STUDENT UNDERSTANDING OF LINEAR SCALE IN MATHEMATICS: EXPLORING WHAT YEAR 7 AND 8 STUDENTS KNOW

By

Michael Richard Arthur Drake

A thesis
submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy

Victoria University of Wellington

2010

Abstract

This thesis set out to undertake a curriculum review. Scale was chosen as the focus of the review because the Mathematics and Science Taskforce (1997) indicated that measurement was an area of the curriculum that needed priority attention, however comparatively little had been done to provide this by 2005.

Scales are a common mathematical object. A search of many (Western) homes is likely to find a variety of scales being used for different purposes. They can be found on car dashboards, kitchen stoves, tape measures, and clocks. They will be found in graphs, newspapers, and magazines. Scales also underpin important mathematical learning over and above their everyday application to measurement and graphing. For example, concepts like gradient, rate of change, functions, and limits all rely on an understanding of scale. But how much do we know about how students learn to use scales? What does learning about scale involve?

The thesis approaches the review through an exploration of student understanding of scale in the context of mathematics. It focuses on answering the question "what understanding of scales do students in Year 7 and 8 at the case study schools have?" A definition of scale is developed and is based on the mathematics to which the curriculum document indicates Year 7 and 8 students should have been exposed. This could be identified as the notion of a linear scale. Students in Years 7 and 8 were chosen because by that age the mathematics curriculum document implies students should have a good understanding of scale; at the same time, their errors and misconceptions are likely to indicate learning barriers that need to be addressed.

The literature and the New Zealand mathematics curriculum were used to define a construct of scale appropriate to explore with Year 7 and 8 students. Two tests were then developed to measure understanding of that construct. Where possible, items were initially developed or adapted from the literature then, when early findings suggested new avenues for exploration, new items were developed to investigate further issues of understanding. Both tests were used at different schools in the format of a cognitive interview; Test 1 was also used at another school as a written test. Additional items were developed to use with groups of teachers in an attempt to challenge early findings; these teacher trials used a third assessment format requiring both a written answer and a written explanation of method.

In Test 1 students were assessed with pairs of items. In each pair one item used a decontextualised number line, the other a measurement or graphing context promoted by the curriculum. During the cognitive interviews, the verbal responses of students were recorded on audiotape, while field notes were used to capture non-verbal data. Follow-up probe questions were used to clarify solution strategies and the understanding underlying these strategies. The

written test was then used to identify if the interpretations made could be transferred. Test 2 repeated the data collection methodology from Test 1 but used a different structure within the test. In total, 45 cognitive interviews and 81 written tests were undertaken with students, while 32 teachers participated in the teacher trials.

Facilities, point bi-serial correlation coefficients, and levels of significance were used to ascertain the fitness for purpose of the developed test items. Data collected during the cognitive interviews were also analysed using both qualitative and quantitative methods to ascertain patterns of response.

The mathematics curriculum in New Zealand had assumed that students would develop understanding about scale from exposure to scales in measurement and graphing situations. This approach might have been appropriate if a scale was a simple (or single) object to be mastered but it is not. A scale is a mathematical tool of vast flexibility that can be applied in numerous forms to a wide range of situations. The results suggest that teaching of scale needs to be more deliberate, and also needs to be considered when curricula are designed.

A high proportion of the students in the study had not developed the expected understanding of scale by Years 7 and 8. A complex series of factors were identified that impacted on how the students worked with scale. These included: their understanding of number and number symbols; their understanding of the measurement conventions that are foundational to scale; the strategies they had developed to partition unmarked intervals; their strategies to decide on the value of a partition in marked intervals; their understanding of the role of the marks and spaces on the scale; and their ability to iterate a unit. These different bodies of understanding needed to be integrated and used in a coordinated manner for the students to become effective users of scale.

Acknowledgements

I would like to thank the following people, without whom this thesis would not have been possible:

- All those teachers and students who gave so freely of their time to help;
- My supervisors Cedric Hall and Jo Higgins for their patience as they taught me the intricacies of the writing process;
- My wife Di and my son Matthew for all those days when I was ‘missing in study’;
- Susan Kaiser for proof reading; and
- All my friends and colleagues who asked about progress and encouraged me to keep going.

I would also like to thank Victoria University of Wellington College of Education for granting me paid leave to help me complete this thesis.

Contents

List of Tables

List of Figures

Chapter 1
Introduction

This study is about students' understanding of scale. By focusing on the understanding of New Zealand Year 7 and 8 students the study adopts a definition of scale relevant to those students (see Chapter 2), so focuses on scale as it is found in number lines, linear measurement scales, and graphs. The research seeks to discover what students know about this form of scale, and how they come to know it. It also seeks to identify understandings that are difficult to master, and answer the question *why*? The findings from this work are then used to suggest teaching approaches that may help address these issues.

1.1 Background: The state of mathematics education in the 1990s

Several series of reports provide significant evidence of the state of mathematics learning in New Zealand during the 1990s. Two of these series, the *Third International Mathematics and Science Study* (TIMSS) reports (Garden, 1996, 1997), the *National Education Monitoring Project* (NEMP) reports (Crooks & Flockton, 1996, 2000; Flockton & Crooks, 1998), along with *Critical Factors in the Implementation of the New Mathematics Curriculum* (Knight & Meyer, 1996), and the *Report of the Mathematics and Science Taskforce* (Mathematics and Science Taskforce, 1997), all attested to system-wide problems. Their findings were supported by the concerns of New Zealand mathematics researchers at the time (e.g., see Higgins 1999, 2001). Also, the Education Review Office (ERO), in their school reviews, were not able to confirm that the curriculum was being effectively delivered in all parts of the country (e.g., see ERO, 1996, 1997, 1998a, 1998b).

TIMSS was a major international comparative study, not only of student performance, but also of the teachers, schools, and systems within which they learn, so was an important review of the effectiveness of mathematics education in New Zealand. The reports (Garden (1996, 1997) identified that the achievement of Standard 2 children (8 to 9 years old) ranked 18th of the 24 countries involved, while Standard 3 children (9 to 10 years old) ranked 20th of 26; these results placed New Zealand on a par with Cyprus, Greece, and Norway, the lowest of any English speaking country, and significantly below the international mean. Children in Forms 2 and 3 (12 to 14 years old) performed just below international means, although the difference was not of statistical significance (Chamberlain & Walker, 2001). In his conclusion, Garden commented that one of the brightest features of student performance in TIMSS was that relative to other countries, mathematics performance improved markedly between Standards 2 and 3, and Forms 2 and 3. He added that "[i]f New Zealand students could be given a better knowledge and skill base in the early years, they might well perform on a par with the best of the Western countries" (Garden, 1997, p. 252).

The TIMSS reports also talked about teacher knowledge and noted that: "[T]he extent of mathematics and science knowledge of trainee teachers at the primary level remains an area of concern" (Clark & Vere-Jones, cited in Garden, 1997, p. 180). "The number of core hours that all students receive on mathematics in college of education primary teaching programme varies between colleges of education, ranging from 50 hours to 100 hours." These are split over three years and "include instruction in both pedagogy and subject content" (Garden, 1997, p. 43). Garden went on to identify that the most direct influence on student achievement in mathematics is the teacher. When teachers lack confidence because of inadequate knowledge "they are likely to have poor attitudes to the subject and to unwittingly communicate these to the students" and "may avoid teaching the subject" (Garden, 1997, p. 250).

Further comments were made regarding the implementation of the curriculum, where a mismatch between the intended and implemented curriculum was identified. According to Garden (1997), this gap was also identified in the 1981 *Second International Mathematics Study* (SIMS), when a different curriculum was in use. He adds that the materials used by teachers indicate they may not have the resources to "assist them in translating the intended curricula into classroom practice" (p. 247) and that detailed teacher-guide material was needed.

TIMSS results are noteworthy given the methodology, the wide-ranging nature of the study and the detailed analyses of the results. The comparative achievement results were especially worrying because, as Garden pointed out, children in New Zealand generally start school at a younger age, thus New Zealand nine-year-olds "have at least one, and in some cases two, more years of schooling than children in most other countries. Additionally, a relatively high proportion of our children (over 80%) experience some form of pre-school education" (Garden, 1997, p. 247). The comparative results also gain credibility from the fact that they supported the general findings of SIMS carried out in 1981.

> …the mathematics performance of New Zealand students, when compared with their counterparts in the other participating countries, is unsatisfactory at the third form level but satisfactory at the seventh form level. (Garden & Irving, cited in Garden, 1996, p. 23)

The causes of the intended/implemented curriculum mismatch were also explored by Knight and Meyer (1996). This study took the position that the intended curriculum *Mathematics in the New Zealand Curriculum* (MiNZC) (Ministry of Education, 1992) had not been implemented, then explored reasons for this with teachers. They found that for the curriculum to be properly implemented, adequate resourcing plus teacher professional development were needed. The report commented that it is hard to change classroom procedures by simply changing the curriculum, and recommended "that in any future curriculum change the financial implications of successful implementation are acknowledged and that research is undertaken to identify priorities for expenditure” (p. 38).

The Mathematics and Science Taskforce supported many of these findings. Their report identified “[a]n issue of real concern about the acknowledged lack of teacher *knowledge* about mathematics, is the difficulty teachers have in knowing *why* a particular piece of learning early in schooling is important for later development” (Mathematics and Science Taskforce, 1997, footnote p. 11). “Teachers are unlikely to have high expectations of students’ achievement … if their own experience and achievement in, and attitude to, these subjects are not good” (p. 6). The Taskforce also commented on the confidence and knowledge of trainees recruited to teachers’ colleges. The issue of attracting and retaining “trainees with experience and knowledge in mathematics” was raised (p. 4) along with the need to devote “extra effort…to training in mathematics” (p. 6). The report also raised the poor perception of mathematics in the community and the effect of past poor classroom experiences on a new generation of learners: “many parents remember their mathematics learning with distaste. Their negative attitudes … are readily ‘caught’ by their children” (p. 7). The Taskforce suggested that immediate priorities involved helping teachers of 5 to 9-year-olds, and providing teacher resources that exemplify good practice; these resources should be designed for non-mathematics specialists, have a focus on number concepts and skills across all curriculum areas, and be accompanied by school-based teacher professional development provided by trained professionals. After number, the next priority area areas were identified as algebra and measurement.

Assessment evidence regarding the knowledge of New Zealand primary school children is also available from NEMP. Annual samples of 1440 students enrolled in each of Years 4 and 8 are taken (about 2.5% of the population), with students being assessed on a particular set of skills on four-year cycles. The NEMP reports on graphs and tables (Crooks & Flockton, 1996, 2000) and mathematics (Flockton & Crooks, 1998) show, not unexpectedly, lower achievement in low socio-economic areas, and amongst Māori and Pacific Island students.

The picture generated is that New Zealand children under-perform in mathematics generally with particular groups being over-represented among low achievers (Garden, 1996, 1997; Crooks & Flockton, 1996; Mathematics and Science Taskforce, 1997; Higgins, 1999). Concern about the mathematical knowledge of primary teachers, identified in the 1980s, has continued to be of concern (ERO, 2002).

1.2 The Numeracy Project

The government’s main response to these issues has become known as the New Zealand Numeracy Project. Following successful work by the Auckland College of Education based on number knowledge frameworks and delivered through a professional development project focusing on Year 3 students, a national pilot for *Count Me In Too* was established in 2000 (Parsons, 2005). This was essentially the New Zealand trial of the New South Wales *Count Me In Too* project, a professional development initiative based on a learning progression outlining a series of cognitive stages through which students' understanding develops as they become more sophisticated users of number (Thomas & Ward, 2001). This was originally developed from a very successful intervention programme for less able students and was based on the work of Wright (see Wright, 1996, 1998; Wright & Gould, 2000). Table 1.1 below provides a summary of the numeracy projects in New Zealand. Note that the projects are based on an extension of the *Count Me In Too* framework called the *New Zealand Number Framework* (Ministry of Education, 2005a).

Table 1.1: The Numeracy Projects

Project name	**Age group**
Early Numeracy Project (ENP)	Years 1 to 3
Advanced Numeracy Project (ANP)	Years 4 to 6
Numeracy Exploratory Study (NEST)	Years 7 to 10
Intermediate Numeracy Project (INP)	Years 7 and 8
Secondary Numeracy Project (SNP)	Years 9 and 10

The *New Zealand Numeracy Project* has a number of strategic objectives (Parsons, 2005):

1) Improving the mathematics knowledge, skills and confidence of all primary teachers;
2) Improving the mathematics achievement of all students in New Zealand;
3) Improving the mathematics achievement of Māori and Pasifika students; and
4) Building the mathematics education community.

