later in CHAPTER 7, showed that they did not find the text easy to read. Consequently, I felt inclined to argue strongly in favour of the look-listen approach but hoped to support my argument with research literature. I found only two relevant studies that considered look-listen in test situations. The later of these dates back to 1986 and, in the intervening period, research in this field does not appear to have progressed to any great extent, at least for children of primary age. Unfortunately, neither study related to mathematics but their main findings are of interest.

In Wilkinson's (1980) study, children in grades 2 (age 7) to 6 (age 11) took two comprehension tests that had vocabulary derived using familiar word lists and answered questions of the *Who?, When?,* and *Where?* categories. Each test was designed to be comparable in reading demand. Children were tested individually and, for one test, they read the text aloud; for the other test, they had the text read for them with the text visible to the children as it was read, i.e. look-listen or in Wilkinson's words 'bimodal reading and listening' (Wilkinson, 1980:563). Children then gave oral answers to the questions. The results were that children had a greater proportion of questions correct in the look-listen mode of testing. Children who read the text for themselves made few errors, but Wilkinson concluded that:

i. the understanding of novice readers who read the provided text 'accurately but laboriously' (Wilkinson, 1980:561) was less than when the comparable text was introduced in 'look-listen' mode;
ii. the understanding of skilled readers, identified as those in grades 5 or 6, was better when they could simultaneously read and listen to a text than when they read aloud a comparable text.

For children who struggled to read the text, Wilkinson's (1980) data suggested that their comprehension was poor even if they could read accurately. Many of the Y2 children in my study, novice readers, read neither fluently nor accurately so comprehension was likely to be minimal or non-existent. For them, whilst such tests remain statutory, look-listen would appear to be the only sensible option since, as Harrison (1980) pointed out, misreading or skipping a word in mathematics text can have a drastic effect on the meaning. For skilled readers too, Wilkinson's (1980) data suggested that the child 'can listen attentively to the examiner's rendition of a passage and can simultaneously scan the passage visually to clarify or review important information' (Wilkinson, 1980:568). This made me wonder if, indeed, look-listen might advantage children taking the test who were above-average readers, which was contrary to the opinion of most teachers whose views I sought. Of course, a piece of running prose such as that used by Wilkinson (1980) that forms the basis of a comprehension test bears little relationship to the content and style of a KS1 mathematics test. Nor is the vocabulary in a KS1 mathematics test controlled using familiar word lists, a matter to which I shall return in my analysis of the booklet in CHAPTER 8. However, it did appear from this study that children at both extremes of reading competence benefited from look-listen. It might even be argued that the most competent readers have the best of both worlds i.e. access to spoken text and more advanced reading skills to cope with the written text. In the mathematics test, teachers may have judged that children could read independently based on their general reading skills and under-estimated those required to read mathematical text

The study by Homan, Hall & Topping (1986) considered look-listen for administration of a test other than for comprehension of running prose. Their view was that '[t]he readability level of a test item can contaminate results' and that 'surprisingly little research has studied the effect of reading of test items on good and poor readers' (Homan, Hall & Topping, 1986:363). They did, however, describe findings from a limited number of studies that were peripheral to my research interests. For example, they distinguished between two approaches to the oral presentation of test questions. First, they described the term 'auding' as referring to 'the type of listening that requires listening comprehension' and explained that '[w]hen test items are read aloud to students, auding rather than listening is required' (Homan, Hall & Topping, 1986:363). They argued that children spend the first few years of their life listening rather than reading and, therefore, their auding skills in those years will be superior to their reading skills. Indeed, their research led them to believe that auding skills were 'superior to reading skills through to sixth grade' (Homan, Hall & Topping, 1986:363). Second, they referred to the look-listen approach where 'the student experiences advantages from both auding and reading' (Homan, Hall & Topping, 1986:364). These researchers set up a study in which they compared the look-listen approach with the silent reading approach in a common test on economic understanding. The test comprised multiple-choice items and was administered to 4400 fifth-grade pupils in Florida who had below-average, average and above-average reading skills (described as level 1, level 2 and level 3 readers in the study). An analysis of their results revealed that 'superior performance was associated with the teacher read treatment over the student read treatment' and that there was 'no clear tendency for the read treatment to be more effective for one level than another' (Homan, Hall & Topping, 1986:365). As a result, they concluded that all pupils benefited from the look-listen approach and that:

> … teachers can be more confident that reading test items aloud to the low attaining group will not negatively affect the average or above average readers. However, the application of the look-listen approach to the test situation cannot be expected to compensate totally for reading level difficulties in test performance. (Homan, Hall & Topping, 1986:365).

In their review of the language demands of the KS2 mathematics test, Shorrocks-Taylor & Hargreaves (2000) commented on the benefits for good and poor readers discussed in the latter study but felt that the look-listen approach 'could not be implemented on a national scale' (p40) although reasons were not given for such a statement. The 'look-listen' teachers in my study proved that, for KS1 children at least, it is possible to implement such an approach and believed in its benefits. Indeed, I remain convinced that look-listen was and still is in common usage nationally for the KS1 test; the Y2 teachers in my study were not atypical.

If the findings from the studies of Wilkinson (1980) and Homan, Hall & Topping (1986) are related to Y2 children taking the statutory mathematics test, it would appear that:

i. the term auding describes the skill children require to answer the first five questions in the mathematics test where text is not provided and the teacher reads the questions aloud. This would not be suitable for whole test since many questions are in written form because they contain too much key information for children to memorise and process without supporting text;

ii. the majority of children taking the test are likely to benefit from look-listen although more research would be needed on its effect e.g. on cut scores;

iii. the look-listen approach, i.e. the combination of auding and supported reading, may offer the best of both worlds for the written questions.

Conclusions on look-listen

The good and the bad

Undoubtedly, there will be critics of the look-listen approach. Teachers themselves raised concerns that the brightest children, working at the pace dictated by the slowest child or not working independently at their own pace, could become frustrated. However, by setting up compatible groups of children for each administration, this would be less problematic. Linked to this would be cost implications, i.e. the need for generous staffing to release the teacher to work with small groups so that the pace could be adapted to suit each group. Other potential concerns include:

i. logistical issues, e.g. the need for a reader with a clear voice, appropriate speed of reading, good acoustics, absence of background noise e.g. children playing outside;

ii. behavioural issues, e.g. children whose attention wanders during reading or whose concentration span is short;

iii. the assumption by teachers that children who are left to read independently, because they are regarded as competent readers, know when to ask for support.

Most of these concerns could be addressed in small group administrations where each child could be observed throughout and through forward planning. I have already made several arguments in favour of the look-listen approach. Other justifications might include that:

i. when the pace is dictated by the teacher, children will be less likely to omit questions or spend too little or too much time on individual questions;
ii. children’s mental energy, diverted to the effort of reading or trying to read so much text, can be redirected to the mathematics;
iii. it has dual advantages since it gives each child the opportunity to concentrate on the spoken or written version of the text. Poor readers will probably rely on auding clues more than reading clues; confident readers may rely more on the text;
iv. the look-listen approach is not uncommon in classroom practice so most Y2 children would be familiar with it.

As I stated earlier in this chapter, Lupetti, Schagen & Sainsbury (1995) found, in their study of the reading demand of the KS1 1995 mathematics test, that seven of the eight children regarded as poor readers had the whole text read to them on an individual basis and they recommended that ‘attention should be given to approaches other than unmediated reading’ (p74). I consider that the look-listen approach is worthy of attention and of further research into its efficacy for younger children. Although further research would not contribute to answering my research questions, I now have further evidence from the look-listen teachers of more variations in the mode of administration of the test and research evidence, albeit limited, that the look-listen approach would be likely to have advantages over the hands-up approach for most children.

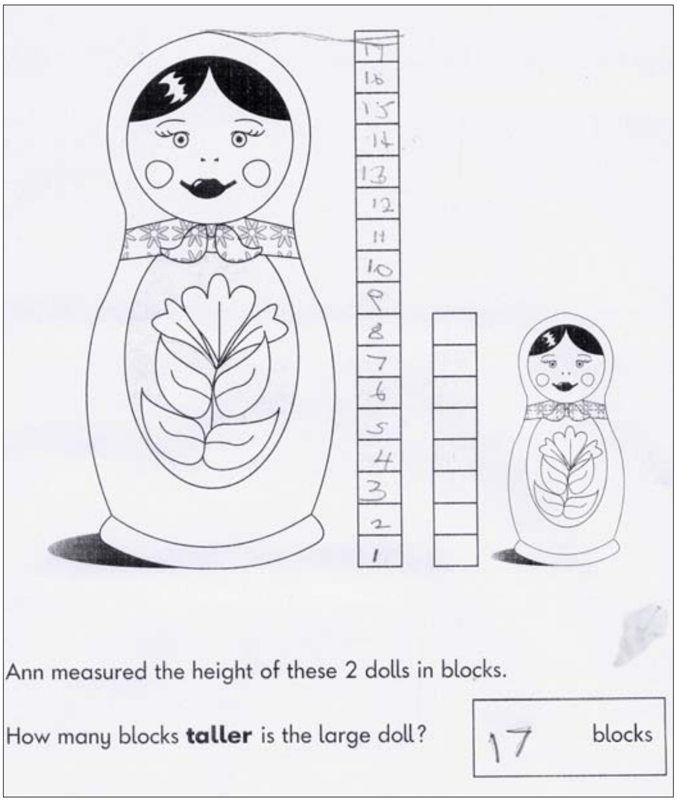
I now return briefly to those of my research questions that are answered, at least partially, in this chapter. *Is teacher support with reading available to children during the test whether they ask for it or not?* and *What are the views of Y2 teachers on the reading demand of the test*?. For children who have the test administered by teachers using look-listen, the answer to the former question is yes since children do not have the responsibility of asking for reading support. For the latter question, all teachers in the look-listen group believed that the reading demand was excessive and that look-listen was the best approach to compensate for children's lack of reading skills. Even amongst teachers who had used the hands-up approach, there was acknowledgment that look-listen might advantage the less competent readers. Indeed, some of these teachers moved to the look-listen approach for poorer readers in the following year.

I now felt that I had sufficient data about teachers' views on the reading demand. Taking into account all the variations in administration used by the 'hands-up' teachers (see CHAPTER 5) and the 'look-listen' teachers, I believed that I was building up a convincing picture of a so-called standard test being administered in many different ways, some of which would appear to advantage some children over others. Of the two modes, I believed that look-listen was undoubtedly fairer for most 6- and 7-year olds since it gave children access to all the text and, thereby, removed the possibility that reading difficulties were an obstacle to the mathematics. I recognise that the look-listen administration is not without its difficulties but these could be overcome by careful planning. The hands-up administration, on the other hand, could never ensure that all children had access to all the text. Only for a test administered by look-listen could I answer in the affirmative to my research question, *Is teacher support with reading available to children during the test whether they ask for it or not?.*

I must stress that none of the teachers using the hands-up approach (see CHAPTER 5) or the look-listen approach were breaking any rules. The teacher's guide allowed for variations in the size of groups for test administration and for the approach used although the text assumed that the hands-up approach was the more prevalent. So, although the test booklet and teacher's guide were standard, the variations in how many children were grouped together to take the test, the arrangements for administering the test and for supporting children were non-standard.

In the next chapter, I move to a further and more comprehensive analysis of children reading from the 2001 test, undertaken in June 2002 to address concerns that I raised at the end of CHAPTER 5.

CHAPTER 7



Amarpreet was working at level 2C in mathematics.

CHAPTER 7

Comparing Y2, Y3 and Y4 readers

Links to earlier research

In CHAPTER 5, I concluded that my evidence, after analysing interviews with Y2 children, showed that children assessed as reading at level 2, but in particular those reading at levels 2C (below average) or 2B (average), were likely to have difficulty accessing text from the KS1 2001 mathematics test, if reading was not supported appropriately by teaching staff. In 2001, 29% of Y2 children gained level 3 in the statutory reading comprehension test (DfES, 2001) in which children read independently. Assuming for the moment that children reading within level 3 were likely to cope with the text of the 2001 mathematics test, this meant that the majority of children entered for the test could find at least some text difficult. Even for children awarded level 3 in reading, the difficulty of the text could be exacerbated by the inclusion of Mathematical English, by contexts shifting from question to question and by the amount of unrelated text, concerns that I raised in CHAPTER 2. Some children, I recognised, were likely to have all items read to them, a phenomenon discussed with teachers in CHAPTER 6, but I had no doubt that many would be expected to read independently.

Data from my pilot study in CHAPTER 3 showed that Y3 children reading at level 3 made few reading errors in the limited number of test items observers heard them read. However, children reading at level 3 are reading with competence comparable to the majority of children in years 3 and 4, with children in year 4 generally reading more securely within level 3. Indeed, as explained in CHAPTER 1, one description of children reading at level 3 is that they can ‘‘read a range of texts accurately and fluently’ (DfEE & QCA, 1999a:56) and independently. This led me to ask a question that I had not considered at the outset of writing my literature review: *What is the reading age required by a Y2 child in order to read the text in a typical KS1 mathematics test fluently and accurately?*

If I could answer this question, I could suggest the chronological age at which independent reading would be unlikely to hinder access to the mathematics. To try to answer this question, I decided on the following research design.

Research design

Timing

I interviewed eight average readers in years 2, 3 and 4 to compare their competence in reading items from the *Key stage 1 2001 mathematics test booklet* (QCA, 2001c). I chose June 2002 to conduct the interviews because:

i. the Y2 children had taken the statutory tests in English and mathematics in May 2002, and I would have access to the levels they obtained reading;

ii. the mathematics test booklet I was using was from May 2001, the year before, and the children in year 2 in 2002 would not, therefore, be influenced by having taken it a few weeks previously;

iii. the Y3 children, who would have taken the 2001 mathematics test live, one year before, were unlikely to remember the content to any extent which might assist their reading;

iv. the years 2, 3 and 4 children were at the same point in the school year.

Choosing children to participate

Through experience recounted in CHAPTER 5, I was now more aware of the time-commitment required to interview children individually then to transcribe and analyse the data from audio-tape. So, as in CHAPTER 5, I followed the advice of Gillham (2000) to keep ‘the number of interviews to a minimum for adequate representativeness’ (p12). Consequently, I contacted the headteachers of four local primary schools in different catchment areas, explained the reason for my research and requested to interview at least two children, preferably of both sexes, in each of the three year-groups. I also asked to interview only children whom their teachers regarded as average readers. I confirmed the purpose of the field work and arrangements by letter and included a consent form of the same design as used in my pilot study (see CHAPTER 3) so that parents or guardians could agree to the interviews. In so doing, I felt that I had met ethical considerations, other than permission from the children to conduct the interviews. I used the same ethical and logistical procedures to interview children as explained in CHAPTER 5. However, unlike my interviews in CHAPTER 5 in which I generally asked children to read items that they had read independently during the 2001 test, I would concentrate on the four chosen items only.

This meant that I would obtain data from 24 interviews for analysis. Even with only eight children from each age group, I felt that the participants would be typical of their peer group nationally. Since such an analysis could be costly in time, I needed to minimise the time involvement without compromising the quality of the evidence I was seeking. One way to do this was to select the same items for each child to read and to concentrate on analysing reading competence. I saw no point in comparing mathematical competence since the years 3 and 4 children would have spent up to two years more on learning mathematics than the Y2 children.

Sample of items to be analysed

I selected from the test booklet (QCA, 2001c) the four items with which Y2 children were most likely to ask for help in reading during the live test in 2001 as explained in CHAPTER 5. I also chose four items that were close to each other towards the middle of the booklet since they assessed different aspects of mathematics, included varied text and diagrams, and were meant to be accessible in mathematical demand to the majority of the children taking the test. These were items 11, 14, 15 and 17 as shown in Figure 7.1 (or see Appendix 4.1 for items in the actual test booklet).

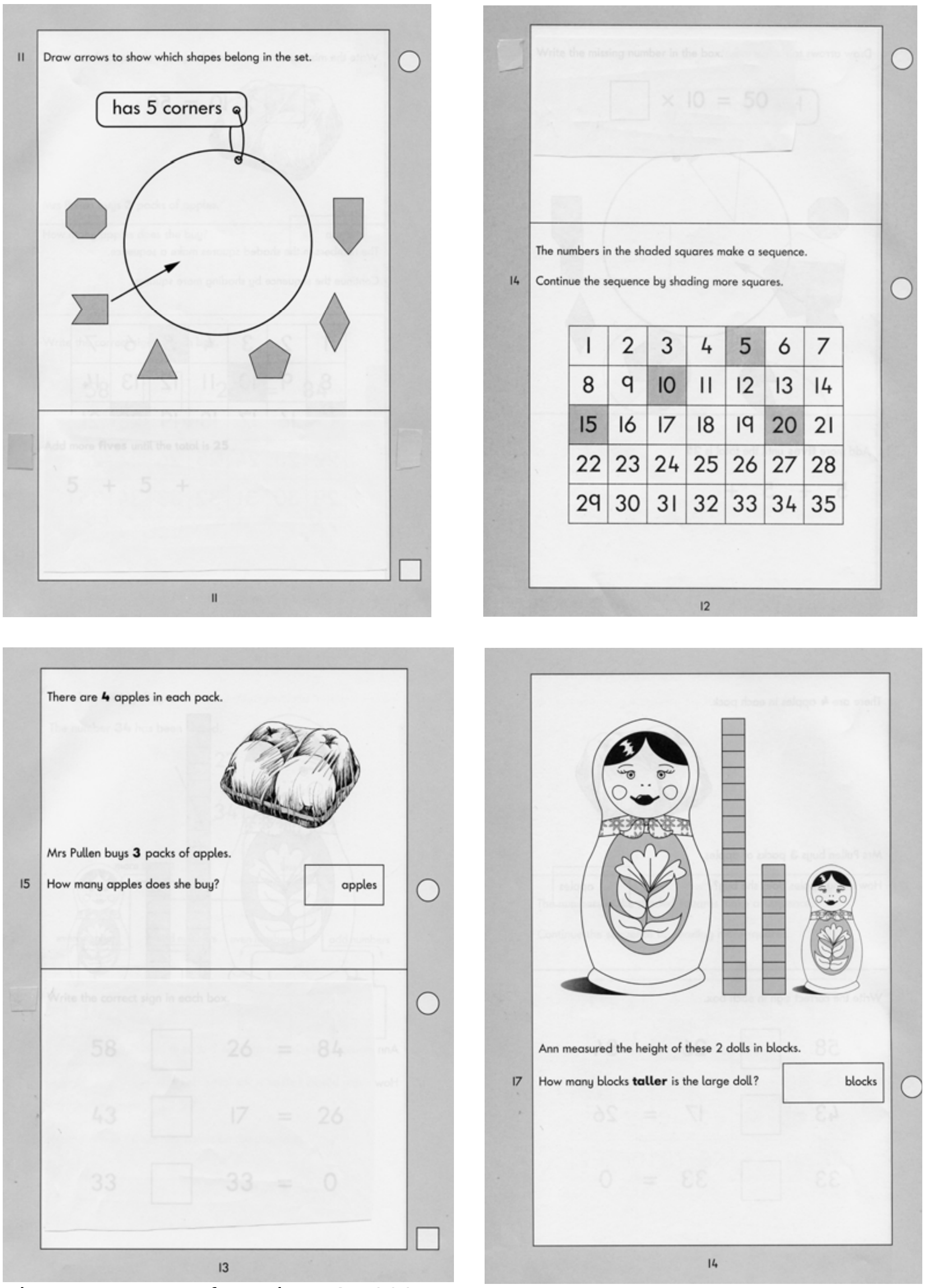


Figure 7.1: Items from the KS1 2001 mathematics test for comparative reading

To get an idea of how long a fluent and accurate reader would take to read the items, I asked three teachers to read the four items to me and to explain what each item was asking them to do to ensure that they were reading for meaning. The average reading time was 6 seconds for item 11 and 8 seconds for items 14, 15 and 17.

Analysing transcripts

Choosing categories for transcript analysis

By limiting my interviewees to average readers and by having set items to be read, I reduced the number of variables I had to manage in the transcription and analysis. As in my interviews with children in CHAPTER 5, I wished to analyse the accuracy, fluency and pace of the children's reading. This I would continue to do but in more detail so that I could better answer the as yet unanswered research question. For this analysis, I was concerned to categorise the reading errors and weaknesses, pauses or hesitations in reading, use of phonics, and the total time taken to read each question. This was because of my concern to reveal the ways in which problems with reading might be masking mathematical performance. I could then compare all these factors within and across the three years groups.

Since the transcripts would even more detailed than in CHAPTER 5, I excluded peripheral conversations with children, e.g. explanations of how they worked out answers, although I would refer to them if they turned out to be important in relation to the reading. This was an acknowledgement that the time taken to transcribe in full and analyse the conversations would outweigh the benefits of these extra data and would be unlikely to help me to answer better the question set at the outset of this chapter. My sole intention was to find the age at which a child could reasonably be expected to read the text fluently and accurately if working independently.

I also had to consider how much detail I could cope with in the transcriptions, because, as I quoted in CHAPTER 5, ‘transcriptions concerned with linguistic features have … often been subjected to very detailed and intricate transcription’ (Powney & Watts, 1987:146). I expected my transcriptions to fall into this category. As in CHAPTER 5, I followed the recommendation of Miles & Huberman (1994) to create ‘a provisional “start list” of codes prior to fieldwork’ (p58). I found the transcribing conventions drawn from Heritage (cited in Silverman, 1993:118) and from the *Key stage 1 English tasks teacher’s guide* (QCA, 2001b:15) particularly useful for this purpose. The former has its origins in conversation analysis whereas the latter sets out a procedure for the teacher to code children’s reading errors and strategies on a running record whilst they read a section from a book. The text I wished to transcribe did not fit exactly into either of these coding systems so I chose to use or adapt a subset of the transcribing conventions from both sources. The conventions, listed in Table 7.1, were chosen for four main reasons:

i. I did not want to use codes that would introduce too many variables for the analysis of the scripts, particularly where I felt that they would not contribute towards the comparisons I wished to make;
ii. any reading difficulties and comparing the extent and detail of these should be clear to anyone interpreting the codes;
iii. the coding conventions adapted from Heritage (cited in Silverman, 1993:118) and from the teacher’s guide (QCA, 2001b:15) provided a means of timing how long children took to read each item and of coding reading errors or strategies that could also affect the time taken to access the text, e.g. use of phonics;
iv. I needed to be able to summarise my findings by decoding the transcripts.

Table 7.1: Symbols used for transcription analysis

Symbol	Explanation of symbol
()	Closed parentheses indicate a brief silence of up to 1 second in duration.
(4)	Numbers in parentheses indicate a silence measured in seconds. Normal expected pauses, e.g. between sentences were not coded except where there was a silence of an exceptional duration. The rare spoken interventions by me were excluded from the timings.
[witch]	A word within parentheses indicates what a word sounded like, e.g. ‘witch’ for ‘which’.
O	‘O’ indicates where a word was omitted and the omitted word is also indicated by a strike-through.
T	‘T’ indicates where I felt compelled to tell the child a word, e.g. where the child was struggling or asked for help. These words are also enclosed in curly brackets e.g. {continue}.
P	‘P’ indicates use of phonics (print symbols or sound patterns).
F	‘F’ indicates knowledge that the word came from the same family, and is related, e.g. ‘belongs’ substituted for ‘belong’.
S	‘S’ indicates a recognisable word substituted for the one being read, e.g. ‘sentence’ for ‘sequence’.
U	‘U’ indicates an unrecognisable word substituted for the one being read.

After listening to samples of transcripts, I noted that children occasionally substituted a word which did not affect the meaning significantly, e.g. reading *Mrs* as *mister* or *miss*. Originally, I had decided not to create a special code for these occurrences since there would be no way of testing whether children made sensible substitutions in live tests. I did, however, add the final three codes after the trial analysis of the transcript in Table 7.2. I felt that the words in parentheses needed to be coded in more detail to make comparisons between children easier to analyse and summarise. As a result, the codes ‘F’, ‘S’ and ‘U’ complement the use of parentheses for substituted words. From the perspective of my research, any kind of word substitution indicated inaccuracies in reading that could affect mathematical performance, if undetected.

I decided to enter the first three codes within the body of the text and the rest above the relevant part of the text. For example, if a child read *shadow* for *shade*, a recognisable word substitution, I would code it thus:

S
[shadow].

If a child read ‘number’ for ‘numbers’, a word in the same family, I would code it thus:

F
[number].

Where necessary, I numbered the occurrences of a particular event, e.g. S1 and S2 to identify two occurrences of substitution in the order made, to make it easier to refer to specific misreads in the comments section of the analysis tables. I recognised that the codes could also be used to provide quantitative data in my analysis, involving timings and the number of occurrences of a coded event for comparisons between and within the year cohorts. As in CHAPTER 5, I decided on the more time-consuming option of transcribing verbatim what children read to provide an exact record of their reading skills.

Putting the codes to the test

I chose to analyse the transcript of one Y3 child, whom I shall name Bola, to trial the codes listed in Table 7.1 and to trial the design of a table whose structure I hoped would assist the analysis. I chose to start with a Y3 child to get an idea of the average reading competence across the range of the three year groups. It was also an opportunity to test how well the codes worked as unique categories. This I did in the expectation that children in year 2 were likely to be less competent and those in year 4 more competent readers. Table 7.2 shows the coded transcript.

Table 7.2: Y3 average reader (girl) from School A: Bola

Item number	Text read (verbatim)	Time taken (secs)	Comments
11	S1 Draw [around] (2) to show which S2 [square] () belongs () in the set. S3 has (1) has five [coins]	14	Three key words substituted (S1-S3). Meaning of instruction lost in reading errors. Very little expression in reading. No awareness of errors or attempts to self-correct. Had no understanding of what to do.
14	F1 O [Co] The [number] (1) in the shaded S1 S2 F2 [squares] (3) [circles] [makes] a (1) P1 S3 [se – k] (1) a [square] *(She looks to me* T *for help.)* {sequence}. S4 [Con] the se()quence by U1 S5 [shirting] more [s] () [sequences].	41	Started to read second line with number 14 alongside in margin but quickly self-corrected. F1 and F2: could have affected her response since the subject and verb are changed from plural to singular. S2: probably the result of reading 'sequence' (S1) incorrectly as 'squares' and knowing that the next word could not also be squares. Code S3: I asked her to have a guess and she read 'sequence' as 'square' again. I then told her the word. Could not explain what to do.

Table 7.2: Y3 average reader (girl) from School A: Bola (continued)

15	There () there are four apples in F1 each [packet]. (2) Mrs Pullen buys three F2 [packets] of apples. P1 How many (1) apples [do-es] (2) F3 does she [buyed]?	19	F1 and F2: showed partial knowledge of the word ie reading ‘packs’ as ‘packets’. In this instance, it does not change the overall meaning. Could not explain what to do.
17	F1 S1 P1 [Annie] () [mister] the () [hi-t] of these () two () dolls () in P2 () [bol-id). U1 S2 How many (1) [bolid) [tell) (2) S3 S4 [tells] () is the [longest] doll?	28	I resisted the urge to help because she was happy to read and her willingness shows the extent of her reading difficulties. Substituted word (F1) unlikely to affect comprehension but the substitutions (S1-S4 and U) together with her incorrect use of phonics (P1-P2) meant that the text was largely meaningless. She lacked fluency, accuracy and pace throughout. Told me that the answer was the longest doll.

Comments on Bola’s transcript

The transcript codes in Table 7.1, although time-consuming to insert, worked well and I was able to categorise Bola’s reading errors and weaknesses, pauses and hesitations as well as measuring the time taken to read each item.

Bola, who was learning English as an additional language, was assessed by her teacher as having fluency in spoken English ‘in most social and learning contexts’. In her reading, Bola demonstrated that she could read simple everyday words, such as *show*, *more*, *apples*, but was dependent on phonics for less familiar words such as the mathematical terms in items 14, 15 and 17, *sequence* and *taller* in particular. Her phonic knowledge was not advanced enough to decode unfamiliar words correctly and she also substituted words, such as reading *square* for *sequence* in item 14, code S3 (14/S3) or invented words i.e. *bolid* for *blocks* (17/U). She also had some ‘near-misses’ such as reading a plural noun and verb in their singular form as in 14/F1 and 14/F2. Such an error, whilst seemingly minor, could influence a child’s response since it changes the meaning of the text. Other ‘near-misses’, such as reading *packet* for *pack* (F1 and F2) would be unlikely to affect a child’s grasp of the mathematics but it was only in hearing Bola read that I could make this distinction. In a live test, a teacher would have no way of knowing whether or not a child’s substitutions would adversely affect a response.

The pauses and hesitations in Bola’s reading further showed the extent of her difficulties. She appeared to have no awareness that what she was reading made little sense and it came as no surprise that she could not explain what to do. Bola read none of the items fluently or accurately but, even discounting the reading weaknesses and errors, the times taken to read items 14 (41 seconds) and 17 (28 seconds) were of particular concern. By comparison, the two Y4 children at the same school read these items fluently in about one third of the time.

What I learned from this transcript

I realised that, if Bola was typical of other average readers in year 3, then the text was still too difficult to read with fluency or accuracy, even with the benefit of one year’s extra reading experience compared with children for whom the test was intended. The total time (about 100 seconds) and effort taken to read the four items, combined with the misread text, made it reasonable to assume that Bola’s access to the mathematics would be hindered by the struggle to read and that the time taken to read the text could exceed the time needed to engage with the mathematics. If this child had taken the test live, she would have needed reading support throughout if her mathematical potential were not to be compromised. Consequently, I decided to transcribe in full and code all of the Y3 interviews to find out if Bola were indeed typical of her peers. I made the same decision for the treatment of the Y2 transcripts, which I analysed last, since they were likely to be the most complicated to code, and, by that time, my coding skills would be better rehearsed.

I also took the decision not to transcribe in full the interviews with the Y4 children. After listening to the tapes several times, I felt that, in general, the fluency and accuracy of the Y4 readers were such that I would have few concerns if they were reading these items independently. However, in Appendix 7.1, as evidence of how unproblematic I found their reading, I offer a coded transcript and analysis of item 14 read by Y4 children, the item that proved to be the most difficult to read in each year group. Additionally, I timed how long these children took to read each of the other items from the audio-tapes, which are available on request. The timings are considered in detail later.

The coding system and tabular layout worked well but I had visions of readers flicking backwards and forwards between Appendix 4.1 to access the text and layout of each item and Table 7.1 to access codes. Consequently, I decided that I would focus on one question at a time using the transcripts of one particular year group, not only to make the data easier for readers to access but also to facilitate comparisons between average readers in different year groups and within and between schools. The analysis of the transcripts follows, starting with year 4 and finishing with year 2.

Comparing transcripts across year groups

As well as including the transcript of Y4 children reading item 14 in Appendix 7.1, I have included, in Appendices 7.2 and 7.3, transcripts of Y3 and Y2 children reading the same item as evidence of the extent of children's reading difficulties. This item was chosen because more children asked for help with reading this item than any other in the 2001 test. Additionally, I transcribed in full and coded the eight Y3 and Y2 children's reading of items 11, 15 and 17. These transcripts are available if required.

I have included, in Tables 7.3 – 7.11 that follow, a sub-set of the coded transcripts for all four items because of its relevance in answering the question whose answer I sought. In these tables, I include the transcripts for each item for the two children who appeared to have had the least then the most reading difficulties to show the range within each year group. Where more than one child appeared to fall into these categories, one has been chosen as being representative.

The Y4 transcripts for item 14

See Appendix 7.1 for transcripts of item 14. It was interesting to note that three of the children read 'shade' for 'shaded' (with one self-correction), but all of them read 'shading' correctly. However, the analysis of the transcript showed that, in spite of some slight hesitations and minor substitutions, all but one of the Y4 children were able to explain correctly what the item was asking them to do thus indicating that they had internalised the meaning of the text. In general, the children read confidently and did not use phonics. The speed of reading was appropriate, with no extended delays or hesitations. Indeed, four of the children read at approximately the same speed as the teachers who were timed reading this item.

Year 4: Comparing the most and least competent readers

Table 7.3: The most and least competent readers of item 14 in year 4

Child	Text read (verbatim)	Time taken (secs)	Comments
B-Y4-G	The numbers in the shaded squares make a sequence. Continue the sequence by shading more squares.	7	Words read quickly but with good intonation, accuracy and confidence. Gave correct response by stating the numbers she would shade.
A-Y4-B	F1 The numbers () in () the [shade] (2) squares make a sequence. Continue the sequence by shading () more () squares.	14	Some hesitations. Read with reasonable accuracy but with little expression. ‘Shade’ read for ‘shaded’ (F1) but ‘shading’ read correctly in next line. Appeared disinterested. Said that this was key stage 1 maths.

By comparing these two transcripts, I concluded that the more competent reader (B-Y4-G) had no reading difficulties and that the less competent reader (A-Y4-B) coped well with the reading of the key words, in spite of minor errors and hesitations, even though he was reluctant to explain what to do. Both children also read at an appropriate pace, with the first child reading at approximately the same pace as an adult.

The qualitative analysis of the transcripts for Y3 and Y2 children in that order starts in the next section with a general analysis of the reading for each item. To support the analysis, transcripts for the two children who appeared to have had the least and most reading difficulties are included as exemplification.

The Y3 transcripts: a qualitative analysis

A full set of coded transcripts for item 14 is included in Appendix 7.2.

Item 11

Three-quarters of the Y3 children read the item fluently and accurately. There were few reading errors and, generally, hesitations were momentary and did not break up the flow of the text. No children used phonics.

The speed of reading ranged from 7 to 14 seconds, with the two children who took longest being the only two who showed any indication of reading difficulties. Of these, one child (A-Y3-G), whose transcript is shown below, struggled with three key words whilst the other (C-Y3-G) read unproblematically.

Year 3: Comparing the most and least competent readers

Table 7.4: The most and least competent readers of item 11 in year 3

Child	Text read (verbatim)	Time taken (secs)	Comments
C-Y3-G	Draw arrows to show which shapes belong in the set. has 5 corners	7	Read accurately and fluently with good intonation. Sounded confident. Could explain what to do.
A-Y3-G	S1 Draw [around] (2) to show which S2 [square] () belongs () in the set. S3 has (1) has five [coins]	14	Three key words substituted (S1-S3). Meaning of instruction lost in reading errors. Very little expression in reading. No awareness of errors or attempts to self-correct. Had no understanding of what to do.

Child C-Y3-G read with confidence, accuracy and speed comparable to the best of the Y4 readers. Whilst the speed of reading of child A-Y3-G was not excessively slow, her reading errors of three key words, ‘arrows’, ‘shapes’ and ‘corners’, made the text meaningless in relation to the context. In summary, the text was too difficult.

Item 14

Three of the eight children read the item with ease and accuracy and could explain what to do. Of the remaining five, two made minor reading errors that did not prevent them from knowing the correct response whilst three misread and hesitated over words, demonstrating that the reading demand was excessive for them.

The word that caused most reading difficulties was ‘sequence’, a mathematical word that is phonetically complex. Three children had to have the word read to them, one of them after an eight second delay, one after a 12 second delay and two unsuccessful attempts at reading it, and the third after reading it as ‘square’. Another child started to use phonics, hesitated briefly then read it correctly. One child was disadvantaged by not being able to read ‘continue’, pronouncing it as ‘con’, ‘contin’ and ‘contune’ in turn. Where children read what they thought a word said, they did not always realise that the word was sometimes meaningless, as shown in the transcript for child A-Y3-B that follows.

The speed of reading varied from 9 to 41 seconds, with the three children for whom the reading demand was excessive taking the longest. The reading speed of the three children who took 12 seconds or less, with no loss of comprehension, was comparable to that of the Y4 readers.

Year 3: Comparing the most and least competent readers

Table 7.5: The most and least competent readers of item 14 in year 3

Child	Text read (verbatim)	Time taken (secs)	Comments
C-Y3-G	The numbers in the shaded squares make a sequence. Continue the sequence by shading () more squares.	9	Read fluently and accurately with good intonation. Sounded confident. Explained clearly what to do.
A-Y3-B	The numbers in the shaded (1) P1 U1 squares make a [s-sq-] (8) [scumen] U2 (4) [scillient] *(He looks to me for help.)* T1 {sequence}. U3 U4 [Contin] the [secrence] by shading more squares.	32	Pace and fluency compromised mainly by unfamiliarity with 'sequence' and 'continue' (U1 to U4). 12 seconds spent trying to decode 'sequence'. Even when told 'sequence', failed to read it correctly in following sentence. Did not know what to do.

The contrast between the reading competences of these two children was pronounced. For child C-Y3-G, the reading was effortless and accurate; for child A-Y3-B, the reading was laboured, inaccurate and punctuated by two delays totalling 12 seconds. The latter child took nearly four times as long to 'read' the item, and, after his struggle, had gained little meaning from the text as shown by the substitutions he made for 'sequence'.

Item 15

In general, Y3 children read this question with few difficulties apart from the abbreviation 'Mrs'. Four of the eight children read 'Mrs' as 'Mistress', with three reading

‘Mister’ and one reading ‘Miss’. The only hesitations of over one second in duration were made by two children as they decoded ‘Mrs’. Fortunately, the substitution in this instance did not alter the context and was unlikely to have caused problems with comprehension. Nevertheless, the evidence showed that half of the Y3 children interviewed did not interpret the abbreviation correctly and, for these children, ‘Mrs’ was too difficult to decode. I was surprised to find the extent to which the layout of this item caused confusion. As shown on page 13 of Appendix 4.1, the first line of text lies above, and the next two, below the diagram. Five children started reading the text below the diagram, with two self-correcting and the other three having to be interrupted and told where to start. It was only by hearing children read that I made this discovery, one which could have a serious impact since the first line provided essential information.

The speed of reading varied from 12 to 19 seconds, excluding the false starts, with five children taking 14 seconds or less, which was comparable to the timings for the slower readers in year 4.

Year 3: Comparing the most and least competent readers

Table 7.6: The most and least competent readers of item 15 in year 3

Child	Text read (verbatim)	Time taken (secs)	Comments
C-Y3-B	Mrs Pullen *(I ask him to start at the top line.)* There are four apples () in each pack. Mrs Pullen buys three packs of apples. How many apples does she buy?	12	Read fluently with good intonation and steady pace. Explained correctly what to do.
A-Y3-G	There () there are four apples in F1 each [packet]. (2) Mrs Pullen buys three F2 [packets] of apples. P1 How many (1) apples [do-es] (2) F3 does she [buyed]?	19	F1 and F2: show partial knowledge of the word ie reading 'packs' as 'packets'. Could not explain what to do.

Child C-Y3-B read confidently, accurately and with good intonation. Child A-Y3-G substituted 'pack(s)' for 'packet(s)' which, in this context, was a sensible substitution but she read less fluently and struggled to read the final line that gave the question to be answered. The difference between the reading competences of the two children for this item was less marked than for the other three items read by this year group. A possible explanation for this was that there were few words that were mathematical English rather than ordinary English. Nevertheless, in a test, I would not have been confident that child A-Y3-G would have coped without assistance even for this comparatively ‘easy read’.

Item 17

Three children read the item accurately, fluently and at an appropriate pace. A fourth child read accurately but appeared to read too quickly which probably resulted in her misunderstanding what the question was asking. Two children made minor reading errors in the second sentence, although one of these still explained correctly how to solve the problem. However, the remaining two children had reading difficulties that would have compromised their access to the mathematics. A minority of children could not read ‘measured’ but all but one of the eight children were able to read ‘height’, both of these being words that they would rarely read in ordinary text and are phonetically irregular.

I was surprised to find that three of the children, including two who read the second sentence correctly, appeared to ignore the comparative ‘taller’ in the sentence, and to interpret it as ‘tall’ with the result that they explained incorrectly what to do.

The speed of reading ranged from 7 to 28 seconds. Six children read the item in 13 seconds or less, which compared with all Y4 children. Two children were slowed down by unsuccessful attempts to read unfamiliar words. Child D-Y3-B made several attempts at reading ‘measured’ and, consequently, took 27 seconds. The final child, A-Y3-G, had extensive difficulties and her transcript follows.

Year 3: Comparing the most and least competent readers

Table 7.7: The most and least competent readers of item 17 in year 3

Child	Text read (verbatim)	Time taken (secs)	Comments
D-Y3-G	S1 Ann [met] *(quickly self-corrects)* measured the height of these two dolls in () blocks. How many blocks taller is the large doll?	9	Read confidently and fluently with good intonation. Pace of reading about normal talking speed. Explained correctly what to do.
A-Y3-G	F1 S1 P1 [Annie] () [mister] the () [hi-t) of these () two () dolls () in P2 () [bol-id). U S2 How many (1) [bolid) [tell) (2) S3 S4 [tells] () is the [longest] doll?	28	I resisted the urge to help because she read willingly and her willingness shows the extent of her reading difficulties. Substituted ‘Annie’ for ‘Ann’ (F). Substitutions (S1-S4 and U) together with her incorrect use of phonics (P1-P2) altered meaning of text. Lacked fluency, accuracy and pace throughout. Thought that answer was the longest doll.

Child D-Y3-G had no problems with reading and comprehending the text. In contrast, child A-Y3-G made frequent hesitations, several word substitutions and used phonics incorrectly. Misreads included several key words whose incorrect decoding completely changed the meaning of the text. In addition, her struggle resulted in her taking three times as long to read the text as the competent reader. She appeared to remember the last words she read since her answer was ‘the longest doll’. This was probably the only part of the text that made any sense to her.

The Y2 transcripts: a qualitative analysis

A full set of coded transcripts for item 14 is included in Appendix 7.3.

Item 11

Of the twelve words to be read, three in particular caused reading difficulties: arrows, shapes and corners. More than one-half of the eight children found ‘arrows’ difficult. Four children read ‘[d]raw around’ instead of ‘[d]raw arrows’, a predictable substitution since the former is a common instruction in written tasks and has the same two starting letters. A fifth child resorted to using phonics to decode the word whilst a sixth paused for eight seconds before reading it. Of the three children who misread ‘shapes’, the substitution ‘shape’ led one of them to believe that only one shape had to be identified. The other two read ‘sheep’ and ‘space’ which have similarities in structure but not in meaning to the correct word. Four children could not read ‘corners’, possibly the most important word in the text since three children, who read the set label correctly after reading difficulties in the lead-in sentence, could explain what to do. Furthermore, two children, after extensive reading difficulties, had ‘corners’ read to them and were then able to explain what to do.

Of the three problem words, the correct reading of ‘corners’, part of the set label, was likely to have the greatest impact, since, together with the diagram and an example of a shape with five corners, a child could make a reasonable guess about what to do.

The speed of reading varied from 7 seconds (D-Y2-B) to 58 seconds (A-Y2-B). In both cases, the children’s reading was flawed. The former child read fluently and confidently but misread ‘corners’ as ‘squares’ thus obscuring the required response; the latter, whose transcript follows, struggled throughout, had help with reading including ‘corners’, and then responded correctly. Even the most fluent and confident reader (B-Y2-G), whose transcript follows, made minor reading errors in the lead-in sentence.

