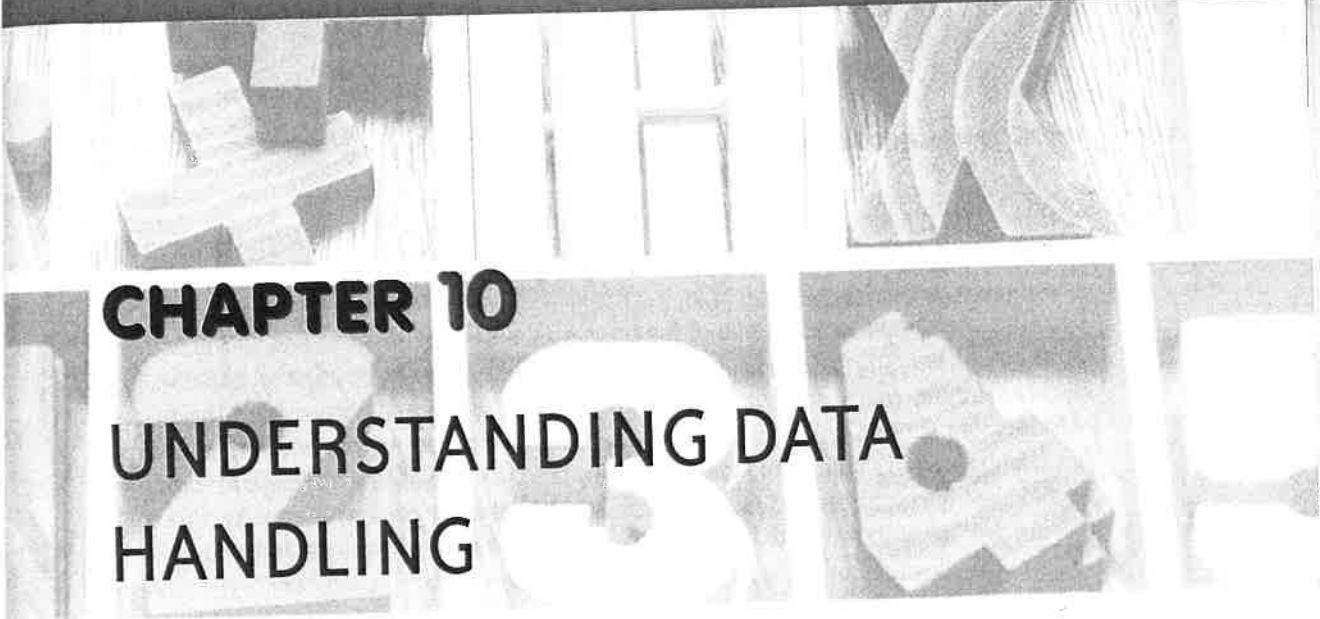




CHAPTER 10

UNDERSTANDING DATA HANDLING

A CONFUSING CATEGORIZATION

I put the children in my class into groups at the beginning of the year – green apples, yellow bananas, and so on – they work in these groups for everything. Jamie never cottoned on to the fact that he was a yellow banana. He was always wandering around lost. Then when we were getting ready for the Christmas play I told all the shepherds to go to one side. Jamie didn't move. 'Go on, Jamie,' I said, 'over there – you're a shepherd'. 'No, I'm not', he said, 'I'm a yellow banana!' (Reception class teacher).

- Think of some other examples of situations when Jamie is likely to get labelled and put into various sets, groups or categories.
- Jamie's problem seems to be the idea of being in two categories simultaneously: being both a shepherd and a yellow banana. What diagram could be connected with this overlapping of categories?

In this chapter

In this chapter we look at some of the basic mathematical ideas involved in the early stages of learning to handle simple statistical data. We start with the process of classification, which is fundamental to all aspects of mathematics – from the first stages of counting and cardinal numbers onwards, and, as we saw in Chapter 9, in making sense of shape. In the lower primary years children will often engage in sorting and classifying objects (including, especially, themselves) according to various criteria. They will learn to record these experiences using a range of pictures and diagrams. From classification into various sets they will move on to considering how many there are in each set. This numerical data can then be organized into frequency tables and presented in simple graphical forms, such as pictograms, block graphs and bar charts.

Meaningful, purposeful and cross-curricular

The work here has the advantage that it is likely to score highly in terms of meaningfulness, purposefulness and cross-curricular links.

Meaningfulness

By 'meaningfulness' we mean that the children are familiar with the context of the tasks in which they are engaged. From their own direct experience they are aware of what is significant, they recognize what is a problem and appreciate what counts as a solution. Clearly, in teaching data-handling there is great scope for the teacher to exploit the child's own direct experience in the questions posed and the data collected – and thereby to make the work meaningful.

Purposefulness

By 'purposefulness' we mean that the tasks have some purpose from the child's perspective. So much of what children do in school has very little genuine and immediate purpose other than to satisfy the demands of the teacher. With this area of mathematics the tasks can have a real purpose, such as: to answer a question; to solve a problem; or to gather, organize and communicate information within an area of enquiry. All our experience suggests that motivation for learning is directly proportional to the degree of purposefulness in the learning tasks.

[When] surveys are related to issues which affect them, they are more interested. Children in one nursery were engaged by the question, 'Do the boys hog the trikes?' and so collected data to investigate this, by tallying boys and girls. (Gifford, 2005, p. 98)

Cross-curricular

Of all the different strands in the mathematics curriculum, handling data is the strand with the greatest potential for linking mathematics with other areas of the curriculum. Science experiments, geography and history topics, performance in physical education, for example, are all sources of data that can be collected and processed in order to address questions, to solve problems and to provide insights. In doing this, the teacher will plan for the children to be involved in all the stages of handling data – being involved in determining what data should be collected; then collecting it and recording it; then organizing it; then representing it in some kind of a graph, and, finally, interpreting it.

An example

For example, as part of a topic on food, one group of 7- and 8-year-olds was involved in all these stages when investigating some of the less familiar fruits available in the supermarket – where they came from, how they got here, what they cost, and so on. They decided to find out which of four exotic fruits were most popular with the children in their class. The children organized a tasting session and recorded in a table each child's favourite fruit from the four options. Next, they produced a tally chart showing how many chose each fruit, converted this into a frequency table and then a block graph (see page 272). Finally they wrote a paragraph summarizing their findings:

The children in our class tasted 4 fruits. These were mango, papaya, star fruit and tamarillo. Most children liked the mango best. We liked the papaya and tamarillo the same. Only one boy chose the star fruit. The mango had 4 more than the tamarillo and the papaya. The star fruit had 6 less than the tamarillo and the papaya. We should tell the supermarket to sell more mangoes and if they keep the star fruit longer it might get sweeter.