When conceived, the Project had several important design features. Firstly, it aimed to draw on international research into mathematics education, effective teaching, teacher learning, effective professional development, and educational change (Parsons, 2005). Secondly, it was intended to be dynamic and reflective, responding to critique and changing circumstances as well as issues raised within schools and by research (Parsons, 2005). Thirdly, the effectiveness of the project was to be monitored each year by research which considered not only the progress of students, but also the impact on teachers' mathematical content and pedagogical knowledge (e.g., Crooks & Flockton, 2002; Higgins, 2001, 2002; Irwin & Britt, 2005; Irwin & Niederer, 2002; Thomas & Ward, 2001, 2002; Young-Loveridge, 2004). Fourth, the project is developmental, progressively incorporating new understandings of how students and teachers learn mathematics. Also, while number was the starting point for the development, the definition of being numerate "to have the ability and inclination to use mathematics effectively – at home, at work and in the community" (Ministry of Education, 2005a) makes it clear that other mathematics strands are not excluded. For instance, work has been done over the last few years on how algebra and algebraic thinking develops from number (e.g., see Britt & Irwin, 2008), and such an approach to algebra was formally introduced in 2005 to SNP for teachers of Year 9. Work has also started on frameworks for the development of understanding of the other curriculum strands (Ministry of Education, 2004).

By the end of 2005, 17,000 teachers and 460,000 students had been involved with what is now called the Numeracy Development Project (NDP) (Parsons, 2005). Before it is finished, it is expected that all mathematics teachers, and consequently their students, will have benefited from sustained school-based professional development. The focus of the professional development is on both the mathematical content and pedagogical content knowledge of teachers. The number framework has also been incorporated into the revised mathematics curriculum (Ministry of Education, 2007).

1.3 The research context

The background to this research provided in Section 1.1 suggests that in the 1990s student outcomes in mathematics were of concern, and that some of this concern was long-standing. Results from assessments undertaken as part of the Numeracy Project confirm that student understanding of number in Years 4 to 8 continued to be an issue in the early years of the 21st Century (Higgins, 2002, 2003; Higgins, Bonne & Fraser, 2004; Irwin, 2003, 2004; Irwin & Niederer, 2002). When this research was started in 2005, New Zealand was in the early phases of a curriculum review. At that point in time it was clear from student achievement results that the official curriculum document for mathematics was not sufficiently focused on the sorts of achievement that were expected, so it was timely to evaluate that curriculum document in some way. It was also clear from the reports cited earlier in this paragraph that the issues in number and algebra noted by the Mathematics and Science Taskforce (1997) were being identified and attempts were being made to address them through the NDP.

This research was conceived within the context of the curriculum review and the NDP. It was hoped the study's findings could inform the review process, so timelines were developed to permit this. It was further hoped that by focusing on scale this research could form part of the evidence base that would allow the NDP to address the issues related to measurement – the third priority mentioned by the Mathematics and Science Taskforce. Scale was chosen as the focus because scales underpin graphing (both statistical and algebraic) and measurement, both important aspects of the curriculum. Within number itself, scales are found as number lines. These help develop our understanding of fractions, decimals, integers, and the concept of infinity. As such, scales are used widely in mathematics so it is important for students to understand and be able to use them. Additionally, it can be argued that scales can be considered as a bridge between number understanding and essential learning in algebra, measurement, and statistics.

1.4 Research aims

For the above reasons, this research seeks to answer the following question:

- What understanding of scales do students in Year 7 and 8 at the case study schools have?

Answering this question involves analysing the curriculum document to identify what students should be able to do, then comparing these expectations with what students actually know and can do. Common student errors will also be considered as they may indicate where the curriculum is not providing adequate direction or support for teachers. Since scales can be used as a way to represent numbers, students' understanding of scale will be compared to their understanding of number to see if these are linked. Finally, the results of these explorations will be collated and possibilities for addressing any identified lack of understanding will be suggested. Years 7 and 8 have been chosen because by that age the mathematics curriculum document implies students should have a good understanding of scale (see Chapter 4); at the same time, their errors and misconceptions are likely to indicate learning barriers that need to be addressed.

The intended audience for this thesis is mathematics educators. However, scales are also used in other subjects (e.g., technology, the sciences, and the social sciences) so the research may also be of interest to people working in fields other than mathematics education. Such readers would benefit from a familiarity with mathematical language.

1.5 About the researcher

During 20 years of secondary teaching I spent considerable time working with students to identify areas of misunderstanding that were impeding their learning. This led me to create assessments, teaching approaches, and resources to identify and address these. One problem I identified was in the junior secondary school where a large percentage of students had difficulty working with number lines, reading measurement scales, and reading and drawing graphs, especially when non-unit scales were involved. However, with explicit teaching a distinct improvement was often noticeable.

In the last seven years I have been the secondary mathematics adviser for the Wellington region. This involves working with mathematics teachers in the region's schools, providing in-service training, advice, support, and up-to-date information. As part of my work I have been involved with the NDP. This aspect of the job has two major components. The first is as a facilitator of the project in intermediate and secondary schools; and the second is as a developer of the project from an exploratory study to an implemented professional development programme. This has introduced me to the theoretical framework underpinning the project and a considerable body of research about developing mathematical understanding. During this work I have continued to be intrigued by the problems students have with scales and numbers lines, and have discovered this is a recurrent theme in the literature. With my dual background as both teacher and adviser I believe I have the experience needed to explore student understanding of this topic; the work will also be my contribution to the growing understanding that New Zealand teachers have of how students develop understanding of mathematics.

1.6 About this research

This thesis began life as master's research for an MEd. At that stage it involved two research questions; the one reported here on student understanding of scale and a second seeking to identify if a teaching intervention could be developed to improve scale understanding. The magnitude of the project and the nature of the findings led to a PhD upgrade and allowed a tighter focus on scale understanding. This change in focus was primarily due to four factors:

1) It became apparent during the research that scale was not a common subject of enquiry. A research base for scale therefore needed to be constructed. It also meant that the research in this thesis was fundamentally exploratory. Both of these factors had impacts on what could reasonably be covered in the word limit.
2) The need to undertake a second set of interviews to explore a number of unresolved issues relating to scale understanding. These issues were identified while assessing the impact of the original teaching intervention.
3) Methodological problems associated with the teaching intervention beyond the control of the researcher. (One school's mathematics programme was regularly interrupted leading to "about a week of good teaching" instead of the planned three. Enquiries with teachers in other schools found that high levels of disruption were not uncommon – a topic which will in itself need to be followed up by further research.)
4) Identification that the exploration of scale understanding was a major undertaking in its own right.

It is intended that the results of the teaching intervention will be reported elsewhere, although a small emphasis on 'implications for teaching' has been retained in this report, in recognition of the existence of this work and its place in the original design.

1.7 Thesis structure

This thesis contains 13 chapters. Chapters 1 to 6 detail the background, the literature and the design of the research; Chapters 7 and 11 address the development of the test items while Chapters 8, 9, 10, and 12 report the results from the three case study schools and the teacher trials. Chapter 13 uses the main findings to answer the research question before considering implications and making recommendations for further study. The part each chapter plays in reporting this study is outlined in Table 1.2.

Table 1.2: Chapter contents

Chapter	Title	Purpose
1	Introduction.	• Sets the scene.
2	What is scale?	• Introduces what this thesis is about – scale.
3	Scale: the literature in review.	• Identifies what the literature says are the issues for students as they develop an understanding of scale.
4	Document analysis.	• Identifies what the students should know and be able to do.
5	Methodology: General research design.	• Addresses general methodological considerations.
6	Methodology: Instrument design and interview protocols.	• Addresses issues associated with developing a valid and reliable diagnostic tool.
7	The development of the items for Test 1.	• Outlines the development of the items in the first version of the diagnostic assessment.
8	Exploring student understanding: Case 1 – the cognitive interviews.	• Reports the trial of the initial diagnostic assessment on scale. • Explores the nature of students' understanding of scale.
9	Exploring understanding: The written tests for teachers.	• Reports the testing of early hypotheses on groups of teachers.
10	Exploring student understanding: Case 2 – the written test.	• Explores the understanding of a second group of students from a different school. • Explores the value of the diagnostic assessment as a written test.
11	Further item development for the cognitive interviews.	• Outlines the development of the items used in the second version of the diagnostic assessment on scale.
12	Exploring student understanding: Case 4 – the cognitive interviews.	• Further explores student understanding with a third school. • Seeks to answer questions about student understanding arising from the work at School 2.
13	Concluding comments.	• Summarises the research findings and establishes a theoretical framework from which the findings can be viewed. • Answers the research question, considers implications and makes recommendations for 'next steps'.

As the number framework did for the teaching of number, I hope this research will provide mathematics educators and researchers with a better understanding of the issues that students face when developing an understanding of scale. I further hope that, while this thesis has been written in the context of mathematics education and has not set out to address issues of inequity of access to learning, the identification of student understanding of scale (and possible learning gaps) may have a positive impact on the development of the mathematical understanding of all students.

Chapter 2

What is Scale?

Chapter 1 signalled that a particular definition of scale has been adopted for this thesis. This chapter discusses how that definition was developed. The discussion demonstrates that there is more than one concept that is called a scale, making it critical to begin the study with a clear definition of what is meant when the term is used. The discussion also demonstrates that scales come in a variety of forms and levels of sophistication, which make scale complex mathematically, both to learn and to study. The chapter opens by focusing on scale concepts, sketching their complexity while identifying some of the widespread uses of the term. Next, a definition (appropriate to Year 7 and 8 students) is developed, the language used for the features of scale is clarified, and the construct is explored. Finally, the chapter revisits the research question and outlines the sub-questions to be answered.

2.1 Introduction: Common usage

In modern (Western) society, scales are met by most people in daily life. For instance, in a weather forecast, the Beaufort scale can be used to describe wind strength. The strength of earthquakes is reported using the Richter scale. Social surveys regularly use rating scales, such as the Likert scale, to gather information about consumer opinions. A quick glance through any daily newspaper will find graphs with scales, while a search through many homes will locate common measurement tools like analogue watches and clocks, weighing scales, tape measures, height charts, and rulers, as well as objects that are based on scales like maps and plans.

Scales are also used in many occupations. Architects use a scale when developing floor plans for houses, nurses use them to monitor pulse rates and temperatures, stockbrokers use them to track the performance of shares, while judges carefully balance the scales of justice. In both the physical and social sciences, masses of numerical data are routinely simplified for audiences by graphing, thus showing trends and relationships at a glance. Scales are also used by mathematicians and statisticians. For example, applied mathematicians may use both logarithmic and linear scales to identify relationships between sets of data. Mathematicians describe enlargements by using a scale factor. The concept of a function includes understanding a graphical representation, while important algebra is based on scales through the graphical display of related sets of data – including rates of change limits, differentiation, and integration. The widespread use and the variety of definitions implied by these examples of scale pose a challenge for curriculum designers, and to those teaching scale.

2.2 What is a scale?

So what is a scale, how is it defined? Section 2.1 suggests that the term has several meanings within mathematics; however, *The Penguin Dictionary of Mathematics* (Nelson, 2003) has no definition of *scale*, nor an entry under *measurement scale* or *scales of measurement*. The online mathematics dictionaries at http://www.mathpropress.com/glossary/glossary.html#S and http://mathworld.wolfram.com also have no definition. There is little help from the mathematics curriculum document (MiNZC) (Ministry of Education, 1992), it too has no definition in its glossary; the closest reference is in the definition of a dot plot: "represents outcomes as dots on a scale" (p. 212). The associated diagram (Figure 2.1) can be used to infer that the scale includes the line, the marks, and the numbers.

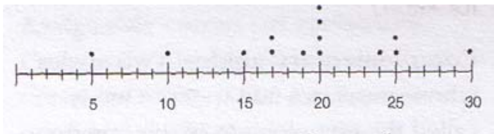


Figure 2.1. Dot plot from the MiNZC glossary of terms

Another possible source of clarification is Stevens' (1946) *On the theory of scales of measurement.* He proposed a classification system "to describe the nature of information contained within numbers assigned to objects" (Wikipedia, 2008b). The four types of scale Stevens identified *nominal*, *ordinal*, *interval*, and *ratio,* are still in common use in the study of the social sciences, including education. These types relate to the form of the variable in the measurement scale and are a property of the unit (Leinhardt, Zaslavsky, & Stein, 1990). They also have implications for which statistics are appropriate for the data (Stevens, 1946). For example, nominal categories like male or female are descriptions that do not imply any sense of order, so one cannot be considered greater or less than the other, and calculating a mean has no statistical value. In an ordinal scale (e.g., a Likert scale) there is a natural order, but if numbers are involved these act as categories in that the intervals between each number are not constant as the meaning of each number can vary from person to person. Interval and ratio scales conform more closely to scales in measurement. In an interval scale the numbers are equal distances from one another as well as ordered. For example, temperatures and time are commonly given examples of interval scales, and illustrate another characteristic of interval scales – an arbitrary zero (a zero chosen for the sake of convenience). Such scales are fundamentally additive in nature; that is why it makes sense to say that 20°C is 10° more than 10°C, it does not really make sense to say that 20°C is twice as hot as 10°C. By contrast, ratio scales have a natural or absolute zero as well as equal intervals between the numbers. With a ratio scale it does make sense to say that a piece of wood 20cm long is twice as long as a 10cm piece of wood. Overall in this usage, the term scale has a specialised meaning, seeming to be the result of a measurement process and appearing to be the term used to describe the set of outcomes generated by this process.