Year 2: Comparing the most and least competent readers

Table 7.8: The most and least competent readers of item 11 in year 2

Child	Text read (verbatim)	Time taken (secs)	Comments
B-Y2-G	S1 Draw [around] to show F1 which shapes [belongs] in the set. has five corners	8	Read fluently with good intonation. Sounded confident. Substituted words (S1 and F1) did not prevent her from knowing what to do. Explained that she had to find shapes with five corners.
A-Y2-B	U1 Draw (4) [drow] (2) draw () draw S1 em (3) [around] to () show () P1 S2 which (2) [sh-ap] (2) [sheep] P2 [b-i-be-beel-be-beel-beel] (2) P3 belong () [b-be-gon] in the (1) set. *(I intervene and read the sentence to him.)* T1 – T10 {Draw arrows to show which shapes belong in the set.} P4 P5 U2 [ha-as] has five (1) [c-o] [colden] *(He points to 'corners' and looks to me for help. I intervene.)* T11 {corners}.	58	Struggled from start to finish to read text but the end result was meaningless. Dependent on an insecure knowledge of phonics (P1-P5) and this slowed down the pace considerably. Spent 14 seconds trying to decode 'belong' (P2-P3). Persevered but his reading lacked accuracy and pace throughout and was full of hesitations. After final intervention (T11), could explain what to do but used word 'ends' instead of 'corners'.

The examples above show a wide variation in reading competence in two children who were both regarded as average readers. The better reader (B-Y2-G), did not read perfectly, but her explanation suggests that the correct reading of the set label cued her into the required response. The poorer reader (A-Y2-B), after a determined but ineffective struggle with the text, also explained what response was required, but only after ‘corners’ was read to him.

Child B-Y2-G’s transcript shows no hesitations in her reading. The same cannot be said about child A-Y2-B whose reading was punctuated by 12 delays, four momentary and eight that could be timed in seconds.

Comparing Y3 and Y2 children

Three-quarters of the Y3 children read the item fluently and accurately with one-quarter making some reading errors. By comparison, all children in year 2 made errors that altered the meaning of the text in varying degrees. In year 3, hesitations were mainly momentary compared with the more frequent and longer hesitations of Y2 children. Whereas no children in year 3 used phonics, half of the Y2 children did, generally with little success.

Only one-quarter of Y2 children read the text in less than 10 seconds compared with three-quarters of children in Y3 and all eight children in year 4. The time taken by the others in year 2 was protracted by silent decoding of words or the use of phonics. Three children in particular were slowed down excessively by their attempts to read unfamiliar words, with the result that they took 32, 39 and 58 seconds.

Item 14

Four words in particular caused reading difficulties: shaded, sequence, continue and shading. No children read ‘shaded’ correctly. Five children substituted either ‘shape’, ‘shapes’, ‘shade’ or ‘shadow’, all of which have visual similarities to ‘shaded’. The child who read ‘shade’ spent eight seconds in silence decoding the word. Three children also used phonics, none of them successfully. Only two children read ‘sequence’ correctly, but both resorted either to phonics or hesitating. Three misread it as ‘square’ and a fourth as ‘question’, perhaps indicating that they thought that they were reading familiar words that included the ‘qu’ phoneme and an ‘s’. Two children had to be told the word. ‘Continue’ caused problems for all the children. Two children read it correctly after delays of five and eight seconds with a third child reading it correctly after reading it as ‘count’, a sensible but misleading guess in a mathematics test, then self-correcting after using phonics. Two other children read the word silently, one taking a three second delay before reading ‘count’ and the other reading ‘con’ after a five second delay followed by another five second delay, at which point I told him the word. Other readings for ‘continue’ were ‘coin’ and ‘contien’. Overall, five children used phonics, only one successfully, and four children experienced delays of several seconds whilst decoding the word mentally.

The speed of reading varied from 21 seconds (C-Y2-B) to 77 seconds (D-Y2-G), with all except the fastest reader taking 30 seconds or more. Only the child who took 21 seconds read with anything approaching fluency and accuracy. For all other children, the time taken was indicative of their difficulties with the text.

Year 2: Comparing the most and least competent readers

Table 7.9: The most and least competent readers of item 14 in year 2

Child	Text read (verbatim)	Time taken (secs)	Comments
C-Y2-B	S1 The numbers in the [shape] (2) squares make a (2) sequence. S2 P1 P2 [Count] [c-] continue the [s-] *(self-corrected 'count' to 'continue')* sequence by () shading more squares.	21	Read with reasonable accuracy and good intonation. Fluency compromised mainly by hesitations before 'squares' and 'sequence' and misread of 'continue'. Substituted 'shape' for 'shading' (S1) but read 'shading' correctly in next sentence. Spent several seconds studying number grid then indicated correctly numbers to be shaded.
D-Y2-G	F1 F2 The [number] in the (8) [shade] F3 square [makes] a (9) *(She looks* T1 *to me for help.)* {sequence}. P1 [C-con-cont-cont] *(I intervene.*) T2 {Continue} the (8) *(I intervene.)* T3 {sequence} by (8) shading (*pronounced with a soft 'a'*) (1) more squares.	77	Pace very slow and text read in a stilted manner. Intonation poor. In particular, attempts to read 'shaded', 'continue' and 'sequence' (F2, T1 and T3) slow down pace. Unable to read 'continue' and 'sequence' (T1-T3) P1: Using phonics to decode 'continue'. F1and F3: subject and verb are changed from plural to singular. (Made the same error in item 11.) By referring to the diagram only, she was able to explain what three numbers to shade. This took only a few seconds.

Child C-Y2-B appeared to have read the text with just enough competence to grasp the essential information. However, his reading was not without effort as indicated by the hesitations in front of unfamiliar words and the overall time taken, which exceeded the time taken by all the year 4 and the majority of Y3 children. Child D-Y2-G fared less well. Apart from taking more than three times as long as the former child, she struggled with the text throughout with four delays of eight seconds or more in front of unfamiliar words. After one of these delays, she read ‘shading’ with a soft ‘a’, but had no strategies for decoding ‘continue’ or ‘sequence’ and had to have the latter word read to her twice. With both children having struggled to varying extents with the text, they gave the correct response within a few seconds by finding the rule for the sequence from the number grid itself. For them, the time taken to read the text, therefore, was far in excess of the time taken to identify the correct answer.

Comparing Y3 and Y2 children

Both year groups had difficulties reading item 14, with difficulties considerably more pronounced in year 2. In year 3, three children found ‘sequence’ and ‘continue’ difficult to read. By contrast, in year 2, no children could read ‘shaded’, six could not read ‘sequence’ and all struggled to read ‘continue’.

In general, however, Y3 children read more fluently and accurately, used phonics less and made fewer prolonged hesitations. Six Y3 children read the text in less than 30 seconds compared with only one Y2 child. In year 3, the slowest reader took 41 seconds compared with 77 seconds in year 2. The timings for both Y2 and Y3 contrast sharply with those of Y4 children, all of whom took 15 seconds or less.

Only three Y3 children read the text with the fluency, pace and accuracy that indicated to me that they understood what they were reading. Whilst this concerned me, my concern was greater for the Y2 children, none of whom appeared to have adequate reading skills to gain meaning from the text.

Item 15

Three words in particular caused reading difficulties: ‘Mrs’, ‘Pullen’ and ‘packs’. Only two of the eight children read ‘Mrs’ as ‘Mistress’, with three each reading ‘Mister’ and ‘Miss’. Seven children had difficulties with ‘Pullen’. Four used phonics to decode it, with one being unsuccessful and having to be told the name. Three children hesitated for several seconds before reading the word with one of these decoding it as ‘Pauline’, which was probably a familiar first name and a sensible substitution. The word ‘pack’ caused fewer difficulties but was read as ‘packet’ and ‘package’ by two of the children. Fortunately, the substitutions did not change the meaning of the text to the extent that it would affect the mathematics. However, the fact that the mathematics was unchanged in this instance was more a matter of luck than judgment and the evidence shows that the children encountered unfamiliar words.

As with the Y3 children, the layout contributed to the children’s difficulties. Five children started reading the text below the diagram, with one self-correcting and the other four having to be interrupted and told where to start. Unanticipated effects of graphics, such as in this item, have been well-documented by Crisp & Sweiry (2006).

The speed of reading varied from 16 to 35 seconds after excluding the false starts made by starting to read at the second line of text.

Year 2: Comparing the most and least competent readers

Table 7.10: The most and least competent readers of item 15 in year 2

Child	Text read (verbatim)	Time taken (secs)	Comments
D-Y2-B	How many apples () oh *(quickly self-corrects and goes to first line.)* There are four apples in each pack. F1 (3) [Mister] Pullen buys three packs of apples. F2 How many apples does [he] buy?	16	Read steadily but with reasonable accuracy and fluency. Substituted words (F1-F2) changed sex gender of featured character. Could explain what to do.
A-Y2-B	F1 F2 [Mister] [Pull] () *(I ask him to start at the top line.)* There are four (2) apples in (2) each F3 [packet]. F1 F1 F2 [Mister] (2) [Mister] [Pull] (2) F1 F2 [Mister] [Pull] (4) *(He makes a blowing out sound as he hesitates.)* S1 P1 [needs] three [p-a-p-a-c] three packs of apples. F3 How many () apples () [did] () she () buy?	35	First line of text missed. Started reading text below diagram. S1 and F1-F3: substitutions unlikely to affect meaning but repetitions showed struggle with reading and slowed progress. P1: resorted to phonics even though 'pack' read as 'packet' in first line. Lack of fluency, accuracy, hesitations and lack of speed in reading very likely to affect comprehension and to be tiring. Could not explain what to do and suggested 'zero' as his answer.

Child D-Y2-B was fortunate that the two word substitutions, F1 and F2, in this instance, had little effect on the meaning of the text. The three-second delay before reading ‘mister’ for ‘Mrs’, in an otherwise fluent exposition, showed that the abbreviation was not familiar. His explanation, however, demonstrated that he knew what to do, perhaps because the substitutions left the mathematics unchanged. Child A-Y2-B fared less well since his reading was punctuated by frequent hesitations, more akin to ‘staccato’ word by word reading, and substitutions that did not affect the meaning significantly but slowed down the pace of reading. Reading ‘Pullen’ caused the greatest problem. As a result of his cumulative difficulties, he took more than twice as long to read the text as child D-Y2-B. Consequently, I was not surprised that the child could not explain what to do.

Comparing Y3 and Y2 children

About half of the children in both year groups missed the first line of text until this was pointed out. Although this was not a reading difficulty, it drew attention to a flaw in the layout that could have deprived children of a mark since the text contained key information.

Generally, Y3 children read this item fairly well although half misread ‘Mrs’ as ‘mister’ or ‘miss’, which was fewer than the three-quarters of Y2 children who made the same error. Although this did not affect the number ‘story’, it caused hesitations in reading, particularly in year 2.

If reading difficulties were similar in both year groups, then this was not the case with the speed of reading. All Y3 children read the item in less than 20 seconds compared with only two Y2 children. Of the remaining six Y2 children, half took between 20 and 30 seconds with the other half taking up to 35 seconds. At this speed, the six Y2 children were disadvantaged by disjointed, laboured reading breaking up the flow of text.

Less than half the children in both year groups read this item with sufficient competence for me to feel that they could read it independently with meaning.

Item 17

No Y2 children read this item accurately although errors varied greatly in extent and impact. In the first sentence, three words in particular caused difficulties: measured, height and blocks. One child read ‘measured’ correctly, four read it as ‘measure’ or ‘measures’ and three children had to be told the word after failed attempts and delays at reading it. Child D-Y2-B spent 10 seconds decoding ‘measured’ but read the remainder of the text with no difficulty. Five children recognised by sight the word ‘height’ but the other three read it correctly only after slowing down to determine the word, perhaps predicting it using the picture. ‘Blocks’ also caused half of the children reading difficulties, with two having to be told the word. In the second sentence, five children found ‘taller’ hard to read, as shown by the delays prior to reading it. Child B-Y2-B took about 10 seconds to decode it, then read ‘tall’, a misread that changed the question completely and resulted in him misinterpreting what to do. When it came to reading ‘large’, it seemed that some children guessed what the question was asking. Only four children read ‘large’ with the rest substituting related adjectives such as ‘longest’, ‘largest’ and ‘long’. Whilst some of the substitutions did little to alter the meaning of the text, some children made several substitutions that collectively did alter the meaning. For this item, children appeared to rely heavily on graphic clues then substituted words that often made sense in the context.

The speed of reading ranged from 16 to 61 seconds, with the responses from the children with these timings shown in the following table. Only three of the children explained correctly what they had to do. It was probably no coincidence that these were the children who read the item the most fluently and accurately, with timings of 16, 22 and 27 seconds.

Year 2: Comparing the most and least competent readers

Table 7.11: The most and least competent readers of item 17 in year 2

Child	Text read (verbatim)	Time taken (secs)	Comments
B-Y2-G	U1 F1 [Ane] [measure] () the height of () these two dolls in blocks. How many blocks taller F2 is the [largest] doll?	16	Read with reasonable fluency and accuracy. Three substituted words close in meaning to original words. Explained what to do and gave correct answer.
D-Y2-G	F1 P1 P2 [Anna] (8) [mes-as-are] [mes-as- U1 are] [me-sure] *(pronounced as two separate words)* T1 {measured} the (1) height P3 of these two dolls in [bl-o-cl] *(She looks to me for assistance.)* T2 {blocks}. P4 How many blocks (4) [t-t-] taller is the (2) large doll?	61	The reading errors, combined with the delays due to meeting unfamiliar sight words and using phonics were likely to mean that the child would have little or no comprehension of what was being asked. Thought that she had to find the height of the taller doll with a ruler.

Child B-Y2-G read fairly fluently but misread three words. She was fortunate that the substituted words had little effect on the meaning of the text and she correctly interpreted what she had to do. By comparison, the codes marked on the script of child D-Y2-G tell the story of her struggle. Her reading difficulties with ‘measured’ alone took 15 seconds when the 8-second delay and time spent using phonics were combined. Further difficulties and delays reading ‘blocks’ and ‘taller’ wasted 61 seconds of her time since she did not understand afterwards what she had to do. She took about four times as long as the former child to complete the reading, but her struggle was in vain.

Comparing Y3 and Y2 children

Half of the Y3 children read the text fluently and accurately compared with none of the Y2 children. The majority of Y2 children found the text difficult to read compared with only a minority of the older children. Only one child in year 2 read ‘measured’ correctly compared with six in year 3.

Seven of the children in year 2, compared with only two in year 3, took more than 20 seconds to read the item. No child in year 3, but five in year 2, took more than 30 seconds to read it. Indeed, the average time taken by year 3 children to read the item, 14 seconds, was two seconds faster than the time taken by the fastest and also the best reader in year 2.

The years 4, 3 and 2 transcripts: a quantitative analysis

Whilst writing up the qualitative analysis of the transcripts, I felt that words alone did not show in sufficient detail the extent of the reading difficulties and the differences in performance between the year groups. As a result, I generated tables that summarised these data. An added advantage was that I had data in the form of field notes from the Y4 transcripts of times taken to read the items and could therefore make comparisons across the three year groups (see Tables 7.14 and 7.15).

Tables 7.12 and 7.13 that follow show the frequency of the codes in the transcripts for year 2 then year 3. Code O for omitted words was not included since omissions occurred so infrequently. On the tables, I classified the hesitations, shown in curved brackets in the transcripts, into four categories of different durations. As a reminder of the codes, P stands for use of phonics; for substitutions, F stands for a word in the same family, S for a recognisable word and U for an unrecognisable word. There is no table showing these data for year 4 since full transcripts were not produced for reasons explained earlier.

An overview

Since a zero indicated that a particular event did not occur, a cursory glance at Tables 7.12 and 7.13 shows that there are more zeroes in the Y3 than the Y2 table. This reflected the greater ease with which Y3 children read the questions. However, the Y3 table also shows that these children exhibited the same reading weaknesses as their Y2 counterparts but to a lesser degree. I knew from my notes on the Y4 transcripts that, had a similar table been prepared for Y4 children, the table would have comprised few numbers other than zeroes, except perhaps for occasional short pauses that would give no cause for concern.

Table 7.12: Y2 - Analysis of code frequencies

Item 11

	Frequency of hesitations				Frequency	Frequency of substitutions		
Child	< 1 sec	1-2 secs	3-5 secs	5+ secs	P	F	S	U
A-Y2-G	6	1	2	0	5	0	0	0
A-Y2-B	4	6	2	0	5	0	2	2
B-Y2-G	0	0	0	0	0	1	1	0
B-Y2-B	0	0	2	0	2	0	4	0
C-Y2-G	4	1	0	1	0	0	2	0
C-Y2-B	4	0	2	0	0	0	1	0
D-Y2-G	4	1	1	0	3	2	1	1
D-Y2-B	0	0	0	0	0	3	0	0
Mean	*2.75*	*1.13*	*1.13*	*0.13*	*1.88*	*0.75*	*1.38*	*0.38*

Item 14

	Frequency of hesitations				Frequency	Frequency of substitutions		
Child	< 1 sec	1-2 secs	3-5 secs	5+ secs	P	F	S	U
A-Y2-G	2	2	0	0	3	0	4	0
A-Y2-B	1	2	1	0	6	1	5	0
B-Y2-G	0	0	0	1	2	2	4	0
B-Y2-B	4	1	3	0	8	1	3	0
C-Y2-G	1	2	0	1	0	0	4	0
C-Y2-B	1	2	0	0	2	0	2	0
D-Y2-G	0	1	0	4	1	3	0	0
D-Y2-B	0	0	0	0	1	0	2	3
Mean	*1.13*	*1.25*	*0.5*	*0.75*	*2.88*	*0.88*	*3.00*	*0.38*

Item 15

	Frequency of hesitations				Frequency	Frequency of substitutions		
Child	< 1 sec	1-2 secs	3-5 secs	5+ secs	P	F	S	U
A-Y2-G	2	1	0	0	1	1	0	0
A-Y2-B	5	4	1	0	1	3	1	0
B-Y2-G	0	1	2	0	0	2	3	1
B-Y2-B	2	0	1	0	1	1	0	0
C-Y2-G	0	1	0	1	0	0	1	0
C-Y2-B	1	0	0	0	1	1	0	0
D-Y2-G	0	1	0	0	1	2	2	0
D-Y2-B	1	0	1	0	0	2	0	0
Mean	*1.38*	*1.00*	*0.63*	*0.13*	*0.63*	*1.5*	*0.88*	*0.13*

Item 17

	Frequency of hesitations				Frequency	Frequency of substitutions		
Child	< 1 sec	1-2 secs	3-5 secs	5+ secs	P	F	S	U
A-Y2-G	8	2	0	0	2	1	3	1
A-Y2-B	3	4	2	0	4	1	0	0
B-Y2-G	2	0	0	0	0	2	0	1
B-Y2-B	2	3	1	1	4	2	0	0
C-Y2-G	0	0	1	1	0	2	1	0
C-Y2-B	0	1	0	0	0	2	1	0
D-Y2-G	0	2	1	1	4	1	0	1
D-Y2-B	0	0	1	0	2	0	1	0
Mean	*1.88*	*1.50*	*0.75*	*0.38*	*2.00*	*1.38*	*0.75*	*0.38*

Table 7.13: Y3 - Analysis of code frequencies

Item 11

	Frequency of hesitations				Frequency	Frequency of substitutions		
Child	< 1 sec	1-2 secs	3-5 secs	5+ secs	P	F	S	U
A-Y3-G	2	2	0	0	0	0	3	0
A-Y3-B	2	0	0	0	0	0	0	0
B-Y3-G	0	0	0	0	0	0	0	0
B-Y3-B	2	0	0	0	0	0	0	0
C-Y3-G	0	0	0	0	0	0	0	0
C-Y3-B	3	0	0	0	0	1	0	0
D-Y3-G	1	0	0	0	0	0	1	0
D-Y3-B	0	0	0	0	0	0	0	1
Mean	*1.25*	*0.25*	*0*	*0*	*0*	*0.13*	*0.5*	*0.13*

Item 14

	Frequency of hesitations				Frequency	Frequency of substitutions		
Child	< 1 sec	1-2 secs	3-5 secs	5+ secs	P	F	S	U
A-Y3-G	2	3	1	0	2	3	3	1
A-Y3-B	0	1	1	1	1	0	0	4
B-Y3-G	0	1	0	0	1	1	0	0
B-Y3-B	3	0	0	0	1	0	0	0
C-Y3-G	1	0	0	0	0	0	0	0
C-Y3-B	2	0	0	0	0	0	0	0
D-Y3-G	2	3	0	1	0	0	0	0
D-Y3-B	2	0	0	0	1	0	1	1
Mean	*1.5*	*1.00*	*0.25*	*0.25*	*0.75*	*0.5*	*0.5*	*0.75*

Item 15

	Frequency of hesitations				Frequency	Frequency of substitutions		
Child	< 1 sec	1-2 secs	3-5 secs	5+ secs	P	F	S	U
A-Y3-G	1	3	0	0	1	3	0	0
A-Y3-B	0	0	0	0	0	2	0	0
B-Y3-G	3	0	0	0	0	0	0	0
B-Y3-B	1	3	0	0	0	1	0	0
C-Y3-G	0	0	0	0	0	1	0	0
C-Y3-B	1	0	0	0	0	0	0	0
D-Y3-G	1	0	0	0	0	0	0	0
D-Y3-B	0	0	1	0	0	2	0	0
Mean	*0.88*	*0.75*	*0.13*	*0*	*0*	*1.13*	*0*	*0*

Item 17

	Frequency of hesitations				Frequency	Frequency of substitutions		
Child	< 1 sec	1-2 secs	3-5 secs	5+ secs	P	F	S	U
A-Y3-G	7	2	0	0	2	1	4	1
A-Y3-B	1	0	0	0	0	0	1	0
B-Y3-G	2	0	0	0	0	1	1	0
B-Y3-B	4	0	0	0	0	0	0	0
C-Y3-G	1	0	0	0	0	0	0	0
C-Y3-B	3	0	0	0	0	0	0	0
D-Y3-G	1	0	0	0	0	0	1	0
D-Y3-B	1	0	0	0	2	0	2	0
Mean	*2.5*	*0.25*	*0*	*0*	*0.5*	*0.25*	*1.13*	*0.13*

Hesitations

Where hesitations did occur in year 3, these were most likely to fall into the first category, i.e. durations of less than one second. Of most concern were the last two categories where hesitations were extended to at least three seconds in duration. Across the four items, all eight Y2 children paused at least twice for three seconds or more, generally before unfamiliar words. Four Y3 children also paused for three seconds or more, but, with one exception in item 15, this was whilst reading item 14. Indeed, reading item 14 led to the most extended pauses in both year groups.

By reading across the rows, I noted that four children in Y2 paused repeatedly in reading most or all items. This was a clear indication that the text was too difficult since the pauses protracted the speed of reading, interrupted the flow of text and left children concentrating on the reading of words rather than the mathematics. In contrast, only one Y3 reader, child A-Y3-G, Bola, struggled in the same way.

Overall, the frequency of hesitations lasting less than one second was fairly comparable between the two years groups. However, Y2 children were more likely to hesitate longer and more frequently between words. They were about twice as likely to hesitate for 1-2 seconds and nearly seven times as likely to hesitate for 3 or more seconds as their Y3 counterparts. The frequency of hesitations led me to conclude that they were a matter of concern for the minority of Y3 readers but of great concern for the majority of the Y2 readers.

Phonics

In Y3, the use of phonics was infrequent, with the exception of child A-Y3-G, Bola, who had the most reading difficulties overall in her age group. Apart from Bola, four other children used phonics once in item 14. In both year groups, but particularly in year 2, phonics were most likely to be used in item 14. In year 2, half the children used phonics in item 11 and more than half used phonics in the other three items. Even the more competent Y2 readers resorted to using phonics at least once, but some of the children were more dependent on phonics, for example, child B-Y2-B who used phonics eight times in item 14.

The use of phonics, more prevalent among Y2 children, was evidence that they were faced with reading unfamiliar sight words. Usually, attempts failed since children generally applied this decoding strategy to words that were phonetically irregular. This meant that time and effort was wasted on reading in a test situation where children's total concentration and energy should have been focused on the mathematics.

The transcripts showed that it was not uncommon for a pause to be followed by an attempt to use phonics and, collectively, these extended reading times considerably.

Substitutions

Substitutions occurred in the three categories (codes F, S and U) in both year groups. In general, occurrences in all categories were more frequent in year 2 than in year 3. In year 2, the total number of substitutions ranged from seven for child C-Y2-B to 17 for child B-Y2-G. In year 3, the number of substitutions ranged from one for child C-Y3-B to 19 for child A-Y3-G, Bola. However, if Bola's results were excluded, the upper end of the range would drop to seven substitutions made by children A-Y3-B and D-Y3-B. No children in year 2 but five children in year 3 made fewer than seven substitutions. After reading the four items, each Y2 child had substituted an average of about 12 words compared with about five words for each Y3 child. When all children in each year group and all items were taken into account, Y2 children overall were more than twice as likely to make substitutions as those in year 3.

An analysis of individual items showed that at least half the children in year 3 made substitutions when reading each item but children in this category varied from item to item. In year 2, all children made substitutions in all questions apart from child A-Y2-G whose reading difficulties in item 11 were in other categories. Children in year 2 were most likely to substitute words in item 14 with an average of 4.3 substituted words. This compared with an average of 1.8 substitutions for that item in year 3. However, the greatest difference was in item 11 where children in year 2 made an average of 2.5 substitutions compared with an average of 0.8 in year 3.

I concluded that the frequency of substitutions was another indication that the text was too difficult to read independently for all children in year 2 and the majority in year 3. For the Y2 children, all identified as average readers, this was of particular concern because there was no way of knowing if other average, below- or above-average readers, who had taken the test live in 2001, had been given or asked for the support in reading that it appears they would have needed.

Speed of reading: comparing year groups

I explained earlier that the average time taken by three teachers to read the items was six seconds for item 11 and eight seconds for items 14, 15 and 17, a total of 30 seconds. Since their reading was fluent and accurate and they could explain how to respond, I was confident that they read at an appropriate and probably, for them, ‘optimum speed’ (Gilliland, 1972:13). Whilst I had no expectations that the reading skills of children in years 2, 3 or 4 would be comparable, the teachers’ results showed speeds of reading at which understanding and success could be achieved. I became increasingly aware, as I coded the transcripts, that the pauses, hesitations and use of phonics of children in year 2 in particular, and to a lesser extent in year 3, were slowing down the reading to a speed that could never be described as ‘optimum’ with the consequence that the text was fragmented.

Using Table 7.14, I was able to compare the reading speeds of children within year groups and across year groups. My main findings from the comparison follow.

Table 7.14: Time taken to read items across year groups (seconds)

			Year 2		
Child	Q11	Q14	Q15	Q17	Total time
A-Y2-G	39	48	23	43	153
A-Y2-B	58	53	35	35	181
B-Y2-G	8	37	34	16	95
B-Y2-B	23	76	22	54	175
C-Y2-G	26	34	29	31	120
C-Y2-B	18	21	16	22	77
D-Y2-G	32	77	35	61	205
D-Y2-B	7	30	16	27	80
Mean	*26*	*47*	*26*	*36*	*135*
			Year 3		
Child	Q11	Q14	Q15	Q17	Total time
A-Y3-G	14	41	19	28	102
A-Y3-B	10	32	13	10	65
B-Y3-G	7	10	12	9	38
B-Y3-B	10	15	17	13	55
C-Y3-G	7	9	12	7	35
C-Y3-B	11	12	12	10	45
D-Y3-G	9	25	14	9	57
D-Y3-B	8	14	16	27	65
Mean	*10*	*20*	*14*	*14*	*58*
			Year 4		
Child	Q11	Q14	Q15	Q17	Total time
A-Y4-G	10	15	15	9	49
A-Y4-B	7	14	13	11	45
B-Y4-G	8	7	12	7	34
B-Y4-B	8	15	12	17	52
C-Y4-G	6	7	13	10	36
C-Y4-B	6	9	8	8	31
D-Y4-G	6	13	9	12	40
D-Y4-B	8	8	9	7	32
Mean	*7*	*11*	*11*	*10*	*39*

(All means rounded to nearest second)

Year 4

Children in Y4 were generally reading at speeds that were slightly slower than the teachers but were of an appropriate pace, as shown by their fluency, accuracy and comprehension of the text, on which I commented earlier. Indeed, half of the children took only a few seconds longer in total than the teachers to read the four items. It did not surprise me that the pace of reading was unproblematic since Table 7.14 confirms that these children hesitated less and made virtually no use of phonics, the main factors that increased reading times for the younger year groups. There was a maximum difference of 10 seconds in the rate of reading an item between the fastest and slowest reader (see item 17 on Table 7.14), with 21 seconds overall separating the speed of reading all four items by the fastest and slowest readers. Overall, the rate of reading of the Y4 average readers was fairly comparable and of little concern.

Year 3

On average, Y3 children took nearly twice as long to read the four items as the teachers and 1.5 times as long as those in year 4 since they hesitated more and used phonics occasionally. Differences in reading competence between children assessed as average readers within the year group were also more marked than in year 4. For example, in item 14, 32 seconds separated the speeds of the fastest and slowest readers compared with 8 seconds in year 4. For reading all four items, 67 seconds separated children in year 3 at the extremes of the range. Three Y3 children took more than a minute to read the items, with child A-Y3-G taking closer to two minutes. Five children took longer to read all items than the slowest reader in year 4.The pace was most likely to be compromised in item 14, with an average reading time of 20 seconds compared with 11 seconds in year 4, although the average reading times for year 3 were slower than those for year 4 for every item. I noted with interest that the two most fluent Y3 readers read the four items at a speed slightly faster than the average for the Y4 children.

The combined data from Tables 7.13 and 7.14 led me to conclude that reading difficulties, manifested in part by a lack of pace in reading, would have prevented the majority of Y3 interviewees from accessing at least some of the items in the test. This concerned me since this group was a year older on average than the target age group for the test. I considered that most of them, particularly those taking more than a minute to read an item, did not read with the optimum speed that Gilliland (1972) and Dale & Chall (1948) associated with success in reading and comprehension. Most of all, the lack of pace, mainly caused by children's struggle with unfamiliar words, probably meant that the effort to read diverted their mental energies at the expense of the mathematics. In addition, their effort to read individual words may have prevented them gaining a sense of meaning of the holistic text, thus greatly reducing the likelihood of knowing what they were being asked to do. This being so in year 3, the situation was of even greater concern in year 2.

Year 2

On average, Y2 children took more than twice as long to 'read' the four items as Y3 children and nearly four times as long as Y4 children. I deliberately placed 'read' in inverted commas since most children did not find the items readable. As with Y3 children, the lack of pace was associated with pauses and, more particularly for year 2, with the use of phonics. Even the fastest reader, child C-Y2-B, who had fewer reading difficulties overall than his peers, as shown in Table 7.14, took 77 seconds to read the items, a pace that suggested that he did not find the reading easy. Only one child in years 3 and 4 exceeded this time. Even more worrying was that five of the eight children took two minutes or more to do so, with two children taking over three minutes.

As with the older year groups, item 14 was the most problematic with children taking from 21 to 77 seconds to read the text, some with essential but minimal assistance from me. Whereas Y3 and Y4 children took an average time of 20 and 11 seconds respectively to read this item, Y2 children took an average of 47 seconds.

Since increased pace of reading was associated with fewer substitutions and less use of phonics, Table 7.14 showed that there was a marked improvement in reading in the move from one year group to the next. Improvement was more marked between years 2 and 3 than between years 3 and 4. I established this by comparing the differences in the average times taken to read each item across each year group and the average time taken to read all items across each year group. This appeared to suggest that, whilst children in year 3 generally found the text difficult to read compared with children in year 4, children in year 2 found the text significantly more difficult to read than those in year 3. Furthermore, it was not until year 4 that children read all the items with comparative ease and accuracy.

In year 4, the children would have had two more years' reading experience and would be reading at approximately one national curriculum level higher than average readers in year 2. For Y2 children this would be level 2B; for Y4 children, this would approximate to level 3B. This led me to conclude that children who were average or below-average readers in year 2 could not be expected to read the text without support. In CHAPTER 5, I also found evidence that children reading at level 2A had some difficulties so this would appear to confirm that the overall reading demand of the test is pitched at level 3. Even if the test had been written for Y3 rather than Y2 children, I would have reached the same conclusion since most average Y3 children are likely to be reading on the border of levels 2 and 3.

Although referring to continuous text, Harrison (1980) described a simple procedure for determining readability, i.e. the '5% rule', which assumes that if a child makes one or more uncorrected errors when reading aloud 20 consecutive words, then the text is too difficult. Each of the four items read by children had fewer than 20 words. By referring back to Table 7.12, the data in the final three columns show that the majority of Y2 children made at least one error in each item and, in some cases, several errors. Transcripts show that errors were rarely corrected. This was further confirmation of the excessive reading demand of these items for a Y2 audience.

Speed of reading: comparing schools

Whilst writing up the transcripts, I became aware that there were noticeable differences in the speed of reading in children identified as average readers within each year group in particular schools. As a result, I constructed Table 7.15 to make it easier to compare these differences. By doing this, I also found some similarities linking the schools. I acknowledge that only 24 children were involved but the differences and similarities did repeat across all four schools.

Table 7.15: Time taken to read items across schools (seconds)

		Item				
School	Child	11	14	15	17	Total time
A	Y2-G	39	48	23	43	153
	Y2-B	58	53	35	35	181
	Y3-G	14	41	19	28	102
	Y3-B	10	32	13	10	65
	Y4-G	10	15	15	9	49
	Y4-B	7	14	13	11	45
B	Y2-G	8	37	34	16	95
	Y2-B	23	76	22	54	175
	Y3-G	7	10	12	9	38
	Y3-B	10	15	17	13	55
	Y4-G	8	7	12	7	34
	Y4-B	8	15	12	17	52
C	Y2-G	26	34	29	31	120
	Y2-B	18	21	16	22	77
	Y3-G	7	9	12	7	35
	Y3-B	11	12	12	10	45
	Y4-G	6	7	13	10	36
	Y4-B	6	9	8	8	31
D	Y2-G	32	77	35	61	205
	Y2-B	7	30	16	27	80
	Y3-G	9	25	14	9	57
	Y3-B	8	14	16	27	65
	Y4-G	6	13	9	12	40
	Y4-B	8	8	9	7	32

Differences within schools were most pronounced in year 2. For example, in schools B, C and D, one Y2 reader took considerably longer than the other to read all four items. By year 4, the two children from each school were reading at speeds that were much more comparable. This reflected the older children's greater consistency in reading competence.

There were also noticeable variations in reading competence in children regarded as average readers between schools. Difference in performance was most noticeable between schools A and C. In school A, children took longer than average to read all four items, compared with school C where the opposite was true. One possible explanation for this was the fact that the interviewees in school A had English as a second language although they had been described by their teacher as being 'confident as a user of English in most social and learning contexts' in spoken English. However, this did not translate to confidence or, indeed, aptitude in reading the mathematical test items in years 2 and 3. Whilst their reading competence was less than that of children in school C, they may indeed have been average readers for that particular school. However, in year 4, the reading of the children in both schools gave me little concern.

In each school, the speed of reading followed the same general pattern with children in year 2 taking considerably longer to read the items than those in year 3, and with the improvement in reading being greater between years 2 and 3 than between years 3 and 4. However, the improvement made by children during year 3 was not sufficient for most of them to read the items competently as explained previously.

The cumulative data obtained in Tables 7.3 – 7.15 complemented each other well. The coded transcripts (Tables 7.3 – 7.11) highlighted in detail the exact nature of the reading difficulties whereas Tables 7.12 – 7.15 summarised the frequency of categories of reading difficulties, thus providing qualitative and quantitative data. Together, the tables reaffirmed the conclusion that I had already reached in

CHAPTER 5 that my concerns about the reading demand of the test for Y2 children were justified. Furthermore, Y3 children demonstrated that my concerns were more than justified because they showed that, even though a year older on average, the improvement in their reading between year 2 and year 3 was not sufficient for them to read the selected text with the ease and accuracy required in a test situation.

Arrangements for administering the Y2 tests

At the four participating schools, I obtained information on how teaching staff administered the Y2 tests. In each case, administrative arrangements varied. School A, in which there was a high proportion of children with English as an additional language, had changed its administrative arrangements two years previously. Before then, the class teacher and one assistant administered the test to one Y2 class at a time, with children requesting help with reading. When the staff became concerned that children who had limited reading skills were not asking for adequate help, they decided to administer the test to groups of four or five children at a time and to read each question to individuals as they worked at their own pace. It was the general view that this strategy had improved the school's overall results for the test. School D used an approach similar to school A but varied the size of group to whom they read all the questions from two up to a maximum of six children depending on their needs. Here, the children worked on the same double-page spread after the text was read by the teacher. School C split the Y2 class into three groups of children of similar mathematical competence, with up to 10 children in a group, and the test was administered separately to each group by the class teacher. Help with reading was given on request although the teacher was not convinced that children asked for all the help they required. In School B, a small one-form entry school, the teacher administered the test to everyone in the Y2 class, usually about 25 children, simultaneously and children requested help with reading. She expressed concern that lack of funding prevented extra staff support during the administration, since priority was given to the administration of the English tests or to children identified as having particular learning difficulties or behavioural problems. On hearing the transcripts for the four Y2 children in her class, she expressed concern and surprise at the extent of the reading difficulties.

In earlier chapters, I have already noted concerns that arrangements between schools vary considerably in how children are assisted with reading during the test. The feedback from the four Y2 teachers in this part of my study further confirmed the disparities. The possible effect of non-standard approaches to test administration is discussed as one issue in my concluding chapter.

Conclusions

To round off this chapter, I considered three aspects in particular:

i. comparisons with Y2 children who took the test live in 2001;
ii. the review of the literature;
iii. the question posed at the start of this chapter.

Comparisons with Y2 children who took the test live in 2001

In CHAPTER 5, I obtained evidence that there were children reading at level 2 who did not always ask for help with reading items during the KS1 mathematics test in 200I. I subsequently found out that these children could not read unassisted. Indeed, there were some children reading at level 2B (average) and below who did not ask for help with reading the four items under scrutiny in this chapter. In interviews soon after the live test, I found children who could not read these items, had not asked for help with reading, but who demonstrated the same weaknesses in reading as the Y2 children taking part in the interviews in this chapter. These children also were considered to be average readers and were typical of the age cohort at the time of the live test. The Y2 children involved in my research in June 2001 and 2002, therefore, typical of their peers, performed similarly. I thought that it was reasonable to assume therefore, that, nationally, there would be children taking the live test for whom reading support was needed but not requested or offered.

Reflecting on the review of the literature

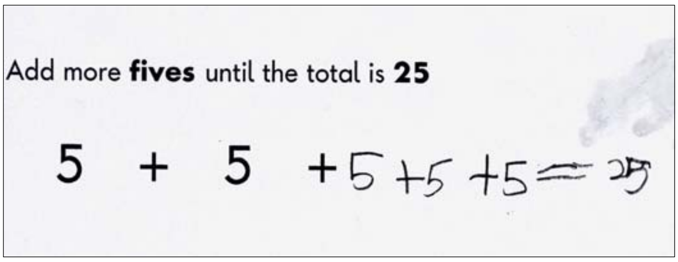
As I wrote up the data analysis for this chapter, my thoughts returned to issues that I had chosen to write about in the review of the literature. I refer back to some of these issues because of their relevance:

Children in Y2 and most children in Y3 did not read to the degree of accuracy or at the 'optimum speed' (Gilliland, 1972:13) that would enable them to understand what they were being asked to do. Most children, particularly in Y2, could not maintain the normal pace of reading and hesitated frequently, characteristics that, according to QCA (2002), showed that they did not recognise words on sight. What was of even greater concern was that, for the Y2 children in particular, the excessive time taken to 'read' the items were for four items only so did not take account of how long it might take to 'read' all the items where support was not given, never mind to engage with the mathematics. When referring to reading books considered suitable for assessing Y2 children, QCA (2001a) concluded that less competent readers analysed words into single letter sounds, were unsuccessful at blending words, made a high level of substitutions and self-corrected infrequently. When reading the four chosen test items, children who were considered to be average readers in both years 2 and 3 exhibited these characteristics showing that they became less competent readers in the test situation, often because of unfamiliar ME words, e.g. 'sequence' or 'measured'. A detailed analysis of the readability of the text in the booklet from which the items were chosen, and which identifies likely reasons for reading difficulties, follows in CHAPTER 8.

The evidence from this phase of my research showed that concerns of writers such as Harrison (1979, 1980); Shuard & Rothery (1984), and Mobley (1986), about the reading difficulties caused by subject-specific text including mathematics, were justified. Harrison (1980), for example, explained that secondary teachers had expressed concern that children could not always cope with the language of examination questions even though they were competent in the subject. Whilst I was not judging the mathematical competence of the children, the transcripts clearly showed that, in general, children considered to be average readers in years 2 and 3 could not cope with reading the text of typical test items to the degree of accuracy or fluency that would reassure me of their competence to read independently. It was not until children were in year 4 that they demonstrated the characteristics of fluent readers as described by Gibson (1989) on page 21 in CHAPTER 2. Faced with unfamiliar words, they spent too much time 'struggling at frustration level' (Committee of inquiry into reading and the use of English, 1975:103) in a test where a misread of one key word could deprive of a mark a Y2 child taking the test live. I agreed with Harrison (1980) who stated that '[i]n some subjects, skipping a word here or misreading a word there will not have a damaging effect on overall comprehension, but in mathematics it may alter the meaning drastically' (p124). In such circumstances, the mathematical attainment of a child in the test could be compromised because of poor reading skills. Fisher-Hoch & Hughes (1996) identify two types of difficulty in mathematics questions, stating that '[*v]alid difficulty* has its source in the mathematical requirements of the question, and is intended by the examiner. *Invalid difficulty* … is caused by features of the question which are not mathematical … and is not intended by the examiner (p2). Since the reading of text in statutory mathematics tests is not a mathematical requirement and not being assessed, any unrecognised reading difficulties could invalidate the tests.