Pictorial representation

Making connections

It is not difficult to recognize the significance of this kind of activity in terms of understanding mathematics, using the underlying idea of making connections that has been one of the themes of this book. The mathematics here starts with a real-life situation – in the example above this was the children's tasting of four exotic fruits. The observations are recorded in symbols, in a frequency table, connecting numbers with the real-life situation. This is then transformed into a graph, connecting the symbols with a picture. And the whole activity promotes mathematical language, starting with the articulation

of a question to be investigated and concluding with the implications drawn from the data and the graph. The strong connections between concrete situations, symbols, pictures and language involved here imply that this is a very significant area of activity for developing children's understanding and confidence with number and number relationships. For example, the representation of the data in pictorial form gives a visual image to support such concepts as 'more than', 'less than', 'greatest' and 'least'.

Differences between numerical and pictorial representations

We asked some of the teachers we worked with what they saw as the main differences between presenting some data in purely numerical form and presenting the same data in pictorial form, as shown in Figure 10.1.

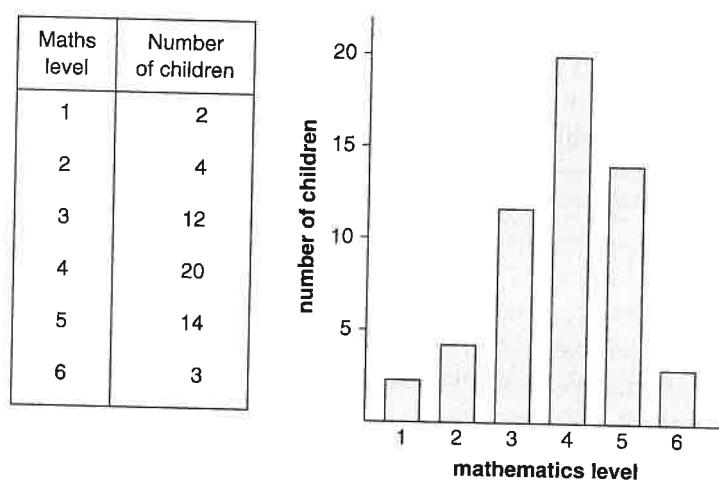


Figure 10.1 Comparing numerical and graphical representation of data

In what ways did they consider one or other format more useful or more helpful? Their responses bring out clearly the significant principles here:

- In the pictorial form the same information is more accessible visually and comparisons like 'bigger than' and 'smaller than' can be made more immediately.
- The bar chart is a more interesting way of presenting the information; it can be understood at a glance.

- You get an instant impression with the graph. It's easier to take in the important facts.
- Level 4 obviously stands out above the others in the bar chart – and it's immediately apparent that there are few children at levels 1 and 6.
- The pictorial representation is better where the trend over the level range is more important than the exact numbers involved.
- It might be difficult to read across if you wanted the exact numbers, so the table would be more accurate.
- The table of numbers is easier to use if you're doing calculations that need specific values – the graph would have to be translated back.

Data in numerical form lends itself to sequential analysis, as you examine the detail of the data bit by bit. You can manipulate numbers, add them up, calculate differences and ratios, find averages and so on. By contrast, the pictorial representation provides an overall, at-a-glance picture of the situation and the relationships involved. The graph provides a global synthesis of all the data, so that you can take it all in – in one go. Some of the detail is harder to get hold of in the picture, but you get an immediate impression of the differences and ratios involved, for example, without any manipulation or calculation.

Both kinds of thinking and representation are significant in doing mathematics. Often, a picture enables us to conceptualize and get an overall view of a problem or a mathematical situation. This gives us some idea of what it's all about and may suggest a strategy for proceeding. Then, actually to solve the problem we usually find ourselves representing it in mathematical symbols (numbers or algebra) that are amenable to manipulation. In the course of problem solving, mathematicians will often move backwards and forwards between pictorial and numerical or algebraic representations – and will be very likely to use pictorial representations to communicate their thinking and conclusions to others. The work that young children do in school in collecting numerical data and representing it in pictorial form is, therefore, an early experience of this moving between two kinds of representation that are central to thinking and working mathematically.

Ways of representing data

In this section we discuss the different forms of representation of data – the different kinds of pictures – that children should learn to use. The examples we use will refer to data collected by a class of 26 children. These examples illustrate the wide range and different types of information that can be collected by children about themselves as the basis for meaningful learning in this area of mathematics.

Representing discrete sets

The first kind of representation we will consider are *set diagrams* – sometimes called *Venn diagrams*. For our purposes in this section we will represent the 26 children in the class by the 26 letters of the alphabet. The most natural way in which we classify children in a class is by putting them into two quite separate sets: boys and girls. Sets that have no members in common like these are called *discrete* sets. They can be represented in a simple set diagram by two enclosed regions that do not overlap, as shown for our 26 children in Figure 10.2(a).

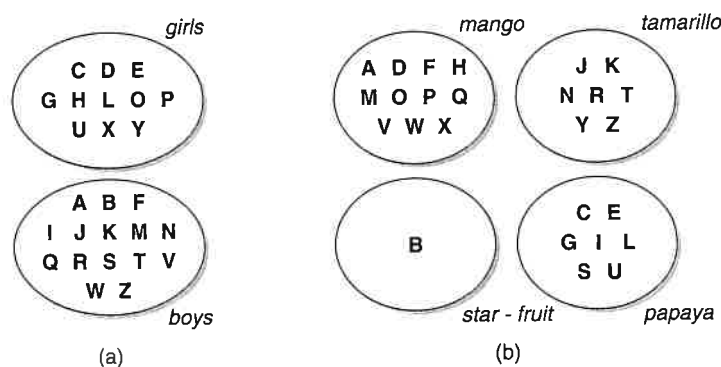


Figure 10.2 Set diagrams showing (a) boys and girls; (b) choice of fruit

This kind of set diagram is the most basic form of pictorial representation of data, showing a simple classification into two sets. The sex of the child is an example of what is called a *discrete variable* – because it is a variable (that is, something that can vary from one member of the class to another) and it separates the class of children into discrete, non-overlapping sets. In Chapter 3 we explain how the union of two discrete sets – putting the two sets together to form one set – is the basis of one structure for addition (see Figure 3.1). So Figure 10.2(a) provides a picture that connects the question 'How many children altogether?' with the addition $11 + 15$.