In the absence of a mathematical definition, by returning to some of the examples from common usage, it is possible to infer what is meant by the term scale; however, this is not consistent from situation to situation. For example, the *Beaufort scale* is a series of descriptors that originally related qualitative wind conditions to effects on the sails of a 'man of war'. Each descriptor was given a number, 0 to 12 (Wikipedia, 2008a), which implies it is the numbers that form the scale. In comparison, a *balance scale* does not need numbers, and a map scale (like a *scale factor*) can be a ratio. Observations of other measuring instruments indicate that the term *measurement scale* can be inferred to mean the numbers and marks shown on a measuring instrument (which may or may not be placed on a line). Meanwhile observation of *graph scales* suggests that a scale is the numbers shown with (and the marks shown on) the axis of a graph. Clearly the term scale is used widely but with little consistency.

Fortunately, there are some sources that do provide mathematical definitions of scale, which are useful for developing a particular definition for use in this thesis. *The Concise Oxford Dictionary* (Fowler & Fowler, 1964) defines a scale as "a simple balance … or weighing instrument" (p. 1121); the ratio of reduction or enlargement on a map; to "represent in dimensions proportional to the actual ones"; and the "[s]et of marks at measured distances on a line for use in measuring or making proportional reductions & enlargements, rule determining interval between these, piece of metal etc. or apparatus on which they are marked" (p. 1122). *A mathematics dictionary for kids* (Eather, 2010) also offers four different mathematical definitions:

- (Measurement) scale – the set of calibrated marks on a measuring instrument;
- Scale (or scales) – weighing devices;
- Scale (in size) – to reduce or enlarge; and
- (Number) scale – the basis for a number system, e.g., the binary scale or the decimal scale.

For the purposes of this thesis, the definitions a "[s]et of marks at measured distances on a line for use in measuring or making proportional reductions & enlargements" (Fowler & Fowler, 1964) and "the set of calibrated marks on a measuring instrument" (Eather, 2010) have been adopted as starting points for defining scale.

2.2.1 Scales or number lines?

The confusing term scale could possibly be replaced with using the phrase *number line* instead, as it can be argued that this term is sometimes used as a synonym. For example, Ernest (1985) identifies that graph axes (statistical or algebraic) and scales on measurement instruments can all be considered forms of number line. *The Penguin dictionary of mathematics* (Nelson, 2003) also has an explanation of a number line (under real number):

> There is a one-to-one correspondence between the set of real numbers and the points of an infinite directed line containing a fixed origin. The positive real

> number +*a* corresponds to the point whose distance from the origin is *a* units measured in the positive direction, and the negative number -*b* corresponds to the point *b* units measured in the negative direction. The number 0 corresponds to the origin. In this context the line is called a number line. (p. 360)

However, MiNZC (Ministry of Education, 1992) appears to make a distinction between a scale and a number line. It identifies that students need to learn mathematics through "applying concepts and skills in interesting and realistic contexts which are personally meaningful to them. Thus mathematics is best taught by helping students to solve problems drawn from their own experience" (p. 11). This is reinforced by stating that students should work through achievement objectives "within a range of meaningful contexts". This distinction can also be found in the literature. For example, Leinhardt et al. (1990), in reviewing the literature on functions, graphs, and graphing, identify that the situation a problem is placed in can be more or less *contextualised* or *abstract*, with an abstract situation being a *naked context* (Monk, 1989, cited in Leinhardt et al.). Abstract situations often relate to developing an understanding of the more formal and abstract concepts of mathematics.

The complexity and ambiguity in the language surrounding scale reflects the challenge of trying to define scale. Fortunately, that Leinhardt et al. identify that a situation can be described as either contextualised or abstract seems to offer a way forward, especially as MiNZC emphasises the importance of using meaningful contexts for learning. It suggests that scale be considered the underlying mathematical concept as it is the scale (rather than the number line) that can be more contextual or abstract. Situations involve scales, and a number line (like a measurement scale or a graph axis) is one of those forms of scale. This approach is consistent with the definition of a number line from *The Penguin Dictionary of Mathematics* (Nelson, 2003) which clearly indicates that the number line is a way of representing numbers. It is also consistent with some of the literature. For example, Goldin and Shteingold (2001) indicate that both the real number line and the Cartesian plane are both representations – the former being a way to represent numbers and the latter a way to depict a data set, a function, or the solutions set of an algebraic equation.

The distinction found between situations involving a context and abstract applications of scale can be represented pictorially (Figure 2.2). The diagram clarifies that for this thesis number lines with no real-world context are a form of scale, but need to be considered separately from scales based on contexts that are (presumably) meaningful for the student. However, as noted by Goldin and Shteingold (2001), adopting this distinction requires that graphs based on the Cartesian plane and without a context also be considered abstract. The next section seeks to further clarify the language of scale, and sets out to define the terms related to a scale that are used in this thesis.

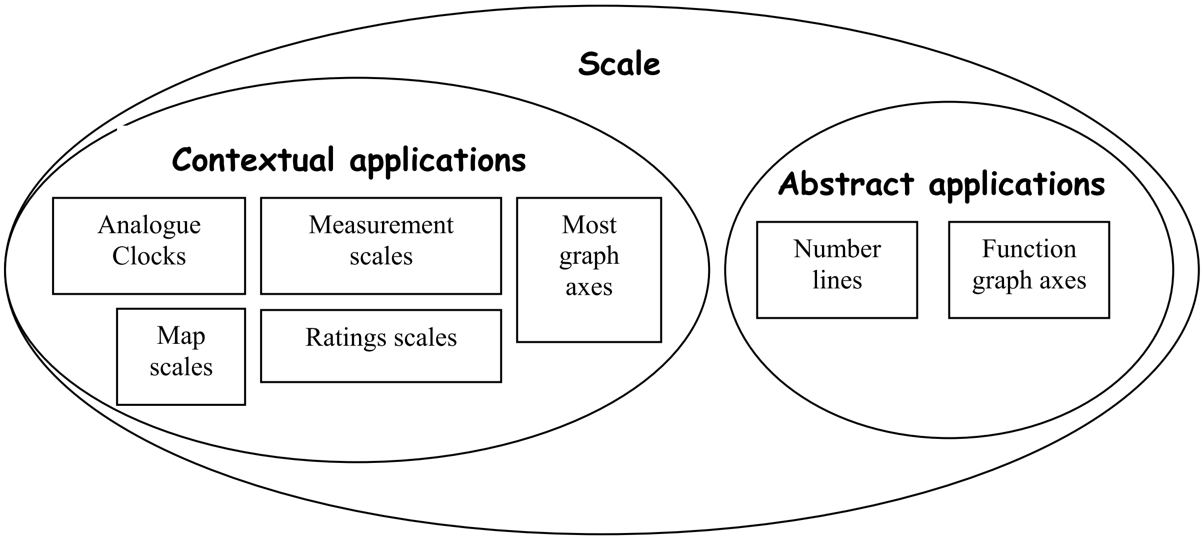


Figure 2.2. Types of scale

2.2.2 The language of scale

Building from Fowler and Fowler (1964) and the analysis of the term scale in common usage, the definition of scale developed for this thesis is: *the set of numbers and marks placed at measured distances on a line, whether it is part of a contextual or abstract application*. This definition provides direction to the study and clarifies that the numbers, the marks, the line, and the interval size all need to be attended to. It further identifies that both contextual and abstract applications are part of the concept, while recognising the distinction about context made by MiNZC (Ministry of Education, 1992). Following common usage, the term *scales* (for example, as in measurement scales) will be used when discussing a type (or types) of scale, although the definition remains the same when using either *scale* or *scales*.

The concept of a scale will also be referred to as a *construct.* This term is consistent with literature (e.g., Goldin and Shteingold, 2001) and is appropriate as it highlights that a scale is a (mental) model *constructed by the human mind*. As a construct, a scale can be compared to other mathematical constructs. A scale can also be *represented* in different ways, as previously shown in Figure 2.2. T*he number lin*e is how we represent scale when working in the number strand of the curriculum; it is useful for emphasising certain aspects of number. Note that in this usage, the definite article invokes the notion of a number line. Typically this involves a marked and numbered line on which any set of numbers can be shown (e.g., Figure 2.3), although an *empty number line* starts out as a simple line on which numbers have yet to be marked (Figure 2.5). By comparison, the term *graph axis* is the name given to a scale in the context of graphing while *measurement scale* is used when discussing the context of measurement.

Finally it is important to note that a term related to scale, *scaling*, is used in the literature (for example: Kaput, 1992; Leinhardt et al., 1990; Rangecroft, 1994). Scaling relates to the development and use of appropriate scales for a graph. Scaling tasks require particular attention to the graph axes and their scales, but also to the unit that is being used (Leinhardt et al., 1990).

2.2.3 Features of scale

Now that the concept of a scale has been defined, the features of scale to be discussed in this thesis need to be outlined. Figure 2.3 shows a number line with a *linear scale*, meaning both the numbers and marks are equally spaced. Such a scale is usually numbered with a *skip count* (e.g., 0, 5, 10, 15, …), with *marks* (*tick marks* or *graduations* in some literature) placed where these numbers are located. In this case the scale involves multiples of five, so could be called a *multiples scale*. If the numbers went up in ones it would be called a *unit scale*. Each mark on the number line is *labelled,* showing the number located there. Between the marks are gaps; these will be referred to as *intervals* or *spaces*.

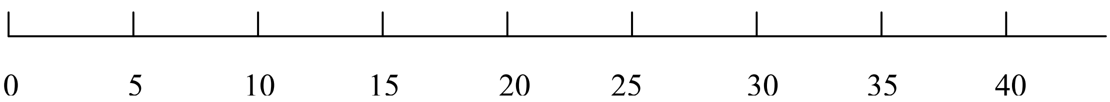


Figure 2.3. A number line with scale in multiples of five

Number lines do not all have the same format. Figure 2.4 shows a number line with *a scale in fives with unit marks,* while Figure 2.5 shows an empty number line with three identified *points* (numbers) named A, B, and C. The location of these points on the number line is shown by the use of *arrows*. In many measuring devices an *arrow* (or *pointer*) is used to indicate where to read from the scale.

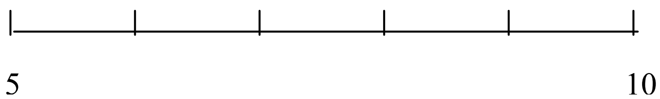


Figure 2.4. A number line with a scale in fives and unit marks

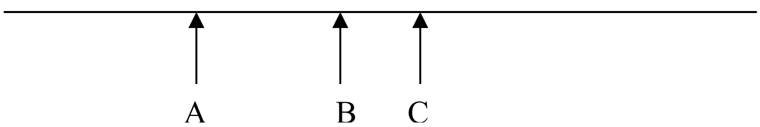


Figure 2.5. An empty number line

Graphs are similarly varied in format, and as shown in Figure 2.2 can be found in both contextual and abstract situations. Figure 2.6 shows a bar graph from a context (minus the *surface features* of title and axis labels) where the horizontal axis uses the convention of placing the bars *between the marks* on the axis. The vertical axis uses the measurement convention of placing the numbers *on the marks* to show a *multiples of twenty scale* (or a *scale in multiples of 20*). *Gridlines* are used to make reading the graph easier. Other graphs may use horizontal and vertical gridlines or *gridpaper*. Gridpaper is commonly used with *algebraic graphs;* that is, graphs based on the Cartesian system with no real life context – an abstract application of scale.

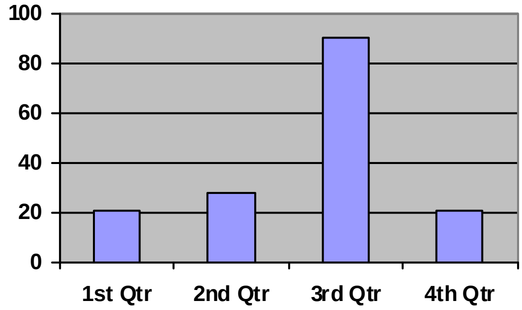


Figure 2.6. A bar graph showing different conventions on each axis

In summary then, a scale is something that does not seem to have a clear mathematical definition, rather it is a term that is commonly used on the assumption that everyone knows what is being talked about. For the purpose of this thesis a particular definition has been developed. The features of scale have also been defined so a language has been established to enable discussion.