More specifically, Lupetti, Sainsbury & Schagen (1995) and Shorrocks-Taylor & Hargreaves (1999, 2000) respectively had concerns about the reading demand of the KS1 and KS2 statutory mathematics tests. They recommended that further research was needed since, to restate the words of Shorrocks-Taylor & Hargreaves (1999), 'there is very little research evidence about the precise issues surrounding the use of language in test contexts' (p123). What the evidence in this chapter suggests, albeit based on results from a small number of children, is that it is not until children are, on average, two years older than the age of children for whom the test is written that they can maintain the normal pace of reading and read the items accurately with no apparent frustration. In other words, a test written to assess the mathematical competence of children in Y2 required reading skills typical of a child in year 4, a worrying imbalance with far-reaching implications for test developers and administrators to minimise the potentially damaging effect of unsupported reading. Newman (1977) had similar concerns. In a mathematics test with carefully controlled vocabulary designed to be appropriate for 9-year olds, it was found that 12% of the 124 low-attaining 11-year olds who took the test made reading errors due to faulty word recognition, especially of key vocabulary. The evidence I obtained also seemed to support Mobley's (1986) view that '[i]f the reading is to be unsupported, it may need to be two years *below* [sic] the pupil's reading level' (p49). Put another way, if adequate reading support cannot be assumed, Y2 children need to be able to read at approximately one national curriculum level higher than average readers in year 2, i.e. at the expected standard of average Y4 children, if they are to cope independently. This conclusion corresponded well with a point I raised in CHAPTER 1 that it is not until children are reading within level 3 that they are expected to 'read a range of texts accurately and fluently' (DfEE & QCA, 1999a:56) and independently. This was the answer to the question I sought to answer at the start of this chapter, i.e. *At what age could a child reasonably be expected to read the text fluently and accurately if working* *independently?* The answer only increased my concerns. It suggested that the reading demand of the test would be appropriate only for Y2 children reading securely within level 3, the expectation by the end of year 4, rather than at the threshold of level 3, the expectation by the end of year 3. Such Y2 children represent a minority of their peers. Confronted with the data from the live test in 2001 (see CHAPTER 5) and from this chapter, I felt that I had to extend my study into an in-depth analysis of the test booklet itself to seek answers as to *why* Y2 children in particular found items hard to read to 'arrive at an informed judgement' by looking 'closely at the styles of writing used in ME' (Shuard & Rothery, 1984:2). This was the final element of my methodological triangulation, anticipated as a minor part of my original research design. However, I hoped that such an analysis would substantiate and confirm what the children's voices had already told me, i.e. that the reading demand was unlikely to be appropriate for most of its target audience.

CHAPTER 8



Jack was working at level 2C in mathematics.

CHAPTER 8

Analysis of test booklet

A review of readability

In CHAPTERS 5 and 7, I produced evidence, through my analysis of diagnostic interviews with Y2 children, that children reading at levels 2B and 2C, and to a lesser extent, those reading at level 2A, had difficulty reading text in some items in the KS1 2001 mathematics test booklet. In CHAPTER 5, I found evidence that children in year 2 had not asked for or been offered reading assistance with at least some of these items in the KS1 statutory mathematics test in 2001. In this chapter, I return to my intention, stated in my methodology, to triangulate the data from the children with an analysis of the booklet to try to determine its readability. With hindsight, I wish that I had started my investigations with this analysis rather than with the interviews with children because I had already found out, from children themselves, many of the words that they had difficulty reading. However, when I started my data collection through interviews, I had not considered such an undertaking. Out of interest, I wished to determine *why* so many or so few children had asked for help with reading certain questions. To avoid bias, I needed to conduct this investigation independently of interviews with children and only then compare results.

In CHAPTER 2, I considered the views of various writers (see, for example, Klare, 1963; Gilliland, 1972; Harrison, 1979; 1980; Perera, 1980; Shuard & Rothery, 1984; and Mobley, 1986) on the strengths and weaknesses of readability formulae. As a result, I concluded that formal measures of readability could not be applied to the KS1 mathematics test because of its unique content and structure, and the age of children for whom it was written. Although the inappropriateness of readability formulae would make my task harder, Harrison (1980) suggested that text difficulty could be evaluated at sentence level using ‘a word-frequency list’ (p55), by examining the syntax and, as I have already done in CHAPTERS 5 and 7, by monitoring the materials as they were used by children. Shuard & Rothery (1984) also believed in the importance of assessing ‘whether a particular piece of ME writing might be ‘easy’ or ‘difficult’ or ‘about right’ for a particular child’ (p2) and recommended that not only hearing children read but also examining ‘the styles of writing used in ME’ (Shuard & Rothery, 1984:2) was important. A similar view was held by Mobley (1986) who stated that ‘readability refers to every aspect of a text which makes it either easy or difficult to read’ (p9). Consequently, I decided to engage in an analysis of text in the booklet that would take into account:

i. word frequency, i.e. how often a word was likely to occur in written form in the everyday experience of Y2 children;
ii. sentence length, i.e. the number of words in sentences;
iii. any other features on an item by item basis that could affect its readability.

For each category, I referred back to my review of the literature. In this, I referenced several sources that provided recent word lists relevant to children taking the test in 2001 (see, for example, Reid, 1989; *National Literacy Strategy*, DfEE,1998; *National Numeracy Strategy*, DfEE, 1999a; *The National Numeracy Strategy: Mathematical vocabulary*, DfEE:1999b; Stuart et al, 2003). I also found guidance on sentence length. For example, as stated in CHAPTER 2, Klare (1974) believed that ‘counts of the 2 [sic] simple variables of word length and sentence length are sufficient to make relatively good predictions of readability’ (p97).

For the third category, I considered features that Mobley (1986) believed were appropriate in the consideration of readability, as listed in my review of the literature (see page 14 in CHAPTER 2) i.e.:

i. legibility, e.g. the clarity and size of the type face, the length and spacing of the lines and the paper type;
ii. visual aspects, e.g. the quality and quantity of the illustrations;
iii. language aspects, e.g. the use of words and language structures which are familiar;
iv. content and conceptual difficulty, e.g. the amount of information introduced at one time;
v. clarity of meaning, e.g. absence of ambiguity;
vi. interest level, e.g. the features which will motivate the reader;
vii. text simplification, e.g. keeping sentences short and the structure as simple as possible.

Collectively, I considered that these sources would enable me to undertake an analysis that would combine objective and subjective assessments of the readability of the test. The word and sentence analyses would provide some objectivity since these are variables used in readability formulae, as discussed in CHAPTER 2. Whilst I would never be able to *measure* the booklet's readability, I would be able to *predict* how readable it was. The checklist provided by Mobley (1986) would provide a basis for more subjective assessments.

Categorising words

Word lists

For manageability, I categorised words into four broad categories:

i. the high-frequency words (approximately 200) that, according to the *National Literacy Strategy* (DfEE, 1988), children should recognise by sight in and out of context by the end of year 2 'to help them get pace and accuracy into their reading' (p60);

ii. those not appearing in the previous list but which were the highest frequency sight words in the lists produced by Stuart et al (2003) or by Reid (1989). Stuart et al (2003) analysed words used in '[t]exts from 685 books from reading schemes and story books read by 5-7 year-old children' (p585) in England. Reid (1989) analysed scripts from 7-year olds in British schools to determine words that they most frequently used in their writing;
iii. those that were unlikely to be identified as in frequent use in children's reading books or in their writing but that Y2 children would be likely to read frequently as part of their classroom experience, e.g. instruction words like write, draw, numbers etc.;
iv. those that were used infrequently in texts read by 5-7 year olds as identified in the lists by Stuart et al (2003) or by Y2 children in their writing as identified in the list by Reid (1989).

This left me with the problem of deciding which words, other than those in the first category, could reasonably by classed as 'high-frequency' sight words. I was unable to find any literature that suggested how many words an average child might recognise on sight, i.e. maintaining 'the normal pace of reading … without evident pause for spoken or silent working' (QCA, 2002:24) towards the end of year 2. In any case, any such data would have had limited value since each 'average' child's bank of high-frequency sight words would have common elements but also vary depending on reading schemes in use, personal interests etc. Not only that, some children taking the test were below-average readers. I decided to err on the side of caution since a child reading independently in the mathematics test should not have to pause to decode unfamiliar words as reading is not being assessed. Consequently, I drew up a list of words, eventually totalling 251, which comprised:

i. all the words from the *National Literacy Strategy* (DfEE, 1988*)* list as in the first category above;

ii. any words from the 100 most frequent words in the studies undertaken by Stuart et al (2003) and Reid (1989) that were not already included in the preceding category;

iii. any words that I felt children would see frequently in written instructions either across the curriculum, e.g. *write* or in mathematics, e.g. *number* that were unlikely to have a high-frequency in lists collated from reading schemes or children's writing.

List 1 in Appendix 8.1 shows a merge of these words, with words from the third category shown in blue since these were based on my own judgment. so that 'objective and subjective assessments can be used side by side' as recommended by Perera (1980:160).

I was reassured to find that all but nine of the top 100 words in Stuart et al's (2003) list and all 100 of the top 100 words on Reid's (1989) list were already on the *National Literacy Strategy* (DfEE, 1998*)* list. However, some words on the latter list, that children were meant to read on sight at the end of year 2, occurred infrequently on the two other lists, e.g. months of the year, so there was some divergence. I too had some concerns about including months of the year since these are likely to be seen frequently in writing only when a particular month is displayed in the form of the date unlike days of the week that are seen regularly. Nevertheless, for my analysis, I regarded the list as a tolerable compromise for the average Y2 reader, recognising that most children's sight vocabulary would be wider or narrower than or slightly different to that indicated on the list. I also noted how quickly the frequencies of occurrences of words dropped in Stuart et al's (2003) and Reid's (1989) lists. For example, in the former list, Stuart et al (2003) explain that their database contained unique 9748 words (defined as types) and 268028 occurrences of words (defined as tokens) (p588). The most frequent word (type), in their database of just over one quarter million occurrences of words (tokens), was *the* that occurred 17422 times compared with 432 times for the 100th most frequent word *off.* Not only that, Stuart et al (2003) found that '[t]he 100 most frequent words …account for 54.1%' (p588) of all words in their database. This further justified my decision to devise a list for sight vocabulary that was not over-ambitious.

Another factor I had to consider was that, because a word did not occur frequently in children's texts or writing, did not mean that it was necessarily hard to read. For example, most average or above-average Y2 readers who could sight-read *big* were likely to read words of a similar structure e.g. *bag* or *dig*, not necessarily by sight but with little difficulty. I ultimately decided on a simple coding frame for each item for analysis.

Coding items

Word frequency and traffic lights

For each word, I assigned 'traffic light' colours to categorise each word into one of three categories:

green: high-frequency sight words (see List 1 in Appendix 8.1) that I anticipated would be sight-read by all but a minority of Y2 children;

yellow: words in the most frequently occurring 250 words (excluding those in the green list) of Stuart el al's (2003) and Reid's (1989) lists that I anticipated would be sight-read by most Y2 children (see Lists 2(a) and 2(b) in Appendix 8.1) or any that had lower frequencies but were phonetically simple or used familiar word structures, e.g. *hat*.

red: words that were not included in the green or yellow categories on the word lists and that I anticipated would be unlikely to be sight-read by average or below-average Y2 readers. I shall now refer to List 1, List 2(a), List 2(b) or Lists 2, as appropriate.

In addition, I felt that I should indicate terms that, according to the Y2 examples in the National Numeracy Strategy *Framework for teaching mathematics* (DfEE, 1999a) children should 'understand, use and begin to read', a phrase that generally precedes references to vocabulary (see, for example, the Y2 column on p3 of the Y123 examples). A single underline indicates a word that was recommended for introduction in year 1 or earlier; a double underline indicates a word that was recommended for introduction in year 2 in the *National Numeracy Strategy: Mathematical vocabulary* (1999b) booklet. It should be remembered that mathematical vocabulary introduced in year 2 may be seen infrequently in written form, particularly in topics that are covered only once or twice a term or introduced in the latter part of the school year. I also remind the reader that Stuart el al (2003:588) cite an earlier study by Stuart, Masterson & Dixon (2000) of word learning experiments with 5-year olds in which they found that '36 experiences of reading a word in text proved insufficient to guarantee that the words became stored in the children's memories to subserve subsequent recognition'. Finally, I counted the number of words in each sentence but also indicated how many other words had to be read e.g. words on graphs. Word counts are indicated in brackets. Figure 8.1 exemplifies how the coding works using the text only from item 17.

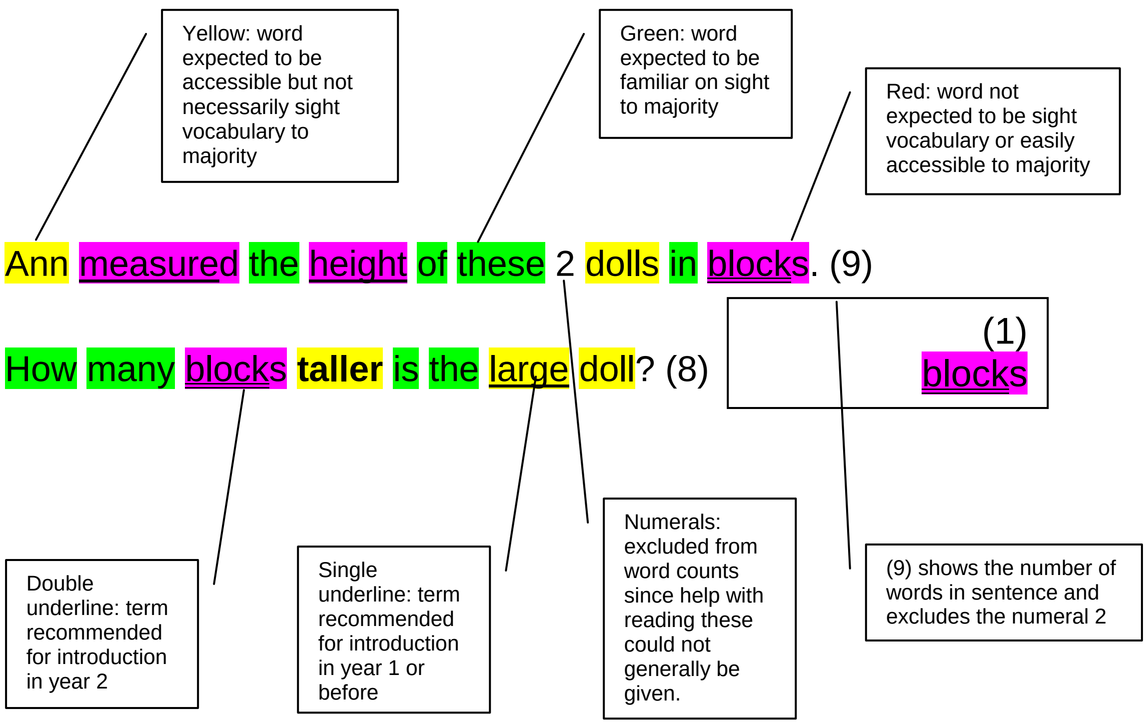


Figure 8.1: Explanation of codes used in text analysis

The next decision was to choose which items to analyse.

Choosing items to analyse

Prioritising

The decision about how many and which items to analyse was made by referring back to Table 5.3 (see page 139) whose data show the percentage of 164 children who asked for help with reading particular items during the live test in 2001. I decided to analyse the twelve items where 25% or more of children in my dataset had requested help with reading but to contrast these with the two contextual items for which children requested little help. My analysis had three foci. First, I analysed in detail, i.e. by coding as in Figure 8.1 and by commentary, the two contextual items where only a small percentage of children requested help since these were text-dependent but tackled independently by most Y2 readers (see Table 8.1). In so doing, I hoped to establish why children requested so little help.

Table 8.1: Two 'easy read' items for analysis in booklet order

Item number	Frequency (children)	Target level	% (children)
6	7	2C	4%
7	9	2C	5%

Second, I also analysed in detail the six items that were targeted at children working at levels 2B and 2C in mathematics since these were items where many of the average and below-average mathematicians would be expected to gain marks (see unshaded items on Table 8.2). To some extent this was a logistical decision, influenced by the time that it would take to conduct such an analysis. But more importantly, I could also compare how my analysis of the items compared with my analysis of the transcripts as reported in CHAPTERS 5 and 7. Third, I analysed in less detail, i.e. by coding as in Figure 8.1 only, the remaining six items for which at least 25% of children had requested help with reading but for which transcript data were not available (see shaded items on Table 8.2). If required, interview data for most shaded items are available on request.

Table 8.2: Twelve ‘hard read’ items for analysis in booklet order

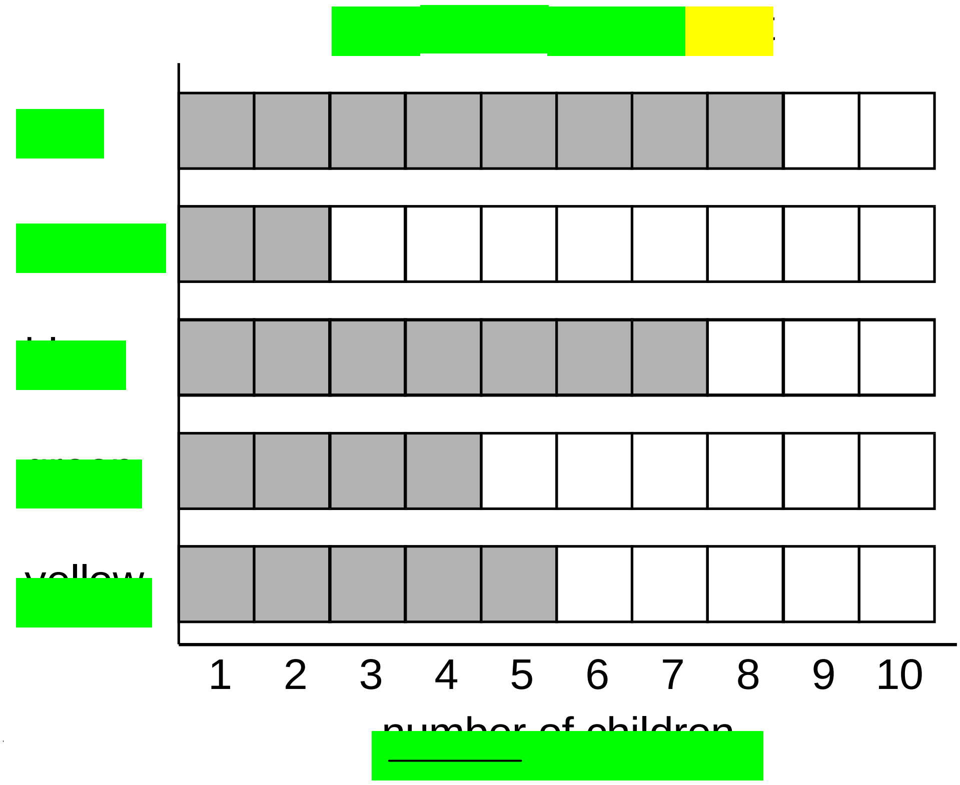
Item number	Frequency (children)	Target level	% (children)
11	57	2C	35%
14	76	2C	46%
15	41	2B	25%
16	42	2B	26%
17	56	2B	34%
18	62	2B	38%
20	48	2B	29%
22	41	2A	25%
27	41	2A	25%
29	55	3	34%
30	42	3	26%
31	43	3	26%

The items are shown and analysed in booklet order since that is how children would have experienced them. As far as possible, the layout and font size replicate those in the test booklet (see Appendix 4.1). The detailed analyses follow but the less detailed analyses for the shaded items are included in Appendix 8.2

Detailed analyses of items in Table 8.1

Analysis of practice question and item 6

Some children made this **graph**.



Practice question

Look at the graph.

Which colour do **5** children like best?

6 Look at the graph. (4)

How many children like **red** best? (6)

children

Figure 8.2: Analysis of practice question and item 6

Vocabulary

The first written question is a practice question that is not part of the test and not marked. The teacher could spend as much time as needed helping the children answer it and this would almost certainly have resulted in the context and the vocabulary on the graph being introduced to the children. Only two words were not coded green. The word *best* was introduced in the practice question but was also 206th in List 2(b) and likely to be familiar on sight to most Y2 children. However, *graph*, introduced in year 2 in *The National Numeracy Strategy*: *Mathematical vocabulary* (DfEE:1999b), would have been seen rarely in written form since handling data accounts for only a small part of the mathematics curriculum. The word was not listed in the databases by Stuart et al (2003) or Reid (1989). However, the opening line of text that includes *graph* was also in the practice question so the teacher would have read it previously and drawn children's attention to the graph to answer the practice question.

Sentence length

The two sentences of four then six words are short and of a structure likely to be familiar to Y2 children. Additionally, the first sentence was rehearsed with the teacher in the practice question leaving only a short question for children to read. This was unlikely to cause reading difficulties, especially as it included three words already read with the teacher in the second line of the practice question.

Other points of interest

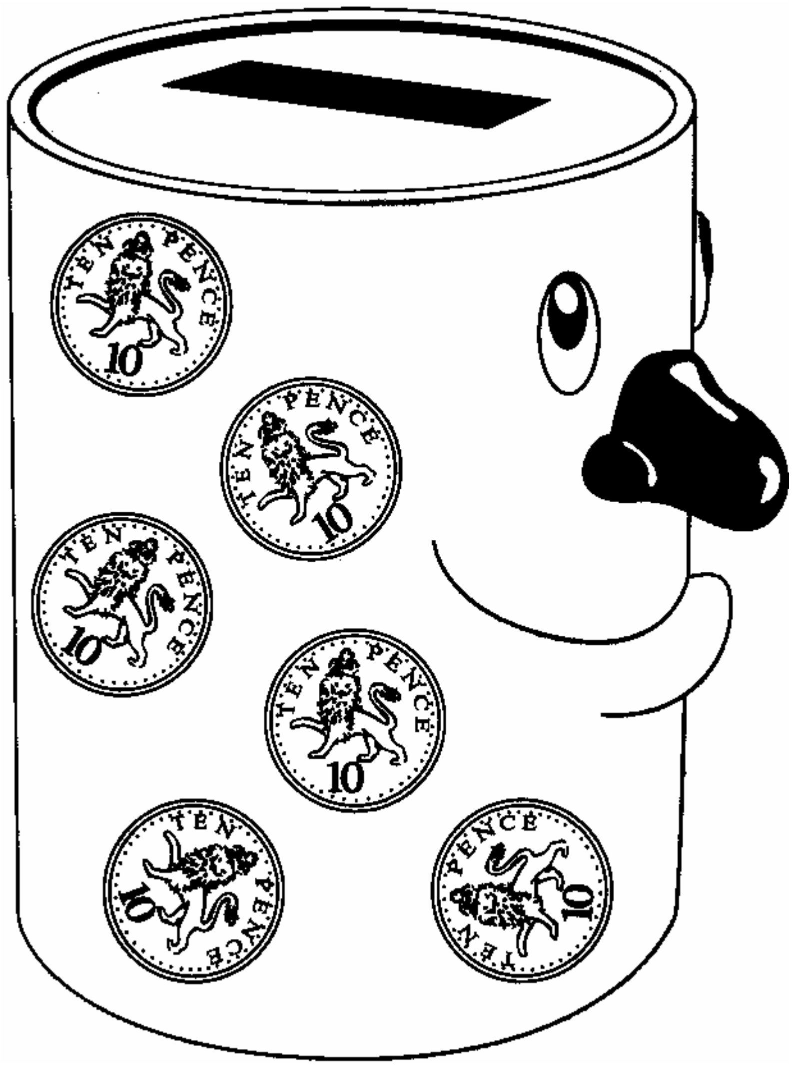
In my dataset, only 4% of the 164 children requested help with reading this. QCA (QCA, 2001e) collected data from about 3500 children who took this test in a nationally representative sample of schools. The mark for this question was awarded to 95% of these children indicating that the mathematical and reading demand were accessible. However, without teacher input, the coding suggests that some Y2 readers would have struggled to read the text independently, with the likely effect that fewer children would have been awarded the mark.

Predicted readability:

Due to teacher input, I predict that children reading at level 2C upwards would generally cope with reading item 6. Without teacher input, the reading demand would be greater because of the amount of text and words likely to be unfamiliar in written form.

Analysis of item 7

7. How much money is in the money box? (8)



p

Figure 8.3: Analysis of item 7

Vocabulary

The word *money* does not appear in any of the prepared lists that would give it a green or yellow code and it cannot be assumed that most children would read it as sight vocabulary. However, the word *box* was ranked as the 172nd most common word in List 2(a) and likely to be sight-read by more Y2 children than *money*. Of the eight words to be read, three may have caused less confident readers to pause, guess or use decoding strategies had it not been for the diagram.

Sentence length

The single sentence of eight words is short and of a structure likely to be familiar to Y2 children.

Other points of interest

In my dataset, 5% of children asked for help with reading, suggesting that children could read the text or guessed what to do without reading the text. In the study undertaken by Lupetti, Sainsbury & Schagen (1995), as discussed in CHAPTER 2, the researchers observed that poorer Y2 readers answering items in a mathematics test booklet scanned the page 'in order to work out what was required before attempting to read the text' (p13), e.g. for illustrations that would avoid them having to read text. Children using this strategy were likely to have guessed correctly what to do, i.e. total the coins, and gained the mark without accessing the text. In the QCA sample (QCA, 2001e) of about 3500 children who took the test, 95% were awarded the mark. However, some children may have gained the mark regardless of the text.

Predicted readability

Level 2B but accessible to less competent readers because of diagram

Items 6 and 7: reasons for few requests for help

Both items required only small amounts of independent reading that were unlikely to overload memory or affect concentration. The reading demand for item 6 was not excessive due to the effect of the practice question, prior discussion of the graph and rehearsal of most of the vocabulary with the teacher. For item 7, the illustration provides a strong prompt and some children may have ‘read’ the diagram only and gained the correct mark. However, such a strategy is not one that children should have to adopt because text is difficult to read. With hindsight, it would have been interesting to have found out if children needed help with reading the text even where they gave the correct response.

Detailed analysis of items in Table 8.2

Analysis of item 11

11 Draw arrows to show which shapes belong in the set. (10)

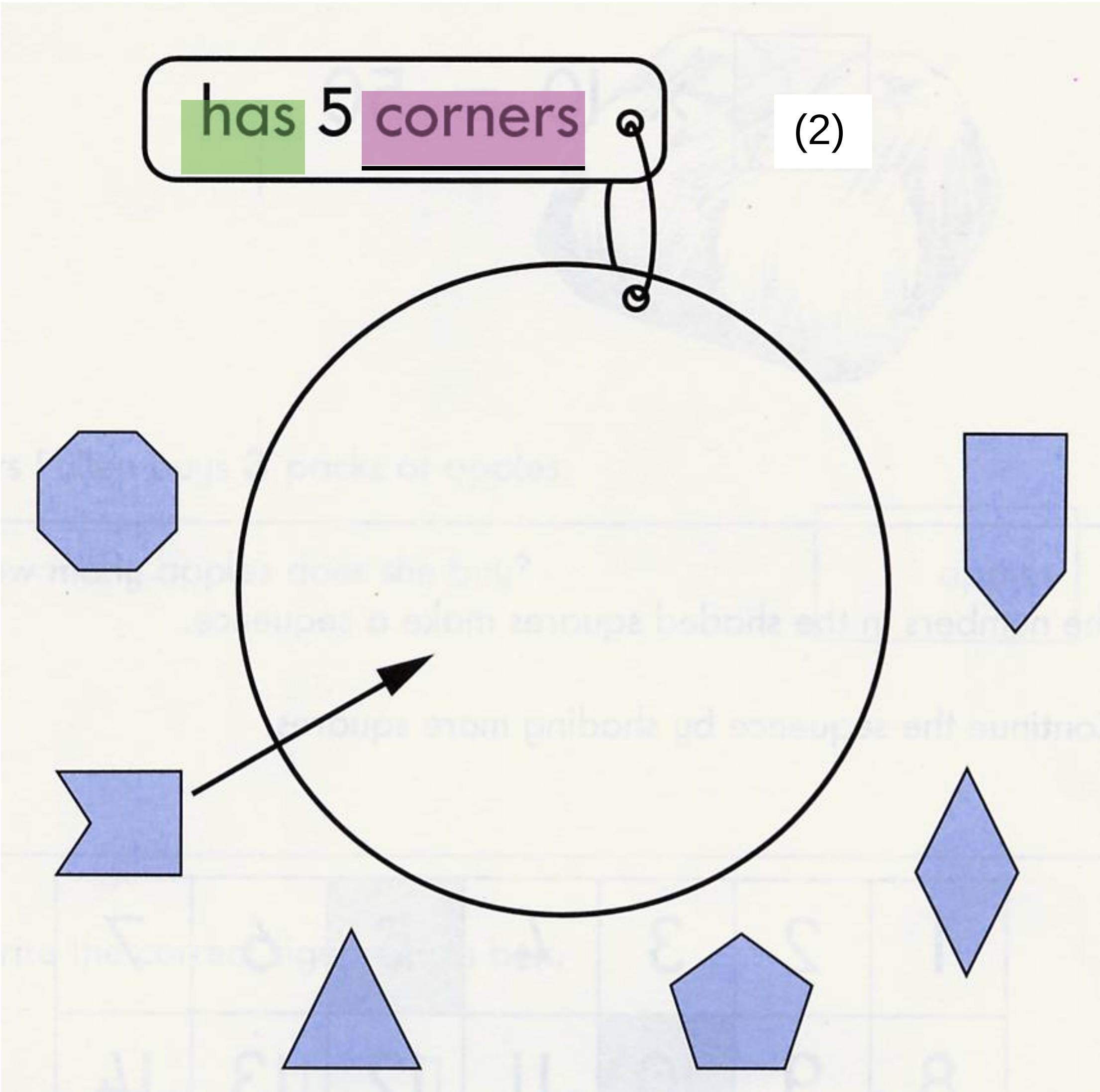


Figure 8.4: Analysis of item 11

Vocabulary

As the single underlines show, the words *arrows* and *corners* are recommended for introduction in year 1 (DfEE, 1999b). However, it is unlikely that children will see them frequently enough in written form for them to be recognised on sight. The words are coded red rather than yellow because they arise infrequently in the 262,028 occurrences of words in Stuart et al’s (2003) database and not at all in Reid’s (1989) database of 2000 words. The third word coded red, *belong,* does not occur on either list. I felt that many Y2 children would read *show* and *set* with little difficulty but, since they do not appear in Lists 2, I have coded them yellow since I cannot assume that children would read them as sight vocabulary. Twelve words have to be read of which I have coded just over one-half green.

Sentence length

The sentence of ten words is a reasonable length but is packed with information and appears to include too many words that could cause reading difficulties. It is difficult to see how the sentence could be made shorter without losing clarity.

Comparing colour coding and interview data

As described in CHAPTER 7, I transcribed interviews of eight Y2 children, regarded as average readers, reading this item towards the end of year 2. From these, I found that six children had difficulty reading *arrows*. Four children substituted *around*, a word that also starts with *ar* and often follows *draw* as a common instruction; one child read arrows correctly at the third attempt and one after an eight second pause. Children fared somewhat better at reading *belong* but two children had to use phonics and two read it as *belongs*. Only four children read it correctly on sight. Two children at the same school struggled so much with reading *corners* that I read the word for them. It was also misread as *squares* and *horners* by two other children with the remaining four

reading it correctly on sight. The words coded yellow, *show* and *set* were generally read with little difficulty although two children read *set* as *seat*, with one self-correcting. Green-coded words, as predicted, were generally read on sight but with occasional errors, e.g. child D-Y2-B reading *the* as *this*, a sensible substitution in the context.
As I hoped, the colour coding appears to suggest why reading difficulties were likely. No child read the text with sufficient accuracy for me to believe that they understood the instruction, even though some children read the text more easily than others, with timings ranging from 7 seconds to 58 seconds. However, children D-Y2-B and B-Y2-G, who read with the greatest ease, made word substitutions that affected meaning. The remaining children struggled throughout. By comparison, the Y3 average readers I interviewed at the same time of year, generally read the item fluently and accurately, although child A-Y3-G misread key words. Year 4 children read confidently and fluently. To me, this suggested that above-average readers in year 2 only, reading at level 2A or above, could reasonably be expected to read the text independently.

Other points of interest

About one third of children in my dataset asked for help with reading this item. It is also likely that many children would have done so during the live test with some less confident or lazy readers likely to have used the diagram only to find out what to do. At the very least, children not being given help with reading may have expended time and mental energy decoding words or interpreting the diagram. Nevertheless, this item was answered well in the live test with 91% of children awarded the mark. Children did not have to draw arrows as teachers could accept any other clear way of indicating the correct response.

Predicted readability

Border of levels 2 and 3

Analysis of item 14

The numbers in the shaded squares make a sequence. (9)

14 Continue the sequence by shading more squares. (7)

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	32	33	34	35

Figure 8.5: Analysis of item 14

Vocabulary

The word *shade* is introduced as a verb in reception (DfEE, 1999b). However, *shaded* and *shading* occur only once in Stuart et al's (2003) database and not at all on Reid's (1989) list. Although likely to be familiar in spoken form, *squares* is not an easy word to read since it is phonetically irregular and not set in the usual context for the word, i.e. properties of shapes. *Sequence,* introduced in year 2 (DfEE, 1999b), is not likely to be seen often in written form and is not present in either Stuart et al's (2003) or Reid's (1989) lists. The verb *continue* was also introduced in year 2 (DfEE, 1999b). In summary, the words *shaded*, *squares*, *sequence*, *continue* and *shading* appeared very rarely or not at all on Stuart et al's (2003) and Reid's (1989) lists. Consequently, not only were they likely to be unfamiliar as sight words but, even if children spent time trying to decode them, their attempts were likely to fail since none is phonetically simple or regular. I concluded that seven of the 16 words to be read could be unfamiliar on sight to most Y2 children and of excessive reading demand. Although it is not unreasonable to expect children to understand *sequence* as a term in classroom activities, *pattern* may have been easier to read and a suitable substitute.

Sentence length

Both sentences are of reasonable length but the text comprises too many words that are coded red.

Comparing colour coding and interview data

I transcribed 17 interviews with Y2 children reading this item, nine after the live test in 2001 (see Table 5.12 on page 164) and eight about one year later. I noted in my evaluation of Table 5.12 that the words *shaded*, *sequence*, *continue* and *shading* caused most problems, with *squares* also not recognised by about half the children reading across the range of level 2. Two children misread *sequence* as *sentence* and two misread *squares* as *question*. The same reading difficulties can be seen in the transcripts in Appendix 7.3. Of the eight average Y3 readers who read this item (see Appendix 7.2), there was some improvement but only three could read the item with acceptable fluency and accuracy. For the others, *sequence* and *continue* caused the most reading difficulties. Indeed, I had to read *sequence* to three of the children. Although the use of phonics was less evident than in year 2, some Y3 children continued to use phonics with one child taking 41 seconds to read the item. Although there were reading errors, there was also more evidence of self-correction in year 3 than in year 2. Unlike children in year 2, those in year 3 generally recognised *shaded* and *shading* on sight. However, the majority of Y3 interviewees found this item difficult to read and it was not until the same item was read by average readers in year 4 (see Appendix 7.1) that I felt that the appropriate skills were in place to support independent reading. The coding complements the results from the interviews and provides supporting evidence that the reading was not accessible to children reading at level 2 and they should not have been expected to read the item independently. I now believe that this item would have been more suitable as an oral question read by the teacher.

Other points of interest

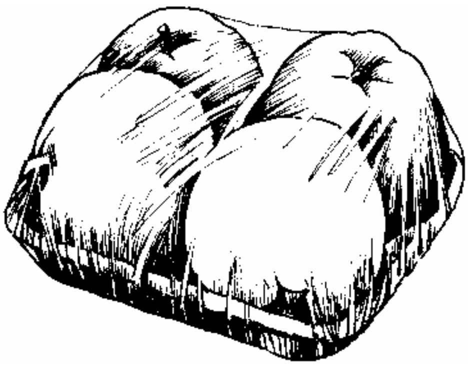
Nearly half the children in my dataset requested help with reading this item in the live test. My analysis of the text shows the likely reason to have been the high incidence of unfamiliar words. It is also likely that some children would have been discouraged by the difficult text and guessed what to do using the diagram only. Fortunately, counting on using number grids is a recommended activity in the numeracy strategy and children may have found enough clues in the diagram to know what to do.
Although the reading demand may lend itself to criticism, the mathematical demand was accessible to the majority with 85% of children being awarded the mark according to the *Key Stage 1 technical appendices* (QCA, 2001e). This suggests that many children asked for help with reading or used the diagram only to decide what to do. The latter option is not unlikely since, as noted in my review of the literature, Lupetti, Sainsbury & Schagen (1995) observed that an approach used by children who found text difficult 'was to scan the page several times in order to work out what was required before attempting to read the text' (p13).

Predicted readability

In view of evidence from Y4 children, this item appears to be suitable for children reading confidently within level 3.

Analysis of item 15

There are **4** apples in <u>each</u> pack. (6)



Mrs Pullen <u>buys</u> **3** packs of apples. (6)

15 How many apples does she <u>buy</u>? (6)

(1)
apples

Figure 8.6: Analysis of item 15

Vocabulary

Only about half of the words appear on Appendix 8.1 List 1, as shown by the green coding. Of those remaining, *apples* occurs outside the top 250 words on Stuart et al's (2003) and Reid's (1989) lists but is coded yellow because of the diagram. The word *buy* is introduced as a verb in reception (DfEE, 1999b) in the context of money. However, it needs to be sight vocabulary because it is phonetically irregular, and, as it does not appear in Lists 2, is coded red. For the same reason, *each, pack and does* are coded red. The green coding suggests that Y2 children are likely to recognise on sight about half the words.

Sentence length

The three sentences of six words excluding numerals are short and familiar in structure.

Comparing colour coding and interview data

I collated ten transcripts of Y2 children reading this item, two in 2001 after the live test (see Table 5.9 on page 159 and Table 5.10 on page 161) and eight in 2002 about a year later, all average readers. Apart from the final word, children generally read the first line of text well. About half the children had difficulties with *pack*, indicated by hesitations, use of phonics, or substituting similar words, e.g. packet or package. The start of the second line of text caused the most problems. Although *Mrs* came in List 1 because of its position in the top 100 words in Stuart et al’s (2003) list, abbreviations are not introduced formally at key stage 1 in the *National Literacy Strategy* (DfEE, 1998) and only three children read it as ‘*missus’* the colloquial form of *mistress*. The remaining children read it as either *mister* or *miss*, showing some awareness of abbreviated forms. However, only one of the 10 children could read *Pullen* on sight, with five more children succeeding only after breaking down the word into two syllables *Pull*- and *–en*, hesitating for up to six seconds or repeating words. Children appeared to be familiar with the word *pull* and, for three children, the surname was substituted to *Pull*. Fortunately, the substitutions for *packs*, *Mrs* and *Pullen* did not affect the overall sense of the text but they did slow down the pace of reading for most children. I certainly expected children to find the title and name less challenging to read than they did. Not only that, it was pure chance they were not key words. With hindsight, an easier alternative, e.g. Miss Hill or simply a first name, e.g. Pat, would have been much more accessible and made no difference to the mathematics. Although coded red, children generally read *buy(s)* and *does* with little difficulty, although *does* became *do* and *did* in two instances.

Unlike items 11 and 14, the colour coding and the results from the children diverged somewhat in this item. For example, the abbreviation *Mrs* did cause reading difficulties whereas *each*, *buy(s)* and *does* were less problematic. However, the transcripts show that children had reading difficulties.

Other points of interest

Hearing Y2 and Y3 children read this item raised a weakness in the layout. About half the children in years 2 and 3 omitted the first line of text that contains essential information and starting reading the first line of text below the diagram. In these instances, I had to point out where to start reading. Two further children in year 3 started reading at line 2, then self-corrected by going back to the opening line. Mobley (1986), in considering the position of illustrations, stated that ‘the attention of young children is attracted more to the picture than the print, and so any writing which *precedes* [sic] the illustration may well be ignored’ (p18), a statement certainly upheld in this instance. The split text is likely to have caused as many problems as reading difficulties.

One quarter of children in my dataset from the live test requested help with reading this item. By referring to my transcripts of Y2 and Y3 children (see CHAPTER 7), I found that Y3 children generally read the item more accurately than their Y2 counterparts but half the Y3 children read *Mr* or *Miss* for *Mrs* and there were other occasional misreads. However, they had no difficulties with *Pullen*, thereby showing more maturity at combining segments of words. Adult readers averaged about eight seconds to read this item, as I explained in CHAPTER 7, compared with an average time of about 14 seconds in year 3 and 26 seconds in year 2 (see Table 7.14 on page 235). It was not until I heard Y4 average readers read item 15 (not transcribed but available on audio-tape) that their reading skills had the fluency and accuracy needed for independent reading.

Predicted readability

The inclusion of *Mrs* and *Pullen* in particular suggests that the reading demand is tipping over from level 2 into level 3. Prior to the transcriptions and colour coding, I would have predicted the text suitable for children reading at level 2B and above.

Analysis of item 16

16 Write the correct sign in each box. (7)

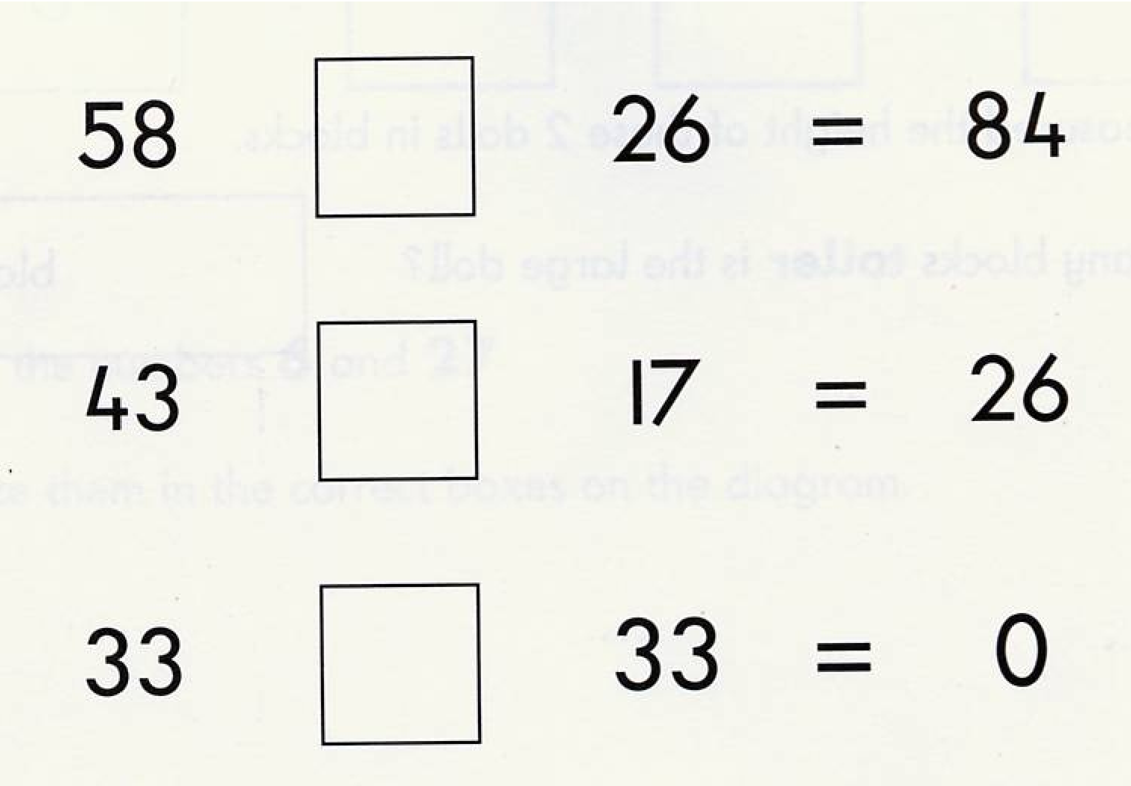


Figure 8.7: Analysis of item 16

Vocabulary

The word *box* is in List 2(a). Children are also likely to see the word fairly frequently on worksheets in contexts similar to this, i.e. where to write an answer. *Correct*, *sign* and *each* were not in Lists 2, although I found in item 15 that average Y2 readers read *each*, a word introduced in year 2 (DfEE, 1999b), with little difficulty. *Correct* is also introduced in year 2 (DfEE, 1999b) with *sign* introduced in the previous year. If not recognised on sight, some children might be able to decode *correct* but this would be unlikely with *sign*.