Often a discrete variable will separate the items available into more than two sets. The children's choice of favourite fruit, for example, separates them into four discrete sets, as shown in Figure 10.2(b). One teacher gave us one of her favourite examples of children collecting discrete data:

I get my Year 2 class to collect the number of letters in the first name of each child. They always enjoy investigating and representing this data. We usually get something like seven discrete sets, ranging from three-letter names to nine-letter names.

Figure 10.3 shows an alternative way of representing data by focusing on just one particular value of the variable in question. To sort the children in Figure 10.3(a) the question asked is, 'Are you a girl? Yes or no?' So the children are categorized as 'girls' (inside the circle) and 'not girls' (outside the circle). In Figure 10.3(b) the question asked is, 'Did you choose mango? Yes or no?' The children are then categorized as 'mango' (inside the circle) and 'not mango' (outside the circle).

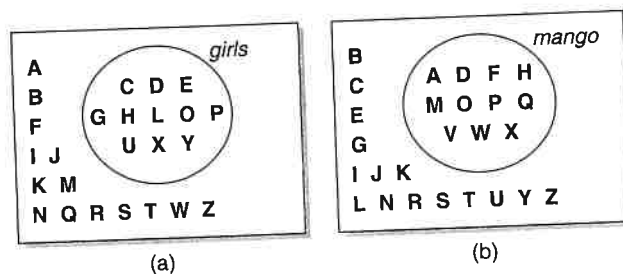


Figure 10.3 Sorting by one value of a variable: (a) girls, (b) choosing mango

The rectangle drawn around the outside in Figure 10.3(a) and (b) represents what is called the *universal set* – this is a mathematical term meaning 'all the things we are considering'. In this case the universal set is the 26 children in the class, all of whom are within the rectangle somewhere. The set of children who are 'not girls' is called the *complement* of the set of girls. In Chapter 3 (see Figure 3.6) we discuss how the complement of a set is one example of a structure represented by subtraction. So, for example, Figure 10.3(a) connects the question 'How many are not girls?' with the subtraction $26 - 11$.

Intersecting sets

Sometimes, of course, sets will overlap, because they have members in common. This is a sophisticated idea that children come to grasp gradually – as is illustrated by the anecdote at the head of the chapter, where Jamie had problems with being both a yellow banana and a shepherd. However, as their experience of classification increases, children will be able to engage in sorting objects (and themselves) according to two variables at a time. For example, they might be interested to show how the split between boys and girls within their class is related to their age in years. Two sets such as 'girls' and '7-year-olds' will intersect, because some girls may be 7-year-olds and some not – and some 7-year-olds may be girls and some not.

A set diagram like that shown in Figure 10.4 provides one way of representing the classification of children by two variables such as those involved here, namely, sex and age in years. One region encloses the set of girls, the other the set of 7-year-olds.

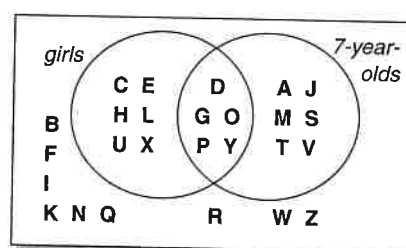


Figure 10.4 Two intersecting sets

The set of children in the overlapping region (the 7-year-old girls) is called the *intersection* of the two sets. Notice that there is a place for every child. The diagram in Figure 10.4 divides the set of 26 children into four subsets:

- those who are girls and 7 years old (the intersection)
- those who are 7 years old but not girls
- those who are girls but not 7 years old
- those who are not girls and not 7 years old.

We should also note that, because the two sets (girls and 7-year-olds) are not discrete, the union of the two sets (all those who are girls or 7 or both) does not correspond to addition: there are eleven girls and eleven 7-year-olds, but there are only 17 in the two sets put together (not 22). So, when we make the connection between addition and the union of two sets we must add the rider that the two sets are discrete, that they do not intersect.

Databases and spreadsheets

In today's technological world all the processing of data can be done quickly and efficiently by a computer. Many examples of simple database and spreadsheet software are now available for children in their earliest years of schooling. In the simplest versions of a database, children enter their answers to various questions and the computer then organizes the data and displays it in an appropriate tabular or visual form. The value of this kind of activity is that the child can focus on the two stages of collecting and interpreting the data, while the computer does the more technical stages of organizing it and representing it. In interpreting the data, children can interrogate the database. For example, having entered the information about sex and age, in some versions of a database children could ask to have listed 'girls and 7' which would produce the intersection of the two sets. Similarly, a request for 'girls or 7' would produce the union of the two sets.

The language of logic

The potential in this kind of classification activity for the development of the basic language of logic is considerable – not just the specific vocabulary of logic (not, all, some, none, or, and, both) but also the characteristic language patterns of logical arguments. For example, in Figure 10.5, one of the subsets from Figure 10.4 has been removed (the girls who are not 7). Once this is done we can make at least these two statements that are true about the children who remain on the diagram:

- 1 All the girls are 7 years old.
- 2 All those who are not 7 are not girls.

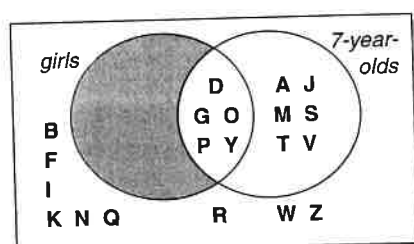


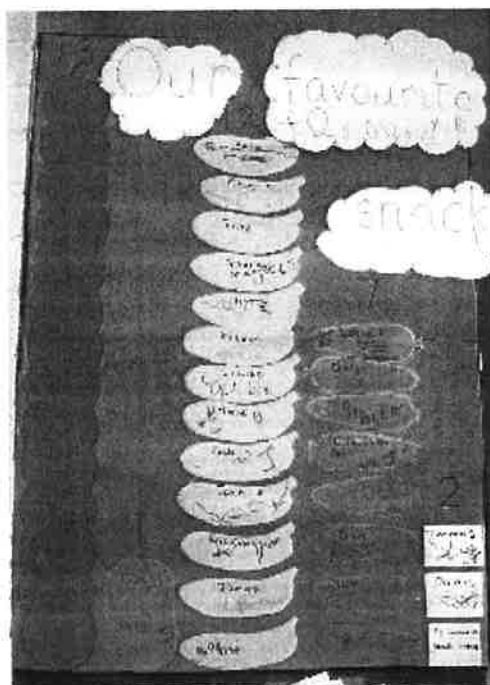
Figure 10.5 Two intersecting sets with one subset removed

You should now pause and convince yourself that these two statements are logically equivalent – they use different words to say the same thing! Similar logical challenges can be posed by removing any one of the subsets in Figure 10.4.