2.2.4 How should we classify a scale?

A scale has already been identified as a mathematical construct that is represented in different ways in different mathematical domains. It is now time to consider other possible classifications of scale. As an example of how a scale might be classified, a number line has been classed as a representation (Lesh & Doerr, 2000), a mathematical structure (Freudenthal, 1983), an inscription (Lehrer, Schauble, Carpenter, & Penner, 2000) and, in the context of fractions on the number line, an interpretation (Kieran, 1976), a model (Hart, 1989), and a subconstruct (Behr, Lesh, Post, & Silver, 1983).

On one level a scale can be classed as a *mathematical tool*, since it is something that is used to undertake mathematical tasks. The scale is a powerful mathematical tool as it has many different uses, being applied in different ways to the concept of number and in measurement and graphing. In addition, however, it also needs to be remembered that a scale can be used to represent numbers, so it must also be considered a *representational object*. That is, it is something used to represent the numbers in the number system. In the next section the impact of choosing a particular definition of scale will be considered from the viewpoint of the mathematics that Year 7 and 8 students should be able to do by using a scale.

2.3 Issues in studying scale

It is important to consider the impact of adopting a definition of scale from a mathematical viewpoint. For example, it can have an impact on the learning of students; too narrow a definition can be responsible for omitting mathematical concepts and issues that students need to address, while too wide a definition can over-complicate learning by introducing issues beyond the understanding of the age group being taught. So does the use of the chosen definition enable major mathematical concepts about scale to be addressed, and can it provide a foundation upon which more sophisticated mathematical concepts can be built? This section seeks to answer this question, as well as clarify which sorts of scales and mathematical domains are most relevant for a study of Year 7 and 8 students' understanding of scale. Here the term *domain* is used in the sense of a specific content area of mathematics, like measurement or graphing, as used by Leinhardt et al. (1990).

Firstly, the chosen definition suggests that scales based on discrete data (such as the Likert and Beaufort scales; ordinal scales according to Stevens, 1946) do not form part of the study. In such scales each number represents a qualitative statement, so the provision of a line, marks, and an interval are not important. The exclusion of these sorts of scale does not compromise the study as little mathematics related to such scales is generated or used by students in Years 7 and 8. In addition, logarithmic and other non-linear scales should be excluded as these are not met until well beyond the curriculum levels to which students in Years 7 and 8 are expected to be exposed (Ministry of Education, 1992).

As this thesis explores the understanding of Year 7 and 8 students, the focus is on linear scale. However, even within linear scales it is important to make a distinction. As a mathematical tool, a number line can be used to solve certain number-based problems. For example, Adjiage and Pluvinage (2007) review the effectiveness of a teaching intervention that uses number lines to solve problems like $7 \div 4$. A double number line is also a tool which can be used to solve various ratio problems relevant to the more able students in the age group. Here it is important to clarify that this thesis is about *understanding* scale – rather than ascertaining whether or not students can follow certain procedures that utilise scale to answer problems.

Secondly, the definition allows for a scale to be represented in both contextual and abstract applications. For example, it can be represented not only as a number line, a graph axis or as a measurement scale like a ruler, but also by a number stick (Figure 10.9), or a rope or an empty number line (Figure 2.5). However, such simplistic representations do not conform to the definition of scale until some numbers are located upon them. This limitation again has little impact on the mathematics to which students are exposed; however, the concept of a variety of possible representations does allow consideration of the issue of recognising the construct of scale as it is used in different situations. It also enables the following question to be asked; 'can students recognise a scale in whatever form it is represented, and apply to it their understanding of how the construct is structured?' This is an important question because MiNZC (Ministry of Education, 1992) emphasises the need to use meaningful contexts when teaching. So, is it true that students can read the speedometer of a car travelling at 34kph more easily than a number line showing a similar number but no context? How students respond to abstract scales and scales in contexts may challenge the heavy emphasis on contexts advocated by MiNZC. The responses may also suggest particular types of learning activity for those with poor understanding.

Thirdly, use of the definition raises the question of whether students in Years 7 and 8 have an understanding of scale that reflects the definition. Do the students understand the properties of the construct? That is, do they understand the conventions, notational systems, and concepts upon which a scale is based? Common errors and misconceptions could be particularly useful in exploring this. They may also indicate learning that can help develop understanding. These properties are explored below by considering how we can interpret scales that conform to the definition from a mathematical viewpoint.

2.3.1 Scale as the linear depiction of number

One interpretation of a scale is as the *linear depiction of number*. By using both a line and numbers, a scale is an object that highlights the natural order of numbers. It also allows us to picture numbers that may otherwise be meaningless (so is visual), can provide a sense of the relative size of numbers, and can allow comparison between numbers.

As a linear depiction of number, the number line shares many of the conventions of linear measurement. For example, both number lines and measurement scales are built around the conventions that the unit is repeated (*iterated*), with the numbers located on the marks, rather than in the spaces. Zero, however is a point of difference: measurement starts at zero, so a measurement scale needs to mark it, whereas a number line does not always show zero (although if needed it can be found – Figure 2.7). The number line is also a rigid structure, whereas linear measurement scales can undergo certain *flexions* (acceptable transformations) (Freudenthal, 1983), like being curved to form the scale of a voltmeter.

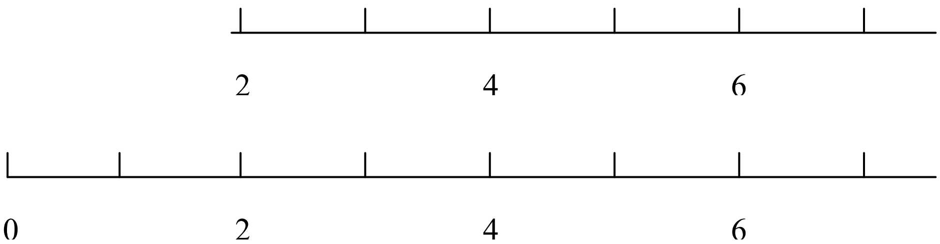


Figure 2.7. A number line fragment, and the use of its conventions to locate zero

Another flexion that separates the number line from the linear measurement scale is the stretch. In linear measurement, the unit is of a fixed size, and is often standardised. On a number line (or indeed a graph axis), the amount of space chosen to represent a unit is arbitrary, so can be stretched to fit a given situation. This stretching concept has some commonality with the scales on certain other forms of measuring instrument, such as an oven temperature dial, where the reduction in scale is such that individual units have been compacted to the point they are no longer shown.

The number line also shares many of the conventions of graph axes; for example, both graph axes and number lines are rigid. Some axes (particularly those used in algebra to show relationships) share with number lines the concept of an iterated unit, have a directional convention whereby numbers get bigger in certain directions, and locate numbers on marks. However, other graph axes use different conventions. For example, the pictogram and bar chart – the graphs to which students are first introduced in MiNZC (Ministry of Education, 1992) – initially follow different conventions; the horizontal axis generally shows categories (nominal data according to Stevens, 1946) that are placed between the marks. In addition, the vertical aspect tends to show stacks of discrete objects, or a vertical axis with numbers placed between the marks (e.g., see Tipler & Catley, 1998).

2.3.2 Scale as the continuous depiction of number

Another concept associated with the linear depiction of number is that scale is the *continuous depiction of number*. This depiction of number encourages us to ask questions like, 'what goes after the biggest number we can think of?', helping develop a visual picture of the infinite. Another important question that it allows us to ask is 'what goes in the gaps between the (counting) numbers?' The ability to identify that fractions and decimals belong in each of these gaps is critical for students to develop if they are to understand the continuity of the number system. Consequently, in any study of scale it is important to examine if students are proficient in locating numbers like 11/4 and 3.2. The concept of numbers being continuous also helps us explain the convention we have of placing a mark to show where a particular number is found and why units are iterated (because each number has a unique location in the number order, labelling an interval rather than a mark identifies an infinite set of numbers, while incorrectly placing a mark invokes a different number).

Furthermore, scale as the continuous depiction of number enables us to consider negative numbers. Once again the development of the construct relies on the continuation of a convention, that smaller numbers are placed to the left of bigger numbers. This makes the number line a useful representation for establishing and explaining the relative size of integers, which from personal experience many students find counter-intuitive.

2.3.3 Discussion

The above exploration of the definition has identified that the study of scale is not straightforward and involves a number of mathematical domains. The exploration also showed that the definition incorporates major mathematical concepts such as linear measurement, fractions, and scaled representations (graphs) all of which are the subject of research and debate. It identified several issues related to the research question, such as: students' understanding of the conventions of a scale; whether students find working with scales in supposedly familiar contexts easier than working with number lines; and factors involved when representing whole numbers, fractions, decimals, and integers on a scale. However, the special case of the analogue clock and how it fits with the definition and other applications of scale now needs to be considered.

According to the scale definition, the analogue clock is a context in which scale is applied, so should be included in this study. However, the clock is a highly specialised tool which utilises scale in a unique manner. For example, a clock may use the numbers 1 to 12 or 1 to 24 to represent hours, but can also use Roman numerals, or dots. Minutes are only shown by the position of the *large hand*, so students need to learn the conversion factor between hours and minutes – the large hand on 1 means 5 (minutes) etc. The language for partitioning the hour is unusual, using *quarter to* rather than *three quarters past* the hour and students are expected to remember the locations of the large hand that signifies the quarters and half. There is also no zero on a clock face. Such numerous differences suggest that studying the measurement of time with an analogue clock may provide more insight into issues with learning about time through this instrument than it does about scale understanding. Therefore the analogue clock is excluded from this study and is suggested as the subject of a separate study.

With the definition of scale now established, the next section revisits the research question, and outlines the process through which this question will be answered.

2.4 Review of the research question

For the reader's convenience, the research question is restated:

- What understanding of scale do Year 7 and 8 students in the case study schools have?

To answer this question, the research needs to identify the understanding that the Year 7 and 8 students in the case study schools should have of the construct, and the understanding they actually have, particularly in relation to issues identified above and in the research literature. This involves answering the following sub-questions.

- What *should* Year 7 students know about and be able to do with scales?
- How do some commonly available mathematics resources present the learning of scale?
- What do students in the case study schools actually know about scale?
- What issues of understanding do the students have? and
- What does this research suggest as ways of addressing these issues?

The first sub-question requires an analysis of the curriculum, so is at the heart of the curriculum evaluation. The second sub-question seeks to explore how students' understanding of scale tends to be developed in classrooms, on the assumption that the available resources have an impact on teachers' approaches to teaching and learning (Leinhardt et al., 1990). Exploring this sub-question will provide a sense of fit between the *intended curriculum* (in MiNZC) and the *delivered curriculum* (what is actually taught in the classroom). The third sub-question requires the development of an assessment tool to determine, for example, if students understand the conventions of a scale, whether scales in contexts are easier to work with than number lines (as implied by MiNZC) (Ministry of Education, 1992), and how students work out where numbers are located on scales. Answering this question provides a measure of fit between the intended understanding that students should develop by following a programme of work based on the curriculum and the actual understanding of the students. The fourth and fifth sub-questions require consideration of student errors, and the implications these may have for changing the curriculum and for teaching; there is little point identifying problems if there is nothing that can be done about them.

In the process of answering the research question, various assessment items were developed. As this research involves an iterative process, student responses have the potential to suggest other avenues of exploration which may help make explicit the understanding a particular student has of scale. For now, the next stage is to explore the literature relating to scale to identify appropriate starting points for this study. This is undertaken in Chapter 3.

Chapter 3

Scale: The Literature in Review

3.1 Introduction

The chapter reviews the literature on scale. As Leinhardt et al. (1990) identify for the study of function, the notion of a scale draws upon a number of different mathematical domains and needs to integrate apparently disparate threads into a unified notion. Thus, like a study of function, tracing the roots of scale understanding involves visiting each mathematical domain. In this chapter, Sections 3.2 to 3.5 outline the background issues related to student understanding of scale. Sections 3.2 and 3.3 consider the debate in mathematics surrounding the use of contexts, word problems, and real practical problems. Sections 3.4 and 3.5 introduce the concepts of a representation and a schema. Section 3.6 explores the research related to the use of number lines as an object upon which whole numbers, decimals, fractions, and integers can be represented. Sections 3.7 and 3.8 cover the literature on the application of scales to measurement, algebraic graphing, and statistical graphing, while Section 3.9 considers differences between the processes of counting and measuring. The chapter closes with a summary. Within each section, the findings of the different studies are reported sequentially to progressively build a richer picture of what is known about student understanding of scale in the domain. A discussion section is then used to reflect on the findings, and where relevant make links between the domains.