Sentence length

The sentence of seven words is short and in a familiar format for an instruction.

Comparing colour coding and interview data

I transcribed only two interviews of this item after the live test although further taped interviews, available, if required, show that the two transcribed interviews are typical. Stephanie (see Table 5.8 on page 157) and Marius (see Table 5.9 on page 159) had difficulties reading *correct* and *sign*, both hesitating and making attempts at segmenting these words. Their efforts did not succeed and neither child asked for help with reading; it is possible that Marius did not read the text since he inserted numbers in the boxes.

Other points of interest

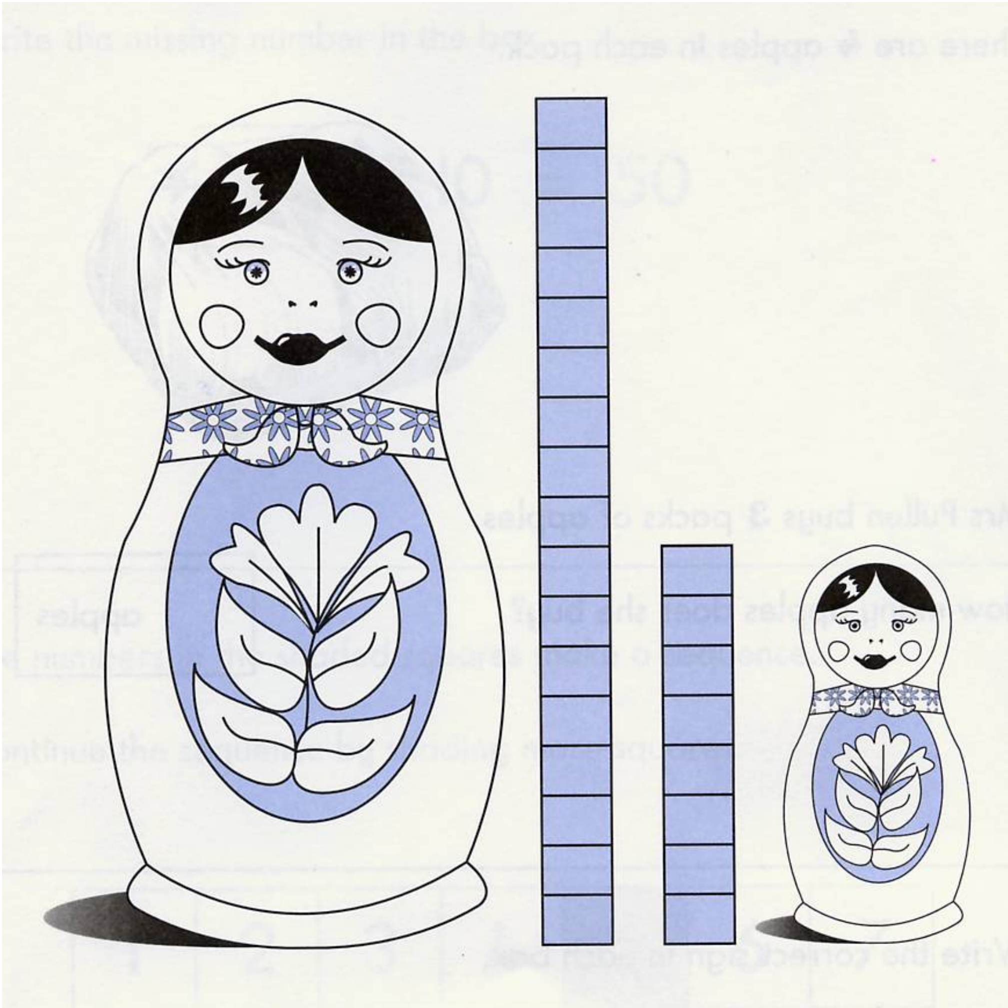
Children may have used one of two approaches to answering this. More confident readers may have read the text first. However, this item does not look text dependent and less confident or lazy readers may have tried to respond using the diagram only (see Lupetti, Sainsbury & Schagen, 1995:13 as discussed in item 14 on page 268). Since signs need to be written in the boxes rather than numbers, which would be more familiar, children may have realised that number answers would not make sense and resorted to the text. Without reading support, many Y2 children are unlikely to have had access to the complete text.

Just over one quarter of the 164 children in my dataset asked for help with reading this item in the live test. About two-thirds of children overall gained the mark (see the *Key Stage 1 technical appendices* (QCA, 2001e).

Predicted readability

Level 2A because of *correct* and *sign*.

Analysis of item 17



Ann measured the height of these 2 dolls in blocks. (9)

17 How many blocks taller is the large doll? (8)

(1) blocks

Figure 8.8: Analysis of item 17

Vocabulary

Of the 18 words to be read, only about half appear to be appropriate sight vocabulary for children reading at level 2. Although not occurring in Lists 2, *Ann*, *doll(s)*, and *large* were words that I considered would be read easily by most children by the end of year 2. *Larger* and *largest* were introduced in Reception (DfEE, 1999b) in the context of numbers and shapes. *Large* was not included although I decided to underline it because the comparative forms had been included. The words coded red had very few or no occurrences on Stuart et al's (2003) and Reid's (1989) databases. *Measure* (rather than *measured*), *height* and *taller* are introduced in Reception (DfEE, 1999b). In the *National Numeracy Strategy* (DfEE, 1999a), it is recommended that measures are covered twice a term in year 2 and this includes all measures, not just length, so *taller* and *height* are not likely to be seen regularly enough to become sight vocabulary. Together with the more complex *measured*, none of these three words lend themselves to phonic decoding. Finally, *block* is introduced in year 2 (DfEE, 1999b) in the context of block graphs, and children may have made 3-D graphs by fitting blocks together as is suggested by the illustration. However, *blocks* occurred three times only in Stuart et al's (2003) database of over a quarter-million words. Some children might decode it successfully, albeit at the expense of time. Because of the illustration, children are likely to be given clues as to how some words should be read, which may assist reading.

Sentence length

The two sentences of nine then eight words are not of unreasonable length. However, not only do they include words that are seemingly difficult to read, they are also packed with information. In this instance, sentence length alone does not suggest the difficulties likely to be associated with reading.

Comparing colour coding and interview data

I transcribed tapes of 14 children in year 2 reading this item, six who took the live test in 2001 (see Table 5.13 on pages 168-170) and eight in June/July 2002 (see

CHAPTER 7). Of these, eleven were regarded as average readers and I comment on these first. Only two children read *measured* correctly although several others came close, reading it as *measure* or *measures*, sometimes after hesitating to think it through. Several children had to be told the word after hesitating or struggling to decode it. There were mispronunciations too such as *mestured* and *me-sure* (two words) and, from another child, *message*, both arrived at after extended pauses. Children coped better than expected with reading *height* although there were some hesitations and stumbling before reading it correctly. Only one child misread it completely as *head.* The diagram may have cued children that measurement of height was required but I did not expect so many to recognise the word. Most children could read *blocks*, sometimes after a short hesitation or blending sounds together but a few children had to be prompted. *Taller* was found difficult to read with most children making more than one attempt to read it or hesitating for up to four seconds before reading it. Some children read it initially as *tall* then self-corrected to *taller* but a few children did not self-correct. Those reading the word as *tall* would have been unlikely to answer correctly since it altered the question to one not requiring a comparison. For example, Joshua (see Table 5.13 on pages 168-170) read *tall* rather than *taller* and gave an answer of 17 blocks which would correspond with what he read. Of the words coded yellow, *Ann* was misread by one Y2 girl but she went on to give the correct answer. I was surprised at all the variations for *large*, most of which were sensible substitutions, perhaps prompted by the diagram. These included *lard*, *longer*, *lady*, *largest* and *long.* One child used phonics successfully and another paused before reading it. Only about half of the 11 children read *large* as sight vocabulary. The diagram provided a prompt for *doll(s)* that was read correctly by everyone although a few children paused before reading it. Children reading at level 2C, such as Priya and Scott, simply did not cope (see Table 5.13 on pages 168-170). Daniel (see Table 5.13 on pages 168-170), reading at level 2A, read all words accurately but his reading was punctuated with momentary pauses. In some instances, children who made minor reading errors could still explain what to do and/or give the correct answer. In year 3 (see CHAPTER 7), children generally read more accurately but words like *measured* were still found difficult by a minority. Only three of the eight Y3 children read every word correctly and with minimal use of phonics or hesitations. Year 4 average readers generally read accurately and fluently as explained in CHAPTER 7.

In general, the colour coding and interview data told the same story. Some children read words coded red better than I had predicted, but not always with fluency or without hesitation. I believe that few children reading at level 2 would have coped with the text without reading assistance, in spite of the illustration.

Other issues of interest

Just over one third of the 164 children in my dataset asked for help with reading item 17 during the live test. However, some children, whose work is transcribed and shows reading errors, did not ask for assistance. It is also possible that some children guessed what to do from the diagram (see Lupetti, Sainsbury & Schagen, 1995:13), particularly as the text is positioned underneath. Because the item is text-dependent, guessing was unlikely to lead to the correct answer.

As explained in CHAPTER 7, teachers could read item 17 with understanding in about eight seconds. For the eight average readers in year 3, the average reading time was 14 seconds with a maximum reading time of 28 seconds. For the eleven average readers in year 2, the average reading time was 36 seconds, with one child taking 61 seconds. The timings in themselves are indicative of the difficulties faced by the Y2 children, and, in some cases, by Y3 children who had an extra year of reading experience.

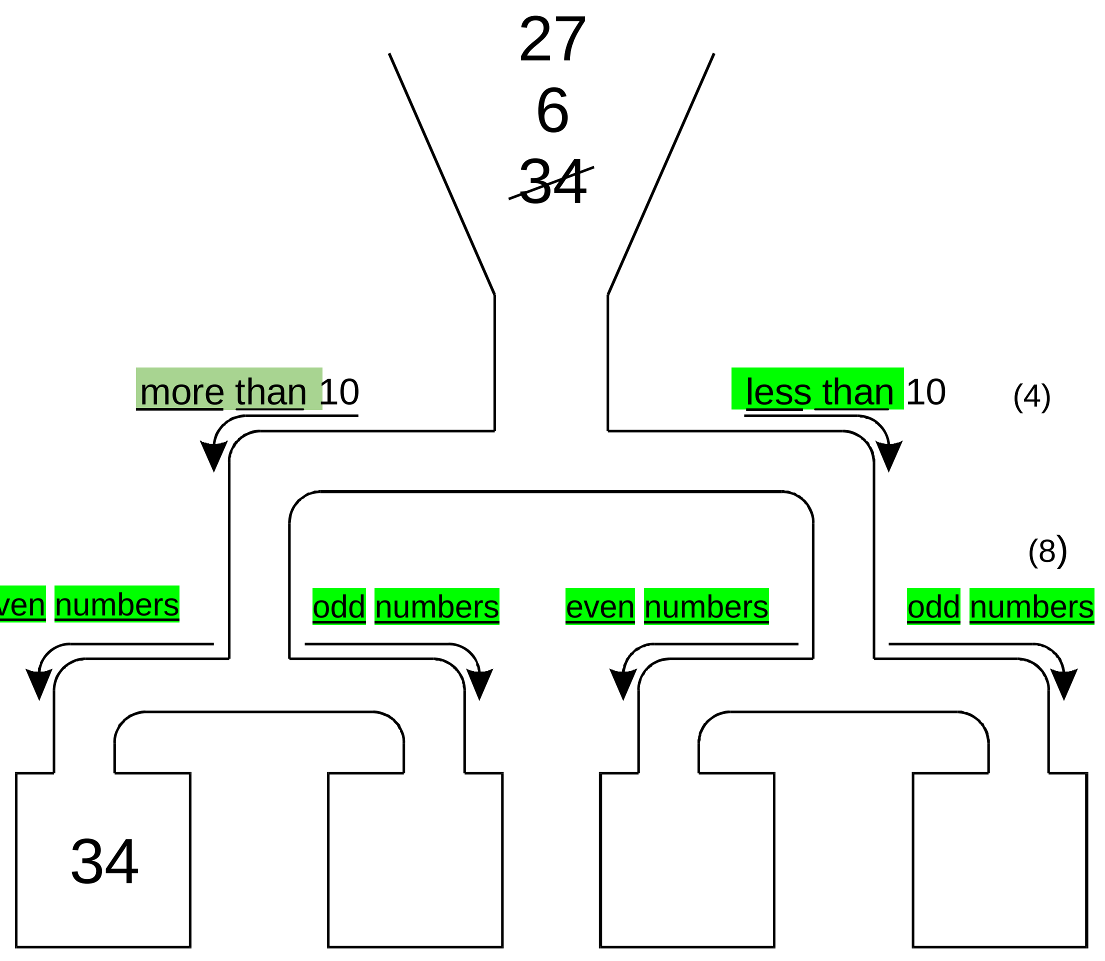
Predicted readability

Even in year 3, less than half the children were reading with the fluency and accuracy necessary for independent reading. Consequently, I feel that the text is suitable for children reading securely within level 3.

Analysis of item 18

Here is a sorting diagram. (5)

The number **34** has been sorted. (5)



18 Sort the numbers **6** and **27** (4)

Write them in the correct boxes on the diagram. (9)

Figure 8.9: Analysis of item 18

Vocabulary

On List 1, I included *sort* because I considered that most Y2 children would read it on sight. Indeed, *sort* is introduced as a verb in reception year (DfEE, 1999b), and children are likely to hear and see it frequently across a range of topics. However, I am less sure that *sorting* and *sorted* will be as familiar on sight. It could be argued, and I would agree, that many Y2 children should be able to decode them with little difficulty by combining *sort* and *–ing* and *–ed*, common endings for many words, but children should not be expected to decode words in a mathematics test. The noun *diagram* is not introduced formally until year 3 (DfEE, 1999b) so may be used infrequently in year 2, and least of all in written form. It occurs only five times on Stuart et al's (2003) database and not at all on that of Reid (1989). Not only that, it is phonetically irregular. The word *box* was included in List 2(a) and I would expect that many Y2 readers would read the plural form with little difficulty. I was surprised to find that *correct* occurred only twice in over a quarter-million occurrences of words in Stuart et al's (2003) list. Even though I considered that it would appear fairly commonly in worksheets etc., I coded it red since I could not assume that it would be sight vocabulary. Overall, there appear to be at least five words that many children are unlikely to have internalised as sight vocabulary.

Sentence length

As well as the 12 words on the diagram, children had to read four short sentences ranging from 4 to 9 words (excluding numerals). Three sentences include words coded yellow or red although most words are coded green as entries on List 1. In total, there are 35 words, although some words are repeated e.g. *odd* and *numbers*, which assists the flow of reading. Nevertheless, this item has about double the number of words of item 17, which, up to its position in the booklet, had more words than any other written item. The second sentence of five words, although short, includes the passive tense, which writers are advised to avoid since ‘they make test questions less accessible’ (Clausen-May, 2001:28). Initially, I could think of no way to avoid this but perhaps an alternative wording could have been, *The number 34 is more than 10 and even.* However, such a change would have helped with interpretation of the diagram, which was part of what was being assessed.

Comparing colour coding and interview data

I transcribed only three interviews of Y2 children reading this item after the live test although further interviews on audio-tape are available. As with some children in item 15, Priya (see Table 5.7 on page 155), a below-average reader, and Stephanie (see Table 5.8 on page 157), an average reader, ignored text above the diagram and started reading at the third sentence. I pointed this out and she then read most words correctly but her main error was to substitute *start* for *sorted* and *sort* after several failed and time-consuming attempts, thereby making two sentences largely meaningless. I was surprised that she could not read *sort*, a verb commonly used in worksheets etc. She took 46 seconds to read the item, mainly because of these two words. She responded correctly in the test without reading support, which suggests that the graphics and the words on the diagram, which she read fluently, may have been sufficient to cue what to do. Priya found such great difficulty in reading the words in the last two sentences, including *number* and all words coded yellow or red, that I did not ask him to start at the beginning. In the live test, he requested help with reading and was awarded the mark. In contrast, Daniel, an above-average reader (see Table 5.11 on page 162) read all text correctly within 20 seconds but he hesitated before both occurrences of *diagram*.

The evidence from transcribed interviews is scant but only the above-average reader coped well. More data would be needed to predict how well the coding is anticipating reading difficulties.

Other points of interest

About 40% of children in my dataset asked for help with reading this item. Children had to read 35 words with text split above and below the diagram. Apart from the split text, children may have been discouraged by the amount of information on the page or by an unfamiliar layout. However, in the *Key Stage 1 technical appendices* (QCA, 2001e), just over three-quarters of children were awarded the mark, suggesting that help was given with reading or that children could gain enough meaning from the diagram to know what to do. Reasons for dependence on graphic cues have been discussed in analyses of earlier items.

Predicted readability

For fluent and accurate reading, I predict that a child would have to be reading on the borders of levels 2 and 3.

Reasons for frequent requests for help

In general, children coped well with words coded green and yellow. However, with one exception in item 18, all sentences contained at least one but more commonly two or three words coded red. Item 14, two sentences long, and item 15, three sentences, long, each had seven words coded red. This suggested that each of the five items had too many low-frequency words, further implying that it was unreasonable to expect most Y2 children to read without hesitation or decoding, actions that would slow down their progress and detract from the mathematics. Fortunately, many children in my dataset recognised the need for help with reading but those who did not were likely to have relied on graphic cues or wasted time and energy using word decoding skills.
I believed that sentence length and syntax were less of a reason for requesting help with reading than the words within the sentences and the number of sentences to be read. In his readability formula for application to short passages with at least three

sentences, Fry (1990) estimated that if sentences were on average between 6.7 and 8.6 words long, then the sentence difficulty (but not necessarily the word difficulty) was suitable for children in grade 2 (7-year olds). In the test, sentences were most commonly seven or eight words long but with a range of from two to 11 words. However, some sentences that were relatively short contained words unlikely to be familiar in written form; indeed some items included more than one sentence that came into this category, for example item 31. The more sentences there are to read, the greater the demands must be on ‘memory span’ (Gilliland, 1972: 91) and on comprehension skills. A child faced with reading several sentences might be discouraged and automatically request help with reading. In my review of the literature, I discussed how even short sentences in subject-specific text could be ‘packed tight with meaning’ (Mobley, 1986:30). Mathematical instructions are often concise but information-rich and children have not only to read them but to unpack their meaning to respond. It is doubtful whether some Y2 children reading independently could unpack the meaning due to reading limitations.

Sentence and question structure too were generally conventional but only caused me concern because Perera (1980) found evidence that ‘children read more easily those sentence structures that they would themselves say or write’ (p156). The language of a mathematics test bears little resemblance to how children would speak or write. As the coding revealed, the test included many words that children used rarely in their writing as confirmed by Reid’s (1989) list.

Another possible reason why children may have asked for so much help could have been the cumulative effect of reading text that stretched or over-stretched their reading capacity. Items 14 to 18, five of the items analysed in detail, were consecutive and included words that were coded red and/or comprised several sentences. Working through this block of items must have been hard work for many children unless reading assistance was requested. The amount to be read is discussed in more detail later in this chapter.

Items colour-coded only in Appendix 8.2 show that the items chosen for detailed analysis are not atypical. There are potential reading difficulties for Y2 readers in these items later in the test when some children will be mentally tiring.

Strengths and weaknesses of the coding system

Comparison with transcripts revealed that the colour-coding was accurate for the majority of words and a useful instrument for identifying words that were likely to cause reading difficulties. I certainly do not claim that every red word would have been beyond the competence of Y2 readers. It is likely that above-average readers would sight read at least some of these, e.g. mono-syllabic words that are phonetically simple such as zoo, *pack* or *need*. What I can claim is that red words were unlikely to be familiar to most children in written form and that is cause enough for concern in a test where reading is not being assessed. Nor can I assume that all children would read all green words as sight vocabulary, for as Perera (2000) stated, '[i]f words on a 'familiar word list' are not really familiar, there is no guarantee that their presence in a text will contribute to reading ease' (p155). Although the colour-coding complements the transcripts, the transcripts provide the more convincing evidence since they show actual rather than predicted reading difficulties and the laborious efforts that some children made to read items. However, this does not mean that such a word-by-word analysis has not been worthwhile. It has provided me with some evidence as to *why* children found items too hard to read. Nevertheless, I would never recommend that test items were written using only the word lists in Appendix 8.1 since these lists include few of the Mathematical English (ME) terms that Y2 children are expected to know and understand by the end of the key stage. I see its use more in the *checking* of new items for potential reading difficulties so that steps could be taken to try to simplify wording or to prevent children being expected to read them independently. An item is only as easy to read as the most difficult word and even one word coded red may make it inaccessible to its reader.

Other aspects of readability

At the outset of this chapter, I included a checklist that summarised points that Mobley (1986) believed should be considered in the study of readability. Most of the points have been dealt with on an item-by-item basis. I now deal with those of a more general nature and start with those that cause least concern. I regarded the legibility of the booklet to be entirely appropriate and have no recollection of any Y2 teachers or children over the years commenting adversely on font size, letter or numeral style or paper quality. The same applies to the quality of the illustrations although I have already reported that children often missed text that preceded illustrations. Of more concern is interest level. I can only surmise that children who are struggling to read or constantly asking for help with reading cannot be highly motivated during the test even though they might enjoy classroom mathematics. However, my greatest concern came under '[c]ontent and conceptual difficulty' (Mobley, 1986:10), one aspect of which is the amount of information presented at one point in time.

Content and conceptual difficulty

I have already suggested that the amount of information to be read and absorbed in some items would have been daunting to less competent readers in particular. In CHAPTER 5, I commented that requests for help by children reading at level 3 seemed to be clustered towards the end of the test at which point the combined demands of the reading and mathematics were likely to have caused mental fatigue. Even these competent readers may have lacked the energy to sustain independent or fluent reading. How much worse it must have been for the majority of children whose reading was below this level, even with intermittent help with reading. I constructed the spreadsheet in Appendix 8.3 because I was curious to find out how many words altogether children were faced with reading and was shocked to find that my estimate of about 250 words was understated. The data showed that a child reading all text independently starting with item 6, would have read 409 words (tokens) by the end of the test, with 169 of these words (types) occurring only once. When colour-coding was added, the summative results were as in Table 8.3.

Table 8.3: Breakdown of words in KS1 2001 mathematics test

Types	Tokens	G types	G tokens	Y types	Y tokens	R types	R tokens
169	409	76	261	16	31	77	117

Key: G – green; Y – yellow; R – red

The summative data for words coded red gave me concern because:

i. of the 409 words (tokens) in the test, 117 (28.6%) were coded red;

ii. of the 169 unique words (types), 77 (45.6%) were coded red;

iii. of the 117 words coded red (tokens), 77 (65.8%) of these occurred only once.

From these data, I predicted that children were faced with low-frequency words that comprised nearly 30% of words in the test, of which about two-thirds occurred only once, as they turned the pages. Almost half of the word types were coded red. These included a mixture of OE and ME words. Even if I assume that children would try to decode some words, since this is what they are encouraged to do in independent reading, many red words are phonetically irregular, e.g. *height* or multi-syllabic, e.g. *millilitres*. Such words can only be read through having been memorised as a result of frequent exposures at regular intervals. However, Stuart, Masterson & Dixon's (2000) study found that 36 exposures to 16 words embedded in stories over a period of several months did not guarantee memorisation for 5-year olds. I have no doubt that few of the red words in the test would have been seen frequently enough to be memorised by all but the most competent Y2 readers. Another factor is that the language of the test in no way resembles natural 'child-speak'. Indeed, as discussed in my review of the literature, *A language for life* (Committee of Inquiry into reading and the use of English,1975:72) stated that 'a number of studies show that a printed text is easier to read the more closely its structures are related to those used by the reader in normal speech' Perera (1980) made the same point.

Out of interest and empathy, I put myself in the place of the child and coded all words in booklet order that a child would read, if assistance was not requested or offered. Of course, this is the worst scenario because most children would have had at least some reading assistance but there were poor readers in my dataset who asked for no or little help. Table 8.4 shows all text in item order and includes word counts for each item.

Table 8.4: Summative analysis of all words in KS1 2001 mathematics booklet (item order)

Item	Text	G	Y	R	Total	%
Practice (read with teacher)	Some children made this graph. The colour we like red orange blue green yellow number of children (Practice question: Look at the graph. Which colour do 5 children like best?)	n/a	n/a	n/a	n/a	
6	Look at the graph. How many children like red best? children	9	1	1	11	4%
7	How much money is in the money box?	5	1	2	8	5%
8	Look at these cards. Use one card each time to make these correct.	8	0	5	13	12%
9	Write the answers.	3	0	0	3	5%
10	Tick (✓) the shape which has more than half blue.	9	0	0	9	17%
11	Draw arrows to show which shapes belong in the set. has 5 corners	7	2	3	12	35%
12	Add more fives until the total is 25	5	0	2	7	18%

Table 8.4: Summative analysis of all words in KS1 2001 mathematics booklet (continued)

Item	Text	G	Y	R	Total	%
13	Write the missing number in the box.	5	1	1	7	12%
14	The numbers in the shaded squares make a sequence. Continue the sequence by shading more squares.	9	0	7	16	46%
15	There are 4 apples in each pack. Mrs Pullen buys 3 packs of apples. How many apples does she buy? Apples	8	4	7	19	25%
16	Write the correct sign in each box.	3	1	3	7	26%
17	Ann measured the height of these two dolls in blocks. How many blocks taller is the large doll? Blocks	9	4	5	18	34%
18	Here is a sorting diagram. The number 34 has been sorted. More than 10 less than 10 even numbers odd numbers even numbers odd numbers Sort the numbers 6 and 27 Write them in the correct boxes on the diagram.	29	1	5	35	38%
19	The bus left at 9 o'clock to go to the zoo. It arrived 1 hour and 15 minutes later. Draw a ring around the time it got to the zoo.	18	1	8	27	21%
20	This table shows the ages of some children. Name Age Fred 7 years 4 months Harriet 7 years 0 months Isla 7 years 10 months Julian 7 years 6 months Kate 6 years 11 months Asim 6 years 11 months How many children are older than Harriet? Children	12	2	22	36	29%
21	Who is the youngest?	3	0	1	4	18%
22	Join two other numbers which total 100	4	1	1	6	25%
23	Write the total	2	0	1	3	9%
24	Shade the correct triangle in the last hexagon starting shape after a half turn	8	1	5	14	23%
25	Match each number in a box to the nearest 10 to the nearest 10 is	9	1	3	13	23%

Table 8.4: Summative analysis of all words in KS1 2001 mathematics booklet (continued)

Item	Text	G	Y	R	Total	%
26	Write the number which is half of 38	7	0	0	7	23%
27	Write 3 of these numbers in the empty boxes. The numbers in the boxes must be in order. largest smallest	14	2	3	19	25%
28	Write the missing number in the box.	5	1	1	7	13%
29	Draw the 2 lines of symmetry on this shape. You may use a mirror.	9	0	4	13	34%
30	Add together 24, 67 and 45. Show how you work it out in the box. The total is	11	2	2	15	26%
31	Sadi needs 26 cartons of juice for her party. There are four cartons in a pack. How many packs does she need to buy? packs	12	0	12	24	26%
32	Write the answer.	3	0	0	3	5%
33	Ellen has a £5 note. She spends £1.99 Draw a ring around each coin she gets in her change.	10	1	6	17	19%
34	Write the number in the box to make this correct.	8	1	1	10	12%
35	This jug has water in it. millilitres Ravi pours 150 millilitres of water out of this jug. How much water will be left in the jug? millilitres	18	3	5	26	20%
	Totals for booklet	261	31	117	409	n/a

Key:
G: frequency of green words
Y: frequency of yellow words
R: frequency of red words
Total: number of words to be read per item
%: percentage of children in dataset who asked for help with reading during live test

Because of the colour coding, the table requires little supporting commentary. I ask the reader to imagine an average or below-average Y2 reader in particular who, left to read independently, but, expected to know when help with reading should be requested, encounters the text for the 30 written items (excluding the practice) that move from one context to another. Only in four items are all words coded green, i.e. items 9, 10, 26 and 32, and three of these are abstract. All other items include at least one red word and I calculate an average of almost four red words for each item. From items 14 to 20 towards the middle of the test, for example, the coding on Table 8.4 suggests that the reading is difficult across this block of items. This was also borne out by data collected during the live test where at least one-fifth of children in my dataset requested help with reading each item in this block. This is not surprising judging by the amount and predicted difficulty of the text.

Although there were about the same number of unique red as green words (types), those coded red were not only less likely to be recognised on sight but also less likely to be repeated since they were often particular to a context, e.g. *symmetry* in item 29. Consequently, children were less likely to benefit from having seen or being helped with reading a red word previously. However, where children asked for help with reading a red word that was repeated elsewhere, it cannot be assumed that they would have remembered that word after one reading particularly where the occurrences were separated by intervening text. Indeed, I noted in my interviews with children that where I told a word to children, e.g. the first occurrence of *sequence* in item 14, they did not always recognise it on its second occurrence in the same item. In contrast, children asking for help with reading *years* and *months* in item 20 would have been more likely to sustain recognition of the words, even if read only once to them, since the words are repeated on consecutive rows of a table. Frequent repetition of words is important in texts for emerging readers if they are to recognise them subsequently as found by Stuart, Masterson & Dixon (2000) but red words had low frequencies so recognition would have been unlikely.

In CHAPTER 1, I explained how teachers could read with the children a list of words in the test that children had found difficult to read in pre-tests. In the light of Stuart, Masterson and Dixon’s (2000) research, it now seems that if these words are unfamiliar in written form at the start of the test, they are likely to remain so during the test. Indeed, JV, an experienced Y2 teacher and observer of the pre-test that was the focus of my pilot study said:

> For the lower ability readers, putting a word bank on the wall before you start the test, is not really a particularly useful strategy because the children have forgotten what the words say by the time they need to access them.
> (Appendix 3.6)

The outcome for children struggling to read and not requesting or being offered sufficient help would surely have been mental fatigue but without ever gaining full access to the text. As Pressley (2000) explains, ‘when readers are first struggling to recognise a word, much short-term capacity must be dedicated to sounding out the word, with the result that comprehension of words is uncertain’ (p33); such dedication would surely be at the expense of easy access to the mathematics for struggling independent readers. Johnstone’s (1988) study found that the key factor that made examination questions difficult was the capacity of working memory in contrast to Mobley (1986) who considered that it was the readability of questions. In view of Pressley’s (2000) preceding comment, perhaps they are both correct. Further evidence of Y2 children’s struggle was the excessive time that they took to read some items as shown by the timings in transcripts in CHAPTERS 5 and 7, because children had to pause to decode or work out the meaning of unfamiliar words. This could only have increased the conceptual burden. I repeat a relevant quote from my review of the literature taken from *A language for life*:

> If children are faced with texts containing more than a very small proportion of unfamiliar words they will spend far too much time struggling at frustration level and will derive neither meaning nor enjoyment.
> (Committee of Inquiry into reading and the use of English, 1975:103)

In summary, the combination of unfamiliar words, the amount of text overall, the variety of vocabulary and the length of the test must have been daunting for some children, even with reading assistance on request.

The wider picture

Equality of opportunity?

I sense that there will be readers who think that I may be painting too pessimistic a picture of the 'loneliness of the long-distance reader' during the test but I do not believe that my concerns are over-stated. The schools involved in my research were sited in varied catchment areas that could be replicated nationally and I have no doubt that the majority of Y2 children nationally were given appropriate reading support. However, my data show that the quality and quantity of reading support varied from school to school as I explained in CHAPTERS 5 and 6. For example, interviews with teachers revealed that some read only words that children indicated whilst others read the complete text for an item; in some schools, children had every item read to them. The amount of exposure given to children, to written forms of vocabulary in mathematical activities, would also vary from teacher to teacher, depending on the status of flashcards, worksheets etc. I have no reason to think that these differing approaches would be unique to my research schools.

The colour-coding shown in Table 8.4 predicts that most Y2 readers would encounter unfamiliar sight vocabulary in most items. However, in my dataset from the live test (see CHAPTER 5), children reading at level 2C or below requested help with reading an average of 12.6 items; those reading at level 2B requested help with reading an average of 6.6 items and those reading at level 2A requested help with an average of 2.6 items. The colour-coding suggests that, for each of these sub-levels, the average number of requests was low compared with the predicted readability of items and, in some cases, children had only part of the text for items read to them. I have already provided evidence in CHAPTERS 5 and 7 that children could not always read items for which they did not request help. I can only conclude that the overall text was too difficult for many Y2 readers and that too much responsibility was put on children to ask for reading support, resulting in some not being given the help required.

In spite of the seemingly difficult vocabulary, some readers may believe that the range of contexts and vocabulary in the test is appropriate, even desirable. I agree, and in my review of the literature, in line with the views of Shuard & Rothery (1984), I argued against reducing the text only to words familiar in everyday speech since interpretation of ME words forms an essential part of children's mathematical development.

Harrison's recommendation to use word frequency lists at the planning stage 'to restrict vocabulary at the authorship stage' (p20) has partial approval from me. Using word frequency lists have certainly helped me to identify words likely to be unfamiliar. However, restricting vocabulary to words likely to be familiar in written form would leave the test with few contextual problems and with a heavy bias towards abstract items. On no account would I wish to argue for a test that comprised mainly abstract items on the grounds that they require few or no words. My concern lies not with the breadth of coverage of the test but with vocabulary that is likely to be unfamiliar in written form.

Some may argue that illustrations in some items supported understanding of the text. Again, I agree but diagrams can mislead as well as prompt if a child picks up the wrong graphic cues and ignores text. Some may argue that a child can have as much help with reading as required. However, the latter argument completely misses the point. If a child is expected to read even some of the text independently, the reading demand should be well within that child's competence. In my review of the literature, I quoted Harrison (1989) who believed that worksheets or instructions for a task 'ought to be in simpler prose than anything else a child reads' (p84), Mobley (1986) who believed that unsupported reading 'may need to be two years *below* [sic] the pupil's reading level' (p49) and Klare (1963) whose research showed that readers generally internalised more information when texts were below their reading level. Newman (1977:252) too found that about 13% of errors in a mathematics test with carefully controlled vocabulary arose because low-attaining 11-year old pupils were unable to read questions accurately enough. Nearly all errors were in reading key words, i.e. words ‘necessary for understanding a passage’ (Fry, 1990:595). It is not possible for a statutory KS1 mathematics test covering the breadth of the programme of study to be written throughout in simple prose or so that, in theory, a 5-year old could read it. However, in CHAPTER 7 I concluded that it was not until I heard average readers in year 4 read some items from the test that I felt confident that they could generally cope independently. To maintain the breadth of coverage of the test and provide equality of opportunity, I believe that the only realistic option is for teachers to read each item with the children, especially those reading within level 2. This form of administration i.e. look-listen was the focus of CHAPTER 6 where I argued that its strengths outweighed its weaknesses. Its greatest strength is that children do not have to read independently and teachers can then have more confidence that incorrect responses are due to mathematical rather than reading difficulties. Having undertaken this laborious but worthwhile colour-coding analysis, I am even more convinced that ‘look-listen’ should be the main method of administration promoted by QCA rather than ‘hands up when you need help with reading’.

In CHAPTER 1, I explained the rigorous procedure for developing a KS1 mathematics test. By the time children take the test, each item will have been through an informal trial and two extensive pre-tests, commented on by teachers, and reviewed by writers and panels of experts in mathematics and special educational needs. During this procedure, language and syntax are taken into account and modifications made at different stages of the test’s development. In spite of everyone’s best endeavours, including my own, the 2001 test included too many words that children were unlikely to recognise on sight. I now believe that the reviewing process could be improved by:

i. having a KS1 literacy specialist on reviewing panels;

ii. having word lists such as those in Appendix 8.1 for guidance for writers and reviewers;

iii. getting a representative sample of teachers taking part in pre-tests to underline text that children needed help with reading so that unfamiliar sight vocabulary could be identified;

iv. getting consultants observing pre-tests to hear a few children of different reading competences read items that had been identified previously as containing unfamiliar sight words to determine the appropriateness of the text for its audience.

Of these, children's involvement in reading items aloud would provide the most telling evidence of whether the reading demand was appropriate.

Multi-method approach

In CHAPTER 4, I explained my intention to and reasons for investigating the reading demand of the KS1 2001 mathematics test using a variety of approaches. The analysis of the booklet provides the final element of my multi-method approach. Figure 8.10 summarises how I triangulated the data collected from children and their test booklet.

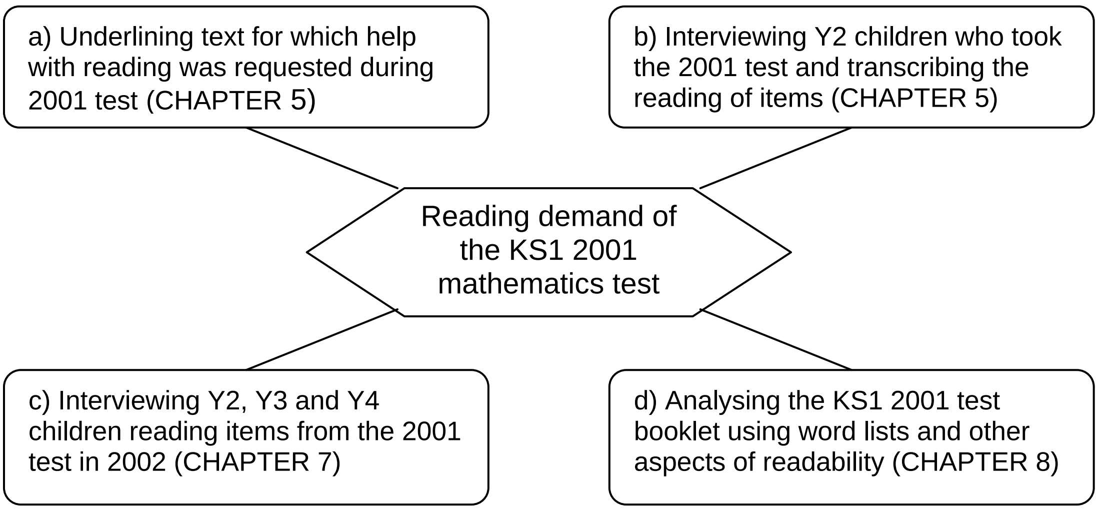
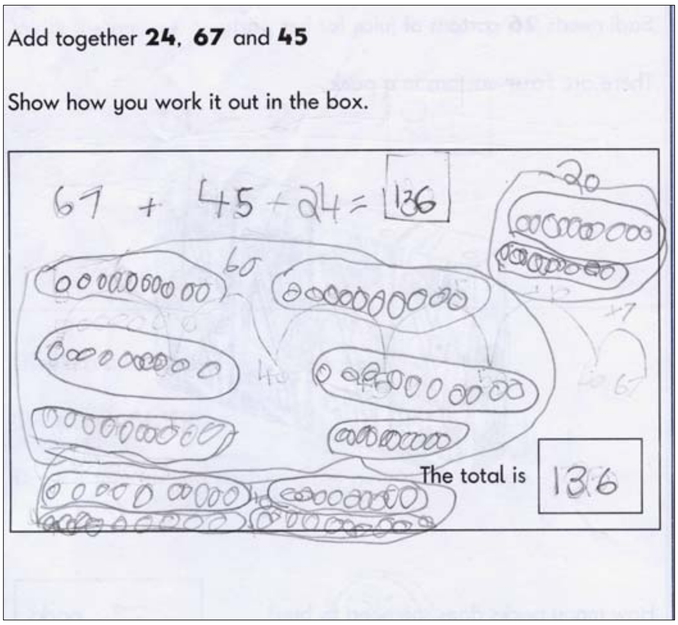


Figure 8.10: Elements of multi-method approach

The main purpose of Approach a) was to identify the extent of children's requests for help with reading during the 2001 test and the items for which most help was needed. Approach b), influenced by the evidence from Approach a), was to audio-tape and transcribe Y2 children reading some items to establish where potential reading difficulties lay and to find out whether they could read items for which reading support had not been requested. Approach c) took a similar approach to Approach b). This was to confirm that reading difficulties that Y2 children had in 2001 were not unique and that children at the end of key stage in 2002 shared these difficulties. At this juncture, the reading competences of average Y2 children reading items were compared with average Y3 and Y4 readers. Finally, wondering *why* Y2 children had difficulties with reading, I undertook the analysis of the booklet explained in this chapter, Approach d). To complement these approaches, I sought teachers views on the reading demand of the test and on their approaches to administration which were often influenced by the limitations of children's reading skills (see pages 142-143 of CHAPTER 5 and pages 179-185 of CHAPTER 6). Although teachers' views were opinions rather than evidence, they supported my conclusions from each approach that the reading demand was almost certainly excessive for many if not most Y2 children. As Clausen-May stated, 'it seems obvious that tests should be accessible, if only to ensure that they are valid' (p35), i.e. test what they are intended to test. A similar view was stated by Homan et al (1994) that 'when readability is not controlled it might impact the validity of test scores' (p349). Taking my multi-method approaches individually or collectively, I can only conclude that the validity of the test was under threat for any child who did not receive as much reading support as required. My evidence suggests that there were such children.

Cohen, Manion & Morrison state that 'multiple methods are suitable where a controversial aspect of education needs to be evaluated more fully' (2000:115). There is no doubt that, since its outset, statutory testing at key stage 1 has had many critics and caused much controversy.

CHAPTER 9



Sasha was working at level 3 in mathematics.

CHAPTER 9

The final chapter of my story

Not just an afterthought

For myself, as well as for my readers, this chapter provides the last chance to review the entirety of this study and to reconsider its implications for Y2 children and their teachers. To get an idea of elements that should be included in a conclusion, I referred to relevant literature (see, for example, Miles & Huberman, 1994; Silverman, 2005 and Wellington et al, 2005). There was general agreement that this should be carefully crafted and 'not written hastily as an afterthought' (Wellington et al, 2005:171). I shall also try to follow Silverman's (2005) advice that:

> An imaginative conclusion will move on from the careful description and analysis of your earlier chapters to a stimulating but critical review of the overall implications of your research. (Silverman, 2005:326)

I leave it to readers to judge whether I meet these criteria, starting with my empirical conclusions.