Other ways of representing classifications

There are many other diagrammatic representations of these kinds of classification. Any Foundation Stage or Key Stage 1 classroom is likely to have examples on display. In one classroom we see the names of children displayed in three columns labelled 'school dinner', 'packed lunch' and 'home'. This is a simple – but nevertheless significant – form of pictorial representation of a classification into three sets. In another classroom we see sets of words beginning with the same letter written inside spaces in the shape of that letter. Another teacher has on the wall an arrow-diagram, with the names of the children listed in a column down the left-hand side, with arrows matching the children to their teams.

I got the children in my nursery class to help me find out how many tomatoes, satsumas, bananas, or carrots we needed to provide each morning for snacks. They each chose a picture of their favourite, we put their names on these and assembled them in a pictogram. We counted how many there were in each column and decided which snacks were most popular and least popular. I was surprised at how much children as young as this seemed to understand what was quite a sophisticated mathematical picture. And how much they engaged with it.



As children get older, the kind of network diagram for sorting shown in Figure 9.27 (see Activity 9.3 in Chapter 9) can be used for classification of two intersecting sets. For example, asking the two questions 'Is the child a girl?' and 'Is the child 7 years old?' at the two stages of the sorting process generates the same four subsets as in the set diagram shown in Figure 10.4. Another diagrammatic representation of this kind of classification is what is called a *Carroll diagram*, as shown in Figure 10.6 (named after Lewis Carroll, 1832–98, who was not just the author of the *Alice* books, but also a mathematician and logician). This format has the advantage of making the identification of the four separate subsets clearer and the logic of the word 'not' more explicit. It also introduces the important procedure of two-way reference, which is the basis of the co-ordinate system. In placing

	girls	not girls
7-year-olds	D G O P Y	A J M S T V
not 7-year-olds	C E H L U X	B F I K N Q R W Z

Figure 10.6 Carroll diagram for two intersecting sets

an object or a name in one of the four cells in the Carroll diagram the child must refer simultaneously to both the vertical label and the horizontal label.

Representing frequency

Frequency tables

From just classifying objects and children into various sets and subsets we then move on to pictorial representations of data that show how many there are in each subset. Using the data about choice of fruit, for example, the children counted how many chose each fruit. The number in each of these subsets is called the *frequency*. An important skill in computing the frequency, particularly with a large amount of data, is *tallying* (see Figure 6.7 in Chapter 6) – with one child calling out the values and another doing the tallying. This can then be converted into a *frequency table*, a simple tabulation of results as shown in Table 10.1. Tabulation, because it encourages systematic processing and careful organization of information, is an important mathematical skill that children should be encouraged to develop and use as often as possible. Now we are moving into the realms of statistical data we can refer to the universal set (in this case, all 26 children) as the *population*.

Table 10.1 Frequency table for choice of fruit

Choice of fruit	Number of children (frequency)
Mango	11
Papaya	7
Star fruit	1
Tamarillo	7
Total	26

Moving towards a bar chart

Figure 10.7 shows four possible stages in the development of pictorial representation of data such as that shown in Table 10.1. Prior to this, of course, the data can be shown using the actual objects (or children) themselves. For example, the children making the various choices of fruit can line up in the hall according to their choice, ensuring that the lines are straight and the children equally spaced in each line – like soldiers on parade. Comparisons of the numbers in each subset can be made with the children looking across the rows to answer questions such as ‘How many more chose mango than chose papaya?’