Conducting this literature review helped identify that, for a variety of reasons, certain forms of scale (e.g., map scales, non-linear scales) should be excluded from the study. These forms are not reported in this chapter. A discussion of why such scales are excluded can be found in Chapter 6.

3.2 The debate about contexts

In recent decades, many influential voices working within mathematics education "have favoured an approach to mathematics teaching and learning which has as a key focus the relating of work in mathematics to some version of the 'real world'" (Cooper & Dunne, 2000, p. 2). The intention is to give meaning to the mathematics so that understanding from outside the classroom can be used "to ground the new mathematics in 'reality' … This approach has developed as a direct challenge to the traditional idea of simply developing (possibly meaningless) algorithms" (Pirie & Martin, 1997, p. 163). In New Zealand the real world approach has been increasingly favoured. For example, in MiNZC (Ministry of Education, 1992) students are expected to complete the achievement objectives by working *within a range of meaningful contexts* (e.g., p. 58); within unit standards based assessment in secondary school mathematics "meaningful contexts" was interpreted to mean that *real, practical contexts* should be used wherever possible, as this quote from Bishton (1998) indicates:

> Solving a problem involves both **choosing** the skills or techniques to apply, AND **applying** them correctly.
>
> Therefore:
>
> - Questions should be in context; these may be real, artificial or mathematical
>
> <u>*Explanation:*</u> *It is important that contexts are accessible to students. Real practical contexts are ideal, but if they are difficult to interpret, artificial contexts are appropriate and acceptable. Mathematical contexts are equally appropriate for unit standards where real contexts are inaccessible and artificial contexts are unduly contrived.*

Two broad rationales are given for the use of real-life contexts; that they impact on student motivation and interest, and that they enhance transfer of learning between school mathematics and the real world (Boaler, 1993). Cockcroft, (1982) also indicate that it will suit their needs better as future consumers or workers. In response to such rationales, Cooper and Dunne (2000) argue the taken for granted assumption that relevance and realism are important aspects of educational practice in mathematics seems to be based "on a general sense of what might plausibly 'work' with children as a whole, or particular groups of children, as on any reference to evidence concerning what actually 'works'" (p. 195). Their research has indicated that such an approach, especially when tied to student assessment, can result in inaccurate measurements of some students' mathematical understanding which is linked to socio-economic status. Boaler (1993) points out that it is inappropriate to assume that all students will be familiar with and have an understanding of a particular context. Watson (2004) adds that offering tasks relevant to students' potential or current social and economic lives "serves to limit the horizons of those whose horizons are already limited by age, ignorance and context" (p. 371).

Van den Heuvel-Panhuizen (2005) highlights that the Realistic Mathematics Education (RME) movement encourages a different approach, emphasising the use of *reality* as a context for learning mathematics since mathematics arose from mathematising reality. The term *realistic* is used to mean that the context of the problem can be imagined by students, so includes the fantasy world of fairy tales, and the formal world of mathematics where this can be experienced as 'real' and meaningful.

Other authors also identify issues with the use of contexts. For example, Leinhardt et al. (1990) talk of the use of contextualised tasks being "based on the assumption that it is easier for students to deal with problems that build on familiar situations (e.g., situations they either have experienced or are able to relate to in meaningful ways) than to deal with abstract situations" (p. 20), but point out that it is not clear that "a real-life kind of context always supports the learning process" (p. 20) and that in the context of graphs and functions few studies have compared tasks based on both a contextual and an abstract situation. Pirie and Martin (1997) state that there is a “major problem attached to the whole idea of ‘real life’ associations in mathematics” (p. 163). They question whose reality is generally being represented, point out that lack of experience with the context being used may mean that the intended image does not come readily to hand, and that the use of too much common sense can cause problems for students. In addition, real situations tend to lead to situations where the solution is something real, and generally a whole number, which introduces barriers when more general cases are considered. Van den Heuvel-Panhuizen cites Donaldson (1978), who showed Piaget’s conservation experiments were context-determined, where a very small change in the context, not affecting the mathematics, led to a large change in the number of successful students. In another cited study, Joffe (1990) showed that when calculations are presented in different situations they are not interpreted as being similar. Van den Heuvel-Panhuizen also indicates that contexts have been shown to limit the thinking of students, citing Mack (1993) and Gravemeijer (1994); that contexts can be ignored (Davis, 1989; Greer, 1983); and that when context knowledge is not evenly distributed or common, there are also issues.

In contrast, research has also shown the power of contexts. Van den Heuvel-Panhuizen (2005) cites Clements (1980) and her own research, showing that contexts and *bare* number problems (ones without contexts) elicit different strategies, and that contexts can tap into the informal methods learnt outside the classroom. Additionally, contexts can contribute to the accessibility of the mathematics, allow different solution strategies, and suggest solution strategies to students.

3.2.1 Discussion

This body of research identifies that the value of context when teaching and assessing understanding of scale is worthy of investigation. Indeed, van den Heuvel-Panhuizen (2005) recommends that “more research is done into the effects of alterations in presentations and comparing context problems with bare problems in order to get a better understanding of these issues” (p. 9). The research also suggests that both ‘bare’ (decontextualised or abstract) number line questions and questions set in ‘real’, ‘familiar’ contexts are needed in the diagnostic assessment of student understanding. In the NDP, informal methods traditionally learned outside the classroom have been researched and included as part of the accepted teaching approach to the development of number understanding. The use of ‘real’, ‘familiar’ contexts may lead to the identification of informal learning about scale which in turn may lead to the development of insights about how students develop their understanding. However, what form should the contextual problems take?

3.3 Word problems or real practical problems?

Boaler (1993) indicates that the acceptance of the assertions about the use of context has led to the widespread adoption in the UK of the practice of putting mathematical content into word problems that supposedly represent real world situations. However, Van den Heuvel-Panhuizen (2005) identifies that there is a big difference between word problems and real practical problems. In a word problem, the context is not essential and can be exchanged for another without substantially changing the problem. Lesh & Doerr (2000) also point out that word problems involve trying to make meaning from a symbolically stated question, whereas real situations require symbolic representations to be made of something that is meaningful in itself.

Verschaffel, De Corte, and Lasure (cited in Cooper & Dunne, 2000) report that students seem to learn that problems using simple numbers but set in a context are to be solved with no reference to real-life considerations. Problems deliberately constructed to provide nonsensical answers (like each child getting 4½ balloons) sometimes provoked a response about the interpretation of the solution in the real world (up to 50% of students noting the inconsistency), but this depended on the problem. Cooper and Dunne argue that such text-based problems intended to represent the real world are often understated and leave out information needed to solve the problem in a 'real' context, and that students are often penalised in assessments if they inappropriately consider real-world issues (that is, refer to real-world considerations that the assessors had not intended). Boaler (1993) also talks of the "the make believe world of school mathematics questions" (p. 14) in which students must believe what is told in the task and ignore reality, while Leinhardt et al. (1990) report that the science-based applications provided by mathematics educators "are not always real problems in terms of how mathematicians or scientists would define a real problem or in the way students perceive them" (p. 20). Thus word-based 'practical' problems are not necessarily something to which students should apply their understanding of the context. Rather, they may often be another artificial way in which school-based mathematics is divorced from how people use mathematics outside the classroom. They require a particular understanding of the 'rules' surrounding the interpretation and answering of such items. Van den Heuvel-Panhuizen (2005) adds that Freudenthal, the father of the RME movement, "has the suspicion that teaching word problems may cause an anti-mathematical attitude in children" (p. 5).

3.3.1 Discussion

When using contexts to assess student understanding of scale the issue of using word problems or real, practical problems to assess measurement understanding is an important one. The research suggests that it is preferable to use real measurement instruments in situations that students are familiar with. However, the assessments for this study need to identify individual understanding. Observations of students working in measurement practicals have indicated to the researcher that it is very difficult to control student interaction. In some situations this is because the measurement task requires more than one individual. In others it is due to the number of available stations in the classroom and/or the fact that other students can be observed manipulating the instruments. Some instruments also need to be inconspicuously reset after each use. Furthermore, some familiar commercial tools used for practical measurement (e.g., kitchen scales) are sophisticated instruments. In combination, these factors indicate that for practical reasons the use of such real instruments would be problematic. However, the assertions about the use of contexts indicate that word problems with pictures of simple scales set in familiar situations should access similar understandings.

3.4 The notion of a representation

In Chapter 2 it was identified that a scale can be *represented* in a variety of ways. However, while the related term *representation* is in widespread use, the notion is one with different meanings so needs definition. In this study, *representation* is used to mean "a way of showing or thinking about a mathematical concept". Thus the number line showing three quarters, and the fraction circle showing three quarters can be considered as two different representations of the same concept, while the symbol ¾ is a third. This interpretation is used elsewhere in the literature. For example, Janvier (1987) discusses the concept of a representation in relation to a single musical note. He identifies the written symbol, the name, the vibrating note (as it is sung and heard), and the played note (as it is played on an instrument) as four different modes of representation of this note. Lesh, Post, and Behr (1987) also imply such an interpretation when talking of translation between representations, that is the process of changing or recognising information represented in one way to how it is represented in another.

A number of researchers identify that it is important to distinguish between *internal* and *external* representations (e.g., Dufour-Janvier, Bednarz, & Belanger, 1987; Goldin & Shteingold, 2001; Kaput, 1998). Internal representations concern more the mental images we construct of reality, which suggests that the mental image we hold of ¾ can be considered another representation of the mathematical concept three quarters. Internal representations can include natural language, particular words and phrases, visual and spatial imagery, action sequences, students' personal symbolisation constructs and assignments of meaning to mathematical notations, formal notational procedures, and problem-solving strategies and heuristics (Goldin & Shteingold, 2001). They are hypothesised mental constructs (Kaput, 1998) since we cannot observe anyone's internal representations directly. To identify a student's internal representations of a concept we can only make inferences on the basis of what is observed when they interact with, produce, or discuss external representations of that concept (Goldin & Shteingold, 2001). The term external representation refers to the symbols, diagrams, spoken words, physical objects, etc. we use to represent a mathematical reality, as exemplified in the previous paragraph.

It should be noted that articles written around the end of the 20th Century go beyond the notion of a representation to discuss notions like that of a *representational system*, the argument being that a representation cannot be understood in isolation or merely in relation to other representations (Goldin & Shteingold, 2001; Kaput, 1998; Lesh & Doerr, 2000). Rather, it only makes sense as part of a wider system within which meanings and conventions have been established (Goldin & Shteingold, 2001). The example given by Janvier (1987) could be considered an example of a representational system, as can the example related to the fraction ¾. Because the notion of a representational system has a connotation of understanding attached to it, the development of understanding in relation to the notion of a representation will be explored in the next section.

It should also be noted that recent research has moved on to discuss *inscriptions* rather than representations (e.g., Lehrer et al, 2000). Such literature promotes an interpretive approach and seems to focus on the active role of the learner in *mathematising* (drawing out the mathematics behind a real life event) and *symbolising* (creating and reasoning with symbol systems). Inscriptions are the students' own creations rather then agreed-meaning notations, and can be the early foundations of representational competence (Lehrer & Schauble, 2000). Using the musical example of Janvier (1987), in this approach it is recognised that the score does not hold all of the message, so needs to be interpreted by the musician. Thus the inscriptions a student develops and learns to use are not static images but are part of a dynamic learning process, and it is the process of symbolising rather than the symbols themselves that need to be the central focus of attention (Cobb, Yackel, & McClain, 2000). While this approach has merit for the actual teaching of scale, the concept is not explored further as this study focuses on what understanding the students in Years 7 and 8 already have of scale.