Empirical conclusions

Data collected during the KS1 2001 mathematics test

In CHAPTERS 5, 7 and 8, the main sources of evidence for my study, I tried to apply faithfully the methodological choices, justified in CHAPTER 4, that I believed gave me the best chance of answering my research questions. In CHAPTER 5, I explained how I was in the fortunate position of being able to collect data from 164 children during the KS1 2001 mathematics test and from a sub-set of these children and their teachers in the following weeks and months. By analysing and interpreting these data, I obtained answers to four inter-related research questions:

Is the reading demand of the text in a typical KS1 mathematics test appropriate for the target children in year 2?

What is the extent of the help needed by children of different reading competences during the test?

What are the views of year 2 teachers on the reading demand of the test?

Is teacher support with reading available to children during the test whether they ask for it or not?

I am quietly confident that my evidence in CHAPTER 5 will have convinced readers that:

i. the reading demand of the 2001 mathematics test was almost certain to exceed the skills of many, perhaps most Y2 children reading within level 2, but especially those reading at level 2B and below;
ii. appropriate teacher support was not always available when children were left to read independently;
iii. the extent of help needed varied widely within and between children of different reading competences.

My strongest evidence leading to these conclusions came from Y2 children themselves.

Annotations in test booklets, made during the test, showed that some children requested frequent help with reading, suggesting that they could not read the text, did not want to read it or lacked confidence in their reading skills. Whatever the reason, a request for help suggested that a child wished to hear the text read by an adult. However, the fact that the average number of requests for help with reading increased as the children's reading level decreased, suggests that reading difficulties were the main catalyst for needing support. For example, children reading at level 2C were almost twice as likely to ask for help as those reading at level 2B and almost four times as likely as those reading at level 2A. This did not surprise me but I was surprised and concerned at the range of requests even within children of similar reading competences, e.g. children reading at level 2C requested help with reading from 0-29 items. For children whose reading skills were reasonably similar, data like these convinced me that some children could not be receiving the reading support that they required.

I also found evidence that children frequently asked for help with particular items. For example, at least one-third of children asked for help with reading items 14 and 17, both ‘wordy’ items and both including several ME words. In contrast, children rarely requested help with reading items involving abstract calculations where text was minimal or which could be interpreted without reading text. The contrast in children’s reaction to these two types of item was further evidence that children needed help with reading items that were ‘wordy’ and text-dependent.

However, the children’s voices literally provided the most compelling evidence. In diagnostic interviews with children reading within level 2, I found that they could not always read items for which teacher support had not been requested or offered. In the worst cases, as can be seen from the transcripts of children’s reading of items (see Tables 5.7 - 5.13), children demonstrated all the characteristics of poor readers as defined by QCA (2001a) (see pages 19-21 in CHAPTER 2). The inadequacy of their reading concerned me greatly since what they read sometimes made no sense and must have prevented access to some mathematics. This showed that teacher support was not always available whether children asked for it or not. To me this suggested that too much responsibility was put on children to ask for support with reading. This was all the more surprising because the seven teachers I interviewed, who expected children to read independently and to ask for support, were unanimous in their views that the reading was demanding, but, with one exception, they were not convinced that children always asked for as much help as they needed. This contrasted with the approach used by teachers who had adopted the practice called ‘look-listen’ for administering the test, as discussed in CHAPTER 6, where each item was read with children. In this situation, teachers assumed responsibility for reading, thus leaving children to concentrate on the mathematics. Teachers explained that they had adopted this practice because of their concerns about the reading demand and, in some cases, also because of the amount of text children had to read.

Comparing readers in years 2, 3 and 4

To triangulate data obtained in 2001, I adapted the procedure a year or so later to include more Y2 readers but also some Y3 and Y4 readers, as is explained in detail in CHAPTER 7. From these data, I was able to confirm that average Y2 readers had difficulties reading four items from the 2001 mathematics test that had also been found difficult by Y2 readers a year earlier. This provided further evidence that the reading demand was not appropriate for children reading at level 2B or below. That left me with one unanswered question, *What is the reading age required by a Y2 child in order to read the text in a typical KS1 mathematics test fluently and accurately?* In CHAPTER 2, I justified in my review of the literature why no readability formula could be applied to the test, so, once again, I returned to the child-centred involvement promoted by Shuard & Rothery (1984) as well as Harrison (1979) who recommended monitoring children’s responses as ‘a perfectly valid method of assessing whether children can cope with a particular book’ (p73). In this chapter, I coded in detail and compared transcripts of average Y2, Y3 and Y4 readers in term 3 reading the same four items from the 2001 test (see Tables 7.2-7.11). I found that average Y3 readers, who would generally be reading on the borders of levels 2 and 3 had some difficulties; average Y2 readers had the greatest, and in some cases, extreme difficulties. This led me to conclude that it was not until children had the reading skills of average Y4 readers, i.e. reading at approximately one reading level higher than those in year 2, that I would be confident of their competence to read the test independently. I would have been less worried about this had I been convinced that all Y2 children received the reading support that they required during test administration, but I obtained evidence in CHAPTER 5 that suggested that this was not the case. In 2001, 29% of Y2 children were awarded level 3 in reading. I would predict that the majority of these would be reading more with the competence of Y3 rather than Y4 children. Consequently, I would not even feel confident that Y2 children reading at level 3 would cope with the reading demand of *every* item in the mathematics test unless teachers knew that children had a reading age at least two years ahead of their chronological age. This conclusion reminded me of literature quoted in CHAPTER 2. Harrison's (1979) view was that worksheets or instructions 'ought to be in simpler prose than anything else a child reads' (p84); Mobley's (1986) view was that '[i]f the reading is to be unsupported, it may need to be two years *below* [sic] the pupil's reading level' (p49). Since a Y2 mathematics test written to correspond with the reading skills of average 5-years old is out of the question, the alternative is that teachers need to find more effective methods to maximise reading support so that all children know what they are being asked to do, regardless of mathematical difficulty. For as long as the KS1 mathematics test remains statutory in England, I believe that the only way to guarantee that children have access to the text is through teachers reading text with all but the most competent readers, the 'look-listen' mode of administration to which I return when I suggest ideas for future research.

Test booklet analysis

By the end of CHAPTER 7, I believed that I had enough evidence to respond to my research questions, although I had always intended engaging in a basic analysis of some text in the test booklet as another means of triangulating earlier data. Wiest (2003) believed that '[the] limitations of determining the readability of mathematical text are particularly great', but because I was curious to know *why* Y2 children found text difficult to read, I extended the boundaries of my study. I used familiar word lists and sentence length as the dominant variables to see if a formal analysis would confirm what children had already told me, i.e. that some of the reading had been too difficult for them. Although such an analysis would always be a predictive tool, I concluded that almost 30% of the 409 words in the written items, occurring in all but a few items across the test, were likely to be unfamiliar on sight and would require children to use decoding skills if they tried to read the text and if reading support was not provided. However, since most of these words were phonetically irregular, I was certain, based on teaching experience, that the success rate for independent readers would be minimal and wasteful of time and effort. Nor did I believe that children should have to use decoding skills in a test that was not assessing reading. I was heartened to find that, in most instances, words that I predicted would not be sight vocabulary were often words that children could not read in their diagnostic interviews. Table 8.4 shows at a glance words unlikely to be familiar on sight coded red. Without the children's interviews to compare with appropriate parts of this analysis, my confidence in the colour-coding system would have been less. This third and final aspect of triangulation of my data did indeed corroborate what children's voices had already told me. However, I believe that an analysis of this type shows that it is possible to make a prediction about reading difficulties that Y2 children are likely to face during the test, and, consequently, would be a useful tool in test development.

The status of the KS1 2001 mathematics test

Test reliability and validity

A test is valid if it assesses what it is intended to assess. At various points in this study, I questioned the validity of the KS1 2001 mathematics test by referring to relevant literature (see, for example, Homan, Hewitt & Linder, 1994; Fisher-Hoch & Hughes, 1996; Shorrocks-Taylor & Hargreaves, 2000; Clausen-May 2001 and

Wiliam, 2003). In CHAPTER 7, I cited Fisher-Hoch & Hughes (1996) who compared two types of difficulty in mathematics questions. First, '*valid difficulty*' (Fisher-Hoch & Hughes, 1996:2) is intended since it relates to the mathematical components of questions. Second, '*invalid difficulty*' (Fisher-Hoch & Hughes, 1996:2) is not intended since it relates to features of questions that are non-mathematical. Wiliam (2003) and Homan, Hewitt & Linder (1994) made similar points. Wiliam (2003) argued that , 'a test may be more valid for one group than another – a particular mathematics test may involve complex language, and would tell us little of the mathematical abilities of poor readers, but might work quite well for fluent readers' (Wiliam, 2003:12). The view of Homan, Hewitt & Linder (1994) was that 'when readability is not controlled it might impact the validity of test scores' (p349). From evidence provided by children and the analysis of the booklet, I concluded that the reading demand was likely to be too complex for many, perhaps most, of its target audience. In the absence of appropriate reading support, some children were faced with '*invalid difficulty*' (Fisher-Hoch & Hughes, 1996:2). Until I undertook my analysis of the readability of the booklet in CHAPTER 8, no formal attempts had been made to predict its readability, nor was I aware of any attempts by QCA to control the readability other than through feedback from professionals involved in pre-test development. Consequently, I have no doubts that there would have been an impact on the validity of test scores of any children who did not request or were not offered appropriate support with reading. In their unpublished report, Lupetti, Sainsbury & Schagen (1995) predicted through a statistical analysis that KS1 children who were better readers were likely to gain more marks in a mathematics test than poorer readers of comparable mathematical competence (see Lupetti, Sainsbury & Schagen, 1995:71). My evidence leads me to the same prediction.

I now turn to test reliability and its relevance for my study. Clausen-May (2001) defines test reliability in her explanation that '[i]f the same person had taken the test for the first time on another occasion, then they would have got the same result' (p6). However, reliability would be under threat if the conditions under which the test was taken on another occasion changed so that the same result was not obtained. For example, imagine a child of average mathematical and reading competence who takes the KS1 mathematics test for the first time by reading independently and asking for occasional help with reading, who then has no recall of taking the test. The child then takes the test for the second time with the teacher reading all text-dependent items with him/her. The evidence from the literature (see, for example, Wilkinson, 1980 and Homan et al, 1986) and from my study suggests that the child would be likely to do better the second time around. However, there would be a contradiction in the results. By varying the conditions for the second administration, test reliability would be lessened but test validity would be increased because the test would be assessing mathematics with no threat to validity due to undetected reading difficulties.

Teachers whom I interviewed (see CHAPTERS 5 and 6), rarely administered the test or supported children with reading in exactly the same way. For the KS1 mathematics tests, the teachers' guides allow teachers considerable flexibility in how they group children and support them during the test. However, I have argued at various points in my study that some arrangements for supporting reading appeared to advantage certain children over others. A test that is supposedly 'standard' nationally and is then administered in non-standard ways can also reduce reliability, since as Clausen-May (2001) points out, administering a test 'in exactly the same way to everybody under 'controlled' conditions improves its reliability' (p7). QCA, in a genuine attempt to encourage teachers to make their own judgments about how to make the tests accessible to children, thereby reduced the reliability of the test by allowing such teacher autonomy. I conclude from this that, whilst the test remains statutory, administering it using 'look-listen' to all but the most competent readers, would increase test validity and reliability. The possible exception to this would be for items that are not text-dependent, e.g. abstract calculations where children of comparable mathematical competence would have the same opportunities to succeed regardless of reading skills. Indeed, the test would be fairer if all children had equal access to the mathematical components. Children deserve nothing less.

Evaluating the test

In CHAPTER 4, I positioned myself as someone undertaking research in order to evaluate 'the effectiveness and appropriateness' (Robson, 2002:205) of a KS1 mathematics test during and after its administration, i.e. the test process and its outcomes for Y2 children. Evaluative research is commonly associated with making formative (process-related) and/or summative (outcome-related) judgments. According to Verma & Mallick (1999), one purpose of formative evaluation is to gather data whilst a programme (such as a test) is in progress, and, if necessary, 'to provide evidence of its fundamental weakness so that it can be ended' (p90). Whilst the power to end the test would never be in my control, I can claim to have found some evidence of 'fundamental weaknesses' in the administrative arrangements for tests, i.e. some children in my study were expected to read independently but, unbeknown to their teachers, did not always receive the amount of help with reading that they required. The importance of supporting children appropriately with reading, encouraged by QCA in teachers' guides, was not always reflected in practice. For children reading independently, there appears to be a discrepancy between QCA's expectations that reading skills should not prevent access to the mathematics and what actually happened. Summative evaluation, on the other hand, is 'designed to measure the effectiveness of a programme on its completion' (Verma & Mallick, 1999:90). One aspect of such an evaluation would be whether the test had achieved its purpose which was to 'provide a standard "snap-shot" of attainment at the end of the key stage' (QCA, 1999:8). I have already argued that the test cannot be regarded as standard since teachers' arrangements for administering it (see CHAPTERS 5 and 6) appeared to advantage some children over others; I would also argue that the test would only provide a snap-shot of attainment for children where it was certain that lack of reading skills had not prevented them knowing what mathematics was involved. My evidence showed that this was not the case and the outcome of the test, i.e. the number of marks obtained, was likely to have been affected adversely for some children where reading support was not maximised. Consequently, for such children, the 'value or worth' (Verma & Mallick, 1999:46) or validity of the test was questionable.

Moving on from 2001

Test development is an evolving process and over the years, changes have been made to statutory arrangements. Because of links to my study, I now explain some major changes that took place for tests administered in 2003 and 2005.

First, for KS1 mathematics tests administered in May 2003, there was a move from a test that included items of levels 2 and 3 demand to separate levels 2 and 3 mathematics tests. I believe that this change advantaged poorer readers in particular. The tests were shortened, with a maximum 30 marks in each test rather than 36 marks in the combined test. This decreased the *amount* of text to be read, a positive change, but not, I predict, the overall *reading demand.* It also meant that children who struggled with both mathematics and reading did not have to try to read items pitched at level 3 where they may never have heard or seen some of the terms required by that level of mathematics. In the same year, more emphasis was given in the teacher's guide to remind teachers that they could read all questions to individuals or groups.

Second, for KS1 mathematics tests administered in 2005, although they remained compulsory for eligible children, the status was changed to underpin teacher assessment rather than to have equal status. The main consequence is that teachers report an overall teacher assessment level for each child to LEAs and to the DfES rather than levels achieved in both statutory and teacher assessments. Teachers also report children's teacher assessment levels to parents/guardians although they can request levels obtained by their children in statutory tests/tasks. The other change for that year was that a choice from KS1 mathematics tests from two years, in this instance 2004 or 2005, could be administered from January onwards, rather than compulsorily during May. The effect of these changes is that the status of the tests is greatly diminished. Consequently, if a child does less well in the mathematics test than expected, perhaps because of undetected reading difficulties, the effect is less dramatic since the teacher can judge the child's mathematical attainment on evidence gained over the key stage that would not take reading competence into account. Even in these more relaxed conditions, the child who has reading difficulties still has to take the test so it is not a reason to become complacent.

For KS1 mathematics tests that were administered in the academic year 2005/6, teachers could again choose to administer the 2004 or 2005 tests, or a combination of these, but at *any* time during year 2. Similar arrangements have been announced for tests to be administered in 2007. I predict that most teachers will continue to administer the tests in the summer rather than the autumn or spring terms but, for those who choose the earlier options, I point out one possible undesirable consequence. Children who take the test in the spring term, for example, will have been reading for approximately one term less than those who take it in the summer term, so even greater consideration will have to be given to supported reading.

I welcome the changes introduced in the tests in 2003 and 2005; most of all I welcome the move towards the dominance of teacher assessment over statutory testing. I have reasonable optimism that government policy in England will move still further and abandon statutory testing for KS1 children at the very least. This has already happened in Wales, which, until 2000, had a KS1 testing regime matching that in England but then abolished it, from 2002, in favour of statutory teacher assessment. Schools in Northern Ireland follow a statutory curriculum similar to that in England. However, the first statutory assessments are administered to children in year P4 (8-year olds), but children are not required to sit formal tests. Scotland has non-statutory guidelines for the curriculum and for the assessment of 5-14 year olds. There, children are assessed formally for the first time in year P3 (7-year olds) but, as in Northern Ireland, do not have to sit formal tests. Within the United Kingdom, England stands alone in administering statutory written tests to rising 7-year olds.

Limitations of the study

I raise five main limitations.

i. About 600 000 Y2 children in England take the statutory mathematics test. Because much of my data analysis involved detailed coding of transcripts, time constraints restricted me to surveying only a small sub-set of these children, too few to constitute a national sample. However, by selecting children from a range of schools and catchment areas, I believe that they typified their peer group for reasons I gave on pages 174 - 175 of CHAPTER 5. I also believe that my interviews with Y2 children were sufficient in number for me to predict how readable children of different reading competences in the national sample were likely to find the 2001 test.

ii. For the comparison of the competence of years 2, 3 and 4 children reading the same test items, I selected only a small sample of items from the year 2 mathematics test booklet (as explained on pages 197-199 in CHAPTER 7). This was a pragmatic decision, taken as a result of the time required to code the transcripts. I recognised that there were other items that children may have found easier or harder to read and could not comment, therefore, on how readable children found the whole test. However, the data from this analysis complemented data from other sources, including my literature review, and enabled me to answer confidently the research question I sought to answer by this comparison.

iii. My research focused on the readability of the KS1 mathematics test booklet for one particular year. Because it is similar in style and content to tests from other years, I have argued that the readability of the tests in general is likely to be comparable to the one that has underpinned this study. However, this is a conjecture; I am unaware of any studies that have made such a comparison.

iv. The data collected during administrations of the KS1 2001 mathematics test could only identify items for which children asked for help with reading. There are almost certain to have been other children in my dataset who did not request the help that they needed. Because a child did not ask for help with reading did not necessarily mean that help was not required. Consequently, my data may have understated the amount of help required.

v. In my diagnostic interviews with Y2 children in CHAPTER 5, I asked children to read items aloud. Reading aloud demonstrated the extent and nature of reading difficulties for an item but could not reveal how they dealt with that text during the test when help was not requested or offered. Nor was there any way of knowing if children had more or fewer difficulties during the live test. Children's reading behaviour during statutory mathematics tests is likely to remain a mystery since intervention in such situations to collect data would be unethical; at best, it may be possible to intervene directly in practice tests.

In spite of these limitations, I feel that I can make claims to originality, as I discuss in the next section.

Making a claim to knowledge

In general, I feel that I have answered my research questions to an extent that satisfies me and addresses the concerns I set out in CHAPTER 1. My concerns led me to specify two general aims for my research project. These were to review the reading demand of a typical test and teachers' arrangements for administering it to take account of children's reading difficulties.

To the best of my knowledge, no previous published research has:

i. studied the effect of the reading demand of the KS1 mathematics test on Y2 children in the test situation;
ii. used Y2 children to assess the readability of test items;
iii. sought teachers' views on the reading demand of a test and on how they supported children with reading;
iv. compared the reading skills of Y2, Y3 and Y4 children to determine the reading age likely to be required to read the test with fluency and accuracy;
v. undertaken an in-depth analysis of each written item to predict the readability of the booklet for its intended audience.

Through investigating these aspects, I believe that I can claim to have contributed to existing knowledge by building on what is already known about the reading demand of national mathematics tests in England. However, much remains to be learned about children's behaviour and attitudes in test situations. I now recommend some areas where further research would be likely to benefit children involved in high-stakes testing.

Suggestions for further research

When recommending how to build on their study of the reading requirements of KS2 statutory mathematics tests, Shorrocks-Taylor & Hargreaves (2000) stressed the need for research that would 'involve pupils themselves in judging the difficulty and comprehensibility of what they are being asked to read in mathematics, especially under the stress of test conditions' (p59). In light of my findings for KS1 children, I believe that such research is urgently needed and long overdue for KS2 and KS3 pupils. Having been to observations of KS2 and KS3 mathematics tests, it is rare for children to request reading support even though poor readers are participating and teachers can give help with reading words or phrases on request. Fortunately, I have made a small beginning to such research for Y2 children but more needs to be done whilst KS1 mathematics tests remains statutory. For example:

i. More research is needed to compare the effects on Y2 children and the scores they obtain when teachers administer the test with the ‘hands up when you want help with reading’ approach and the ‘look-listen’ approach.
ii. Although not a special focus of my study, children learning English as an additional language, including some regarded by teachers as fairly fluent in English, appeared to have a particular struggle with the text of items they read to me. More research is needed on the effect of the reading demand of the test on these children.
iii. Further research is needed into the advantages and disadvantages of computer-assisted testing as an alternative to or to complement the written test. For example, an advantage might be that children could hear text read over earphones; a disadvantage might be lack of keyboard skills.

Is my research sound?

Any researcher has to be able to defend their research proposal and to convince readers that their research is trustworthy. I shall try to do this briefly using four criteria identified by Miles & Huberman (1994) in their discussion of drawing and verifying conclusions and against which readers can evaluate epistemological integrity. These are confirmability, reliability, credibility and transferability.

Confirmability

I have tried to provide enough detail in my methodology, and then in even more detail on a chapter by chapter basis, for readers to follow ‘the actual sequence of how data were collected, processed, condensed/transformed, and displayed for specific conclusion drawing’ (Miles & Huberman, 1994:278). I could not retain all children’s test booklets that showed evidence of requests for help with reading but I retain a sub-set of these that can be related to data presented in CHAPTER 5. However, by retaining audio-recordings of children and teachers to match against transcripts in my study, I can confirm the accuracy of all interview data. I am particularly pleased that I have the children's voices in audio form as a means of showing that the transcripts do not overstate their difficulties and that my concerns about the reading demand were justified; if necessary, the results could be confirmed by others from this evidence.

Reliability

At all stages of my study, I tried to focus on my research questions so that data were collected, analysed and interpreted in the search for answers. In a sense, I had no option; time was at a premium. The detailed transcripts of children reading and the booklet analysis were particularly time-consuming and complicated but necessary, in my opinion, to investigate the reading demand from different perspectives. I was fortunate that each data source used in my multi-method approach led to the same answers to my research questions. Reliability was also helped because, after using multiple observers in my pilot study (see CHAPTER 3), in which my attempts to standardise interviews with children failed, I decided that I would collect all interview data personally for consistency. For example, children interviewed in a standard way in the months following the KS1 2001 mathematics test demonstrated similar characteristics, i.e. they had reading difficulties, did not always recognise their difficulties and did not always obtain appropriate support; Y2 readers demonstrated similar reading difficulties one year later. The consistency of approach that I applied led me to conclude that I could rely on the data to draw conclusions about the reading demand of the test for Y2 children.

Credibility

I remember reading the phrase ‘thick description’ somewhere. I have tried to evaluate and describe in detail a small part of the world of Y2 children and to make this micro-world come alive for the reader. If a reader has failed to make sense of my study so that they do not empathise with the children, then I have failed as a writer who wished to represent their voice. I have tried to tell the children’s story as credibly and plausibly as I can through rigorous and transparent management of data that they provided. Their data, i.e. the amount of help they requested with reading during the 2001 test and the data from their diagnostic interviews led me to the same conclusions and confirmed the concerns that I set out in CHAPTER 1. I know that credibility can be enhanced when rival explanations for phenomena are considered and rejected. I cannot suggest alternative explanations to those given in the relevant chapters as to why children requested so much help with reading or why they found text difficult to read. Similarly, credibility can be enhanced by asking informants if conclusions are accurate. Clearly this was not possible with Y2 children but readers can, if they wish, compare my conclusions with the considerable evidence I have retained to judge my credibility.

Transferability

One aspect of transferability is how far the conclusions of a study can be generalised to a wider audience. I have already argued that I expect the ‘empirical generalisations’ (Bassey, 1995:97) discussed earlier in this chapter to be applicable to most children in England who take the KS1 mathematics test. For each child who provided me with data, there would be other children across the country showing similar patterns of behaviour during tests or having similar reading difficulties. For each teacher whom I interviewed, there would be other teachers up and down the land who administered the test or supported reading in broadly similar ways. Bassey (1995) describes this as a ‘normative conclusion’ (p97) since it is based on value judgments. I am confident that data collected and analysed in the same manner from Y2 children and their teachers at other schools for the same purpose would lead to the same conclusions. I also hope that I have described the data management in sufficient detail so that it could be replicated by other researchers and would expect them to reach similar conclusions. This would be the ultimate test of the degree of transferability.

General conclusions

I entered the research community as a person who was fortunate enough to have worked with young children and shared the world of the classroom with them. Even when I became a consultant in test development, I carried with me knowledge and experience gained from life as a teacher. For this study, my intention was to use that knowledge and experience to tell a story from the children's perspective, namely 'to work *for* children rather than *on* them and to describe their social worlds' (France, Bendelow & Williams, 2000:51). If the findings from my study could influence social change to benefit children, that would be a bonus; an even greater bonus would be to influence national policy. I never set out to make a political statement but was reminded of a quote from Clough & Nutbrown (2002) that I cited in my methodology chapter but is repeated here for its relevance:

> Research that changes nothing – not even the researcher - is not research at all. And since all social research takes place in policy contexts of one form or another research itself must therefore be seen as inevitably political.
> (Clough & Nutbrown, 2002:12)

Without a doubt, the research has changed me. I now have evidence to support what started out as professional hunches and I will try to use that evidence to be a voice for KS1 children. Without the intention of it ever being so, I can see that my research is 'inevitably political' (Clough & Nutbrown, 2002:12) since government policy led to a statutory national curriculum and assessment arrangements. Griffiths (1998) believed that '[t]he purpose of educational research is to improve the education of children' (p95). This study did not aim to improve the quality of education but it has drawn attention to how children's needs can unintentionally be overlooked in test situations against which the quality of education is frequently judged. Griffiths (1998) also believed that '[a] main reason for doing the (educational) research is to get improvement in social justice' (p95). The powerful impact of children's voices that has given me confidence in my study will, I hope, lead to improvements in test administration that do not disadvantage poorer KS1 readers. Such improvements would make a happy ending to my story although I remain optimistic that KS1 statutory testing will be abolished or, at the very least, demoted to optional status. Now that would be a happy ending!

REFERENCES

Allan, S., McGhee, M. & Krieken, R. van (2005) *Using readability formulae for examination questions*. SQA research and information services. [online]. www.qca.org.uk/downloads/allan_et_al_using_readability_formulae_for_examination_questions_pdf_05_1607.pdf (accessed March 2006).

Bassey, M. (1995) *Creating education through research.* Newark: Kirklington Moor Press.

BERA (2000) *Ethical guidelines for educational research.* [Online] Available at: www.bera.ac.uk/guidelines.html (accessed January 2001).

Bormuth, J. R. (1966) Readability: a new approach. *Reading Research Quarterly,* **1,** 79-182.

Burton, L. (2002) Methodology and methods in mathematics education research: where is "The Why"? In Goodchild, S. & English, L. (Eds.) *Researching mathematics classrooms : A critical examination of methodology.* Westport: Praeger Publishers.

Charmaz, K. (2006) *Constructing grounded theory: a practical guide through qualitative analysis.* London: Sage Publications.

Clausen-May, T. (2001) *An approach to test development.* Slough: National Foundation for Educational Research.

Clements, M. A. K. (1980) Analyzing children's mathematical errors on written mathematical tasks. *Educational studies in mathematics,* **11,** 1-21.

Clough, P. & Nutbrown, C. (2002) *A student's guide to methodology.* London: Sage Publications.

Cohen, L., Manion, L. & Morrison, K. (2000) *Research Methods in Education.* 5th ed. London: RoutledgeFalmer.

Committee of Inquiry into reading and the use of English (1975) *A language for life: report on the Committee of Inquiry appointed by the Secretary of State for Education and Science under the chairmanship of Sir Alan Bullock.* London: HMSO.

Costley, D. (2000) Collecting the views of young people with moderate learning difficulties. In Lewis, A. & Lindsay, G. (Eds.) *Researching children's perspectives.* Buckingham: Open University Press.

Crisp, V. & Sweiry, E. (2006) Can a picture ruin a thousand words? The effects of visual resources in exam questions. *Educational Research,* **48,** (2), 139-154.

Crozier, J. & Tracey (2000) Falling out of school. In Lewis, A. & Lindsay, G. (Eds.) *Researching children's perspectives.* Buckingham: Open University Press.

Dale, E. & Chall, J. S. (1948) A formula for predicting readability. *Education Research Bulletin,* **27,** 11-20; 28.

Denzin, N. K. (1970) *The research act in sociology: a theoretical introduction to sociological methods.* London: Butterworth.

De Vaus, D. (2001) *Research design in social research.* London: Sage Publications.

DfEE (1998) *The National Literacy Strategy: Framework for teaching.* London: DfEE.

DfEE (1999a) *The National Numeracy Strategy: Framework for teaching mathematics from Reception to Year 6.* DfEE: London.

DfEE (1999b) *The National Numeracy Strategy: Mathematical vocabulary.* Watford: DfEE.

DfEE & QCA (1999a) *English: The National Curriculum for England.* London: DfEE & QCA.

DfEE & QCA (1999b) *Mathematics: The National Curriculum for England.* London: DfEE & QCA.

DfES (2001) *Key stage 1 national summary results 2001*. [online]. www.standards.dfes.gov.uk/performance1/pdfKS1_2001,pdf?version=1 (accessed April 2003).

DfES (2005) *Primary National Strategy: Understanding reading comprehension: 1.* [online]. www.standards.dfes.gov.uk/primary/publications/literacy/1162245/pns_nls131005readcomp1.pdf (accessed September 2006).

Drum, P., Calfee, R. & Cook, L. (1981) The effects of surface structure variables on performance in reading comprehension tests. *Reading Research Quarterly,* **16,** 486-514.

Fisher-Hoch, H. & Hughes, S. (1996) *What makes mathematics exam questions difficult?* [online]. www.cambridgeassessment.org.uk/research/confproceedingsetc/BERA1996HFSH/file/ (accessed March 2006).

Fraenkel, J. R. & Wallen, N. E. (2006) *How to design and evaluate research in education.* 6th ed. New York: McGraw Hill.

France, A., Bendelow, G. & Williams, S. (2000) A 'risky' business: researching the health beliefs of children and young people. In Lewis, A. & Lindsay, G. (Eds.) *Researching children's perspectives.* Buckingham: Open University Press.

Fry, E. (1990) A readability formula for short passages. *Journal of reading,* **33,** 594-597.

Gibson, L. (1989) *Literacy learning in the early years: Through children's eyes.* London: Cassell Educational Limited.

Gillham, B. (2000) *The research interview.* London: Continuum.

Gilliland, J. (1972) *Readability.* London: Hodder and Stoughton Educational.

Gitlin, A., Siegel, M. & Boru, K. (1993) The politics of method: from leftist ethnography to educative research. In Hammersley, M. (Ed.) *Educational research: current issues.* London: Paul Chapman Publishing. 191-210.

Griffiths, M. (1998) *Educational research for social justice: getting off the fence.* Buckingham: Open University Press.

Griffiths, M. & Davies, C. (1995) *In fairness to children*. Primary Curriculum Series. London: David Fulton Publishers.

Harrison, C. (1979) Assessing the readability of school texts. In Lunzer, E. & Gardner, K. (Eds.) *The effective use of reading.* London: Heinemann. 72-103.

Harrison, C. (1980) *Readability in the classroom.* Cambridge: Cambridge University Press.

Hart, C. (1998) *Doing a literature review.* London: Sage Publications.

Heritage, J. (1984) *Garfinkel and Ethnomethodology.* Cambridge: Polity.

Homan, S., Hewitt, M. & Linder, J. (1994) The development and validation of a formula for measuring single-sentence test item reliability. *Journal of Educational Measurement,* **31,** (4), 349-358.

Homan, S. P., Hall, B. W. & Topping, M. (1986) Does teacher reading of oral test items affect the performance of students of varying reading ability?. *Journal of Educational Research,* **79,** 363-365.

Johnstone, A. H. (1988) *Meaning beyond readability.* SEC, Newcombe House, London W11 3JB.

Kane, R. B. (1967) The readability of Mathematical English. *Journal of Research in Science Teaching,* **5,** 296-8.

Kane, R. B. (1970) Readability of mathematics textbooks revisited. *The Mathematics Teacher,* **63,** 579-581.

Kane, R. B., Byrne, M. A. & Hater, M. A. (1974) *Helping children read mathematics.* New York: American Book Company.

Klare, G. R. (1963) *The measurement of readability.* Iowa: Iowa State University Press.

Klare, G. R. (1974) Assessing readability. *Reading Research Quarterly,* **1,** 62-102.

Lewis, A. & Lindsay, G. (Eds.) (2000) *Researching children's perspectives.* Buckingham: Open University Press.

Linden, N. (2002) Special educational needs in mathematics - a problem developed in school. In Goodchild, S. & English, L. (Eds.) *Researching mathematics classrooms - a critical examination of methodology.* Westport: Praeger Publishers. 67-89.

Lloyd-Smith, M. & Tarr, J. (2000) Researching children's perspectives: a sociological dimension. In Lewis, A. & Lindsay, G. (Eds.) *Researching children's perpectives.* Buckingham: Open University Press. 59-70.

Lupetti, S., Sainsbury, M. & Schagen, I. (1995) *The language and reading demands of mathematics tests*. Unpublished report presented to SCAA: NFER.

Marshall, C. & Rossman, G. B. (1995) *Designing qualitative research.* 2nd ed. California: Sage Publications.

McCallum, B., Hargreaves, E. & Gipps, C. (2000) Learning: the pupil's voice. *Cambridge Journal of Education,* **30,** (2), 275-289.

Miles, M. & Huberman, A. M. (1994) *Qualitative data analysis.* Beverley Hills: Sage Publications.

Mobley, M. (1986) *Readability.* London: SCDC Publications.

Newman, M. A. (1977) An analysis of sixth-grade pupils' errors on written mathematical tasks. In Clements M. A. & Foyster J. (Eds.) *Research in mathematics education in Australia,* **1,** 239-258.

Noonan, J. (1990) Readability problems presented by mathematics text. *Early Child Development and Care,* **54,** 57-81.

Paul, D. J., Nibbelink, W. H. & Hoover, H. D. (1986) The effects of adjusting readability on the difficulty of mathematics story problems. *Research in mathematics education,* **17,** (3) 163-171.

Perera, K. (1980) The assessment of linguistic difficulty in reading material. *Educational Review,* **32,** (2), 149-161.

Powney, J. & Watts, M. (1987) *Interviewing in educational research.* London: Routledge & Kegan Paul.

Pressley, M. (2000) Comprehension instruction in elementary school: a quarter-century of research progress. In Taylor, B. M., Graves, M. F. & van den Broek, P. (Eds.) *Reading for meaning: fostering comprehension in the middle grades.* New York & Delaware: Teachers College Press & The International Reading Association. 32-51.

QCA (1998) *Standards at key stage 1: English and mathematics – Report on the 1998 national curriculum assessment for 7-year olds.* London: QCA.

QCA (1999) *Assessment and reporting arrangements 2000: Years 1, 2 and reception.* London: QCA.

QCA (2001a) *Standards at key stage 1: English and mathematics: A report for headteachers, class teachers and assessment coordinators on the 2000 national curriculum assessments for 7-year olds.* London: QCA.

QCA (2001b) *Key stage 1 2001 English tasks: Teacher's guide.* London: QCA.

QCA (2001c) *Key stage 1 2001 mathematics booklet.* London: QCA.

QCA (2001d) *Key stage 1 2001 mathematics test: Teacher's guide.* London: QCA.

QCA (2001e) *Key stage 1 technical appendices*. [online]. www.qca.org.uk/3473.html (accessed January 2006).

QCA (2002) *Standards at key stage 1: English and mathematics: A report for headteachers, class teachers and assessment coordinators on the 2001 national curriculum assessments for 7-year olds.* London: QCA.

QCA (2004) *Providing effective learning opportunities for all pupils*. [online]. www.nc.uk.net/nc_resources/html/inclusion.shtml (accessed June 2006).

Raiker, A. (2002) Spoken language and mathematics. *Cambridge Journal of Education,* **32,** (1) 45-60.

Reid, D. (1989) *Word for word: the top 2000 words used by 7 and 8-year olds.* Wisbech: Learning Development Aids.

Robinson, V. M. (1993) *Problem-based methodology: research for the improvement of practice.* Oxford: Pergamon Press.

Robson, C. (2002) *Real World Research.* 2nd ed. Oxford: Blackwell Publishing.

Rye, J. (1985) Computing readability. *Reading,* **19,** (2) 110-116.

Rye, J. & Sjölander, P. (1982) Readability: the importance of the child. *Reading*, **16**, (2) 99-108.

Shorrocks-Taylor, D. & Hargreaves, M. (1999) Making it clear: a review of language issues in testing with special reference to the National Curriculum mathematics tests at key stage 2. *Educational Research,* **41,** (2) 123-136.

Shorrocks-Taylor, D. & Hargreaves, M. (2000) Measuring the language demands of mathematics tests: the case of the statutory tests for 11-year olds in England and Wales. *Assessment in Education,* **7,** (1) 39-60.

Shuard, H. & Rothery, A. (Eds.) (1984) *Children reading mathematics.* London: John Murray.

Silverman, D. (1993) *Interpreting qualitative data.* London: Sage Publications.

Silverman, D. (2005) *Doing qualitative research.* 2nd ed. London: Sage Publications Ltd.

Sticht, T. G, Beck, L. J., Hauke, R. M., Kleiman, G. M. & James, J. H. (1974). *Auding and reading: A developmental model.* Alexandria, VA: Human Resources Research Organisation.

Stuart, M., Dixon, M., Masterson, J. & Gray, B. (2003) Children's early reading vocabulary: description and word frequency lists. *British Journal of Educational Psychology*, **73**, 585-598.

Stuart, M., Masterson, J. & Dixon, M. (2000) Spongelike acquisition of sight vocabulary in beginning readers?. *Journal of Research in Reading,* **23,** 12-27.

Taylor, B. M., Graves, M. F. & van den Broek, P. (Eds.) (2000) *Reading for meaning: fostering comprehension in the middle grades.* New York: Teachers College Press.

Thomas, R. M. (1998) *Conducting educational research: a comparative view.* Westport: Greenwood Publishing Group.

Verma, G. K. & Mallick, K. (1999) *Researching education: perspectives and techniques.* London: Falmer Press.

Wellingon, J. (2000) *Educational research: contemporary issues and practical approaches.* London: Continuum.

Wellington. J., Bathmaker, A-M., Hunt, C., McCulloch, G. & Sikes P. (2005) *Succeeding with your doctorate.* London: Sage Publications.

Wiest, L. (2003) Comprehension of mathematical text. *Philosophy of mathematics education,* [online] **17**. www.ex.ac.uk/~PErnest17/lwiest.htm (accessed October 2003).

Wiliam, D. (2003) Constructing difference: assessment in mathematics education. In Burton, L. (Ed.) *Which way social justice in mathematics education?* International perspectives on mathematics education. Westport CONN/London: Praeger.

Wilkinson, A. C. (1980) Children's understanding in listening and reading. *Journal of Educational Psychology,* **72,** (4) 561-574.

Wolcott, H. F. (1995) *The art of fieldwork.* London: Sage Publications.

Wood, E. (2003) The power of pupil perspectives in evidence-based practice: the case of gender and underachievement. *Research papers in Education,* **18,** (4) 365-384.

Typical development of a KS1 mathematics test

The KS1 mathematics tests take at least 14 months to develop during which:

i. A team of consultants, experienced at working with Y2 children, draft items using the KS1 programme of study for mathematics, as detailed in *Mathematics: The National Curriculum for England* (DfEE & QCA, 1999c:16) as the key reference. They also make informal reference to the *Framework for Teaching Mathematics* (DfEE, 1999a) and its related publication, *The National Numeracy Strategy: Mathematical vocabulary* (DfEE, 1999b) since '[t]he Framework provides a detailed basis for implementing the statutory requirements of the programme of study for key stage 1 in mathematics' (DfEE & QCA, 1999c:16).

ii. Writers arrange for Y2 children to trial the draft items in local schools.

iii. Advisory groups and teachers who are to take part in the first pre-test review the draft questions, taking into account the mathematical demand, validity and accessibility for Y2 children working at level 2 or above.

iv. Approved items are compiled into test booklets for the first pre-test and trialled in a nationally representative sample of schools. The children's responses to the items, teachers' views from questionnaire responses, advisory groups and markers are taken into account together with statistical data. Items are selected for the next stage.

v. These items are then sent out for the second pre-test and standardisation trial. As for the first pre-test, children's responses, teachers' questionnaire responses and advisory groups' views are taken into account. The test is then finalised.

vi. After statistical analysis, thresholds are set for the award of levels, leading to cut scores for each level or grade.

Appendix 3.1

Copy of correspondence for tracking sheet and items for interviews

Contents

- Guidance for completing the *Tracking sheet* (1 page);
- copy of *Tracking sheet* (1page);
- copy of items from previous KS1 mathematics tests for interviews with children (4 pages).

Guidance for completing the *Tracking sheet*

The purpose of the *Tracking sheet* is to record the speed at which each child in the focus group works through the pre-test booklet.

'Child A (–)' means the child who is a below-average reader (level 2C)
'Child B' means the child who is an average reader (level 2B)
'Child C (+)' means the child who is an above-average reader (level 2A or level 3)

Before the pre-test

Enter the information missing from the school and name boxes.

During the pre-test

Enter the time the children start item 7 (first written item);

Write the item number each child is working on at five-minute intervals.

This example shows that a child is working on item 11 after five minutes and item 15 after 10 minutes.

Time (minutes)

Finally, enter the time each child completes the booklet.

During or after the pre-test

If appropriate, record any additional information relating to children A to C, which you find interesting or relevant on the reverse of the *Tracking sheet*.

Tracking sheet

School

Time Q7 started:

Child A (-) Name | Time pre-test completed

5 10 15 20 25 30 35 40 45 50 55 60

Time (minutes)

Child B Name | Time pre-test completed

5 10 15 20 25 30 35 40 45 50 55 60

Time (minutes)

Child C (+) Name | Time pre-test completed

5 10 15 20 25 30 35 40 45 50 55 60

Time (minutes)

Look at these cards.

2 6 3

4 5

You can use two cards to make a number **less** than **30**.

24

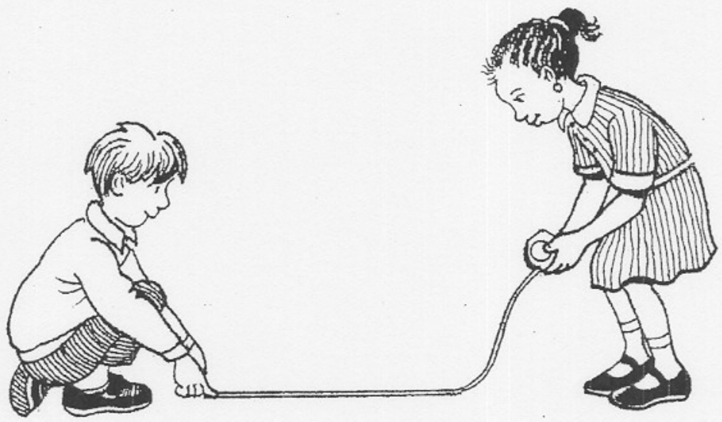
1 Use two of the cards to make a number **between 30** and **40**.