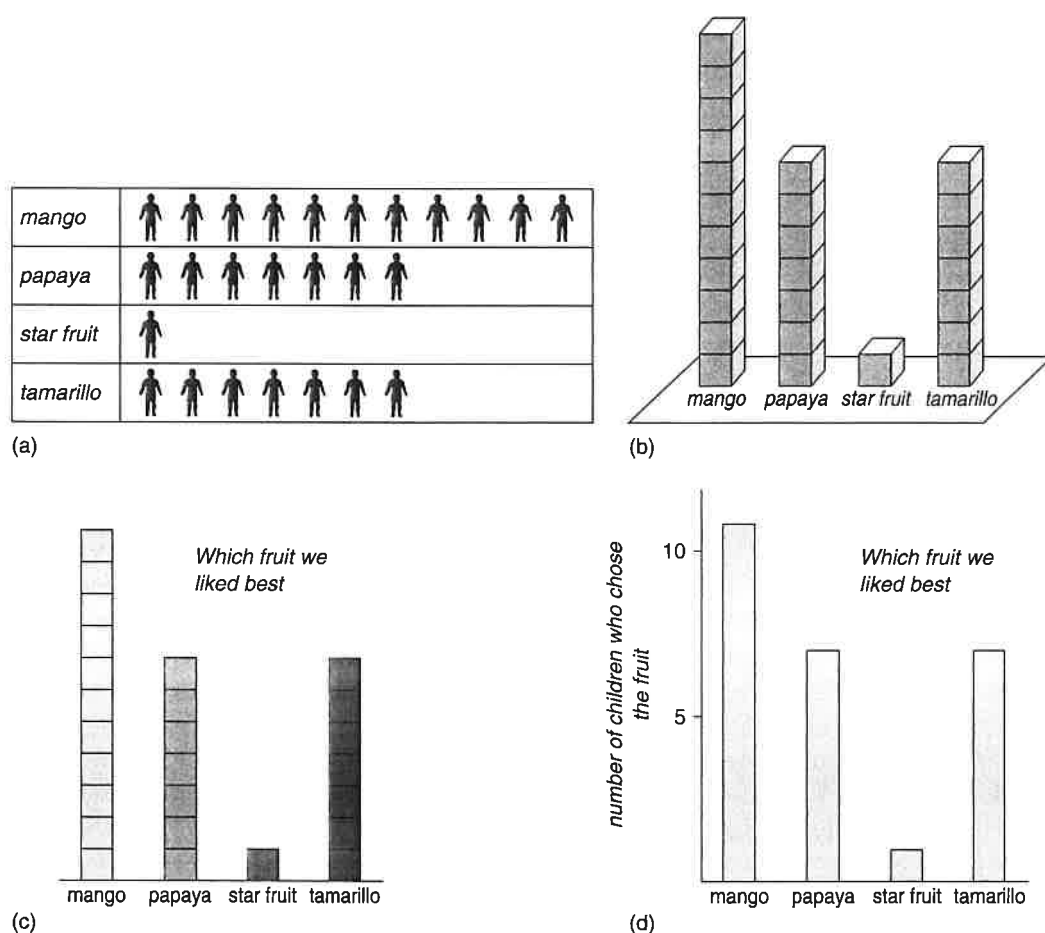


Figure 10.7 Stages in the development of pictorial representation

From using the actual objects or children themselves we move on to using a picture to represent the object or child, as shown in Figure 10.7(a). This kind of representation is called a *pictogram*. The rows of pictures can be horizontal or vertical – but what is important is that they are orderly, with the pictures equally spaced. This can be achieved by children drawing pictures on equal-sized squares of paper. Later on, with a larger population, children will learn to use one picture in a pictogram to represent, say, 10 objects – and then an appropriate part of the picture to represent less than 10. From using a picture to represent each item children can move on to using something more abstract, such as a cube – thus producing a three-dimensional *block graph* as shown in Figure 10.7(b). An alternative to this would be to represent the data with beads on wires or strings. Again, these arrangements could be either horizontal or vertical, although the teacher may wish to start establishing the convention of vertical columns, which is usually associated with block graphs and bar charts.

From these three-dimensional arrangements of solid objects we can move to what is possibly a slightly more abstract representation – colouring individual squares on squared paper, as shown in Figure 10.7(c), producing a two-dimensional block graph. Each square, coloured separately, corresponds to one object or one child in the population in question. At this stage the individual contribution of each child is identifiable – and the number in each subset can be found simply by counting the number of squares.

A significant step is moving from columns of individually coloured squares to the representation in Figure 10.7(d) – the conventional *bar chart*. Now the individual contributions are no longer identifiable and the frequency in each subset is represented by the *heights* of the columns. In order to determine the height of the column a scale must be provided along the vertical *axis*. The following observations from teachers raise an important point:

- I'm never sure whether the labels on the axis should go on the lines or in the spaces.
- Sometimes you see it done one way and sometimes another. Is there any reason for this?
- Sometimes the children label the spaces and then read off the wrong value because they don't know whether it's the number above the line or below the line that they should use.
- They find it tricky learning to label the lines – especially when they're used to labelling the squares.

Note that the labels on the vertical axis in Figure 10.7(d) are correctly located on what would be the lines on the paper – if we were using squared paper – not in the spaces between these lines. This is because we are measuring the heights of the columns, not counting the number of squares in each column. Because of this it is unnecessary and actually unhelpful to get children to draw and label a vertical axis

showing the frequency when they are doing diagrams of the form shown in Figure 10.7(c). Because they are counting the squares here they would tend (rightly) to label the spaces along the axis (that is, the spaces between the lines on their squared paper). This then becomes something they have to unlearn when they move on to representations based on the height of the column rather than the number of squares, when the labels on the axis have to line up with the tops of the columns.

Finally, in this section, we should note again that the actual organization of the data and production of the graphs could nowadays be handed over to a computer, by using an appropriate database or spreadsheet application designed for primary schools. For example, a number of sets of data about the children in the class (age, sex, choice of fruit, number of letters in first name, and so on) could be entered on one database. The information about each child constitutes one *record* and each set of data is called a *field*. The computer can be instructed to display the records for any given field – and appropriate frequency charts and bar charts appear like magic. This allows children to focus on the interpretation of the data rather than on the mechanics of producing the graphs – which can be a very time-consuming process. Having said this, we do think there is considerable value in children having opportunities to draw their own graphs in various ways, in order to develop their understanding of how they work before handing over the job to the technology.

PAUSE TO REFLECT Some chance!

Imagine trying to draw a bar chart to show the frequency of tickets winning various prizes in the National Lottery. In a typical week one ticket might win the top prize of around £2 million. Say this one ticket was represented by a column 1 cm high. The number of tickets that win the next size of prize would be represented by a column about 6 cm high. But how high would you expect the column to be that represents the number of tickets (over thirteen and a half million) that do not win a prize at all? What do you reckon? Higher than the tallest building in the world? (The answer is at the end of the chapter.)

Different kinds of variable

Unordered and ordered discrete data

We explain above that, because they separate the population into discrete sets, variables such as the choice of fruit and the number of letters in the first name are called *discrete variables*. There is, however, a difference between these two examples of

discrete variables. In the case of the choice of fruit the order in which the data is displayed in the frequency table is arbitrary. We might decide to arrange them in alphabetical order, as shown in Figure 10.7, but this is not necessary – we could equally decide to arrange them from smallest frequency to largest frequency, once we had counted up how many in each set, or in some other order. The variable here is non-numerical: it does no more than assign a label (mango, papaya, star fruit or tamarillo) to each individual member of the population. But in the case of the number of letters in the child's name the variable is numerical: each child is associated with a number. So, because of the ordinal aspect of number, it would be perverse – or at least unusual – to display the data with the subsets appearing in anything other than their numerical order, starting with 3-letter names, then 4-letter names, and so on, as shown in Figure 10.8.

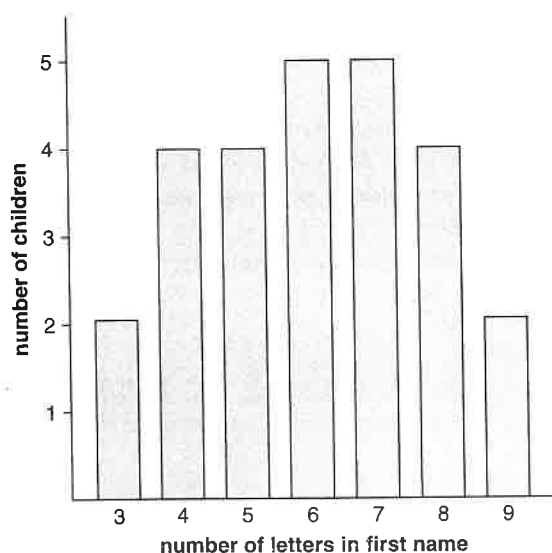


Figure 10.8 Bar chart for ordered discrete data

So, we may sometimes classify by using an unordered, non-numerical, discrete variable, such as your sex, your favourite television programme chosen from a list of six possibilities, your method of transport to school, or the kind of shoes you are wearing. And, at other times, we may classify using an ordered, numerical, discrete variable, such as how many children in your family, how many pets you have, or the size of the shoes you are wearing. (In a few cases of non-numerical data – such as the month of birth – there is also a specific order that would normally be used.)