3.4.1 Developing understanding of a representation

When a teacher uses a representation, they bring with them a set of rich understandings. These understandings allow the teacher to apply, interpret, and manipulate the representation in a manner appropriate to the situation in which each is employed and the limitations of the representation's structure. Teachers tend to expect students to be able to do the same. This presupposes that the learner perceives a conventional representation as a mathematical tool; has 'grasped' the representation; knows its possibilities, limits, and effectiveness; and how it relates to other equivalent representations (Dufour-Janvier, Bednarz, & Belanger, 1987). However, novices may not have the necessary skills and understanding, needing to develop "the rules of usage, the conventions, the symbols, and the language linked to the representation" (p. 118). This lack of familiarity with the representation is often overlooked by teachers. "Most people who encounter a child with difficulty in resolving a problem focus on the lack of understanding of the mathematical concept involved, rather than on the representations that are utilized" (p. 116). They further suggest that the premature use of a representation, or its application in an inappropriate context, can lead to the development of misconceptions that can affect later learning. The following number line example illustrates this point:

> First, when children use the number line during the learning of positive integers, they develop the notion of the number line as a series of "stepping stones." Each step is conceived as a rock, and between two successive rocks there is a hole! This is very far from the concept of the density of the real numbers as illustrated by the number line. It is hardly surprising that at the secondary school so many students say that between two whole numbers there are no numbers or at most one. Nor should there be much surprise that they also have difficulty placing a number if they cannot associate it with the graduation already given on the line. Or again if in their reading of graphs they focus only on the values given by the graduations.
> Secondly, many children see the numbers on the number line as though they were "number-panel-signs" placed along a road. They don't see any necessity for placing these billboards at equal distances, saying only "well, usually they are equally spaced." Here this is far from the measure concept, also illustrated by the number line. (p. 117)

Dufour-Janvier et al. (1987) also noted that children seem to perceive representations as separate mathematical objects, so cannot use them as interchangeable tools. Often they do not even recognise that a representation provided to scaffold a problem is related, rather "we have observed that children do not even bother to look at them" (p. 115). Dufour-Janvier et al. exemplify this by talking of an experiment where a child solves a problem using a particular representation. When shown the work of another child using a different representation (with a different answer) they report a number of children found this quite natural: one is done one way and the other is done differently!

Kaput (1987) talks of the recognised difficulty students have when translating between representations. Dufour-Janvier et al. (1987), however, suggest translation is possible where students 'grasp' each representation and the role of language in the process is given due weight. Janvier (1987) supports the place of language.

> Language or words play a central role. Actually we noted verbal tags were given to the relevant elements and the execution of the process was carried out through an efficient handling of those verbal tags. In a way source and target were verbally simplified. Even in a "graphical context", the speech appeared to be essential. (p. 30)

When the role of language is considered alongside the difficulty students have with recognising the equivalence of different representations, and with translating between representations, this suggests that to develop understanding of a mathematical concept attention needs to be paid to both the different representations used for the concept and the development of a language to link those representations. Lesh, Post, and Behr (1987) bring further clarity to what is meant when we say we 'understand' from this perspective.

> Part of what we mean when we say that a student "understands" an idea like "1/3" is that: (1) he or she can recognize the idea embedded in a variety of qualitatively different representational systems, (2) he or she can flexibly manipulate the idea within the representational systems, and (3) he or she can accurately translate from one system to another. (p. 36)

Furthermore, from this perspective, limitations in understanding are related to the existence of partially formed internal systems of representation which leave in place long-term cognitive and affective obstacles (Goldin & Shteingold, 2001). If such an interpretation of the term 'understanding' is accepted, there are implications for the research undertaken within this thesis, as the following discussion outlines.

3.4.2 Discussion

The literature on representations suggests that it is useful to consider scale as a mathematical concept that can be represented in a variety of ways. In number, it is represented as a number line, in graphing as a graph axis, while a tape measure is one representation used to measure length. From this perspective, to *understand scale* students need to be able to recognise scale within these different situations, know the conventions and rules for using scale, have a language to discuss scale, and have adequate knowledge of the embedded symbol systems. However, as discussed in Chapter 2, a scale can also be considered a representational object, one that is not used to represent just one concept. In number, the number line can be used to represent whole numbers, fractions, decimals, integers, or additions. In graphing, both statistical and algebraic graphs are visual representations based on scales that allow other concepts to be explored (e.g., rate of change and data clustering). In measurement, a measurement scale is a comparative tool where numbers are used to represent quantity. At its simplest level, the concept of scale being a representational object suggests that another part of identifying understanding of scale is to consider if students understand how to use a scale to represent different types of number. It also suggests that understanding of scale includes understanding both the object and the processes that are invoked when using this object. The assessment designed to measure student understanding therefore needs to provide situations that allow students to show their grasp of each of these elements of understanding scale.

In the next section attention is turned to the concept of a schema, and the schema theory of learning. This helps further define how the term *understanding* is used in this study, and what understanding of scale involves.

3.5 Schemas and the development of conceptual understanding

While the term *schema* was first used by Piaget in 1926, it was Skemp's (1971) simple description of the notion that has made it valuable in mathematics education (Tall & Thomas, 2002). "The general psychological term for a mental structure is a schema" (Skemp, 1971, p. 39). A schema is a representational structure and can be thought of as a network of connected concepts, so our schemas as a whole form our mental model of reality (Olive & Steffe, 2002). Skemp's view is that we learn by creating schemas. These help us organise knowledge into categories, and are our way of explaining how the world works. They are built on a process of identifying patterns and similarities. They form a system through which we interpret new information. For example, once a person has constructed a schema for what a chair is, new forms of chair can be compared to this concept. This may result in their identification as a chair, or as something else. In some cases, the schema for what a chair is may be expanded to include samples not previously considered, like a swivel chair with a single pole as a leg. Skemp points out that information which fits with an existing schema is readily integrated and remembered, while information that does not "is largely not learnt at all" (1971, p. 43). Skemp provides supportive evidence on learning retention from an experiment on the learning of sign language through a schematic approach and through rote learning. He also warns that once formed, schemas do not readily accommodate new (contradictory) data, so inappropriate or non-existent early schemas make the assimilation of later ideas much more difficult, "perhaps impossible" (p. 51). However, schemas are not static; rather they are in a constant process of growth and need to meet the challenges faced by interacting with the world. As such, new learning builds on existing knowledge. As adults, we cannot therefore assume that our schemas are the same as those of children, so we need to construct models of their reality if we are to find a rational basis for what children do (Olive & Steffe, 2002).

3.5.1 Discussion

The notion of a schema, as discussed in this section, has parallels with the notion of a representational system discussed in the previous section, in that it is the individual's attempt to make sense of a concept. In this light, a schema can be considered to be the sum of the internal representations the person has, along with the external representations that are recognised as part of the concept. Schema theory suggests that in trying to map student understanding of scale the current study is trying to identify how the ideas relating to scale are organised and connected in the minds of Year 7 and 8 students. Having a model of the child's schemas is also important if we are to devise ways of developing those schemas. In this context the term *understanding* is not static; rather something dynamic that interacts with the environment and grows from interactions with that environment. It also relates solely to the individual, as different individuals can have different schemas, or schemas in different stages of development. The purpose of studying students' understanding of scale is therefore to attempt to identify the range of knowledge and skills that students utilise when dealing with scale, as well as how they are organised and connected, thus shedding light on individual understanding.

Attention is now turned to examining what is known about the understanding that students have of scale, starting with understanding scale as a representational object. This involves exploring how the number line (the application of scale within number) is used to represent different number sets.

3.6 The number line

3.6.1 Whole numbers and the number line

The research into the understanding of scales involving whole numbers appears to be limited; however some material has been found relating to the use of the number line to represent additions and subtractions. Ernest (1985) identifies that one practice to evaluate student understanding of addition and subtraction was to see if a student could represent a problem like 3 + 5 on the number line. He identifies that the 9 and 13 year old students in the US NAEP 2 study (the second National Assessment in Educational Progress in Mathematics), who were in other ways successful with these operations, were generally unable to do this. One problem was deciding 'where to start'. Of students showing both numbers as skips, some started at zero and others at one. Another group showed only the second number as a skip, with some also confused whether to start at 3 or 4. As learning the process of addition did not necessarily involve working with number lines, so the number line representation was not testing understanding of addition and subtraction, rather the ability to understand the representation and translate to and from it.

Carr and Katterns (1984) followed the NAEP items with a New Zealand study of 179 9-year-olds and 352 13-year-olds. They found higher success levels (from a much smaller sample) but noted that students from all abilities used the number line erroneously, rather than only those from the low ability group as expected. One common error was counting the digits (marks) rather than the spaces between the digits.

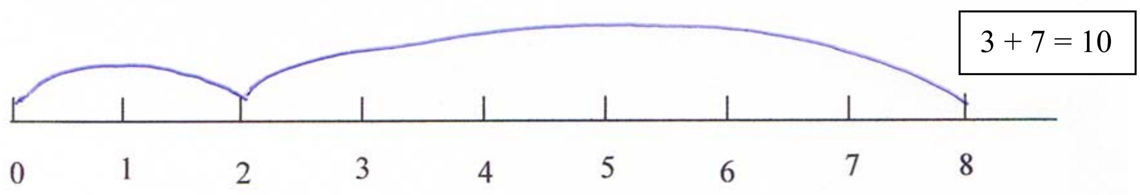


Figure 3.1. Illustration of the type of error 'counting digits rather than spaces' found by Carr and Katterns (1984)

Carr and Katterns identify that the number line is a symbolic representation of measurement rather than a concrete illustration. Emphasis on counting digits (marks) does not help students understand that the number line is fundamentally a measurement model or the significance of the spaces (or units). Implications for teachers include the need "to observe and regularly talk with their pupils to find out how they perceive and operate with the number line model" (p. 34).

Discussion

This research suggests that the number line is based on the conventions of measurement, and that some students do not understand the conventions of the number line, so are unable to represent additions in this manner. The findings of Ernest (1985) support the work reported above on the use of representations.

In the next section the attention turns from representing whole numbers on number lines to fractions, a topic well covered in the literature.

3.6.2 Fractions and the number line

Fraction understanding is the subject of much research, possibly due to the complex ways in which we can interpret them and the recognised difficulty students have with fractions. For this study on understanding scale, the findings relating to understanding fractions on the number line are important because many of the identified issues echo findings from other domains where scales are involved. They also point to directions for enquiry that relate to scale but are beyond the realm of fractions.

The measure sub-construct of fractions

Kieren (1976) identifies that one way to interpret a fraction like ¾ is as a measure (when marked as a point on a number line). He contends that various cognitive structures are necessary for dealing with this interpretation. These build on the concept of conservation of length and substance, and the notion of ordinal numbers. These are:

1) the unit and its arbitrary division, where the unit can be partitioned into any number of equal parts;
2) the part-whole relationship in this context, and recognising equivalence arising from unit partitioning; and
3) an order relation – including the ability to order physical objects and use symbolic ordering statements.

He further contends the measure interpretation is the “natural locus for considering order” (p. 124) for fractions, and “by failing to take into consideration the unique salient features of each interpretation, teachers and researchers have encountered difficulties that could have been logically anticipated” (p. 103). Mack (1993) adds that “Kieren and Behr et al. proposed that for students to develop a complete understanding of rational number, they must construct meaning for rational number symbols and procedures in a way that integrates the various subconstructs” (p. 86).

Issues with understanding the measure sub-construct of fractions

Research indicates that students have difficulty with using the number line to represent fractions (Baturo & Cooper, 1999; Behr et al., 1983; Bright, Behr, Post, & Wachsmuth, 1988; Lesh, Post & Behr, 1987; Larson, cited in Ni, 2000; Novilllis, cited in Ni, 2000). Hart (1989), in talking of the *Concepts in Secondary Mathematics and Science* (CSMS) study, indicated that scrutiny of all the easy CSMS items “suggested many secondary-aged children worked entirely within the set of whole numbers” (p. 46). Scores derived from test items with number lines have also been shown to be insensitive as an assessment tool to reveal what children know about rational number (Ni, 2000).

Hart (1989) describes an English teaching experiment with three classes of 10 to 12-year-olds. The teaching sequence for the 10 to 11-year-olds focused on developing the understanding that students had of fractions on the number line. Hart indicates that most students, when asked where the fraction ¾ could be found on the number line, marked a point greater than 1 (Figure 7.6). More generally she notes a number of children were seen inventing their own methods for solving the problems and that the children seemed to interpret what was being taught in relation to their own understanding of number. While a rule being taught for finding equivalent fractions was based on multiplication or division, some children were using counting, additive, or doubling based strategies (which were only sometimes successful), and treating the numerators and denominators separately. She suggests that the children were not seeing a fraction as formed from two parts that only make sense when interpreted together.

Pearn and Stephens (2004) had similar findings when interviewing 12 Year 8 Australian students (14-year-olds). In one item they sought to determine the students' strategies for identifying the bigger of two non-unit fractions. They found three types of responses. One group was described as proficient multiplicative thinkers; the second tended to revert to inappropriate whole number thinking when faced with new or unfamiliar situations; and the third tended to treat fractions "in ways that ignore the fundamental ratio between numerator and denominator" (p. 432). When asked to show the fractions ¾ and 3/5 on a number line from 0 to 1, one student placed ¾ one 'gap' away from 1, and 3/5 two 'gaps' away from 1. Another roughly divided the unit into four pieces for ¾, but for 3/5 placed five marks between 0 and 1 and marked the midpoint.