Choose a word from the box to finish each sentence.

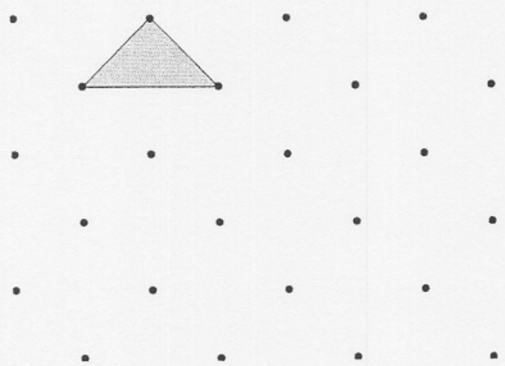
kilograms
litres
metres
hours

2 I can measure the length of the classroom

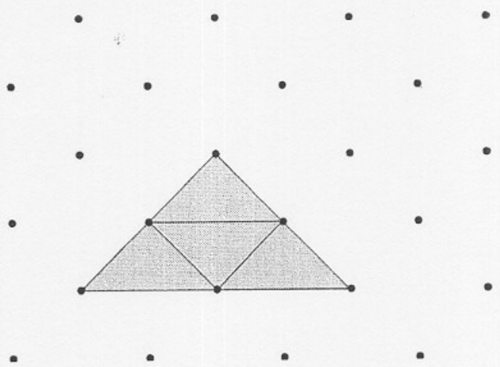
in ______________________ .



Here is one triangle.

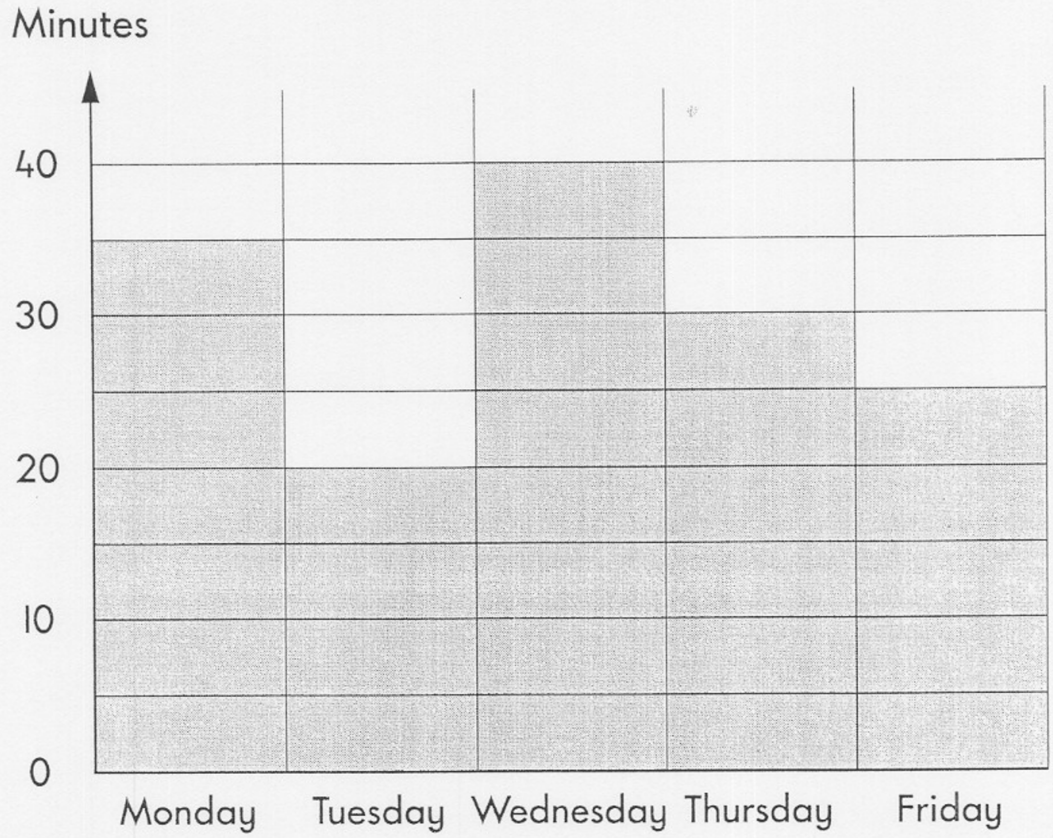


Here are 4 small triangles.
They make a bigger triangle.



3 Use 9 of the small triangles to make a bigger triangle.

This graph shows the time it took Tim to travel to school in one week.



4 How long did it take Tim on Monday? minutes

5 How much longer did Tim take on Monday than on Friday? minutes

KS1 mathematics pre-test results (School H)

No	Child	Oral marks	Written question number																													Total	Ma sc	Ma L	Re L	EAL
		001- 006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035					
1	Liam	5															1															1	19	2C	2B	
2	David	5	1															1									1					3	23	2A	2C	
3	Kieran	5						1		1			1	1		1	1															6	21	2C	2C	
4	Danny	4		1			1	1		1		1	1	1			1	1		1												10	14	2B	2C	
5	Katie	5			1												1	1														3	12	2C	2C	
6	Jack	5			1		1	1				1	1	1		1	1	1														9	15	2C	2C	
7	Joshua	6				1		1	1				1	1		1	1															7	17	2C	2A	
8	Jessica	3				1							1		1	1		1														5	7	2C	2A	
9	Laura	4			1	1		1	1				1			1					1			1		1				1		10	13	2C	2A	
10	Sinead	5			1	1		1					1				1							1								6	14	2C	2A	
11	Chersea	5	1		1			1									1	1		1					1							5	20	2B	3	
12	Emma	5															1															1	22	2A	3	
13	Sean	2																															27	3	2A	
14	James	5																															23	2B	3	
15	Jim	6																															34	3	3	
16	Roseanne	6																															27	3	3	
17	Matthew	6																															28	3	2A	
18	Jamie	4			1			1	1			1	1	1		1	1			1				1	1							11	18	2A	2A	
19	Amy	4																															18	2B	2B	
20	Chelsea	5																															32	3	2C	
21	Joel	5																															26	2A	2C	
22	Aqsa	6																															33	3	2C	
23	Hannah	5																															18	2B	2C	
24	Josie	5																															29	3	2C	
25	Ryan	5																															16	2B	2A	
26	Paul	6																															21	2C	2B	
27	Martin	6																								1		1				2	22	2C	2B	
28	Danielle	3																															14	2C	2B	
29	Ben	3	1	1	1	1	1	1			1	1	1	1		1		1	1	1	1		1									16	10	2C	1	
30	Anna	3						1																								1	13	2C	2C	
31	Kevin	5																															20	2A	2B	
32	Laura	4																															22	2A	2A	
33	Reannan	4																															28	2A	2A	
34	Lawrence	5																															26	2A	2A	
35	Devon	6																															30	3	2A	
36	James	5																															23	3	2A	
37	Jordan	5	1		1	1		1				1		1	1	1	1	1		1		1		1								13	24	2A	2C	
38	Fiona	6																															32	3	2A	
39	Reece	6																															35	3	2A	
40	Jasmeena	6																															35	3	2A	
41	Jessica	5																															30	2A	2A	
42	Thomas	6																															32	2A	2A	
			4	2	8	6	3	11	3	2	1	5	9	7	2	8	11	8	1	5	2	1	1	4	2	2	1	1		1		109				

Ma sc: score in mathematics test

Ma L: level achieved in mathematics test

Re L: level achieved in reading task/test

EAL: has English as an additional language

Each 1 indicates a request for help with reading

Appendix 3.3

Tracking sheets for children in focus groups at Schools A, B and H

Tracking sheet

School SCHOOL A

Time Q7 started: 9.50

10.06 goes back to do those omitted

Child A (-) (L2c) **Name** LUKE [redacted] **Time pre-test completed** 10.27

Missed out some early question

12 19 27 13 32 34

5 10 15 20 25 30 35 40 45 50 55 60

Time (minutes)

10.06

Child B (L2b) **Name** ROSIE [redacted] **Time pre-test completed** 10.25

Went back to do questions omitted

13 19 23 30 34 15 21

5 10 15 20 25 30 35 40 45 50 55 60

Time (minutes)

Child C (+) (L3) **Name** NISHA [redacted] **Time pre-test completed** 10.40

– Still checking – suggested she stop now

12 16 21 22 23 27 29 32 34

5 10 15 20 25 30 35 40 45 50 55 60

Time (minutes)

LIB3\52338\1

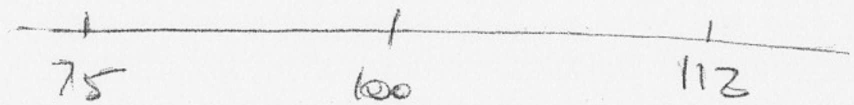
Intense concentration

(Test: 1 hour)
... took one look at some questions & immediately decided not to attempt. — 11, 13, 17, 22 31–34

He "finished" his first run through questions after 16 mins and went back to those he had skimmed over.

Rosie – appeared very confident and worked through systematically omitting only a few She went back to complete questions

Nisha worked very methodically giving a lot of thought to all over the questions asking occasionally for help to read odd words. Used vertical method for calculations although Q28 She checked with the blank number line and changed her answer She showed she was not proficient in the standard algorithm for + & −



Tracking sheet School SCHOOL B (Test: 1 hour + 20 mins oral)

Time Q7 started: 10. 06

Re-start 11. 33

Child A (-) **Name** Katy **Time pre-test completed**

level 2c reader

	Q11	Q15	Q18	Q20	Q22	Q23	Q25	Q27	Q30	Q31	Q35	Finished Stopped
Time (minutes)	5	10	15	20	25	30	35	40	45	50	55	60
	10.11	10.16		Break	11.38	11.43	11.48	11.53	11.58	12.03	12.08	12.13

(51 mins to complete test 7 min break accounted)

Child B **Name** Miles **Time pre-test completed**

level 2b reader

	Q13	Q18	Q20	Q24	Q27	Q30	Q33	Q35	Checking	Stops		
Time (minutes)	5	10	15	20	25	30	35	40	45	50	55	60
		Break 11.33	11.36	11.41	11.46	11.51	11.56	12.01	12.06	12.11		

Child C (+) **Name** Izzy **Time pre-test completed**

level 2a reader or level 3

	Q15	Q20	Q25	Q27	Q30	Still Q30.	Still Q30	Stops at 12.02				
Time (minutes)	5	10	15	20	25	30	35	40	45	50	55	60
		Break	11.33	11.43	11.48	11.53	11.58	12.03	12.08	12.13		

LIB3\52338\1

(H)

Tracking sheet

School SCHOOL H

Time Q7 started: 11.20 (11.25)

Actual times in blue

Estimated times in red – movement between ~~[illegible]~~ two classes

Child A (-) (L2c) **Name** Jack. (CLASS 2)

Time pre-test completed Not completed

Estimated start time 11.25

Stopped for dinner after 50 minutes

7 8 9 16 17 19 20 22 23 23

5 10 15 20 25 30 35 40 45 50 55 60

Time (minutes)

(Jack in 2nd class)

Child B (L2A) **Name** Matthew (CLASS 1)

Time pre-test completed 11.50

Start Time 11.20.

7 11 15 24 34 35 Finished

5 10 15 20 25 30 35 40 45 50 55 60

Time (minutes)

Child C (+) (L3) **Name** Roseanne. (CLASS 1)

Time pre-test completed 11 55

Start Time 11.20.

7 12 14 22 31 33 35 Finished

5 10 15 20 25 30 35 40 45 50 55 60

Time (minutes)

L1B3\52338\1

Jack – had to wait for help with reading on occasions
– Several hands up at same time led to delays in reading support.

Appendix 3.4

Copy of Jack’s interview items with underlined text and observer’s comments

[redacted] Jack. [redacted]

(Re : L2c)

Look at these cards.

2	6	3
4	5	

You can use two cards to make a number **less** than **30**.

2	4

1 Use two of the cards to make a number **between 30** and **40**.

3	2

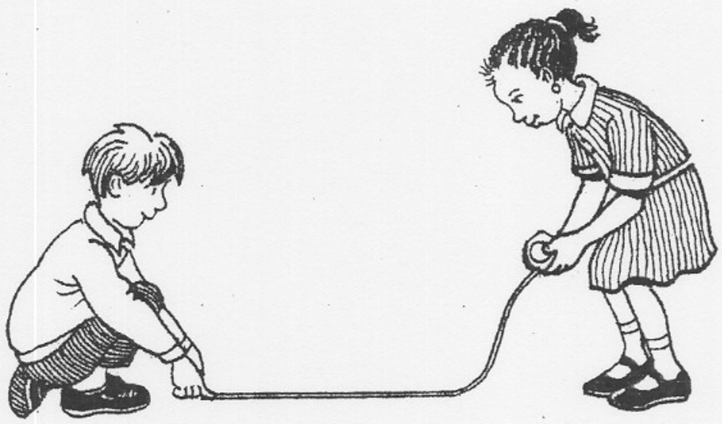
Some problems with reading the words. (underlined in pencil)

Choose a word from the box to finish each sentence.

kilograms
litres
metres
hours

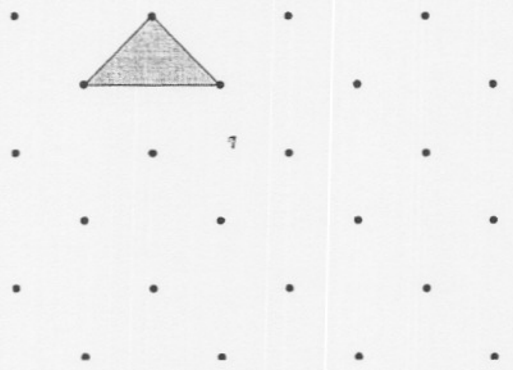
2 I can measure the length of the classroom

in ___metres___ .



'Here' read as 'how'

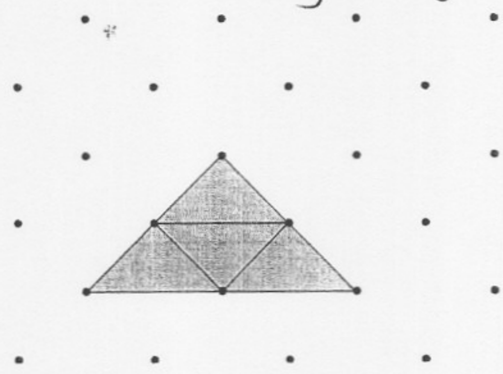
Here is one triangle.



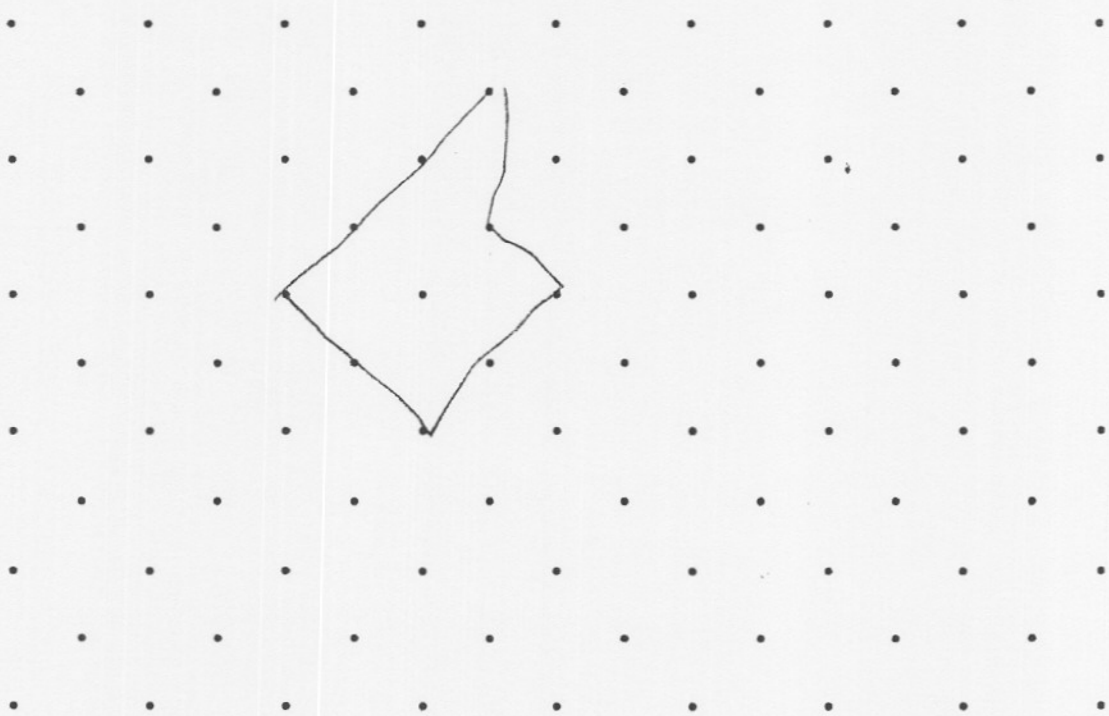
'Small' read as 'solid'

Here are 4 small triangles.
They make a bigger triangle.

Read initially as 'big'



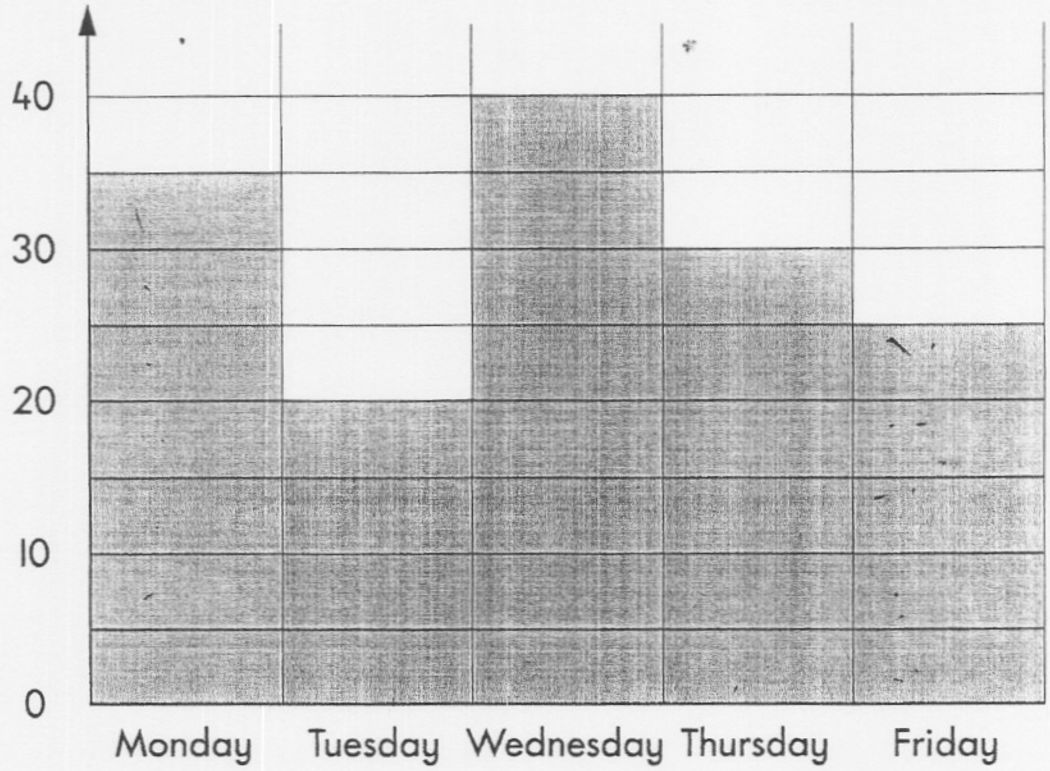
3 Use 9 of the small triangles to make a bigger triangle.



pronounced as 'grap'   pronounced as 'tell'

This graph shows the time it took Tim to travel to school in one week.

Minutes ← 'minutes' read as 'metres'



4 How long did it take Tim on Monday?

30 minutes

read as 'long'

5 How much longer did Tim take on Monday than on Friday?

20 minutes

I have adopted the following conventions.

1. The numbers are from the tape-counter.
2. AW identifies myself as interviewer.
3. I have ignored short pauses in the conversation since these do not appear to be significant in the context of the transcript.
4. Asides are in brackets and indicate occurrences such as a point of clarification for the reader, indistinct speech, etc.
5. Relevant comments are highlighted as follows:

   yellow: comments on reading demand;
   green: comments on administration
   blue: comments on reading support

Transcript 1

Circumstances of interview: Four Year 3 teachers who had administered the first pre-test of the KS1 mathematics test items attended a meeting organised by QCA in which issues arising from the pre-test could be discussed and each item reviewed in light of the classroom trial. Before the meeting started, the teachers were discussing, unprompted, the ways they organised children to support reading. With their permission, an audio-tape recorder was set up and their comments recorded. Comments have been transcribed in full except where indicated. Two QCA observers of the pre-test were also present, AW (myself) and PW. The teachers are referred to as Teachers A-D.

Date of interview: 30 November 2000

001 AW When I came in this morning, some of you had automatically started talking about the reading demand of the test booklets during the pre-test, and maybe the live test as well. I wonder if you'd like to talk about the problems you have had or the problems you think there will be with the reading of the test booklets. Then we can go on, possibly, to look at some solutions. So, any problems? What problems are there at the moment?

003 Teacher A I found that some of my pupils whom I expected to be able to read and answer the questions were not reading the questions properly, and so, I think they were guessing some of the answers, and, where they could have asked for help with reading, they weren't. They were having a guess, and I would have expected them to be able to read the questions properly.

005 Teacher B — When you said (Indicating teacher C) that you read all the questions out to all the children, it made me think, well, maybe that would have been better if I'd done that.

007 AW — So did you actually read out all the questions?

008 Teacher C — We read all the questions out. There were three other invigilating staff in the class - and all the children took the test - just to make sure that they were starting to read from the correct place. We talked about numbers not being aligned where the questions actually started, and that was one of the things we had children doing – going straight to the number which was next to the question but some of the questions had started higher up on the page so children were missing some of the elements and we were having to go back and make sure they were always starting reading – that they were in the right place.

011 AW — Why did you choose to use that method of administration? Was it because of past experience?

012 Teacher C — Yes. We had trialled the year 3 SATs (non-statutory year 3 mathematics tests) previously and again we found a similar situation where given children's choice of having to ask for questions to be read, not only was the more able quite confident and felt they were understanding questions which perhaps they needed a bit more explanation with but the lower achievers as well, needed to have, you know, direct reading and making sure that they were following the questions as well.

014 AW — Yes. What about the live test in year 2? Do you know what method of administration was used?

015 Teacher C — They take out groups of children and, again, we test with them but the groups are such they're in ability groups that they take them out with, so although higher achievers will go together so that one particular child isn't slowing down the pace, if you like, of the working, and we're trying to move towards more independence. We have been for quite a few years now in terms of the reading. We're certainly hoping that the higher achievers are trying to get more independent in terms of their reading so we are trying to move towards that but again we are still finding the difficulty of crucial parts of the questions that children sort of skim over and that's why we are going back and having to actually read it. But it's trying to get that balance between them becoming independent to read the maths questions and understand the vocabulary, but, at the same time, ensuring that they have got access to the question.

30	Teacher D	I notice that the Year 3s doing the test this time (end of November 2000), were far more competent at getting on by themselves than how many months ago when they did the Year 2 SATs. Maybe that's just because they've done a SAT once and maybe they understand the format. I do know that they need an awful lot less help with reading now than in year 2 just a few months ago.
37	AW	They're six months older, aren't they?
45	Teacher C	Mmm. It makes a difference. Then the other observation we made, with the lower achievers, when they were reading the test, em, I found that because circumstances prevailed which meant that we had to administer it to all of them with a break in-between rather than on alternative days, which was what we had originally planned, we found that some of them needed to be in a smaller situation even when we were reading, needed to be in a smaller group to really keep them focused on what they were doing even though we were there invigilating and we had sort of eight or nine children to look over, I still felt some of the less able were sort of, if you like, needing that little bit more time and a little bit more space as well.
059	Teacher D	All the children feel far more comfortable in that first part of the test, the oral part, because you are reading, everything is clear to them.
063	Teacher B	I feel the anxiety when they are on their own, I feel they do like to be spoon-fed a little bit.
064	Teacher C	It's like how you read a question, your voice, your emphasis on the words, I think that helps so much. (Phrase indistinct) their reading, their skill at that stage you, know, they might read the words but whether they understand what they are reading (phrase indistinct) when they hear your emphasis of words.
073	AW	I know that you did this pre-test with year 3 children but did you notice any children sounding out letters phonetically at all to try to get to the word?
074	Teacher D	Yes, they were. Because when I went round the ones that had put their hands up to ask for help, and I put my finger to get them to read it and they were sounding it out, and then I took over because I thought, 'This is going to take ages.' But yes, they were using phonics.
078	Teacher B	Yes, definitely. I only have one on the SEN register and he has made a lot of progress since he has moved into the juniors but he is still using the phonic way of sounding through but his actual maths computation is quite sound. His problem is language and it is a barrier to his actually completing a piece of maths work unsupported, definitely.

086	Teacher A	As you say, the emphasis when you read, even I find this with my higher achievers, who say, 'What does this mean? I don't understand.' until you read it to them, and, again, the way that you emphasise it, immediately they say, 'Oh, right.', and you think, 'But you've just read that.' 'But right, I know what you mean now.'
092	Teacher B	If I had read that chocolate question out to them (tape indistinct) whereas that would have immediately linked the chocolate in, whereas I think some of my (phrase indistinct), if I had read it out, that would have been automatic.
097	Teacher D	Yes, drawing their attention to it.
098	Teacher C	Yes.
099	Teacher A	I saw a couple of children quit the chocolate question, and I said, 'Have you read the question?' 'No'. (Interviewees laugh together.)
104	PW	But the children I observed were quite often more able children who were stuck and said, 'I don't know how to do this.' As soon as you said, 'Tell me what it …read for me what it is asking you to do.', and just in reading it back aloud to someone else made it clear – made it clear to them. (Sound of agreement from others.) And if they didn't read it correctly, you'd come in and read it correctly for them. And it does make a difference.
012	AW	So it's the hearing of the question as well as the reading.
113	Teacher A	They see all the words (the quantity of words). I notice it in day to day work. All of the words used in a test book seem to be overwhelming and the actual computation they have to do is really usually only a one- or a possible two-operation. (Next sentence unclear). I mean, that is part of it, the problem solving aspect of it, isn't it, where they have to decide what to do.
117	AW	Where they have to recognise the operation.
120	Teacher A	They have to unpick the words to work out, but, yes, it is a problem.
121	PW	You want it to be a bit more than unpicking the words; you want it to be unpicking the context.
122	Teacher A	It's the understanding of the context, of course. The bright ones here, as well, I mean, I've got my top two in mind, they are fairly quickly off their seats to seek help if it doesn't immediately hit them, and I say, 'You've just got to read it again and give yourself a little bit of time to think about it', and you can see them, 'Oh, yeah.' It's like they want an immediate – well, they want to understand it immediately and, of course, that does not happen. Yes, the words do act as a barrier.

133	AW	There seems to be some agreement that the language can be a barrier to the maths if you are reading it on your own. Bearing in mind that there is going to be a big review of the test in 2002, what style of test do you think could be trialled that could overcome some of the difficulties that you have been talking about?
137	Teacher B	That's a tough one.
140	AW	How do we get over the difficulties caused by reading?
141	Teacher B	You know, being able to facilitate it and administer it in small groups and that, you know, you have the staff to do it, which, I think, is very hard-pushed in most primary schools. I mean, both schools I have worked in, it's, with the year 3 test anyway because they are voluntary SATs, em, there's only me and the thirty children and because, in my past school, it was like everyone did it the week of the SATs. It's not as if you can do it over a long period of time (phrase indistinct) but I think to have both, you know, being able to hear the teacher say it as well as having the visual input as well on the paper.
153	AW	So you read the text but the child can follow it with you.
154	Teacher B	Yeah. (Sounds of agreement from others also.)
155	Teacher C	You know, it's like some people are more visually-orientated than (indistinct phrase) and, you know, for me, if I'm listening, it helps me to be writing or to be reading something. That's a personal thing but other people could be the other way around – if they were reading something, it would help them more to hear it. So it's like giving as much help as possible. It's like with younger children you say and touch and actually hold it, don't you, as well?
163	PW	When you say about year 3, of course, you don't have the opportunity to administer the test over several sessions. What about in your schools in year 2? Is it all the children being tested on their maths test at the same time?
165	Teacher C	No, that's for the year 6 SATs. That's where the help is put. No, the year 2s are taken out a group at a time.
167	Teacher A	I do all my year 2s at the same time but there are only usually 16 or 17 and the help does go to the rest of my class but they have to go somewhere else.
171	Teacher C	Our year 2 teachers currently use the additional supply money to put that towards a supply teacher to come in and take the class whilst groups are then taken out over a period of time by the class teacher and they get through their papers in that way.

176	PW	In each mixed-age class, of course, the supply teacher is looking after the rest.
178	AW	Do you think it would be successful if it was suggested that the questions were read out then and the children follow the text with the teacher? Do you think that would be a help or a hindrance for the majority of children?
181	Teacher A	It would be interesting to do a trial with one test where the children are using the papers in one way; other children do the same test but they are read out.
186	PW	You can’t do that with the live thing – with the statutory test.
187	Teacher B	But I think also to do it in smaller groups because, if all did it that way, my fear would be that there’d be some children are waiting ages, that the kind of looking at what someone else had written, there’d be more time for that kind of thing, whereas, when they all get to their own stage, they can’t copy because they are on different questions.
191	AW	Yes, that is one of the drawbacks, every child is on the same page and the same question at the same time and it does open up the potential for copying.
193	Teacher B	And someone might say, ’Oh, yes, it’s so and so’.
194	AW	I think that has been quite useful, P W. Is there anything else you think it would be useful to ask at this point in time? (directed at PW)
197	PW	No, I think that’s about it. You are allowed now to do that reading to children.
199	AW	But it not made explicit in the handbook though, is it?
200	PW	No, it’s not in there. It’s in the assessment and reporting arrangements booklet; it says you can do it, doesn’t it. It’s not actually specified when it comes down to the maths part of it. It’s not said again.
203	Teacher A	And also then, you would read out the maths words, wouldn’t you? Would you read out words like kilograms and all that stuff?
204	AW	Well, I was hoping that something you might have said, and my personal view is, there may be some children who could be disadvantaged and I think that is what you were getting at. If there were any children who were going to be disadvantaged by reading every question to the children, what children do you think it would be? Who would be disadvantaged by that method of administration?

209	Teacher A	The very slow learners.
210	AW	The slow workers do you think?
211	Teacher A	No, I think the more able ones that would slowly get bored with it.
214	Teacher B	The slow ones just generally write down an answer – they haven't got it yet. They ask you to read it to them to make sure.
215	Teacher A	That's why it would be good doing it, like, in ability groups but taking a group at a time (next phrase unclear).
216	Teacher B	I mean that does work in school. We've done it for quite a few years now.
218	PW	But even with ability groups, I was watching one girl in this last test who was working on one of the problems – that one where you have to explain your working – and she actually worked on it for 12 minutes and she was one of the brighter mathematicians.
233	AW	Was that because she was doing a lot of explanation?
234	PW	She was doing a lot of sorting out how she was going to answer it. She started off on a trial and improvement thing and adjusting her answer but she started adjusting it the wrong way. Then she then went to, 'I could do a subtraction.', but she wasn't very confident about how to do the subtraction so then she went back to the trial and improvement. I'm not actually sure she got there in the end but she was actually solidly working on it for that amount of time because she was actually one of the children I was observing so I was actually going back to see which question she was on. Five minutes later, she was still on the same question; five minutes later, the same question! If you had read it to her and the others, would you have waited that amount of time for her to finish trying it or would you say, 'Come back to it later'?
238	Teacher C	That's exactly what we did this time if we still say somebody, you know. And you got quite a clear impression of who was still working on a question and who wasn't and it generally considered to be one or two and we did have to say, 'Well, we're going to give you x-amount of time at the end, so just remember that.'
242	AW	Do you tell the children in advance that, you know, 'If you don't have enough time, you'll have time for checking afterwards so you can go back'.

243	Teacher C	This time, you know, we just said the majority we saw – that included the range from the higher-achievers to the lower-achievers – if we saw that they had completed what they had to do, and there were one or two (next phrase unclear) who had stopped by then anyway, we said, 'We are going to move on to the next question. Get ready to read that. If you haven't completed it, you can come back.'
249	AW	I think you'd have to have a system like that, wouldn't you? You couldn't wait 12 minutes for a child to answer a question! (laughter and sounds of agreement)
252	PW	Because you'd be there forever! You'd be out of year 3 before you knew what you were doing.
253	AW	That's been quite useful – a few pointers! Well, I think we'd better let everyone go now.

Transcript 2

I have adopted the following conventions:

1. The numbers are from the tape-counter.
2. AW identifies myself as interviewer.
3. I have ignored short pauses in the conversation since these do not appear to be significant in the context of the transcript.
4. Asides are in brackets and indicate occurrences such as a point of clarification for the reader, indistinct speech, etc.
5. Relevant comments are highlighted as follows:

   yellow: comments on reading demand;
   green: comments on administration
   turquoise: comments on reading support
   grey: views on look-listen

Circumstances of interview: JV was a member of the KS1 mathematics writing team, and has experience of writing test questions and observing pre-tests. She is an experienced Y2 teacher, a mathematics co-ordinator and has several years experience of administering the KS1 statutory tests. She had administered the KS1 2000 mathematics test in May 2000 to her Y2 class.

Date of interview: 29 December 2000

002 AW What are your thoughts about the reading demand of the key stage 1 maths test?

003 JV Really, I think it was quite demanding with only level 3 readers able to cope without help, and this was the opinion of the teachers that I spoke to in the schools where I did the pre-test observations as well as what I think from my own experience.

005 AW How do you cope with less able readers during the test?

006 JV Well, I administer the test in groups and they are asked to put their hand up if they require help with reading. Where there is more than one group, you have a mixture of abilities in each group so that some can get on independently. Where all are the lower ability group for reading or for maths or often for both, the teacher, if she has no help, is really stretched to get round them all. For the lower ability readers, putting a word bank on the wall before you start the start the test, (as recommended in the teacher's guide) is not really a particularly useful strategy because the children have forgotten what the words say by the time they need to access them.

012 AW Are there any other issues arising from the language?

013 JV I haven't got a copy of the test (pre-test booklet) in front of me because I returned all the materials to the QCA, but some questions seemed to have a lot of language or instructions before you could address the mathematics, the one about the bread in the shop, for example. I know this was aimed at the level 3, but I think there were some level 2 questions, which I can't quite recall at the moment, which would have quite a lot of writing to access before you could get down to answering them.

018 AW Would oral delivery of the whole test (look-listen) advantage more children than it would disadvantage?

019 JV I think it probably would. At present, the poorest readers need everything to be read individually so oral delivery would save the teachers rushing around and also the children getting stressed or having to wait. It would also help those level 2 readers who think they are better readers than they actually are or do not read carefully enough or do not wish to put up their hand to ask for things to be read because they think that this is the mark of somebody who is not a very good reader and the description does not apply to them

027 AW What children might it disadvantage and why?

028 JV Those children who may be disadvantaged, I think, are the secure (level) 3s and the borderline (level) 4s for whom the pace would be too slow and boredom could set in with the result that they wouldn't do their best perhaps. The speed of delivery would have to be at the speed of the slowest and it would take a long time. Presumably you would also have the printed text in the test booklet, and, if you did that, how could you stop the children moving on whilst you are waiting for the slowest to complete. You can see this happening already if watch children doing the oral part of the test. If they finish what they have done, and they are waiting for the slowest to make up his mind what he is going to write, you can see it, they flick through the rest of the book; it would be very difficult to stop children actually doing this if you insisted on reading everything out.

041 AW What organisational problems might there be?

043 JV I think one of the things that would be most serious problem would be the problem with copying. I have seen used various strategies, spreading children out in the classroom, using picture books to shield work of children sitting opposite, using blank sheets to cover completed pages, but one of the strategies we use at school is to seat a less able child next to a more able child. You find that the more able child whizzes on through the booklet and is, therefore, not doing the same page as the less able child is doing. Now where everybody is on the same question at the same time, I think copying is going to be much more an issue unless you have got the opportunity to spread the children really well out. I don't think that a formal examination situation is appropriate for key stage 1 testing, especially for low achievers, who need, I think, more personal interaction between the teacher and the pupil.

052 AW Thank you, JV.

Copy of the KS1 2001 mathematics test booklet

(filed in pocket in inside back cover)

The test booklet is not available in the web copy of this thesis

Interview prompt-sheet

1 What are your views on the reading demand of the KS1 maths test?

2 How do you cope with less able readers during the test?

3 Any there any other issues arising from the language?

4 Would oral delivery (look-listen) of the whole test advantage more children than it would disadvantage?

5 Which children might it disadvantage and why?

6 What organisational problems might there be?

Key stage 1 Mathematics Test 2001
(Key to column headings)

No	Child	Oral marks	Written question number																														Total	Ma sc	Ma L	Re L	EAL
		001- 005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035					
1																																					
2																																					
3																																					
4																																					
5																																					
6																																					
7																																					
8																																					
9																																					
10																																					
11																																					
12																																					
13																																					
14																																					
15																																					
16																																					
17																																					
18																																					
19																																					
20																																					
21																																					
22																																					
23																																					
24																																					
25																																					
26																																					
27																																					
28																																					
29																																					
30																																					
31																																					
32																																					
33																																					
34																																					
35																																					
36																																					

Child's first name

Marks obtained in oral questions 001 - 005 read aloud by teacher (max. 5 marks)

006 - 035 are written question numbers. Any 1s in these columns indicate a request for help with reading at least one word. (max. 31 marks)

National curricuum levels appearing in columns Ma L and Re L:
W: no level awarded
L1: level 1(below-average)
L2C: level 2 (below-average)
L2B: level 2 (average)
L2A: level 2 (above-average)
L3: level 3 (above-average)

Total indicates frequency of child's requests for help with reading (max. 30 written questions)

Ma sc indicates child's score out of 36 in test overall. (35 questions with question 30 having 2 linked marks)

Ma L indicates child's level in statutory KS1 mathematics test

Re L indicates child's level in statutory KS1 reading task/test.

1 indicates has English as an additional language.