Grouped discrete data

In one investigation a class collected data about how many pages there were in each child's library book. This data also arises from a discrete, numerical variable. The problem here is that it separates the children's books into too many subsets. The smallest number of pages was 16 and the largest was 87. If we had one subset for each possible value of the variable, from the smallest number of pages in their reading books to the largest we would have 72 subsets, many of which would be empty! This is an example where the data is best *grouped*, so that it can be represented pictorially in a way that makes an appropriate visual impact. We recommend that, normally, grouping should be used when there are potentially more than twelve values of the variable and that it should group the population into between five and twelve subsets. Any less than five subsets and too much information will be lost. Any more than twelve subsets and the graph will contain too much detail to be of any use – you might just as well present the original data. Figure 10.9 shows a bar chart for the data about library books. The data has been grouped in intervals of 10 pages, since this produces eight subsets in total. One principle to be remembered when grouping data is that the intervals used must be equal. We would give a very distorted picture of the distribution in this example if we chose to put, say, children in the range 10–29 in one group and those in the range 30–34 in another.

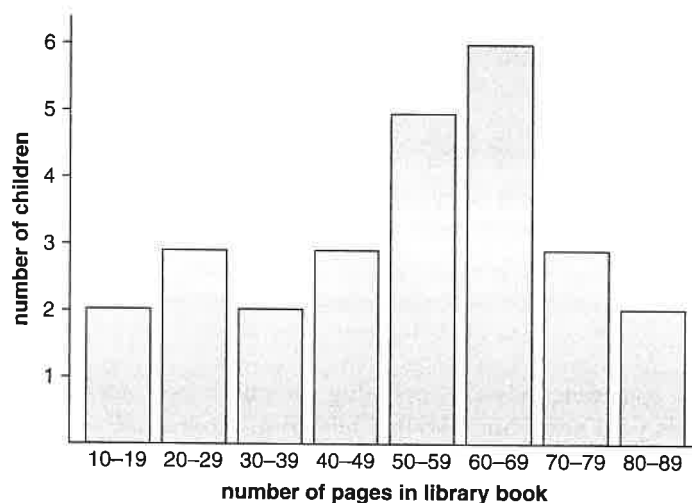


Figure 10.9 Bar chart for grouped discrete data

Continuous data

As part of their work on measurement of length a class of children may collect data about their heights. The height of a child, like any measurement of length, is an example of what is called a *continuous variable*. This differs from a discrete variable in that it is possible for the variable to take any value on a continuum within some range. So, for example, a person's height during their lifetime passes through every real number from their height at birth to the tallest height that they achieve. It does not jump suddenly from 125 cm to 126 cm, for example, but moves continuously through all the heights in between, including 125.1 cm, 125.367 cm, and so on. This kind of variable contrasts with a discrete variable, which can only take particular, separate values and which goes from one value to another in jumps. So, for example, a child might have 7 letters in his or her name or 8 letters, but cannot have 7.456 letters! Most measurements – such as length, mass, capacity, time – are continuous variables. Now, in practice, we always make these measurements 'to the nearest something'. For example, children might measure their heights to the nearest centimetre. This effectively changes the variable to a discrete variable, because now we are only allowing whole numbers of centimetres! So, in practice, the processing of continuous data rounded to the nearest something is no different to that of grouped discrete data. This means that – with an appropriate structure provided by the teacher – the collection, organization, representation and interpretation of continuous data, such as heights and masses, are well within the competence of some children in the age range we are considering in this book. Figure

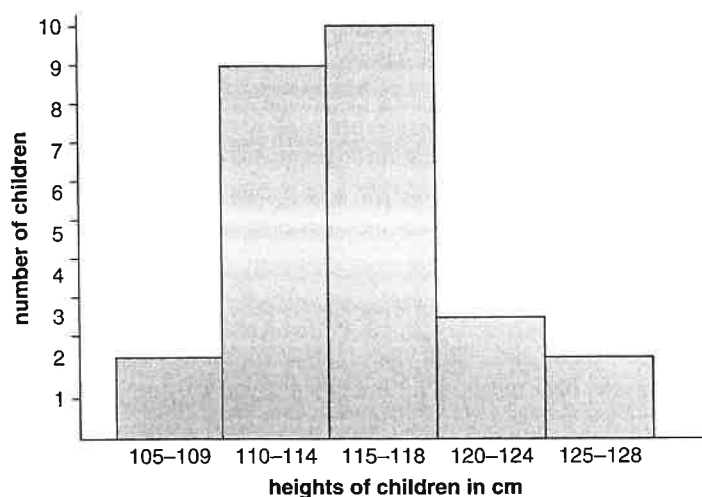


Figure 10.10 Bar chart for data from a continuous variable

10.10 shows an appropriate way of grouping and representing some data about children's heights. This uses 5-cm intervals to produce five subsets.

Notice that in Figure 10.10 we have drawn the columns touching each other, whereas in the previous graphs we have left gaps between them. This is to emphasize that the original variable in Figure 10.10 is a continuous one that moves without gaps from one value to another. When the variable is discrete – that is, when it divides the population into distinct, separate subsets – it is more appropriate to represent the data with a graph that shows the subsets as separated columns with gaps between them.

RESEARCH FOCUS Mark-making and graphicacy

The subject of this chapter has been the ways in which we use pictures in mathematics to represent data. To balance this emphasis on children learning to use pictures in the ways that the curriculum prescribes, we thought it would be useful to reflect a little on the significance of children's own idiosyncratic pictures and scribbles in their early years of engaging with mathematical ideas. This is often referred to as mark-making: a term that has been used to describe the ways in which young children create and use their own marks and pictures to represent and aid their thinking.