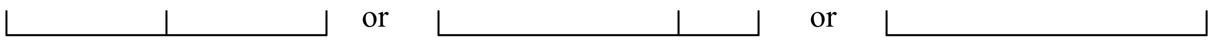
Pearn and Stevens (2007) report a further study into students' understanding of fractions as numbers; in particular "how number lines involving whole numbers can be used to develop fractional language and to articulate fractional concepts" (p. 604). They start by providing information on previous studies by the authors with a Fractions Screening Test for Year 5 and 6 students, and weaker students in Years 7 and 8. Results highlighted that many students had difficulty when working with number lines. For example, many students confused the number 3/5 with three fifths of the number line; and when locating fractions other than halves and quarters, guess work seemed to be used rather than any systematic division of the number line. The same screening test was used initially for the 2007 study. Pearn and Stevens found that at one school where considerable professional development had been given, 50% of students located three fifths of the number line rather than the number 3/5. Overall at the three schools in their study no more than 44% could state and accurately locate a fraction between 0 and ½, and less than 59% could locate the number 1 given the location of 0 and 1/3. Additional number line tasks were also discussed. For example, one task involved an unmarked interval from 0 to 100 upon which students had to locate the number 50. Another task used the same interval marked in twenties; students were asked to identify the numbers that should go on the marks. They found that the "successful students used number knowledge, accurate skip counting and multiplication facts to partition the number line" (p. 608). Students also knew the relationship between wholes and parts, could attend to equal parts, and could apply fraction terms to these equal parts. Other students who did not have the whole number knowledge to draw on frequently guessed where their answer should go, often had difficulty with intervals that lacked a mid-point, and tended to look at the lines rather than the parts. Some students also needed assistance to see the connections between halves, quarters, and eighths, and/or to identify the number of pieces the interval had been divided into (that is to count spaces instead of counting marks). Many of these findings support the initial findings of this study reported in Drake (2007), who suggested that the strategies students used with scales seemed to form a hierarchy linked to number understanding (see Chapter 8).

Baturo and Cooper (1999) also comment on students' number understanding. In their study of 24 Year 6 and 10 Year 8 students they found some had difficulty understanding fraction symbols (improper fractions and mixed numbers) when working with the number line. Some students interpreted the six in 6/3 as a whole number. For example, 'six and a bit more' ($6\frac{1}{3}$), or 'six and you have got to put in thirds' where six is counted as the first third so the mark goes on $6\frac{2}{3}$, or 'six and three partitions for thirds' (6¾). They also found students made errors by counting marks not spaces, sometimes counting just the marks between whole numbers, so identifying that a unit is cut into thirds not quarters. Others counted zero as the first mark, or did not count numbers marked as wholes. Some students seemed to have a holistic view of the number line, so when asked to place 2¼ on the number line marked 'a little past two' (at $2\frac{1}{8}$) – ignoring the quarter sub-divisions provided. Overall students seemed to have their own idiosyncratic understanding of fraction symbols and how to locate the numbers on the line. Achievement also declined with age.

Behr et al. (1983) report on the Rational Number Project (RNP) which focused on the development of rational number understanding in Grades 2 to 8. In one of a series of written tests reported, Bezuk assessed 77 students in three Grade 4 classes and looked at visual distractors to see if the students could ignore these and deal with the tasks on a logical-mathematical level. The items used the different fractions sub-constructs identified by Kieran (1976) and four sorts of visual cues. These forms of cue for locating ¾ on a number line are shown in Figure 3.2.

Complete cues – where all the information is provided

Incomplete – where no cues or partial cues are provided



Irrelevant – where all the cues contain extraneous but neutral information which needs to be ignored

Inconsistent – where all the cues need to be ignored

Figure 3.2. Types of cue identified by Behr et al. (1983)

This test found:

> a disproportionate number of errors in number line problems across all categories and specific fractions. This is true despite the fact that for 3 years the students' text series had employed the number line model for whole-number interpretations of addition and subtraction. Children in the sample were generally incapable of conceptualizing a fraction as a point on a line. … These results are consistent with other findings (e.g., Novillis-Larson, 1980) that suggest number line interpretations are especially difficult for children. (pp. 111-113)

The percentage of errors increased as the type of perceptual cue changed from *complete* to *incomplete* to *irrelevant* to *inconsistent*, a result consistent across the four representations used: continuous rectangular bars, circles, the number line, and discrete sets of objects. Children also differed in how they identified the unit on the number line.

In another experiment, after instruction 11 fourth graders were given a series of number line tasks with different cues. Behr et al. (1983) found differences in success levels, the strategies used and the amount of assistance needed to solve the problems. For example, while all 11 were able to locate 2/3 on a four-unit number line partitioned into thirds, only four were successful with locating 2/4 on a similar number line with inconsistent cues. In Lesh, Behr, and Post (1987) (reporting the teaching experiments in the Rational Number and Proportional Reasoning projects) visual distractors such as extra partitions on number lines were identified as the apparent cause of some students changing their approach when answering mathematically similar problems. Minor changes on any one of a number of dimensions – "number size, relationship of numerator to denominator, perceptual characteristics of figural models, etc" (p. 43) – frequently produced dramatic changes in students' performance, suggesting that their underlying understanding of rational number was unstable.

Bright et al. (1988) report three teaching experiments, including a follow up to the work on visual distractors. They identify that the number line is different from other fraction representations. First, a length represents a unit, so the number line model not only suggests unit iteration but also unit subdivision, allowing it to be treated as a ruler. Second, units are not visually separated as in the region and set models. Third, the number line uses fraction symbols to convey part of the intended meaning so requires the integration of two forms of information, visual and symbolic. They hypothesise that this third requirement causes difficulty.

The first reported experiment was an 18-week clinical experiment with five fourth graders. The interview excerpts seem to indicate that students used a variety of strategies to reach their answers, some addition-based, others multiplication-based, and that not all students were able to plot their points accurately. The second involved a 30-week clinical experiment on eight fourth grade students. Some errors identified were: students using the wrong unit (the whole number line instead of the interval from 0 to 1); students counting marks instead of intervals; and students representing the inverse of the fraction. The third study transferred the work of the clinical experiment to a whole class of 34 fourth graders. Post-test results were significantly higher than pre-test ones. The researchers found that:

- number lines of length one were easier to work with than those of length two (it was reasoned that for lines of unit length one, the length of the unit could be assumed);
- items where the fractions were given were easier than items where the representation was given;
- problems where no repartitioning was required were easier than those requiring repartitioning;
- the relative difficulty of the four cue types (see Figure 3.2). changed from pre-test and post-test.

They note that the shifts in error patterns suggest "the instruction at least sensitized students to the need to attend to some characteristics of the number line model" (p. 224).

As background to the Rational Number Project, Behr et al. (1983) outline the results of Novillis-Larson (1980), who reported a study of seventh graders locating fractions on number lines either one or two units long. In these number lines the number of segments in each unit equalled or was double the denominator of the fraction. Novillis-Larson also found that:

- associating proper fractions was significantly easier with a number line of length one and where the number of segments equalled the fraction denominator; and
- almost 25% used the whole line as the unit with a number line of length two.

Mack (1993, 1995) identifies three issues that are of direct relevance to this study. The first issue, that of symbol understanding, is reported in this paragraph, the others are reported later.

Research into how we construct meaning for mathematical symbols suggests we match formal symbols to other representations that are meaningful to us, such as specific real-life situations or concrete representations. She also identifies that research has documented that the knowledge gleaned through a person's interaction with real-life situations and problems is self-constructed, can be context related, and may be right or wrong. Mack calls this body of self-constructed knowledge 'informal knowledge'. For fractions, most informal knowledge involves the part-whole construct, where students frequently *reunitise*. For example, they solve problems like dividing a pizza by cutting the pizza into slices, then use the slice as a 'new unit' rather than thinking of it as a fraction of the whole pizza. As such, informal strategies convert many fraction problems into whole number problems. Thus she contends students' informal knowledge limits their understanding of rational number.

Siegal and Smith (1997) also argue that students taught with the fraction circle reunitise (but use other language). In doing this they are creating counting-based problems instead of learning that fractions are found between the whole numbers, which relies on understanding measurement conventions. They argue further that there may be a clash of mathematical worlds when teaching fractions. The child's world may be restricted to an understanding of number based on the discrete counting numbers used for computation, while teachers might assume that circle parts represent parts of a whole, and that numbers are also used in continuous measurement. The child's concept may become increasingly resistant to change as they become older. Thus some children, when exposed to fraction-based teaching not clearly relevant to measurement, "may misinterpret the physical situation and resist generalising their theory of number to accommodate for fractions" (p. 20).

The second issue of relevance reported by Mack (1993) is that the concept of the unit is a significant issue for students learning about rational number. She contends that students need to realise a unit is not a fixed entity, but can change in nature, for example, from being a discrete quantity that can be counted to a continuous quantity that can be measured. She notes that research shows students have difficulty identifying appropriate units. Students often arbitrarily shift the unit to include all the elements in the problem, in turn limiting their ability to deal with fractions greater than one. Students need to be able to *reconceptualise the unit* – accept what is identified as the unit in a problem and work from there to subdivide and combine units, if they are to successfully understand fractions.

A third issue identified by Mack (1993) is the lack of connection between students' informal knowledge and formal symbols, procedures, and pictorial and concrete representations. She cites two experiments in which students, when asked verbally, were able to correctly solve a fraction problem, but in one case when presented with a pictorial representation of the problem the student was unable to do so. In another, a student was not able to solve the problem when presented with symbols.

> Students often develop separate systems of arithmetic that operate independently of one another, an informal system that they use to solve problems that are meaningful to them and a school arithmetic consisting of procedures that they apply to symbols or artificial story problems they are given in school. (Hiebert & Carpenter 1992, pp. 83-84, cited in Mack 1993, p. 94)

Hiebert and Carpenter also caution that students attempting to construct meaning for new symbolic representations "may draw on prior knowledge of other mathematical symbol systems, consequently overgeneralising and constructing inappropriate meanings for the new representation" (cited in Mack, 1995, p. 423).

Early fraction learning is the focus of Hunting (1989). He suggests fraction understanding develops from an early age and that the number one half is special for children. "Children in the middle and upper years of primary school use it as a reference number when comparing other fractions" (p. 2). He also suggests a child's understanding of the concept of 'one half' develops through a number of stages before it can be represented precisely. For tasks like cutting lengths of string, these stages represent one half as a:

- multiple sequence of subdivisions (where preschoolers do not know when to stop cutting…);
- single subdivision with gross inequality in each part (the whole is simply cut into two);
- single subdivision with remainder (some attention to equality is made with the first cut, then 'a remainder' is cut off the large piece and discarded); and
- single subdivision into two equal parts that uses all the material (eye movements and/or finger movements are used to check the relative size of the pieces before cutting, and the size of the pieces are rechecked after cutting).

Unequal parts "may as likely be the result of poor technique as immature conception of equality" (p. 3). Hunting identifies more generally that when children are subdividing continuous quantities they halve where the symmetry of the material can be used. When dividing a paper strip into three equal pieces, they use informal measurement processes involving the estimation of a unit, its reproduction then 'check and adjustment'. These processes, he claims, are difficult to develop into further mathematical knowledge.

Mitchell (2005) also identifies that students have to start working in the continuous realm when learning about fractions and decimals, as there are an infinite number of fractions between two points on a number line. The part-whole fraction model, when based on fraction strips, is length-based yet "the part-whole aspect of fractional understanding is often taught as counting and matching" (p. 545). Mitchell goes on to add that teachers:

> often use area models or length models in fraction tasks without ascertaining that each child is confident in the applications of measurement that they will need to use as a tool to help them think conceptually about the task at hand. (p. 545)

For tasks like labelling regions in a fraction where one half is cut into quarters, and the other half is cut into sixths Mitchell found that those who were successful on such tasks "made successful use of measurement principles" (p. 551). For example, students successfully identifying a sixth showed verbally that they had the measurement understanding of the unit, and demonstrated how to re-unitise into smaller (equal) units with no gaps and no overlaps.

Finally in this section, Izsák, Tillema, and Tunç-Pekkan (2008) describe the classroom observation of the teaching of addition and subtraction of fractions on number lines, and the impact this had on one student. Interviews showed that the teacher considered the number lines and fraction strips as temporary aids for visualising amounts, and that while using them provided a different method to using numeric algorithms, it was the development of the algorithm in which she was really interested, and which was really important. The drawings were used when the topic was first introduced and tended to be used when students were confused.

Izsák et al. note that the teacher had access to two different partitioning strategies, iterating and recursive partitioning, but only emphasised one of these (iterating) with the class. Iterating (based on Olive & Steffe, 2001, cited in Izsák et al., 2008) involved taking a segment and joining copies end on end to fill an established unit. Recursive partitioning (based on Steffe, 2003, cited in Izsák et al., 2008) involved partitioning partitions, for example, creating twelfths by first partitioning the unit into four, then partitioning each of these pieces into three. The teacher used fraction strips and number lines to illustrate the process of fraction addition, partitioning the unit by "eyeballing it and trying to space them out" (p. 43), emphasising that each piece needed to be the same size. Where necessary, she adjusted the location of the 1 so the last piece was the same size as the others.