Key stage 1 Mathematics Test 2001 (School K)

(ordered by reading level)

Written question number

No	Child	Oral marks 001- 005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035	Total	Ma sc	Ma L	Re L	EAL
4	Zeeshan	3		1			1	1						1	1	1	1	1	1			1	1	1		1	1	1		1			16	8	2c	1	1
1	Ashley	2			1						1		1				1	1			1												6	4	W	2c	1
2	Lakhbir	1	1							1					1	1	1				1					1	1	1		1			10	6	1	2c	1
6	Sutinder	4			1		1		1			1	1																				5	9	2c	2c	1
10	Priya	3		1					1	1	1	1	1	1	1				1			1											10	12	2c	2c	1
13	Sophia	2						1	1						1			1			1				1								6	14	2b	2c	1
15	Liam	3			1	1		1	1		1	1																					6	14	2b	2c	
16	Shereen	4	1		1			1	1	1	1	1			1																		8	14	2b	2c	
18	Dwayne	3		1			1	1		1	1	1		1	1	1	1		1	1		1			1	1	1	1					17	15	2b	2c	
19	Yasmin	3			1		1	1		1	1	1		1	1	1	1		1		1					1	1	1	1			1	17	16	2b	2c	
3	Mica	2		1				1			1	1	1	1	1	1	1		1	1	1	1	1		1	1	1			1			18	8	2c	2b	1
8	Gurvinder	2					1		1		1			1					1							1	1						7	12	2c	2b	1
9	Stephanie	1					1	1				1					1		1	1		1						1				1	9	12	2c	2b	1
11	Jaspreet	2						1			1	1														1	1		1				6	14	2b	2b	1
20	Chloe	4							1						1	1							1	1									5	17	2b	2b	
22	Noly	3						1	1					1	1	1	1						1										7	20	2a	2b	
24	Nyaradzo	5																															0	21	2a	2b	
25	Jumaane	5					1	1				1		1		1	1	1	1				1										9	25	2a	2b	
27	Keisha	4																															0	27	3	2b	
5	Curtis	3				1					1		1	1					1		1	1	1										8	9	2c	2a	
7	Monica	2																									1						1	10	2c	2a	1
14	Ellie	3						1					1		1								1	1									5	14	2b	2a	
17	Usman	3									1	1		1	1		1	1					1		1	1	1	1		1		1	13	14	2b	2a	1
21	Saba	2									1				1								1										3	18	2b	2a	1
26	Sanjay	3																															0	26	3	2a	1
29	Omar	5																															0	32	3	2a	1
12	Kajal	2																					1		1				1		1	1	5	14	2b	3	1
23	Adam	4											1		1		1						1		1		1				1		7	21	2a	3	
28	Bethan	4																															0	30	3	3	
			2	4	5	2	7	12	8	5	12	11	7	10	14	8	11	5	9	3	6	6	11	3	6	8	10	6	3	4	2	4	204				

Red total: number of times help was requested for particular questions

Blue total: number of times help was requested by particular children

Average number of requests for help: 7.03

Key stage 1 Mathematics Test 2001 (School N)

(ordered by reading level)

No	Child	Oral marks																	Written question number														Total	Ma sc	Ma L	Re L	EAL
		001- 005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035					
17	Daniel (1)	2			1			1	1		1	1	1	1	1	1	1		1	1	1	1	1	1	1		1	1					19	15	2b	1	
1	William	0					1	1		1	1	1	1	1	1	1	1				1					1							12	3	W	2c	
4	Scott	1																															0	10	2c	2c	
5	Katrina	0					1	1			1			1			1																5	11	2c	2c	
20	Joshua	4								1				1								1			1	1	1	1		1			8	15	2b	2c	
21	Kirsty	4						1											1					1		1							4	17	2b	2c	
47	Daniel (3)	5									1	1	1	1	1	1					1	1				1	1	1		1		1	13	31	3	2c	
2	Sharon	1					1	1	1		1			1																			5	8	2c	2b	
6	Abbey	2						1				1	1	1	1						1					1						1	8	11	2c	2b	
8	Connie	1						1			1	1	1	1	1																		6	12	2c	2b	
9	Olive	3							1		1							1		1													4	12	2c	2b	
10	Nicole	3			1			1	1		1	1	1		1	1	1		1	1	1			1	1		1	1					16	14	2b	2b	
11	Beatrice	2						1	1		1	1	1				1																6	14	2b	2b	
15	Georgina	3				1							1				1	1	1									1					6	14	2b	2b	
18	Aarondeep	3																															0	15	2b	2b	
22	Luke	4												1																			1	17	2b	2b	
23	Laura (1)	3			1						1				1	1						1		1		1							7	18	2b	2b	
26	Bethany	4											1		1																		2	19	2b	2b	
28	Jade	3						1			1	1	1		1	1	1							1		1							9	20	2a	2b	
31	Katie	4																															0	22	2a	2b	
33	Lewis	4									1		1					1		1				1		1						1	7	22	2a	2b	
34	Jordan	4			1						1							1															3	22	2a	2b	
36	Charlotte (1)	4													1		1	1	1		1	1								1	1	1	9	23	2a	2b	
3	Jessica	2						1		1	1	1	1	1	1	1	1		1			1	1	1		1	1			1			16	10	2c	2a	
7	Sophie	3								1	1															1							3	11	2c	2a	
12	Jack	3									1	1	1										1		1	1		1					7	14	2b	2a	
13	Jodie	4																															0	14	2b	2a	
14	Matthew	3																															0	14	2b	2a	
16	Kathryn	1																								1							1	14	2b	2a	
19	Ellis	3													1																		1	15	2b	2a	
24	Laura (2)	3													1											1							2	18	2b	2a	
25	Nicola	4												1					1				1										3	18	2b	2a	
27	Daniel (2)	3									1															1							2	20	2a	2a	
29	Harriet	3						1			1			1				1			1			1			1	1					8	20	2a	2a	
30	Charlie	4																															0	21	2a	2a	
32	Sarah	5																															0	22	2a	2a	
38	Jaspreet	5									1						1													1	1		4	24	2a	2a	
39	Rebecca	4									1								1											1			3	25	2a	2a	
40	Charlotte (2)	4											1	1																			2	26	3	2a	
42	Sasha	5						1			1				1	1																	4	26	3	2a	
44	Liam	5																															0	28	3	2a	
35	Elie	5																															0	23	2a	3	
37	Lucy	5									1																						1	24	2a	3	
41	Mariah	5																															0	26	3	3	
43	Olivia	5									1																						1	27	3	3	
45	Danesh	5																															0	30	3	3	
46	Hayley	5																															0	30	3	3	
			0	0	4	1	3	13	5	4	23	10	14	13	14	8	10	6	8	4	7	6	4	8	4	14	6	7	0	6	2	4	208				

Red total: number of times help was requested for particular questions

Blue total: number of times help was requested by particular children

Average number of requests for help: 4.52

														Key stage 1 Mathematics Test 2001										(School O)											
																	(ordered by reading level)																		
No	Child	Oral marks																Written question number													Total	Ma sc	Ma L	Re L	EAL
		001- 005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035					
1	Michael	1		1			1	1		1	1	1	1				1	1	1			1	1	1		1	1	1		1			17	8	2c	2c	
2	Carlon	1						1		1	1	1	1					1		1			1		1	1				1			11	10	2c	2c	
4	Aaliya	2			1		1	1		1	1	1	1	1		1	1		1			1				1	1	1		1		1	17	12	2c	2c	
10	Kenneth	3							1		1			1	1			1	1					1		1	1	1				1	11	22	2a	2c	
11	Elijah	3						1	1		1			1			1	1						1		1		1					9	23	2a	2c	
3	Marlon	0									1	1		1	1	1	1		1				1		1	1	1	1		1			13	11	2c	2b	
5	Munitta	1																				1	1			1				1			4	14	2b	2b	1
6	Leon	3											1	1					1	1			1		1	1	1				1		9	15	2b	2b	
7	Monica	2									1									1				1			1	1					5	19	2b	2a	1
8	Josh	3							1								1	.								1				1			4	20	2a	2a	
13	Caria	4																															0	26	3	2a	
18	Sasha	4																															0	33	3	2a	
9	Raveena (1)	3																															0	21	2a	3	1
12	Kellie	4																															0	23	2a	3	
14	Raveena (2)	5																															0	27	3	3	1
15	Ashleigh	4																															0	29	3	3	
16	Jade	5																															0	31	3	3	
17	Hannah	5																															0	32	3	3	
.			0	1	1	0	2	4	3	3	7	4	4	5	2	2	5	4	5	3	0	3	5	4	3	9	6	6	0	6	1	2	100				

Red total: number of times help was requested for particular questions
Blue total: number of times help was requested by particular children

\verage number of requests for help: 5.5

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
1																	**Key stage 1 Mathematics Test 2001**										**(School P)**											
2																				**(ordered by reading level)**																		
3																																						
4	**No**	**Child**	**Oral marks**																		**Written question number**													**Total**	**Ma sc**	**Ma L**	**Re L**	**EAL**
5			**001- 005**	**006**	**007**	**008**	**009**	**010**	**011**	**012**	**013**	**014**	**015**	**016**	**017**	**018**	**019**	**020**	**021**	**022**	**023**	**024**	**025**	**026**	**027**	**028**	**029**	**030**	**031**	**032**	**033**	**034**	**035**					
6	1	Nicholas	1						1	1			1		1	1		1		1		1	1		1			1				1	1	13	4	W	1	
7	2	Daniel	0			1			1	1		1		1		1							1	1					1					9	5	1	1	
8	3	Lewis	2		1				1			1				1	1	1				1	1		1			1						10	8	2c	1	
9	4	Joanna	2					1			1	1				1		1	1					1				1	1	1	1		1	12	8	2c	1	
10	5	Frankie	2			1		1				1				1		1	1				1	1	1			1				1		11	9	2c	2c	
11	7	Thomas	2			1		1							1							1												4	10	2c	2c	
12	8	Charlotte	2	1		1	1					1		1		1				1		1										1		9	10	2c	2c	
13	11	Andre	2										1		1																1			3	11	2c	2c	
14	6	Elle-Mae	1						1																								1	2	9	2c	2b	
15	9	Jay	2											1																				1	10	2c	2b	
16	10	Jade	2									1			1																			2	11	2c	2b	
17	12	Jason	4									1																						1	12	2c	2b	
18	15	Ashley	4																															0	16	2b	2b	
19	19	Lee	4																															0	22	2a	2b	
20	13	Amber	3							1																							1	2	13	2c	2a	
21	16	James	4									1			1												1							3	16	2b	2a	
22	17	Alice	4												1																			1	19	2b	2a	
23	18	Jessica	3																															0	21	2a	2a	
24	20	Jake	5																															0	22	2a	2a	
25	21	Emma	4																															0	23	2a	2a	
26	22	Bethany	5																						1								1	2	24	2a	2a	
27	26	Emily	5													1																	1	2	27	3	2a	
28	14	Chloe	3																															0	14	2b	3	
29	23	Cherrie	5													1								1								1	1	4	24	2a	3	
30	24	Callum	5																															0	26	3	3	
31	25	Alexandra	5																								1							1	27	3	3	
32	27	Adam	5																															0	33	3	3	
33	28	Nathan	5																															0	33	3	3	
34				1	1	4	1	3	4	3	1	8	2	3	6	8	1	4	2	2	0	4	4	4	4	0	2	4	2	1	2	4	7	92				
35																																						
36																																						
37																																						
38										Red total: number of times help was requested for particular questions																												
39										Blue total: number of times help was requested by particular children																												
40																																						
41										Average number of requests for help: 3.29																												
42																																						

Key stage 1 Mathematics Test 2001 (School Q)
(ordered by reading level)

No	Child	Oral marks	Written question number																														Total	Ma sc	Ma L	Re L	EAL
		001- 005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035					
1	Portia	1	1	1			1	1			1		1	1					1			1	1										10	5	1	1	
4	Shannon	1									1		1	1	1		1	1								1	1					1	9	8	2c	1	
5	William	3					1	1	1	1	1			1	1		1	1	1		1	1	1	1		1	1	1		1			18	9	2c	1	
2	Charlene	1			1		1	1	1				1		1	1											1	1					9	5	1	2c	
3	Carl	1				1		1	1		1	1	1	1	1				1		1	1				1	1						13	6	1	2c	
6	Jamel	1			1	1	1	1	1			1	1	1	1	1	1		1		1	1	1	1	1	1	1	1		1	1	1	23	9	2c	2c	
10	Jade	1					1	1			1	1		1	1	1	1	1	1		1	1	1	1		1	1	1		1			18	15	2b	2c	
7	Jenna	2						1			1		1	1	1	1	1	1			1						1						10	11	2c	2b	
8	Kuranveer	2						1			1			1									1	1	1								6	11	2c	2b	
9	Marius	4						1						1	1				1			1					1	1					7	14	2b	2b	
11	Saffron	2				1		1				1	1		1		1	1			1	1	1	1				1		1	1	1	15	15	2b	2b	
18	Emily	3						1			1			1			1	1			1	1			1	1		1		1		1	12	22	2a	2b	
12	Leanne	3		1																			1	1		1		1					5	16	2b	2a	
15	Laura	4						1			1			1	1	1						1		1				1				1	9	18	2b	2a	
16	Aneeka	3																										1					1	21	2a	2a	
13	Victoria	3						1			1							1			1			1							1		6	18	2b	3	
14	Nicole	4																								1							1	18	2b	3	
17	Patricia	4									1													1		1	1	1					5	21	2a	3	
19	Nicholas	5									1																						1	27	3	3	
20	Francesco	3									1															1	1						3	27	3	3	
21	Melissa	4																															0	28	3	3	
22	Sean	3																1															1	29	3	3	
23	Alex	5																								1							1	31	3	3	
24	Richard	5																															0	34	3	3	
			1	2	2	3	5	13	4	1	13	4	7	11	10	5	7	8	6	0	8	9	7	9	3	11	10	11	0	5	3	5	183				

Red total: number of times help was requested for particular questions
Blue total: number of times help was requested by particular children

Average number of requests for help: 7.63

Key stage 1 Mathematics Test 2001 (School R)
(ordered by reading level)

No	Child	Oral marks	Written question number																														Total	Ma sc	Ma L	Re L	EAL
		001- 005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035					
17	Jordan	4	1				1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1									18	16	2b	1	
18	Clive	2					1	1		1	1	1		1	1	1	1	1	1	1		1	1	1		1	1	1	1	1	1	1	22	12	2c	1	
1	Ashley (1)	1					1	1	1	1	1	1		1	1	1	1		1		1	1	1	1	1	1	1	1	1		1	1	22	5	1	2c	
3	Kirsty	4	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	28	17	2b	2c	
4	Sonia	4	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	29	22	2a	2c	
7	Ashley (2)	4										1		1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	19	19	2b	2c	
13	Whitney	4			1		1	1		1	1	1		1	1	1			1		1	1		1		1	1	1		1	1	1	19	9	2c	2c	
16	Mick	4			1		1	1	1		1			1	1		1		1							1		1		1		1	13	14	2b	2c	
6	Joshua	4											1		1	1			1		1		1	1				1					8	15	2b	2b	
10	Billy-Jo	2									1			1			1				1	1		1				1		1		1	9	14	2b	2b	
11	Lauren	3						1				1	1		1				1		1			1		1						1	9	15	2b	2b	
14	Danny	4					1	1	1		1		1	1	1	1	1				1			1		1	1	1		1	1	1	17	19	2b	2b	
15	Aiden	3						1			1			1	1	1	1		1		1	1	1	1	1	1		1				1	15	13	2c	2b	
2	Natalie	3						1			1	1			1		1									1							6	12	2c	2a	
9	Naomi	3																															0	18	2b	2a	
12	Charlotte	2									1																						1	9	2c	2a	
5	Dean	5									1	1	1		1									1									5	27	3	3	
8	Luke	5																			1												1	30	3	3	
			3	1	3	2	8	11	6	6	13	10	7	11	14	10	11	5	11	5	12	9	6	13	5	11	6	11	5	8	7	11	241				

Red total: number of times help was requested for particular questions
Blue total: number of times help was requested by particular children

Average number of requests for help: 13.67
(28 children in class - 10 took level 1 maths task rather than test)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
1		**Key Stage 1 Mathematics Test 2001**																																				
2			**(combined data: 164 children)**																																			
3																**Arranged by reading level**							**(Re L)**															
4																																						
5	**No**	**Child**	**Oral marks**																															**Total**	**Ma sc**	**Ma L**	**Re L**	**EAL**
6			**001- 005**	**006**	**007**	**008**	**009**	**010**	**011**	**012**	**013**	**014**	**015**	**016**	**017**	**018**	**019**	**020**	**021**	**022**	**023**	**024**	**025**	**026**	**027**	**028**	**029**	**030**	**031**	**032**	**033**	**034**	**035**					
7	4	Zeeshan	3		1			1	1						1	1	1	1	1	1			1	1	1		1	1	1		1			16	8	2c	1	1
8	46	Daniel (1)	2			1			1	1		1	1	1	1	1	1	1		1	1	1	1	1	1	1		1	1					19	15	2b	1	
9	95	Portia	1	1	1			1	1			1		1	1					1			1	1										10	5	1	1	
10	98	Shannon	1									1		1	1	1		1	1								1	1					1	9	8	2c	1	
11	99	William	3					1	1	1	1	1			1	1		1	1	1		1	1	1	1		1	1	1		1			18	9	2c	1	
12	135	Jordan	4	1				1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1									18	16	2b	1	
13	136	Clive	2					1	1		1	1	1		1	1	1	1	1	1	1		1	1	1		1	1	1	1	1	1	1	22	12	2c	1	
14	137	Nicholas	1						1	1			1		1	1		1		1		1	1		1			1				1	1	13	4	W	1	
15	138	Daniel	0			1			1	1		1		1		1							1	1					1					9	5	1	1	
16	139	Lewis	2		1				1			1				1	1	1				1	1		1			1						10	8	2c	1	
17	140	Joanna	2					1			1	1				1		1	1					1				1	1	1	1		1	12	8	2c	1	
18	1	Ashley	2			1						1		1				1	1			1												6	4	W	2c	1
19	2	Lakhbir	1	1							1					1	1	1				1					1	1	1		1			10	6	1	2c	1
20	6	Sutinder	4			1		1		1			1	1																				5	9	2c	2c	1
21	10	Priya	3		1					1	1	1	1	1	1	1				1			1											10	12	2c	2c	1
22	13	Sophia	2						1	1						1			1			1				1								6	14	2b	2c	1
23	15	Liam	3			1	1		1	1		1	1																					6	14	2b	2c	
24	16	Shereen	4	1		1			1	1	1	1	1			1																		8	14	2b	2c	
25	18	Dwayne	3		1			1	1		1	1	1		1	1	1	1		1	1		1			1	1	1	1					17	15	2b	2c	
26	19	Yasmin	3			1		1	1		1	1	1		1	1	1	1		1		1					1	1	1	1			1	17	16	2b	2c	
27	30	William	0					1	1		1	1	1	1	1	1	1	1				1					1							12	3	W	2c	
28	33	Scott	1																															0	10	2c	2c	
29	34	Katrina	0					1	1			1			1			1																5	11	2c	2c	
30	49	Joshua	4								1				1								1			1	1	1	1		1			8	15	2b	2c	
31	50	Kirsty	4						1											1					1		1							4	17	2b	2c	
32	76	Daniel (3)	5									1	1	1	1	1	1					1	1				1	1	1		1		1	13	31	3	2c	
33	77	Michael	1		1			1	1		1	1	1	1				1	1	1			1	1	1		1	1	1		1			17	8	2c	2c	
34	78	Carlon	1						1		1	1	1	1					1		1			1		1	1				1			11	10	2c	2c	
35	80	Aaliya	2			1		1	1		1	1	1	1	1		1	1		1			1				1	1	1		1		1	17	12	2c	2c	
36	86	Kenneth	3							1		1			1	1			1	1					1		1	1	1				1	11	22	2a	2c	
37	87	Elijah	3						1	1		1			1			1	1						1		1		1					9	23	2a	2c	
38	96	Charlene	1			1		1	1	1				1		1	1											1	1					9	5	1	2c	
39	97	Carl	1				1		1	1		1	1	1	1	1				1		1	1				1	1						13	6	1	2c	
40	100	Jamel	1			1	1	1	1	1			1	1	1	1	1	1		1		1	1	1	1	1	1	1	1		1	1	1	23	9	2c	2c	
41	104	Jade	1					1	1			1	1		1	1	1	1	1	1		1	1	1	1		1	1	1		1			18	15	2b	2c	
42	119	Ashley (1)	1					1	1	1	1	1	1		1	1	1	1		1		1	1	1	1	1	1	1	1	1		1	1	22	5	1	2c	
43	121	Kirsty	4	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	28	17	2b	2c	
44	122	Sonia	4	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	29	22	2a	2c	
45	125	Ashley (2)	4										1		1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	19	19	2b	2c	
46	131	Whitney	4			1		1	1		1	1	1		1	1	1			1		1	1		1		1	1	1		1	1	1	19	9	2c	2c	
47	134	Mick	4			1		1	1	1		1			1	1		1		1							1		1		1		1	13	14	2b	2c	

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
48	141	Frankie	2			1		1				1				1		1	1				1	1	1			1				1		11	9	2c	2c	
49	143	Thomas	2			1		1							1							1												4	10	2c	2c	
50	144	Charlotte	2	1		1	1					1		1		1				1		1										1		9	10	2c	2c	
51	147	Andre	2										1		1																1			3	11	2c	2c	
52	3	Mica	2		1				1			1	1	1	1	1	1	1		1	1	1	1	1		1	1	1			1			18	8	2c	2b	1
53	8	Gurvinder	2					1		1		1			1					1							1	1						7	12	2c	2b	1
54	9	Stephanie	1					1	1				1					1		1	1		1						1				1	9	12	2c	2b	1
55	11	Jaspreet	2						1			1	1														1	1		1				6	14	2b	2b	1
56	20	Chloe	4							1						1	1							1	1									5	17	2b	2b	
57	22	Noly	3						1	1					1	1	1	1						1										7	20	2a	2b	
58	24	Nyaradzo	5																															0	21	2a	2b	
59	25	Jumaane	5					1	1				1		1		1	1	1	1				1										9	25	2a	2b	
60	27	Keisha	4																															0	27	3	2b	
61	31	Sharon	1					1	1	1		1			1																			5	8	2c	2b	
62	35	Abbey	2						1				1	1	1	1						1					1						1	8	11	2c	2b	
63	37	Connie	1						1			1	1	1	1	1																		6	12	2c	2b	
64	38	Olive	3							1		1							1		1													4	12	2c	2b	
65	39	Nicole	3			1			1	1		1	1	1		1	1	1		1	1	1			1	1		1	1					16	14	2b	2b	
66	40	Beatrice	2						1	1		1	1	1				1																6	14	2b	2b	
67	44	Georgina	3				1							1				1	1	1									1					6	14	2b	2b	
68	47	Aarondeep	3																															0	15	2b	2b	
69	51	Luke	4												1																			1	17	2b	2b	
70	52	Laura (1)	3			1						1				1	1						1		1		1							7	18	2b	2b	
71	55	Bethany	4											1		1																		2	19	2b	2b	
72	57	Jade	3						1			1	1	1		1	1	1							1		1							9	20	2a	2b	
73	60	Katie	4																															0	22	2a	2b	
74	62	Lewis	4									1		1					1		1				1		1						1	7	22	2a	2b	
75	63	Jordan	4			1						1							1															3	22	2a	2b	
76	65	Charlotte (1)	4													1		1	1	1		1	1								1	1	1	9	23	2a	2b	
77	79	Marlon	0									1	1		1	1	1	1		1				1		1	1	1	1		1			13	11	2c	2b	
78	81	Munitta	1																				1	1			1				1			4	14	2b	2b	1
79	82	Leon	3											1	1					1	1			1		1	1	1				1		9	15	2b	2b	
80	101	Jenna	2						1			1		1	1	1	1	1	1			1						1						10	11	2c	2b	
81	102	Kuranveer	2						1			1			1									1	1	1								6	11	2c	2b	
82	103	Marius	4						1						1	1				1			1					1	1					7	14	2b	2b	
83	105	Saffron	2				1		1				1	1		1		1	1			1	1	1	1				1		1	1	1	15	15	2b	2b	
84	112	Emily	3						1			1			1			1	1			1	1			1	1		1		1		1	12	22	2a	2b	
85	124	Joshua	4											1		1	1			1		1		1	1				1					8	15	2b	2b	
86	128	Billy-Jo	2									1			1			1				1	1		1				1		1		1	9	14	2b	2b	
87	129	Lauren	3						1				1	1		1				1		1			1		1						1	9	15	2b	2b	
88	132	Danny	4					1	1	1		1		1	1	1	1	1				1			1		1	1	1		1	1	1	17	19	2b	2b	
89	133	Aiden	3						1			1			1	1	1	1		1		1	1	1	1	1	1		1				1	15	13	2c	2b	
90	142	Elle-Mae	1						1																								1	2	9	2c	2b	
91	145	Jay	2											1																				1	10	2c	2b	
92	146	Jade	2									1			1																			2	11	2c	2b	
93	148	Jason	4									1																						1	12	2c	2b	
94	151	Ashley	4																															0	16	2b	2b	
95	155	Lee	4																															0	22	2a	2b	

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
96	5	Curtis	3				1					1		1	1					1		1	1	1										8	9	2c	2a	
97	7	Monica	2																									1						1	10	2c	2a	1
98	14	Ellie	3						1					1		1								1	1									5	14	2b	2a	
99	17	Usman	3									1	1		1	1		1	1					1		1	1	1	1		1		1	13	14	2b	2a	1
100	21	Saba	2									1				1								1										3	18	2b	2a	1
101	26	Sanjay	3																															0	26	3	2a	1
102	29	Omar	5																															0	32	3	2a	1
103	32	Jessica	2						1		1	1	1	1	1	1	1	1		1			1	1	1		1	1			1			16	10	2c	2a	
104	36	Sophie	3								1	1															1							3	11	2c	2a	
105	41	Jack	3									1	1	1										1		1	1		1					7	14	2b	2a	
106	42	Jodie	4																															0	14	2b	2a	
107	43	Matthew	3																															0	14	2b	2a	
108	45	Kathryn	1																								1							1	14	2b	2a	
109	48	Ellis	3													1																		1	15	2b	2a	
110	53	Laura (2)	3													1											1							2	18	2b	2a	
111	54	Nicola	4												1					1				1										3	18	2b	2a	
112	56	Daniel (2)	3									1															1							2	20	2a	2a	
113	58	Harriet	3						1			1			1				1			1			1			1	1					8	20	2a	2a	
114	59	Charlie	4																															0	21	2a	2a	
115	61	Sarah	5																															0	22	2a	2a	
116	67	Jaspreet	5									1						1													1	1		4	24	2a	2a	
117	68	Rebecca	4									1								1											1			3	25	2a	2a	
118	69	Charlotte (2)	4											1	1																			2	26	3	2a	
119	71	Sasha	5						1			1				1	1																	4	26	3	2a	
120	73	Liam	5																															0	28	3	2a	
121	83	Monica	2									1									1				1			1	1					5	19	2b	2a	1
122	84	Josh	3							1								1	.								1				1			4	20	2a	2a	
123	89	Caria	4																															0	26	3	2a	
124	94	Sasha	4																															0	33	3	2a	
125	106	Leanne	3		1																			1	1		1		1					5	16	2b	2a	
126	109	Laura	4						1			1			1	1	1						1		1				1				1	9	18	2b	2a	
127	110	Aneeka	3																										1					1	21	2a	2a	
128	120	Natalie	3						1			1	1			1		1									1							6	12	2c	2a	
129	127	Naomi	3																															0	18	2b	2a	
130	130	Charlotte	2									1																						1	9	2c	2a	
131	149	Amber	3							1																							1	2	13	2c	2a	
132	152	James	4									1			1												1							3	16	2b	2a	
133	153	Alice	4												1																			1	19	2b	2a	
134	154	Jessica	3																															0	21	2a	2a	
135	156	Jake	5																															0	22	2a	2a	
136	157	Emma	4																															0	23	2a	2a	
137	158	Bethany	5																						1								1	2	24	2a	2a	
138	162	Emily	5													1																	1	2	27	3	2a	
139	64	Elie	5																															0	23	2a	3	
140	66	Lucy	5									1																						1	24	2a	3	
141	70	Mariah	5																															0	26	3	3	
142	72	Olivia	5									1																						1	27	3	3	
143	74	Danesh	5																															0	30	3	3	

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
144	75	Hayley	5																															0	30	3	3	
145	12	Kajal	2																					1		1				1		1	1	5	14	2b	3	1
146	23	Adam	4											1		1		1						1		1		1				1		7	21	2a	3	
147	28	Bethan	4																															0	30	3	3	
148	85	Raveena (1)	3																															0	21	2a	3	1
149	88	Kellie	4																															0	23	2a	3	
150	90	Raveena (2)	5																															0	27	3	3	1
151	91	Ashleigh	4																															0	29	3	3	
152	92	Jade	5																															0	31	3	3	
153	93	Hannah	5																															0	32	3	3	
154	107	Victoria	3						1			1							1			1			1							1		6	18	2b	3	
155	108	Nicole	4																								1							1	18	2b	3	
156	111	Patricia	4									1													1		1	1	1					5	21	2a	3	
157	113	Nicholas	5									1																						1	27	3	3	
158	114	Francesco	3									1															1	1						3	27	3	3	
159	115	Melissa	4																															0	28	3	3	
160	116	Sean	3																1															1	29	3	3	
161	117	Alex	5																								1							1	31	3	3	
162	118	Richard	5																															0	34	3	3	
163	123	Dean	5									1	1	1		1									1									5	27	3	3	
164	126	Luke	5																			1												1	30	3	3	
165	150	Chloe	3																															0	14	2b	3	
166	159	Cherrie	5													1								1								1	1	4	24	2a	3	
167	160	Callum	5																															0	26	3	3	
168	161	Alexandra	5																								1							1	27	3	3	
169	163	Adam	5																															0	33	3	3	
170	164	Nathan	5																															0	33	3	3	
171				7	9	19	9	28	57	29	20	76	41	42	56	62	34	48	30	41	15	37	37	37	41	21	55	42	43	9	31	19	33	1028				
172																																						
173																																						
174																																						
175																																						
176																																						
177																																						
178																																						
179																																						
180																																						
181																																						
182																																						
183																																						
184																																						

Red total: number of times help requested for particular questions
Blue total: number of times help requested by particular children

Range of requests for help with reading: 0 - 29
Range for L2C readers or below: 0 - 29
Range for L2B readers: 0 - 18
Range for L2A readers: 0 - 16
Range for L3 readers: 0 - 7

Mean number of requests for help overall: 6.3
Mean number of requests for help for L2C readers or below: 12.6
Mean number of requests for help for L2B readers: 6.6
Mean number of requests for help for L2A readers: 2.6
Mean number of requests for help for L3 readers: 1.6
Mean number of marks for oral questions: 3.2
Mean number of times help was requested for a test item: 34.3

Categories used in transcripts for teachers at Schools K, N, P and Q

Category	**Category description**	**Identifying box in transcript**	**Colour coding used for most relevant text**
Reading demand	Teachers' views on the reading demand of the test	Demand	yellow
Administration	How children were organised for the test, e.g. a whole class setting or in groups; also availability of staff etc.	Administration	green
Support	How staff supported children with reading during the test	Support	turquoise
Look-listen	Teachers' views on look-listen	Look-listen	gray
Other	Any other relevant comments that did not fall into other categories	Other	pink

Appendix 5.4

School N: administration of the KS1 2001 mathematics test

Interviewer:	Anne Woodman (AW)
Interviewee:	Y2 teachers (MA and WK)
Date:	12 October 2001
School N:	Primary school in a catchment area that includes a private houses, council houses and flats. Few children have English as an additional language and no children in the two Y2 classes who took the mathematics test in were considered by their teachers to lack fluency in spoken forms of English. (See Appendix 5.1 page 3 for School N's results.)
Background:	This interview took place five months after the two year 2 teachers had administered the test. Due to staffing difficulties, I had only 10 minutes to interview the teachers.

010	AW	MA, what are your views on the reading demand of the 2001 key stage 1 maths test?
011 Demand Support	MA	I did think it was quite demanding. The language wasn't necessarily mathematical-based in the questions. There was a lot of rushing from child to child to try and read some of the language for them, and some of them, if they were quite timid children, weren't keen to ask. They would sit and just quietly try and get on. It was only if you would say to them, 'Are you all right? Do you need any help? Do you understand what you are doing?'
015	AW	So you had to intervene with some of them?
004	MA	Yes.
005 Demand	WK	I think that the actual reading was more difficult than the maths that was required in quite a few of them. There was way too much reading for the amount of work that was required in a mathematical way and very often the language was too hard.
019	AW	A lot to read to get one mark? Is that what you mean?
020	WK	Yes, one in particular was three or four sentences, wasn't it? So I thought that was inappropriate, and a lot of them were put off. They didn't attempt to read it, I felt, because they were just either weary or it was too daunting.
022	AW	Could you just explain how you actually administered the test? Did you have them class by class or did you use groups or .. .
024	MA	I did mine half and half. I did half the class at a time.
025	AW	I think there were about 25 children?

025	MA	Yes, it was about 13 or 14, and there was the nursery nurse doing something with the rest.
Administration		
026	AW	Right. And the instructions were just to put their hands up if they needed help with reading?
027	MA	I would say to them, you know, ‘If you can’t read it, put your hand up.’ or ‘If you need any help, put your hand up and I will read it and help you.’
028	AW	Are you convinced that all the children did that?
029	WK	You always get the quiet and tentative ones who look as though they are doing something but they’re not really confident, I am sure.
030	AW	Did you administer (directed at WK) ..?
031	WK	Well, in the same way, exactly the same. We just highlighted any words that the children asked us to read to them.
Support		
032	AW	Yes. When it came to the highlighting, did you just underline what the children specifically asked for help with or did you automatically read the whole question to them, can you remember?
033	WK	Just what they asked for help with, wasn’t it (to MA) because I remember being very conscious that it is very strict guidelines, isn’t it, about how you should administer and what you should say? I remember being scared that I might say something that I shouldn’t have done, as well.
Support		
036	MA	I am always conscious that there were words you weren’t allowed to say to them but I think I did read – I highlighted the whole thing, actually on mine even though it included words that I wouldn’t have said.
038	AW	Yes, I think you made a note to that effect on the booklets.
039	MA	I put that note so that you would understand but I just highlighted the whole lot for quickness.
040	AW	At the moment, I am finding that at least half the research schools in my study are using look and listen for at least some of the children, and, in some cases, for all the children. You know, where you actually read every question but the children follow it with their finger. What would your views on that be? I’ll ask your views first this time (to WK).
Look/listen		
044	WK	I believe that has a big influence on the results. I’m not convinced .

045	AW	Positive or negative?
045	WK	Positive. But I'm not convinced that it's the best thing for the more able children because they get frustrated because they can't work at their own speed. I have experienced that. So I'm not convinced it is the best way to do it for them.
048	AW	Which children do you think it would advantage then?
048	WK	The ones that are less able readers. Yes.
048	MA	Yes. Yes.
049	AW	So, who would those be? Would it be your 2c readers?
050	WK	Yes.
050	MA	Yes.
050	AW	(Level) 2b readers?
051	WK	Possibly. I mean, I think you have a mixture there of, you know, sort of, any results that you've got if you do the reading (statutory test) first, which we do, and then your own teacher assessment and your own experience of children's personal confidence.
053	AW	Do you think it is fair nationally if some schools are using look and listen and some are not?
054 Other	WK	It's not a fair test, is it, due to it being administered in so many different ways. It should perhaps be more clearly defined how we should administer all tests, really.
056	MA	Yes, that stands true at key stage 2 and key stage 1.
057	AW	Is there anything else you feel I haven't asked about the reading demand of the test that you'd like to get off your chest?
060	WK	I think the length of the test is too long, way too long. The children run out of steam, definitely, even the able ones run out of steam, and obviously the amount of reading in that, that's a factor of it.
061	AW	If I could provide you with evidence that a lot of children whom I would have expected, based on the data you gave me, to ask for more help (with reading), hadn't asked for help, would it influence how you administered the test next year?
063	MA	Yes, I think it would.
063	WK	Yes.

063	AW	I'd be interested in your views when you've had a look at that data analysis sheet that I've given to Tom (the headteacher).
064	WK	I mean, I think I'd already thought beforehand of saying to them that maybe we should perhaps administer it differently this time.
066	MA	I agree. Maybe even if it was just with the half that was the less able, if other schools are doing it that way.
067	WK	We really ought to be batting on the same wicket.
067	AW	Thanks very much.

School U: administration of the KS1 2001 mathematics test

Interviewer:	AW
Interviewee:	Y2 teacher (TC)
Date:	5 July 2001
School U:	Primary school in an area of mainly private housing on a new estate.
Background:	This interview took place about two months after the teacher had administered the KS1 mathematics test.

096 AW Can you tell me your views on the reading demand of the key stage 1 mathematics test?

096 TC With the children who are better readers, they really didn't have a problem with it but the poorer – our average to the poorer end – although they can read the individual words, most of them, they can't understand what they are reading.

Demand

097 AW They can't string it together.

098 TC That's right, when they've read it, they can't then say, 'Oh, this means so and so.', so they need to have the oral – have it all read to them. It makes it all so much easier for them.

Support

100 AW So, in terms of how you organised the children this year, how did you manage it?

101 TC I'd done the 1997 onwards tests as preparation for them and after the first couple of tests, I could see there were children who were quite happy and able to go at their own pace which was a group of children who - seven children. The rest of them ..

Administration

104 AW .. out of ..

104 TC .. out of 33 .. the rest of them, they needed the oral input. They needed me to read the question to them for them to do the answer. So, when it came to the aTCual SATs paper, the seven at the top end who were able to work on their own, we went into a small room. We started off and I read the first few questions for them to give them a sense of security and then they went on on their own, so they felt they were able to continue, so they did.

Support

109 AW So did you tell then in advance that, ‘I will read the first few with you then you can go on on your own, if you like.’?

110 TC That’s right, and one or two in the middle of the test said, ‘Can I carry on?, and I said, ‘Yes, that’s fine. You carry on.’ After a bit, they were all doing it at their own pace. But that was the only group. The following groups, all in groups of between four and five, I read each question to them.

113 AW Those, I assume, would be your (level) 2c and 2b readers?

114 TC A mixture because although I have got some who have attained quite high levels in reading, they are not confident. They are not confident enough.

115 AW Yes, so it is not only reading ability, it’s a confidence thing as well. They want reassurance that what they are reading is correTC.

117 TC Yes.

118 AW What advantages do you consider there to be in the look and listen method of administration from the children’s point of view?

Look-listen

119 TC Many advantages, because they can then say to me, ‘What did you say? I don’t understand that.’ It (the question) can be repeated, and also, for our children, it is a lot easier for them to listen and to look. They don’t have to have all these words in front of them. I hold the paper up and we look at where we are going. I have got my finger on the aTCual sum that they are doing and they follow the words with me and so, that situation, they feel more confident, they are happier, they are not on their own, and our children do like lots and lots of support. They do like lots of people around them, ‘Am I doing this right? Do I turn over?’

Support

127 AW Which is what happens day to day.

127 TC Absolutely. That is how they are taught. That is what they are used to.

128 AW In terms of management, do you think it is easier, harder or about the same?

129 TC It’s easier because I come out of the class room, the head gives me over a week to do all the SATs papers, not the reading tests, with both key stage 1 and key stage 2. So, for over a week, that is my briefing, solely to be concerned with SATs. It means that the children can come out in their little groups into a nice little room and it is very cosy for them. They don’t feel threatened or anything like that. But as far as management is concerned, as long as the head is prepared to pay for somebody to take my class, which he is, then

Administration

I feel this is the way, the best way of doing it.

136 AW Do you think it is more costly at the moment than when you used to administer it and the children just put their hands up? Did you have small groups then?

137 TC I didn't do it in those days. I have only been doing it for four years now.

139 AW And have you used look and listen strategy ..?

139 TC The first year I did it, having had year 6, and gone through their SATs, the first year we sat and they asked me questions, you know, 'What does this say?' etc., and eventually I got into, 'Well, I'll read the whole thing to you.' So it really snowballed from there.

143 AW So it really started because so many children were asking for so much help? It was easier to do ..

269 TC That's right, and it's easier to have children of similar ability with you as well so you are working at their level.

146 AW Does it help with the pace when you have children of similar ability?

147 TC Yes, you can't do it with a big group – a group of four or five depending on the ability. I think this year I had one child who was waiting for us and all I kept doing was going over it with him, and when you have just got a few children, you can do that.

Look-listen

152 AW Do you think that look and listen should be recommended as the preferred mode of administration? What would your views on that be?

154 TC Absolutely and utterly, I would agree. I have been talking to other year 2 teachers. One in particular this year who'd never done it

156 AW Never done the test or never done look and listen?

157 TC Never done the look and listen. I was telling her, 'Oh, it's much, much better.' and I saw her after the test and she had done it with much bigger groups and she said it was a 100% better and she felt much happier with it as well.

159 AW That's interesting. Thank you.

Transcript analysis of Year 4 children reading item 14

Text from item 14: The numbers in the shaded squares make a sequence. Continue the sequence by shading more squares. (See Appendix 4.1 for artwork.)

Key: A-Y4-G means school A, average reader in Year 4, girl.

Child	Text read (verbatim)	Time taken (secs)	Comments
A-Y4-G	The numbers (2) the numbers in the shaded squares make a [s] (1) make a sequence. (2) Continue the sequence () by shading () more squares.	15	Self-corrected promptly. Some slight hesitancies. Item read accurately with good intonation. Explained clearly what to do.
A-Y4-B	F The numbers () in () the [shade] (2) squares make a sequence. Continue the sequence by shading () more () squares.	14	Some hesitancy. Read with reasonable accuracy but with little expression. ‘Shade’ read for ‘shaded’ (F) but ‘shading’ read correctly in next line. Appeared disinterested. Said that this was key stage 1 maths.

B-Y4-G	The numbers in the shaded squares make a sequence. Continue the sequence by shading more squares.	7	Words read quickly but with good intonation, accuracy and confidence. Gave correct response by stating the numbers she would shade.
B-Y4-B	The numbers (1) the numbers in P the shaded squares () make a [s-] () sequence. Continue the sequence () by shading () more squares.	15	Slight hesitancies but read with good intonation. Explained clearly what to do.
C-Y4-G	F1 The [number] in the shaded F2 squares [makes] a sequence. Continue the sequence by shading more squares.	7	Read confidently and with good intonation. Although she made the subject and verb singular (F1 and F2), she self-corrected herself when explaining what to do.
C-Y4-B	The numbers in the () shaded squares make a sequence. Continue the sequence by shading	9	Read accurately, fluently and confidently with good intonation. Sounded confident. Explained correctly what to do.

	more squares.		
D-Y4-G	F1 F2 The [number] in the [shade] shaded squares (1) make () a () sequence. Continue the sequence by shading more squares.	13	Read with reasonably accuracy and fluency. Made subject singular at F1. Self-corrected error in reading at F2. Explained correctly what to do.
D-Y4-B	F The numbers in the [shade] squares make a [s] () sequence. Continue the sequence () by shading more squares.	8	Read accurately and fluently with good intonation. Sounded confident. Misread F unlikely to affect understanding since ‘shading’ read correctly in next line. Explained correctly what to do.

Appendix 7.2

Transcript analysis of Year 3 children reading item 14

Text from item 14: The numbers in the shaded squares make a sequence. Continue the sequence by shading more squares. (See Appendix 4.1 for artwork.)

Key: A-Y3-G means school A, average reader in Year 3, girl.

Child	Text read (verbatim)	Time taken (secs)	Comments
A-Y3-G	F1 O [Co] The [number] (1) in the ~~shaded~~ S1 S2 F2 [squares] (3) [circles] [makes] a (1) P1 S3 [se – k] (1) a [square] *(She looks to* T *me for help.)*{sequence}. F3 [Con] the se()quence by U1 P2 [shirting] more [s-]() sequences.	41	Started to read second line with number 14 alongside in margin but quickly self-corrected. F1 and F2: could have affected her response since the subject and verb are changed from plural to singular. S2: probably the result of reading S1 (sequence) incorrectly as squares and knowing that the next word could not also be squares. S3: I ask her to have a guess and she reads ‘sequence’ as ‘square’ again. I then told her the word. Could not explain what to do.

A-Y3-B	The numbers in the shaded (1) P1 U1 squares make a [s-sq-] (8) [scumen] U2 (4) [scillient] *(He looks to me for help.)* T {sequence}. U3 U4 [Contin] the [secrence] by shading more squares.	32	Pace and fluency compromised mainly by unfamiliarity with 'sequence' and 'continue' (U1 to U4). 12 seconds spent trying to decode 'sequence'. Even when told 'sequence', failed to read it correctly in following sentence. Did not know what to do.
B-Y3-G	F1 The [number] in the shaded P1 squares make a [s-s-] (1) sequence. Continue the sequence by shading more squares.	10	Slight hesitancies but read with good intonation. Change of subject to singular (F) did not affect her response which was correct. Explained clearly what had to be done.
B-Y3-B	P1 The numbers () in the [shed] shaded () [co] [s] squares make a sequence. Continue the sequence by () shading more squares.	15	Some hesitancies but self-corrected reading errors. Read with reasonably accuracy but lacked fluency. Explained what to do.

C-Y3-G	The numbers in the shaded squares make a sequence. Continue the sequence by shading () more squares.	9	Read fluently and accurately with good intonation. Sounded confident. Explained clearly what to do.
C-Y3-B	Continue the sequence by shading more squares. *(I ask him to start with the top line.)* The numbers in () the shaded squares make () a sequence. Continue the sequence by shading more squares.	12	Read fluently and accurately with good intonation and pace. Starting reading by the line with the question number written alongside. Explained what to do.

D-Y3-G	The numbers in the (1) shaded (1) squares make a (8) (*I intervene* *as she looks to me for help.*) T1 {sequence}. (2) *(She looks to me for help with* *‘continue’ and I ask her to guess.)* Continue the () sequence () by shading more squares.	25	Asks if she has to start with the top line of text. Pace, accuracy and fluency compromised mainly by unfamiliarity with word ‘sequence’ (T). Explained what to do and pointed to numbers to be shaded.
D-Y3-B	The numbers in the shaded S1 squares make [each] () P1 [she-quence]. U1 [Contune] () the sequence by shading more squares.	14	Child points to first line of text and asks if he has to start there. Reasonable pace but accuracy compromised mainly by unfamiliarity with ‘continue’ (U1). Explained what had to be done by naming the numbers to be shaded.

Appendix 7.3

Transcript analysis of Year 2 children reading item 14

Text from item 14:

The numbers in the shaded squares make a sequence.
Continue the sequence by shading more squares.

(See Appendix 4.1 for artwork.)

Key: A-Y2-G means school A, average reader in Year 2, girl.