In his Foreword to the second edition of Elizabeth Carruthers and Maulfry Worthington's (2006) *Children's Mathematics: Making Marks, Making Meaning* John Matthews writes:

This is one of the most important books on emergent mathematical thought in infancy and early childhood ever written. Those of us who have devoted our lifetimes attempting to understand the origin and development of expressive, representational and symbolic thought in infancy and childhood, and how best to support it, quickly came to realise that the beginnings of linguistic and mathematical thought are embedded in rather commonplace actions and drawings made by the infant and young child. Developmentally, these beginnings are of the most profound importance. They form the child's introduction to semiotic systems without which her life in the symbol-rich society of humans will be dangerous if not impossible. Tragically, these crucial beginnings of expressive, representational and symbolic thought are often discounted completely and receive little or no support from the pedagogical environment. (p. xiii)

This extract is so passionate and, in our view, sadly, accurate, that we thought it important to include it. If you refer back to Figure 1.1 in Chapter 1 you will note that we think pictures are a crucial component to a sound understanding of mathematical concepts, and, indeed, this has been a theme of our book. In Chapter 9, in discussing 2- and 3-dimensional shapes, for example, we made the point that making marks to go with their words is a key way in which young children develop mastery of a concept. And yet, as Matthews so eloquently puts it, this key component of learning mathematics is all too often neglected as children become increasingly absorbed into formal schooling.

Carruthers and Worthington (2006) build on the work of others – most notably Martin Hughes (1986) – to present a wealth of research-based material including insights from their own extensive observations of young children of children's mark-making. Their book goes beyond this, however, to demonstrate – using examples from their own early years' classrooms – how adults can engage with children and encourage them to use both their natural mark-making and the mathematical symbols and pictures they have been taught, such as the number line. Thus, for example, using a combination of her own drawings, numerical jottings and a number line, 7-year-old Ann worked out how many packets of nectarines (3 per pack) would be required for 26 children to have one each.

In their later work, (see, for example: www.Childrens-Mathematics.net/graphicacy.htm) Carruthers and Worthington have suggested that the term 'mark-making' fails to do justice to young children's use of graphics in all forms to record and aid their thinking. They have therefore found that the term *graphicacy* encompasses better all the aspects of visual representations that children use. Aldrich and Sheppard (2000) define graphicacy as:

The ability to understand and present information in the form of sketches, photographs, diagrams, maps, plans, charts, graphics and other non-textual two-dimensional formats... The information can be directly representative of what we see (as in photographs or drawings), or more abstract – for example, information which is spatial (as in maps, plans and diagrams) or numerical (as in tables and graphs). (p. 8)

A useful summary of Carruthers and Worthington's research is provided in their chapter entitled 'Children's mathematical graphics' in Thompson (2008). One of their key points is that

... all young children's mathematical marks and representations are significant: their scribble-marks and early drawings, for example, have as important a role in children's development as later personal strategies they use in their calculations. (p. 147)

A prominent theme in their work is that the child's representation of the *process* of their thinking is more significant in terms of quality and depth of learning than making a record of the *product* of their thinking.

For further research in this area we would refer you to the work of van Oers (2010) in the Netherlands (mentioned in the Research Focus in Chapter 3) and that of Jaslow and Jacobs (2009) in the US.

Some activities to use with children

Children will benefit from plenty of practical experience of posing a question, deciding what data to collect, collecting that data, organizing it, representing it and interpreting it. We repeat here that the best opportunities for developing the skills in handling data are those where the context is meaningful, the task is purposeful and the data is drawn

from other areas of the curriculum. Below we give just three examples of this kind (Activities 10.3, 10.4 and 10.5). The actual activities that a teacher might use may be similar in structure to these, but should be determined by the interests and topics being studied by their class. We also provide two examples of activities (10.1 and 10.2) focused on sorting and logical thinking.



ACTIVITY 10.1 Sorting the Children

3-4	4-5	5-6	6-7	7-8
	☐	☐	☐	

Objective

To develop the language and visual imagery of sorting into sets according to one or more criteria.

Materials

A large area, such as the playground or the school hall, with some means of marking out regions in which the children can stand.

Method

Draw large regions on the playground or lay out ropes on the floor of the hall to produce large-scale versions of set diagrams like those shown in Figures 10.2, 10.3 and 10.4. Then get the class to sort themselves according to various criteria, such as age in years, boy or girl, which village they live in, how they travel to school, whether they stay for school lunch, month of birthday. The best way to do this is to ask a question related to each set to which the answer is clearly yes or no for each child. For example: 'Did you walk to school today?'; 'Is your birthday in a month beginning with a J?'

The most interesting classifications will be those where there are two overlapping sets (as in Figure 10.4) with some children in each of the four regions. For example, there may be some children who walked to school who have a birthday in a J-month, some who walked to school whose birthday is not in a J-month, some who did not walk to school who have a birthday in a J-month, and some who did not walk to school and whose birthday is not in a J-month. With the children sorted in this way, ask each subset in turn to describe themselves and then to sit down.

Repeat all this, sorting the children by using large-scale Carroll diagrams (Figure 10.6).



ACTIVITY 10.2 Negating

3-4	4-5	5-6	6-7	7-8
			☐	☐

Objective

To develop logical thinking and the language of logic associated with classification.

Materials

A set of logic blocks (48 blocks, covering every combination of: red, blue, yellow; triangles, oblongs, squares, circles; large, small; thick, thin).

Method

A group of four children sits around a table with the logic blocks shared equally between them. Each child places one block at random in the centre of the table. The children first agree on a statement that is true about this set of four shapes of the form 'All the ... shapes are ...' For example, if the set consists of the red large thick triangle, the red small thick hexagon, the blue large thin square and the yellow large thin square, they might say, 'All the red shapes are thick.' One child must then select a shape from his or her collection to add to this set to make the statement no longer true. For example, he or she could put down the red large thin triangle. The children must now agree on another statement that is true about the set of five shapes – for example, 'All the square shapes are large' – and then negate this by adding a sixth shape. The activity continues in this way until they are no longer able to make a statement of the required form. The activity is not a competition between the four children. They should co-operate with one another, with the aim being to keep going as long as possible.



ACTIVITY 10.3 Voting

3-4	4-5	5-6	6-7	7-8
			□	□

Objective

To provide purposeful experience of all the stages of handling discrete data in a meaningful context, by collecting data to answer a question.