Child	Text read (verbatim)	Time taken (secs)	Comments
A-Y2-G	The numbers in the (1) [sh-a-d- P1 S1 shad-shad-shad-ed] (2) [corners] S2 () make a [square]. P2 [C-c-o-n-t-in-i-o-e] *(Whispers, trying to say the sounds to make the word.)* S3 T1 [Coin] *(I intervene.)* {Continue} S4 P3 the [square] by () [sha-shade] shading more squares.	48	Made a valiant attempt to read the question but was dependent on the use of phonics (P1 – P3) which was not always appropriate or accurate and slowed down the pace of the reading. Together with the substituted words (S1 – S4), the sense and flow of the text was lost in the struggle to read. Reading was stilted and neither accurate nor fluent. In spite of the time spent trying to read, she indicated within a few seconds the correct numbers to shade by referring to the information on the grid.

A-Y2-B	F1 S1 The numbers the [number] [is] P1 P2 the [s-s-sh-a-d-ee-d] [sheed] P3 [squ-i-ars] squares make a () S2 [square]. P4 S2 P5 (3) [C-con-] [count] the [s-e-qu- S3 S4 squ-] [squares] (1) by [shape] P6 S5 [sh-ape-ing] [shaping] more (1) squares.	53	Read determinedly from start to finish. Heavily dependent on phonics (P1-P4) but could not apply them correctly. In particular, struggled with ‘shaded’, ‘sequence’, ‘continue’ and ‘shading’. Confused between ‘squares’ and ‘sequence’. Explained that he would shade odd and even numbers.
B-Y2-G	F1 S1 P1 The [number] [is] in the [shade-s] *(self-corrected ‘is’ to ‘in’)* F2 P2 S2 squares [made] a [s-] [square] S3 [circle]. S4 (5) Continue the [square] by shading more squares.	37	Read slowly with little expression. Lacking in pace, fluency and accuracy. Read ‘sequence’ as ‘square’ twice (S1-S2) then read ‘square’ as ‘circle’ (S3). Did not know what to do.

B-Y2-B	S1 The numbers () in the [shape] (1) P1 P2 F [sh-sh-] (3) [sh-a-sha-d] [shade] P3 squares make () a () [s] sequence. P4 P5 [C-] (5) [C-con] (5) (*I intervene*.) T P6 {Continue} the [sh-] sequence by P7 P8 (3)[sh-sh-sh-a-d] () [sha-d] shading (*pronounced with a soft ‘a’*) S2 S3 [move] [shapes].	76	Reading lacking in fluency and accuracy but persevered to the end. Heavy dependence on phonics eg various attempts to read ‘shaded’ and ‘shading’ (P1, P2, P7, P8). None of these attempts was successful and the time taken was excessive. The substituted words at the end (S2- S3) further obscured the meaning of the instruction. Did not know what to do.
C-Y2-G	S1 The numbers (1) in the [shad-ow] S2 squares make (1) a [question]. S3 (8) Continue the () [squares] by S4 [shadowing] more squares.	34	Reading slow and stilted. Pace further compromised by unfamiliar word ‘continue’. Substitutions of key words ‘shaded’, ‘sequence’ and ‘shading’ (S1-S4) obscure meaning of instruction. Knew that squares shaded grey were important but did not know what to do.

C-Y2-B	S1 The numbers in the [shape] (2) squares make a (2) sequence. S2 P1 P2 [Count] [c-] continue the [s-] *(self-corrected 'count' to 'continue')* sequence by () shading more squares.	21	Read with reasonable accuracy and good intonation. Fluency compromised mainly by hesitations before 'squares' and 'sequence' and misread of 'continue'. Substituted 'shape' for 'shading' (S1) but read 'shading' correctly in next sentence. Spent several seconds studying number grid then indicated correctly numbers to be shaded.

D-Y2-G	F1 F2 The [number] in the (8) [shade] F3 square [makes] a (9) *(She looks* T1 *to me for help.)* {sequence}. P1 [C-con-cont-cont] *(I intervene.*) T2 {Continue} the (8) *(I intervene.)* T3 {sequence} by (8) shading (*pronounced with a soft ‘a’*) (1) more squares.	77	Pace very slow and text read in a stilted manner. Intonation poor. In particular, attempts to read ‘shaded’, ‘continue’ and ‘sequence’ (F2, T1 and T3) slow down pace. Unable to read ‘continue’ and ‘sequence’ (T1-T3) P1: Using phonics to decode ‘continue’. F1and F3: subject and verb are changed from plural to singular. (Made the same error in question 11.) By referring to the diagram only, she was able to explain what three numbers to shade. This took only a few seconds.

D-Y2-B	S1 The numbers in the [shapes] S2 P squares [are] make a [se-] *(Self-corrects ‘are’ to ‘make’. Says, ‘Miss, I can’t read this sentence’.)* T1 T2 T3 T4 T5 {The numbers in the shaded T6 T7 T8 T9 squares make a sequence.} U1 U2 [Contien] the [cleetras] by U3 [tanning] more squares.	30	Lacking in fluency and accuracy. Pace fast with poor intonation and no apparent awareness in the second sentence that the text is meaningless. Appears to prefer to try to read unfamiliar words as a whole rather than to use phonics (U1-U3). When asked what to do, said, ‘I dunno’.

Text from item 15: There are 4 apples in each pack. (See Appendix 4.1 for artwork.) Mrs Pullen buys 3 packs of apples. How many apples does she buy?

Key: C-Y2-G means school C, average reader in Year 2, girl.

Child	Text read verbatim	Time taken (secs)	Comments
A-Y2-G	F1 [Miss] *(I ask her to start at the top line.)* There are four apples in each pack. F1 P1 F1 [Miss] [Pull-en] [Miss] Pullen (1) buys three packs of apples. How many apples () does () she buy?	23	First line of text missed. Started reading text below diagram. Substitution unlikely to affect mathematics in this instance. Read with reasonable accuracy but lacking in pace. Could not explain what to do and suggested ‘4 apples’ as her answer.

A-Y2-B	F1 F2 [Mister] [Pull] () *(I ask him to start at the top line.)* There are four (2) apples in (2) each F3 [packet]. F1 F1 F2 [Mister] (2) [Mister] [Pull] (2) F1 F2 [Mister] [Pull] (4) *(He makes a blowing out sound as he hesitates.)* S1 P1 [needs] three [p-a-p-a-c] three packs of apples. F3 How many () apples () [did] () she () buy?	35	First line of text missed. Started reading text below diagram. S1 and F1-F3: substitutions unlikely to affect meaning but repetitions showed struggle with reading and slowed progress. P1: resorted to phonics even though ‘pack’ read as ‘packet’ in first line. Lack of fluency, accuracy, hesitations and lack of speed in reading very likely to affect comprehension and to be tiring. Could not explain what to do and suggested ‘zero’ as his answer.

B-Y2-G	Mrs (3) Pullen *(I ask her to start at the top line.)* S1 F1 [They] are four [apple] in each pack. S2 U1 Mrs Pullen (3) [packs] (2) [pooks] three packs of apples. F2 S3 How many apples [do] [you] buy?	34	Read slowly and deliberately. It is not clear why she read ‘packs’ and ‘pooks’ (S2 and U1) since she read ‘packs’ and ‘buy’ correctly elsewhere. Lack of fluency, accuracy, hesitancies and lack of speed in reading very likely to affect comprehension and to be tiring. Suggested ‘4 apples’ as her answer since there were four apples in the picture.
B-Y2-B	There are four apples in each pack. F1 P1 [Miss] (3) [Pull-] Pullen buys three packs of apples. How many () apples () does she buy?	22	Apart from ‘Mrs Pullen’, (F and P) the text was read confidently and with good intonation. Comprehension unlikely to be affected in this instance. Explained incorrectly what to do and suggested ‘one’ as his answer.

C-Y2-G	Mrs (2) *(I ask her to start at the top line.)* There are four apples in each pack. S1 Mrs (6) [Pauline] buys three packs of apples. How many apples does she buy?	29	Read slowly but with reasonably accuracy and fluency. Substituted alternative surname for featured character after 6 second delay. Almost one quarter of the time taken was caused by hesitancies due to an unfamiliar name. Could not explain what to do.
C-Y2-B	There are four apples in each pack. F1 P1 [Miss] [Pull-en] Pullen buys three packs of apples. How many apples does () she buy?	16	Read fluently, with good intonation and reasonable accuracy. Substituted word (F) unlikely to affect comprehension. Explained incorrectly what to do and suggested ‘six’ as his answer.

D-Y2-G	Are there (1) *(self- corrects)* There are four apples in each pack. *(She tries to read next word under her breath.)* F1 P1 [Mister] [Pull-] *(She looks to me for* T1 *assistance.)* {Pullen} Pullen buys F2 three [package] of apples. S1 S2 How many [are] [in] *(self-corrects ‘are in’ to ‘apples’)* apples does she buy?	35	Reading slow and deliberate with poor intonation. Struggled to read with a slow pace. She could not explain what to do but suggested ‘eight’ as her answer.
D-Y2-B	How many apples () oh *(quickly self-corrects)* There are four apples in each pack. F1 (3) [Mr] Pullen buys three packs of apples. F2 How many apples does [he] buy?	16	Read steadily but with reasonable accuracy and fluency. Substituted words (F1-F2) changed sex gender of featured character. Could explain what to do.

Text from item 17:

(See Appendix 4.1 for artwork.)

Ann measured the height of these 2 dolls in blocks.
How many blocks **taller** is the large doll?

Key: C-Y2-B means school C, average reader in Year 2, boy.

Child	Text read verbatim	Time taken (secs)	Comments
A-Y2-G	P1 S1 Ann (2) [m-e-e-s-o-r-e-d [sore] (*I* T1 *intervene.*) {measured} the () height of () these () two () dolls () S2 P2 dolls in () [pack] in [b-l-o-c-k-s] U T2 [bocks] *(I intervene.)* {blocks}. F1 How many blocks (1) [tall] *(self-corrects ‘tall’ to ‘taller)* S3 taller is () the () [longest] doll?	43	Reading very stilted, lacking in sense of meaning, fluency, accuracy and pace throughout. Her valiant attempt to use phonics for two key words (P1 –P2) was ineffective and I felt it necessary to intervene. Thought that she had to give height of ‘longest’ doll in blocks.

A-Y2-B	Ann (2) Ann (3) Ann measured P1 the (2) [hi-e] height of () these () P2 two () dolls in (2) [bl-bl-] (1) blocks. P3 F How many blocks (3) [t-] [tall] P4 (2) taller is the [la-r-] large doll?	35	Showed perseverance and some attempts at self-correction. First word, 'Ann' perhaps repeated twice whilst considering what next word said. Dependence on phonics for four key words. Six silences of two or more seconds whilst considering next word. Thought that he had to give height of taller doll in blocks.
B-Y2-G	U1 F1 [Ane] [measure] () the height of () these two dolls in blocks. How many blocks taller F2 is the [largest] doll?	16	Read with reasonable fluency and accuracy. Three substituted words close in meaning to original words. Explained what to do and gave correct answer.
B-Y2-B	P1 Ann (1) [m-m-me-m-m-m-] F1 P2 [measure] the () [h-heigh-hei-] height of these two () dolls in (3) P3 [bl-] blocks.	54	Reading very stilted, lacking in sense of meaning, fluency, accuracy and pace throughout. Speed, fluency and accuracy compromised mainly by

	P4 How many blocks (7) [t-t-] (2) F2 S1 [tall] is the (2) [longer] doll?		unfamiliar sight words and attempts to use phonics for them (P1 – P4). Meaning of question changed by substitution of ‘tall’ (F2) for ‘taller’ and ‘longer’ (S) for ‘large’. Explained that he would count the blocks for the longer doll.
C-Y2-G	F1 Ann [measures] the height (6) *(She looks to me for reassurance.)* F2 of [those] two dolls (1) in blocks. How many blocks (3) taller S1 is the [long] doll?	31	Read slowly and deliberately. Three substituted words unlikely to affect comprehension. Could not explain what to do.
C-Y2-B	F1 Ann [measure] (1) Ann measured *(self-corrects tense)* F2 the height of [the] these two dolls *(self-corrects ‘the’ for ‘these’)* in blocks. How many blocks taller S1 [in] the large doll?	22	Read fluently, with good intonation and reasonable accuracy. Used self-correction for F1-F2.Minor reading error (S1) did not appear to affect comprehension. Indicated the blocks to be counted to find answer.
D-Y2-G	F P1 P2 [Anna] (8) [mes-as-are] [mes-as-	61	The reading errors, combined with the delays due to

	U1 are] [me-sure] *(pronounced as two separate words)* T1 {measured} the (1) height P3 of these two dolls in [bl-o-cl] *(She looks to me for assistance.)* T2 {blocks}. P4 How many blocks (4) [t-t-] taller is the (2) large doll?		meeting unfamiliar sight words and using phonics were likely to mean that the child would have little or no comprehension of what was being asked. Thought that she had to find the height of the taller doll with a ruler.
D-Y2-B	How many blocks taller is the large doll? *(I ask him to start at the top line.)* P1 P2 Ann [meas-ur] (3) [me-sured] S1 [message] *(looks to me for assistance)* T1 {measured} (*starts again)* Ann measured the height of these two dolls in blocks. How many blocks taller is the large doll?	27	Started to read the second line first, where the position of the question number was placed. Read confidently, accurately and with good intonation apart from unfamiliarity with ‘measured’ (P1, P2 and S). This broke up the fluency of his reading. 10 seconds were spent on trying to read ‘measured’. Explained correctly what had to be done.

High frequency word lists

List 1: Words treated as high frequency for word recognition for analysis of KS1 mathematics test 2001 (arranged alphabetically)

Order	Word	Order	Word	Order	Word	Order	Word	Order	Word
1	a	51	dog	101	July	151	on	201	the
2	about	52	don't	102	jump	152	once	202	them
3	add	53	door	103	June	153	one	203	their
4	after	54	down	104	just	154	or	204	then
5	again	55	draw	105	last	155	orange	205	there
6	all	56	eight	106	laugh	156	order	206	these
7	am	57	eighteen	107	left	157	our	207	they
8	an	58	eleven	108	less	158	out	208	thirteen
9	and	59	end	109	like	159	over	209	this
10	another	60	even	110	little	160	pattern	210	three
11	answer	61	February	111	live(d)	161	people	211	Thursday
12	April	62	fifteen	112	long	162	play	212	tick
13	are	63	fill	113	look	163	pull	213	time
14	as	64	first	114	looked	164	push	214	to
15	at	65	five	115	lots	165	put	215	too
16	August	66	for	116	love	166	ran	216	took
17	away	67	four	117	made	167	red	217	top
18	back	68	fourteen	118	make	168	right	218	tree
19	ball	69	Friday	119	man	169	ring	219	Tuesday
20	be	70	frog	120	many	170	said	220	twelve
21	because	71	from	121	March	171	same	221	twenty
22	bed	72	get	122	match	172	Saturday	222	two
23	been	73	girl	123	may	173	saw	223	up
24	big	74	go	124	May	174	says	224	us
25	black	75	going	125	me	175	school	225	very
26	blue	76	good	126	Monday	176	see	226	want
27	boy	77	got	127	more	177	seen	227	was
28	brother	78	green	128	Mr	178	sell	228	water
29	brown	79	had	129	Mrs	179	September	229	way
30	but	80	half	130	much	180	seven	230	we
31	by	81	has	131	mum	181	seventeen	231	Wednesday
32	call(ed)	82	have	132	must	182	shape	232	week
33	came	83	he	133	my	183	she	233	went
34	can	84	help	134	name	184	short	234	were
35	can't	85	her	135	new	185	should	235	what
36	cat	86	here	136	next	186	sister	236	when
37	children	87	him	137	night	187	six	237	where
38	clock	88	his	138	nine	188	sixteen	238	which
39	colour	89	home	139	nineteen	189	so	239	white
40	come	90	house	140	no	190	some	240	who
41	cost	91	how	141	not	191	sort	241	will
42	could	92	I	142	November	192	start	242	winter
43	count	93	if	143	now	193	stop	243	with
44	cross	94	in	144	number	194	sum	244	work
45	dad	95	into	145	October	195	summer	245	would
46	day	96	is	146	odd	196	Sunday	246	write
47	December	97	it	147	of	197	take	247	wrong
48	did	98	it's	148	off	198	ten	248	yellow
49	dig	99	January	149	oh	199	than	249	yes
50	do	100	join	150	old	200	that	250	you
								251	your

List 2(a): Top 250 words in Stuart et al's (2003) list by descending order of frequency

Rank	Word	Frequency	Rank	Word	Frequency	Rank	Word	Frequency	Rank	Word	Frequency
1	the	17422	**63**	them	671	**125**	who	329	**188**	let	187
2	and	7985	**64**	looked	662	**126**	boy	323	**189**	still	187
3	a	7221	**65**	back	648	**127**	us	319	**190**	say	185
4	to	6019	**66**	very	637	**128**	play	318	**191**	tell	185
5	said	5515	**67**	came	626	**129**	tree	316	**192**	only	184
6	he	4147	**68**	big	609	**130**	well	316	**193**	sat	184
7	I	4060	**69**	Mr	609	**131**	monster	306	**194**	because	183
8	in	3365	**70**	got	608	**132**	more	306	**195**	Tom	183
9	it	3306	**71**	now	608	**133**	know	305	**196**	couldn't	181
10	was	3248	**72**	too	582	**134**	think	304	**197**	there's	181
11	you	3105	**73**	home	579	**135**	I'll	299	**198**	shouted	179
12	of	2818	**74**	it's	574	**136**	wanted	298	**199**	cried	178
13	on	2325	**75**	house	569	**137**	or	297	**200**	last	178
14	they	2259	**76**	your	563	**138**	father	292	**201**	baby	176
15	she	2076	**77**	old	556	**139**	people	292	**202**	soon	176
16	is	2054	**78**	children	546	**140**	again	288	**203**	told	175
17	his	1798	**79**	don't	533	**141**	has	286	**204**	he's	172
18	for	1644	**80**	day	527	**142**	Kipper	278	**205**	please	172
19	with	1603	**81**	Mrs	519	**143**	round	278	**206**	witch	172
20	that	1556	**82**	put	510	**144**	water	277	**207**	another	170
21	at	1504	**83**	says	507	**145**	our	271	**208**	happy	169
22	up	1504	**84**	good	500	**146**	right	270	**209**	comes	168
23	but	1368	**85**	saw	497	**147**	gave	269	**210**	miss	168
24	can	1357	**86**	help	483	**148**	take	267	**211**	night	168
25	we	1347	**87**	just	477	**149**	door	264	**212**	really	166
26	my	1312	**88**	could	476	**150**	that's	264	**213**	girl	165
27	all	1303	**89**	time	468	**151**	long	259	**214**	fast	164
28	had	1244	**90**	did	467	**152**	everyone	258	**215**	work	163
29	her	1206	**91**	from	467	**153**	mouse	257	**216**	Amanda	162
30	out	1201	**92**	dog	466	**154**	eat	255	**217**	car	162
31	went	1088	**93**	want	460	**155**	stop	255	**218**	tea	162
32	not	1078	**94**	over	444	**156**	make	254	**219**	window	162
33	little	1054	**95**	frog	443	**157**	red	253	**220**	any	161
34	what	1052	**96**	away	440	**158**	next	251	**221**	jumped	159
35	no	1041	**97**	about	437	**159**	way	247	**222**	king	159
36	me	1011	**98**	oh	433	**160**	school	245	**223**	sun	159
37	have	1006	**99**	by	432	**161**	bed	241	**224**	gran	158
38	this	978	**100**	off	432	**162**	called	240	**225**	fish	157
39	there	974	**101**	if	431	**163**	new	240	**226**	much	157
40	one	951	**102**	asked	428	**164**	must	238	**227**	inside	155
41	are	945	**103**	ran	418	**165**	never	225	**228**	eyes	154
42	go	936	**104**	toad	413	**166**	chip	222	**229**	May	154
43	so	932	**105**	I'm	411	**167**	things	221	**230**	William	154
44	be	930	**106**	where	409	**168**	been	220	**231**	Thunder	153
45	down	925	**107**	yes	405	**169**	find	218	**232**	blue	152
46	were	893	**108**	bear	403	**170**	three	216	**233**	even	152
47	will	853	**109**	going	395	**171**	thought	213	**234**	looks	152
48	do	845	**110**	mother	391	**172**	box	212	**235**	under	152
49	then	845	**111**	man	390	**173**	magic	212	**236**	Wilf	152
50	like	825	**112**	an	384	**174**	Ben	210	**237**	around	149
51	him	815	**113**	how	382	**175**	other	208	**238**	begin	149
52	come	812	**114**	cat	377	**176**	first	207	**239**	goes	148
53	as	791	**115**	can't	376	**177**	head	207	**240**	I've	148
54	look	785	**116**	their	369	**178**	run	206	**241**	before	147
55	here	770	**117**	Biff	346	**179**	began	204	**242**	moon	147
56	see	769	**118**	didn't	344	**180**	let's	204	**243**	prince	147
57	get	740	**119**	made	344	**181**	something	203	**244**	white	147
58	mum	731	**120**	took	338	**182**	fox	202	**245**	morning	146
59	when	731	**121**	would	335	**183**	why	202	**246**	give	143
60	into	701	**122**	after	333	**184**	found	191	**247**	thing	143
61	some	690	**123**	am	333	**185**	shop	191	**248**	duck	142
62	Dad	671	**124**	two	332	**186**	room	190	**249**	every	142
						187	through	188	**250**	wasn't	140

List 2(b): Top 250 words in Reid's (1989) database ranked by descending order of frequency

Rank	Word	Frequency	Rank	Word	Frequency	Rank	Word	Frequency	Rank	Word	Frequency
1	and	6204	**63**	time	228	**125**	now	100	**188**	opened	62
2	the	6079	**64**	big	226	**126**	long	98	**189**	park	62
3	a	3366	**65**	house	222	**127**	looked	97	**190**	giant	61
4	I	3195	**66**	called	209	**128**	too	97	**191**	gone	61
5	to	2650	**67**	would	205	**129**	thought	94	**192**	room	61
6	was	2153	**68**	dad	200	**130**	by	91	**193**	sister	61
7	it	1689	**69**	their	192	**131**	walk	91	**194**	asked	60
8	he	1569	**70**	has	188	**132**	cat	90	**195**	blue	60
9	we	1453	**71**	can	183	**133**	upon	89	**196**	or	60
10	in	1438	**72**	be	179	**134**	who	87	**197**	outside	60
11	went	1352	**73**	could	179	**135**	way	86	**198**	had	59
12	my	1246	**74**	going	178	**136**	dragon	84	**199**	brother	59
13	they	1065	**75**	bed	176	**137**	red	84	**200**	sleep	59
14	then	1050	**76**	do	176	**138**	round	84	**201**	cave	58
15	on	1049	**77**	after	175	**139**	mummy	83	**202**	trees	58
16	of	1038	**78**	what	172	**140**	well	83	**203**	woke	58
17	said	863	**79**	as	170	**141**	where	83	**204**	never	57
18	had	831	**80**	dog	170	**142**	gave	82	**205**	tried	57
19	is	688	**81**	of	170	**143**	lots	82	**206**	best	56
20	got	670	**82**	see	169	**144**	want	82	**207**	hit	56
21	she	656	**83**	people	168	**145**	friend	81	**208**	dark	56
22	when	646	**84**	two	168	**146**	children	80	**209**	end	56
23	you	611	**85**	come	166	**147**	make	80	**210**	always	55
24	so	607	**86**	our	166	**148**	tea	80	**211**	baby	55
25	there	604	**87**	school	166	**149**	through	80	**212**	boat	55
26	one	581	**88**	once	165	**150**	car	79	**213**	lot	55
27	but	551	**89**	if	159	**151**	another	78	**214**	wood	55
28	me	543	**90**	door	156	**152**	heard	78	**215**	daddy	54
29	up	518	**91**	ran	155	**153**	king	78	**216**	green	54
30	for	508	**92**	no	152	**154**	more	78	**217**	it's	54
31	with	494	**93**	next	151	**155**	playing	78	**218**	lady	54
32	day	489	**94**	took	150	**156**	fire	77	**219**	soon	54
33	out	457	**95**	good	148	**157**	white	77	**220**	fair	53
34	that	457	**96**	an	147	**158**	garden	76	**221**	its	53
35	some	456	**97**	about	143	**159**	nice	76	**222**	men	53
36	go	444	**98**	night	143	**160**	friends	75	**223**	Mr	53
37	his	434	**99**	name	142	**161**	don't	74	**224**	only	53
38	have	433	**100**	made	140	**162**	oh	74	**225**	snowman	52
39	came	423	**101**	tree	136	**163**	take	74	**226**	suddenly	52
40	were	408	**102**	over	135	**164**	hair	73	**227**	wind	52
41	saw	395	**103**	again	134	**165**	three	73	**228**	dinner	5 I
42	all	378	**104**	yes	134	**166**	help	72	**229**	find	5I
43	at	371	**105**	from	133	**167**	here	72	**230**	sad	51
44	her	371	**106**	us	130	**168**	how	72	**231**	run	5I
45	home	349	**107**	boy	129	**169**	played	71	**232**	turned	50
46	not	325	**108**	away	128	**170**	eyes	71	**233**	clothes	49
47	like	323	**109**	this	124	**171**	shop	70	**234**	football	49
48	very	302	**110**	old	122	**172**	balloon	69	**235**	top	49
49	are	294	**111**	found	120	**173**	black	69	**236**	wanted	49
50	get	290	**112**	lived	119	**174**	Christmas	69	**237**	why	49
51	him	285	**113**	play	111	**175**	look	69	**238**	sunflower	48
52	down	282	**114**	girl	110	**176**	eat	68	**239**	around	47
53	back	279	**115**	told	108	**177**	things	68	**240**	bird	47
54	mum	279	**116**	fell	105	**178**	witch	68	**241**	head	47
55	them	279	**117**	morning	105	**179**	something	66	**242**	sea	47
56	because	273	**118**	started	102	**180**	know	65	**243**	thing	47
57	put	244	**119**	other	101	**181**	think	65	**244**	gold	46
58	into	240	**120**	water	101	**182**	give	64	**245**	hole	46
59	will	237	**121**	your	101	**183**	story	63	**246**	walking	46
60	did	236	**122**	am	100	**184**	walked	63	**247**	ever	45
61	man	236	**123**	first	100	**185**	castle	62	**248**	let	41
62	little	233	**124**	just	100	**186**	didn't	62	**249**	lost	45
						187	food	62	**250**	mother	45

Analysis of items 20, 22, 27, 29, 30 and 31 for readability

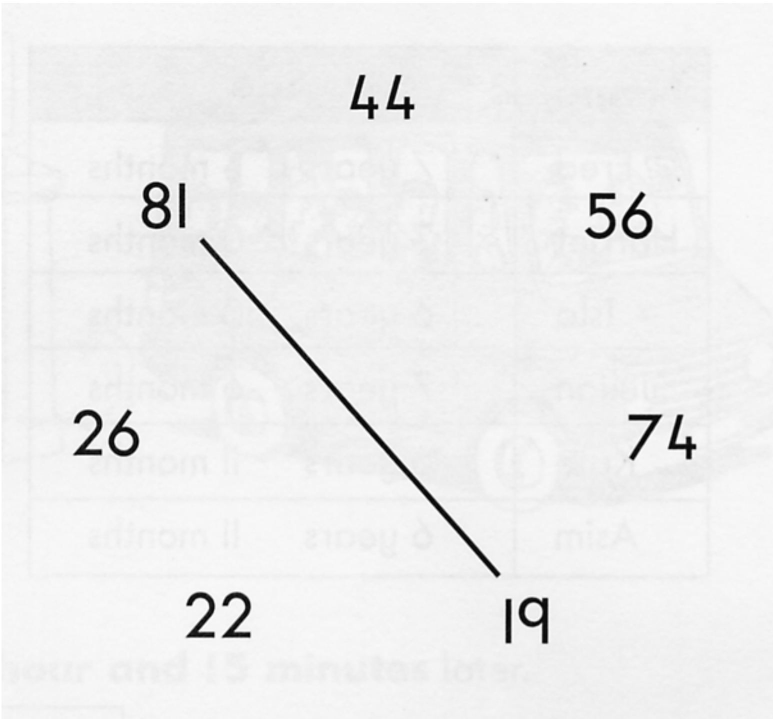
This table shows the ages of some children.

Name	Age	
Fred	7 years	4 months
Harriet	7 years	0 months
Isla	6 years	10 months
Julian	7 years	6 months
Kate	6 years	II months
Asim	6 years	II months

20 How many children are older than Harriet?

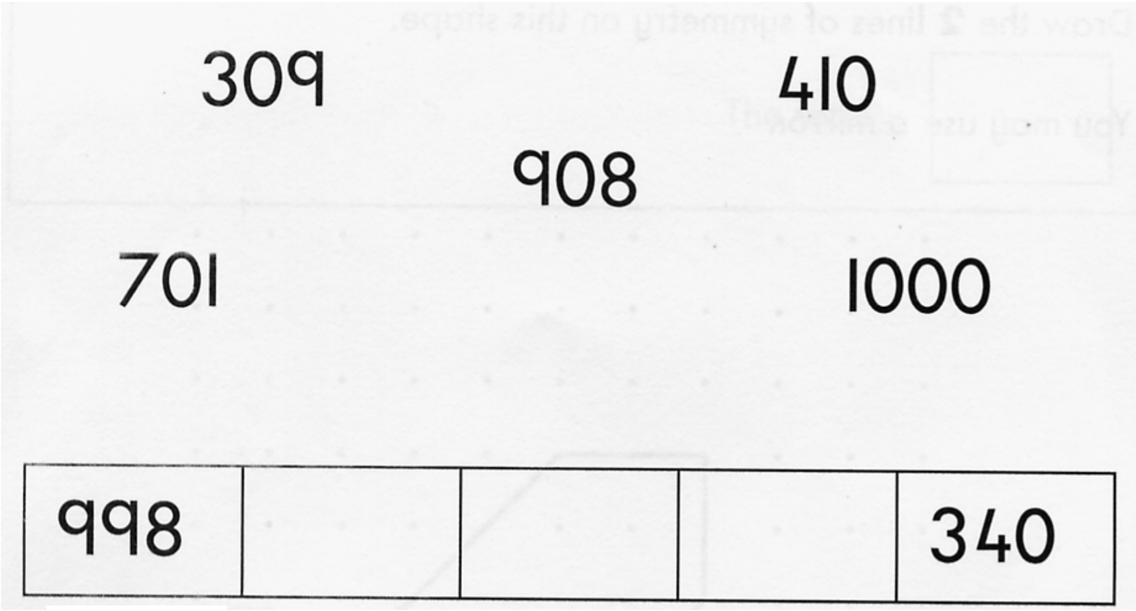
children

22 Join **two** other numbers which total **100**



27 Write **3** of these numbers in the empty boxes.

The numbers in the boxes must be in order.

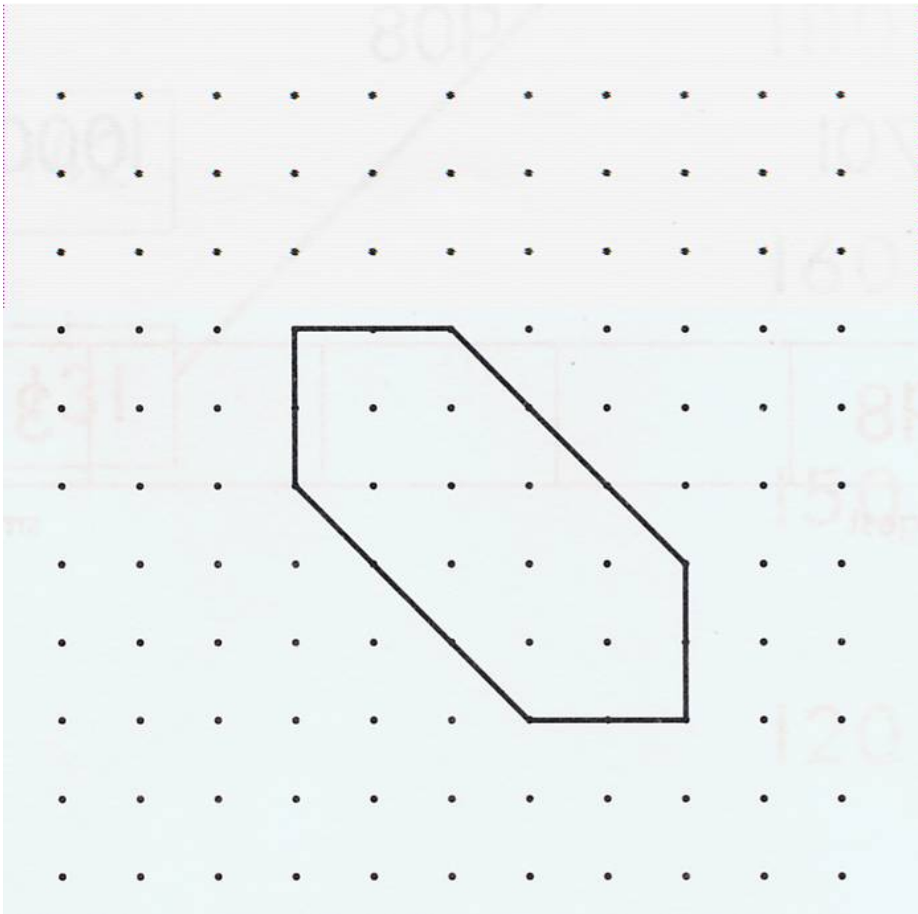


largest

smallest

29 Draw the **2** lines of symmetry on this shape.

You may use a mirror.



30 Add together **24**, **67** and **45**

Show how you work it out in the box.

The total is

Sadi needs **26** cartons of juice for her party.

There are **four** cartons in a pack.

31 How many packs does she need to buy?

packs

Summative analysis of all words in KS1 2001 mathematics booklet (alphabetical order)

Count	Word	W occur	W frequency	G occur	G frequency	Y occur	Y frequency	R occur	R frequency
1	a	1	9	1	9				
2	a								
3	a								
4	a								
5	a								
6	a								
7	a								
8	a								
9	a								
10	add	1	2	1	2				
11	add								
12	after	1	1	1	1				
13	age	1	1					1	1
14	ages	1	1					1	1
15	and	1	3	1	3				
16	and								
17	and								
18	Ann	1	1			1	1		
19	answer	1	1	1	1				
20	answers	1	1	1	1				
21	apples	1	4			1	4		
22	apples								
23	apples								
24	apples								
25	are	1	3	1	3				
26	are								
27	are								
28	around	1	2			1	2		
29	around								
30	arrived	1	1					1	1
31	arrows	1	1					1	1
32	Asim	1	1					1	1
33	at	1	3	1	3				
34	at								
35	at								
36	be	1	2	1	2				
37	be								
38	been	1	1	1	1				
39	belong	1	1					1	1
40	best	1	1			1	1		
41	blocks	1	3					1	3
42	blocks								
43	blocks								
44	blue	1	1	1	1				
45	box	1	7			1	7		
46	box								
47	box								
48	box								
49	box								
50	box								
51	box								
52	boxes	1	3			1	3		
53	boxes								
54	boxes								
55	bus	1	1					1	1
56	buy	1	2					1	2
57	buy								
58	buys	1	1					1	1
59	by	1	1	1	1				
60	card	1	1					1	1
61	cards	1	1					1	1
62	cartons	1	2					1	2
63	cartons								
64	change	1	1					1	1
65	children	1	5	1	5				
66	children								
67	children								
68	children								
69	children								
70	coin	1	1					1	1
71	continue	1	1					1	1
72	corners	1	1					1	1
73	correct	1	5					1	5
74	correct								
75	correct								
76	correct								
77	correct								
78	diagram	1	2					1	2
79	diagram								
80	does	1	2					1	2
81	does								
82	doll	1	1			1	1		
83	dolls	1	1			1	1		
84	draw	1	4	1	4				
85	draw								

Count	Word	W occur	W frequency	G occur	G frequency	Y occur	Y frequency	R occur	R frequency
86	draw								
87	draw								
88	each	1	5					1	5
89	each								
90	each								
91	each								
92	each								
93	Ellen	1	1					1	1
94	empty	1	1					1	1
95	even	1	2	1	2				
96	even								
97	fives	1	1	1	1				
98	for	1	1	1	1				
99	four	1	1	1	1				
100	Fred	1	1					1	1
101	gets	1	1	1	1				
102	go	1	1	1	1				
103	got	1	1	1	1				
104	graph	1	1					1	1
105	half	1	3	1	3				
106	half								
107	half								
108	Harriet	1	2					1	2
109	Harriet								
110	has	1	5	1	5				
111	has								
112	has								
113	has								
114	has								
115	height	1	1					1	1
116	her	1	2	1	2				
117	her								
118	here	1	1	1	1				
119	hexagon	1	1					1	1
120	hour	1	1					1	1
121	how	1	8	1	8				
122	how								
123	how								
124	how								
125	how								
126	how								
127	how								
128	how								
129	in	1	20	1	20				
130	in								
131	in								
132	in								
133	in								
134	in								
135	in								
136	in								
137	in								
138	in								
139	in								
140	in								
141	in								
142	in								
143	in								
144	in								
145	in								
146	in								
147	in								
148	in								
149	is	1	8	1	8				
150	is								
151	is								
152	is								
153	is								
154	is								
155	is								
156	is								
157	Isla	1	1					1	1
158	it	1	4	1	4				
159	it								
160	it								
161	it								
162	join	1	1	1	1				
163	jug	1	3			1	3		
164	jug								
165	jug								
166	juice	1	1					1	1
167	Julian	1	1					1	1
168	Kate	1	1					1	1
169	large	1	1			1	1		
170	largest	1	1					1	1
171	last	1	1	1	1				

Count	Word	W occur	W frequency		G occur	G frequency		Y occur	Y frequency		R occur	R frequency			
172	later	1	1								1	1			
173	left	1	2		1	2									
174	left														
175	less	1	1		1	1									
176	like	1	1		1	1									
177	lines	1	1								1	1			
178	look	1	2		1	2									
179	look														
180	make	1	3		1	3									
181	make														
182	make														
183	many	1	5		1	5									
184	many														
185	many														
186	many														
187	many														
188	match	1	1		1	1									
189	may	1	1		1	1									
190	measured	1	1								1	1			
191	millilitres	1	3								1	3			
192	millilitres														
193	millilitres														
194	minutes	1	1								1	1			
195	mirror	1	1								1	1			
196	missing	1	2								1	2			
197	missing														
198	money	1	2								1	2			
199	money														
200	months	1	6								1	6			
201	months														
202	months														
203	months														
204	months														
205	months														
206	more	1	4		1	4									
207	more														
208	more														
209	more														
210	Mrs	1	1		1	1									
211	much	1	2		1	2									
212	much														
213	must	1	1		1	1									
214	name	1	1		1	1									
215	nearest	1	2								1	2			
216	nearest														
217	need	1	1								1	1			
218	needs	1	1								1	1			
219	note	1	1								1	1			
220	number	1	6		1	6									
221	number														
222	number														
223	number														
224	number														
225	number														
226	numbers	1	9		1	9									
227	numbers														
228	numbers														
229	numbers														
230	numbers														
231	numbers														
232	numbers														
233	numbers														
234	numbers														
235	o'clock	1	1								1	1			
236	odd	1	2		1	2									
237	odd														
238	of	1	9		1	9									
239	of														
240	of														
241	of														
242	of														
243	of														
244	of														
245	of														
246	of														
247	older	1	1					1	1						
248	on	1	2		1	2									
249	on														
250	one	1	1		1	1									
251	order	1	1		1	1									
252	other	1	1					1	1						
253	out	1	2		1	2									
254	out														
255	pack	1	2								1	2			
256	pack														
257	packs	1	3								1	3			

Count	Word	W occur	W frequency	G occur	G frequency	Y occur	Y frequency	R occur	R frequency
258	packs								
259	packs								
260	party	1	1					1	1
261	pours	1	1					1	1
262	Pullen	1	1					1	1
263	Ravi	1	1					1	1
264	red	1	1	1	1				
265	ring	1	2	1	2				
266	ring								
267	Sadi	1	1					1	1
268	sequence	1	2					1	2
269	sequence								
270	set	1	1			1	1		
271	shade	1	1					1	1
272	shaded	1	1					1	1
273	shading	1	1					1	1
274	shape	1	3	1	3				
275	shape								
276	shape								
277	shapes	1	1	1	1				
278	she	1	4	1	4				
279	she								
280	she								
281	she								
282	show	1	2			1	2		
283	show								
284	shows	1	1			1	1		
285	sign	1	1					1	1
286	smallest	1	1					1	1
287	some	1	1	1	1				
288	sort	1	1	1	1				
289	sorted	1	1					1	1
290	sorting	1	1					1	1
291	spends	1	1					1	1
292	squares	1	2					1	2
293	squares								
294	starting	1	1					1	1
295	symmetry	1	1					1	1
296	table	1	1					1	1
297	taller	1	1					1	1
298	than	1	4	1	4				
299	than								
300	than								
301	than								
302	the	1	42	1	42				
303	the								
304	the								
305	the								
306	the								
307	the								
308	the								
309	the								
310	the								
311	the								
312	the								
313	the								
314	the								
315	the								
316	the								
317	the								
318	the								
319	the								
320	the								
321	the								
322	the								
323	the								
324	the								
325	the								
326	the								
327	the								
328	the								
329	the								
330	the								
331	the								
332	the								
333	the								
334	the								
335	the								
336	the								
337	the								
338	the								
339	the								
340	the								
341	the								
342	the								
343	the								

Count	Word	W occur	W frequency	G occur	G frequency	Y occur	Y frequency	R occur	R frequency
344	them	1	1	1	1				
345	there	1	2	1	2				
346	there								
347	these	1	4	1	4				
348	these								
349	these								
350	these								
351	this	1	5	1	5				
352	this								
353	this								
354	this								
355	this								
356	tick	1	1	1	1				
357	time	1	2	1	2				
358	time								
359	to	1	9	1	9				
360	to								
361	to								
362	to								
363	to								
364	to								
365	to								
366	to								
367	to								
368	together	1	1					1	1
369	total	1	4					1	4
370	total								
371	total								
372	total								
373	triangle	1	1					1	1
374	turn	1	1			1	1		
375	two	1	1	1	1				
376	until	1	1					1	1
377	use	1	2					1	2
378	use								
379	water	1	3	1	3				
380	water								
381	water								
382	which	1	4	1	4				
383	which								
384	which								
385	which								
386	who	1	1	1	1				
387	will	1	1	1	1				
388	work	1	1	1	1				
389	write	1	10	1	10				
390	write								
391	write								
392	write								
393	write								
394	write								
395	write								
396	write								
397	write								
398	write								
399	years	1	6					1	6
400	years								
401	years								
402	years								
403	years								
404	years								
405	you	1	2	1	2				
406	you								
407	youngest	1	1					1	1
408	zoo	1	2					1	2
409	zoo								
	Totals:	**169**	**409**	**76**	**261**	**16**	**31**	**77**	**117**