Materials

Whatever is necessary for producing a good-sized graph (pictogram, block graph, bar chart), appropriate to the skills of the children.

Method

The teacher looks for an opportunity for the class to make a democratic decision on an issue of some concern to the children. For example: 'What shall we sing in the class assembly?' or 'Where shall we go for our class outing?' The question is posed and the children discuss how it might be decided fairly. This could lead to the preparation of a simple voting sheet that is duplicated for completion by each child. This activity works best if the teacher manages the process in some way to ensure that only about six options are available to be voted on. One group of children can then organize and present the data in an appropriate graphical form. The class discusses and interprets the graph, and the decision is made. In some cases it may be possible and appropriate for the children to enter the data on to a computer database and to process it that way.

**ACTIVITY 10.4 Litter Survey**

3-4	4-5	5-6	6-7	7-8
			☐	☐

Objective

To provide purposeful experience of all the stages of handling discrete data in a meaningful context, including tallying, by collecting data to provide information related to a cross-curriculum study.

Materials

Whatever is necessary for producing a good-sized bar chart or pictogram, appropriate to the skills of the children.

Method

As part of their study of the quality of the local environment children might carry out a survey on litter in various roads near the school. Discussion might lead to the question, 'What kinds of things cause the most litter?' A group of children, escorted by an adult, could undertake a preliminary survey to determine the various categories of litter, such as empty drink cans, snack packets, cigarette packets, and so on. We suggest that no more than ten categories should be used, otherwise the data collection will become too cumbersome. The teacher may have to guide the children here and lead them to design an appropriate record form for their survey and explain to them how to fill it in, using tallying. Escorted groups of children would survey their chosen road, recording (but not collecting!) the litter observed. This can then be organized, presented in graphical form and interpreted. In some cases this might even lead to positive action, such as the children writing politely to a local shopkeeper suggesting a litter-bin outside the shop. The completed graphs could form part of a display of work related to the topic. As with the previous activity, a computer database may also be used to organize and process the data.

**ACTIVITY 10.5 Pulse Rates**

3-4	4-5	5-6	6-7	7-8
				☐

Objective

To provide experience of all the stages of data-handling in a meaningful context, by collecting and presenting grouped discrete data arising from a science investigation.

Materials

A seconds-timer; appropriate materials for the production of good-sized bar charts.

Method

Much of the investigative work that young children carry out in science inquiries will generate data that can be organized, presented and interpreted purposefully. For example, in their study of the

human body and the importance of exercise, a standard investigation is to measure and record pulse rates before and after exercise. This generates discrete data from a class of children that is ideal for experience of grouping into appropriately sized intervals. The 'before' and 'after' data should be processed and represented in graphs separately. In each case show the numbers of children with pulse rates in various ranges, such as 50–59, 60–69, 70–79 ... beats per minute. The two graphs can then be presented side by side for comparison, discussion and interpretation. Once again, a computer application might be used to process this data.

Summary of key ideas

- 1 Mathematical activities are meaningful for children if they are familiar with the context and aware of what is significant.
- 2 Activities are purposeful if they have some purpose from the child's perspective.
- 3 Activities related to handling data have the potential for high levels of meaningfulness, purposefulness and for cross-curricular links – particularly if the data is collected to answer a question or to provide information related to an area of interest.
- 4 Stages in handling data include: posing the question, determining what data should be collected, collecting it, organizing it, representing it, interpreting it.
- 5 Mathematical problem-solving often involves moving backwards and forwards between numerical/algebraic representations and pictorial representations of the problem.
- 6 The purpose of pictorial representation of data is to provide an at-a-glance, overall impression of the distribution and the relationships within the data.
- 7 The universal set is the set of all the things being considered (in the contexts of surveys also called the population).
- 8 Discrete sets are those that have no members in common.
- 9 A discrete variable is one that divides the universal set into a number of discrete subsets.
- 10 The union of two sets A and B consists of all those elements that are in A or B (including those that are in both A and B).
- 11 The intersection of two sets A and B consists of those elements that are in both A and B.
- 12 The complement of a set A is the set of all those elements not in the set A.
- 13 The basic language of logic includes: *not, all, some, none, or, and, both*. This is developed by discussion of various diagrams showing the classification of different elements into sets.
- 14 Diagrams to represent classification into two intersecting sets include: set diagrams with two overlapping regions; network diagrams; Carroll diagrams.

(Continued)

(Continued)

- 15 Examples of ways to display pictorially the number of items in each subset (the frequency) are: an orderly arrangement of the objects (or children) themselves; a pictogram; a three-dimensional block graph with one cube representing each object (or child); a two-dimensional block graph, colouring in one square for each object (or child); a bar chart where the heights of the columns represent the numbers in the subsets.
- 16 Computer databases and spreadsheets allow children to focus on the collection and interpretation stages in handling data. Key notions in using databases are records and fields.
- 17 Some discrete variables (such as choice of fruit) are non-numerical and the order in which the subsets are arranged is arbitrary. Others (such as age in years) are numerical and ordered.
- 18 Data arising from numerical, discrete variables with, say, more than twelve possible values should normally be grouped into equal-sized intervals, to produce between five and twelve groups before being represented in pictorial form.
- 19 A continuous variable is one, such as height, which can take all possible values on a continuum over an appropriate range, not just specific values going up in jumps. Since measurements are made to the nearest something (for example, heights to the nearest centimetre), in practice, such data is usually processed in the same way as discrete data.



Suggestions for further reading

Three of Derek's books provide further information on the material discussed in this chapter: Chapter 27 of Haylock (2010) deals with data-handling techniques in more detail; many of the check-ups in Haylock (2001) will help teachers who want to improve their own skills in handling and interpreting data presented in official documents such as government education and Ofsted reports; Haylock with Thangata (2007) has entries on 'Meaningful Learning', 'Purposeful Activity' and 'Cross-Curricular Mathematics'.

Section 4 of Pepperell et al. (2009) provides some good examples of purposeful activities for teaching data-handling in primary classrooms.

For those working with younger children, a chapter entitled, 'What is your favourite colour?' by Jared and Thwaites in Anghileri (1995) is a thoughtful discussion of the early stages of sorting and pictorial representation.

Note: See Pause to Reflect earlier in this chapter.

The column would have to be about 84 miles high, which is about the distance from London to Southampton! The tallest building in the world (at the time of writing), the Burj Khalifa in Dubai, is not much taller than half a mile.