The student demonstrated different understandings, while using a similar language to that of the teacher. For example, when asked to draw a strip to show half and another for one third, the student wanted to add another piece on to the diagram showing halves – which would have the impact of keeping the pieces the same size, but changing the location (and size) of the whole. In another example, the student iterated used 12 tick marks for twelfths "because that's what the denominator is" (Izsák et al., p. 56). For this example, the student (and her friend) were also comfortable with 12/12 and 1 being located in different positions, even though they knew them to be the same number. The researchers conjecture that this was a consequence of the way the teacher flexibly changed the size of the unit to fit equal pieces. However, the student was also able to use recursive partitioning correctly on a fixed unit, doing so when she was asked during an interview to partition a unit into quarters and another into sixths. The authors conclude:

> little research has attended to teachers' interpretations and uses of fine-grained representational features such as tick marks. When discussions of pedagogical content knowledge and mathematical knowledge for teaching have turned to representations, they have not done so at a grain size that would distinguish between Ms. Reese's two partitioning methods. … Discussions of pedagogical content knowledge and mathematical knowledge for teaching should treat resources teachers have for eliciting and adapting to students' interpretations and uses of drawn representations at a comparably fine grain size." (p. 59)

Discussion

Research on fractions indicates that there are five essential elements to consider when introducing the measure sub-construct. These are:

1) Understanding that fractions are numbers representing parts of a whole, and are found between the whole numbers;
2) Understanding the symbol notation;
3) Conceptualising a fraction as a measurement;
4) Relating the measure sub-construct to other fraction sub-constructs; and
5) Attending to the fine-grained representational features of the sub-construct.

The need to understand the symbol notation and be able to relate the number line to other fraction sub-constructs were also noted as important in the literature on representations.

In relation to element 1:

- Students' knowledge of whole numbers, and attempts to interpret fractions with this, can interfere with their understanding of how a fraction operates.
- Students may have an understanding of the number line as a series of 'stepping stones', on which fractions have no place.

In relation to element 2:

- Students need to understand that fraction symbols are formed from two parts that only make sense when interpreted together. This understanding needs to include fractions greater than one.

In relation to element 3:

- Students need to have a clear understanding of measurement principles.
- Students need to link the measure construct of fractions to their understanding of linear measurement rather than to their understanding of counting.
- Students must be able to identify and reconceptualise the unit, then both subdivide and aggregate it as a length.

In relation to element 4:

- Students need to relate what is represented in the measure construct of fractions to the same fraction represented in other ways.

In relation to element five:

- Students need to attend to how the marks and other features of the number line are used to convey information.

In addition to these five elements, Mack (1993) states that building on existing, informal, knowledge of fractions is important, but does not outline what knowledge may be related to the measure sub-construct of fractions. However, Hunting (1989) suggests that the number ½ is used by students as a reference point to compare other fractions, while Hart (1981) calls it an honorary whole number. This suggests students developing an understanding of fractions in relation to scale may have some useful knowledge involving halves that should be looked for when assessing student understanding.

This section has focused purely on fractions. The next section continues examining the use of the number line as a representational object for different forms of number, this time considering research involving the related set of numbers, decimals (or as they are sometimes called, decimal fractions).

3.6.3 Decimals and the number line

Three studies that use the number line to assess decimal understanding are reported here. In the CSMS study of place value and decimals, Brown (1981) describes the performance of students aged 12 to 15 on several number line items. The simplest of these looks like Figure 3.3 below.

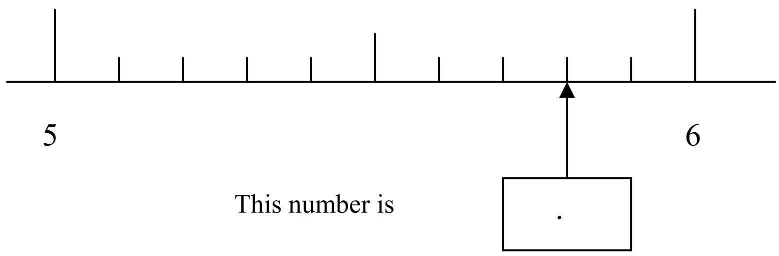


Figure 3.3. Illustration of a decimal item from the CSMS study

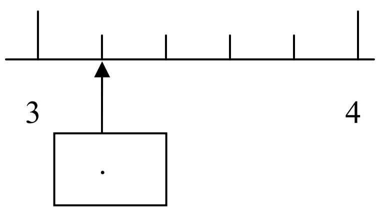
Results showed that the percentage of students reading the scales correctly increased with age. This was recorded as *facilities* (percentage of students with correct answers). The facilities for the item in Figure 3.8 were:

Age	12	13	14	15
Facility	62	74	83	85

The CSMS team placed items in levels according to their difficulty. For the place value and decimal understanding test, this item was placed at level 2, while a similar item involving a number line fragment from 2.7 to 2.8, requiring 2.74 to be labelled, was placed in level 3 (a harder level); only 31% of 12-year-olds successfully answered the item, rising to 71% of 15-year-olds.

A third item looked at a number line fragment from 14 to 15, marked in tenths. Students were asked to identify a point half way between two marks (14.65). This was judged harder than the other hundredths item so was placed at level 4 (only 24% of 12-year-olds were correct). Both items involving hundredths were relatively more difficult than the other items in their level. The final item of relevance was placed at level 6. This involved asking the question "how many different numbers could you write down which lie between 0.41 and 0.42?" (Brown, 1981, p. 62). Even including answers like 'lots' and 'hundreds', only 12% of 12-year-olds and 20% of 15-year-olds provided answers deemed correct.

Swan (1983) builds on the CSMS research. His pre-test of 12 and 13-year-olds prior to a teaching experiment revealed students "were confident at reading scales subdivided into tenths, but were less successful at reading scales in hundredths, fifths, or at estimating a reading which lies between marked calibrations" (p. 53). A common error involved students focusing on the number immediately to the left and counting along in tenths, regardless of the number of marks provided.



	Age and percentage correct			
Response	**12**	**13**	**14**	**15**
3.2 (correct)	23	37	50	58
3.1	48	42	33	28

Figure 3.4. CSMS item reported in Swan (1983)

NEMP also examines student understanding of decimals on the number line. Working with a number line from 0 and 1 marked in tenths (see Figure 7.9), only 54% of Year 8 students could successfully locate 0.1. Results for the decimals 0.5, 0.7 and 0.25 had a success rate below 50% (Crooks & Flockton, 2002).

Discussion

The assessment items mentioned represent different ages of mathematics learning on opposite sides of the world, England and New Zealand. The English CSMS items look at a pre-calculator age based on 'new' mathematics. The New Zealand NEMP items look at a calculator-integrated age with a curriculum built upon problem solving (the "curriculum statement assumes that both calculators and computers will be available and used in the teaching and learning of mathematics at all levels" Ministry of Education, 1992, p. 14). Each item shows a large proportion of students having difficulty with decimal number lines where only a fragment of the line was provided.

The number line items described in Brown (1981) and Swan (1983) require students to have an understanding of how the number symbols relate to positions on a number line if they are to be successful, which suggests a link may exist between number understanding and an understanding of scale. That the concept of a number line as an infinite set of points was identified by Brown (1981) as much harder than any other number line item also supports the existence of such a link. Students identifying that the interval between two whole numbers can be split into tenths, while the intervals between the tenths can also be split into ten, creating hundredths, seem to have gone a long way towards developing the concept of the infinitely small. Perhaps this nesting of decimals within decimals has not been properly explored in teaching. This idea will be explored in Chapter 4.

In the final section on the number line as an object for representing different forms of number, student understanding of how integers are represented is examined.

3.6.4 Integers and the number line

There appears to be little research on students' understanding of the number line model of integers. White (1994) provides a summary of the methods used in the teaching of integers, and difficulties with each of these. She outlines five hurdles in the historical treatment of directed numbers as identified by Hefendehl-Hebeker (1991). Four of these are relevant:

- the initial use of two separate and oppositely oriented half lines instead of a single continuous number line;
- the view of zero as absolute with nothing below it;
- the problem of assigning a concrete sense to integers and operations between them; and
- the elimination of the Aristotlean notion of number which elevated magnitude over the notion of number, making it hard to accept that that $-8 < 1$.

When discussing the number line model, White identifies that the general practice is to start with thermometers and other scales to overcome the first two hurdles, then use the model for addition and subtraction. The first step in this process is generally to establish the convention that when adding, movement is to the right, and when subtracting movement is to the left. Later, when reflecting on the use of the number line, she notes that it needs introductory work to clear the last hurdle, with regular reference to other models.

Kuchemann (1981) reports the CSMS study on understanding of positive and negative numbers, conducted with 302 14-year-olds. Predominantly, students worked on manipulations but the study has some relevant findings about integers and scale. For a problem similar to that in

Figure 3.5, 95% correctly identified the number for the triangle, and 94% the number for the box.

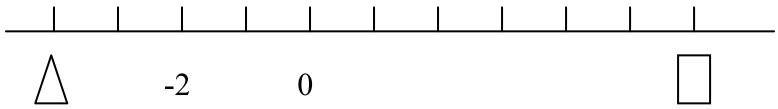


Figure 3.5. Integer identification problem of the style in Kuchemann (1981)

Kuchemann also identifies that students generally seemed to model the addition items in terms of ‘moves’ or ‘shifts’ on a number line. This approach was seen to be “a highly effective concrete model” (p. 84), with over 80% of the pupils correctly answering the items. Overall, problems with negative answers were seen to be more difficult than those with positive ones, although the difference was not great except when there was a ‘shift’ across the origin, “which may have been ignored” by some students (p. 83). The item in Figure 3.5 was placed at facility level 1 (the simplest level) while the majority of the addition items (those involving a negative number somewhere) were placed at level 2. When it came to subtraction items, most students seemed to swap from the number line to a ‘rule based’ approach.

Kerslake (1981) also reports an integer item in the CSMS graphing test. Here students had to choose their own scales to plot (20, 15)(-14, 3) and (5, -12) on graph paper. She notes that while 43.7% of 13-year-olds were able to create a scale for this task, only 27.2% created correct axes. Some students had different scales on the positive and negative parts of the axis, others omitted zero or did not have zero at the origin, and some left no space for the negative numbers.

Discussion

That some students may ‘omit’ the origin (zero) when crossing from negative to positive in an addition problem is consistent with the research on students using the number line to represent additions or subtractions, where students are not sure ‘where to start’. The item developed for integers should therefore include an aspect requiring students to show their understanding of the significance of zero on a scale.

Overall, in this section on the use of scale as an object to represent numbers, issues have been found relating to representing whole number additions and subtractions, fractions, decimals, and integers on the number line. As identified by Dufour-Janvier et al. (1987) in the section on representations, in some cases these issues relate to an understanding of the symbols; in others, to understanding how the number line itself works. The ability of students to relate the number line to an understanding of measurement was also raised. As this literature review moves on to considering contextual applications of scale it is therefore appropriate to first turn our attention to the use of scale in measurement. Here scale is first considered in relation to linear measurement as it is the simplest form of measurement. The small amount of research found on developing students' understanding of other forms of measurement scales is reported at the end of the section.

3.7 Linear measurement

A sizable body of research examines the understanding children have of linear measure, that is, the concepts that students need to master if they are to develop an understanding of length and how we measure it. A selection is reported below. Each study reported highlights different aspects of interest, building a progressively richer picture of students' understanding of scale through linear measurement.

Piaget, Inhelder, and Szeminska (1960) identify that two different concepts come under the umbrella of linear measure; length and distance, but that there is a difference between the two. Length is about the linear size of objects, while distance is the linear separation between objects. When learning about length they identify that before children can construct a measurement unit, they need to understand the principle of conservation of length. This involves understanding that items do not change their size when their position is changed or when they are subdivided. For example, a child who identifies that two lines are equal when both have four matches but changes their answer when a teacher breaks a match does not understand this principle. Piaget et al.'s experiments identified that three quarters of children have developed this concept by the age 7½ to 8½. Some early errors were found to be that:

- things shrink/grow as they are moved;
- as long as the start and end point are aligned, items are the same length; and
- having more items, not necessarily of all the same size, means a longer length.

Overall, measurement concepts develop later than number concepts due to the difficulty of dividing a continuous whole into equal sub-units as opposed to counting a set of discrete objects. Piaget et al. also suggest that children do not necessarily understand the reading they take from a ruler, where the number read indicates the number of units from the start point (0) to where the number is written.

Hiebert (1984) also writes about learning to use a ruler. He identifies that Grade 1 children tend to focus on the endpoint when measuring. He suggests using a ruler with zero a short space from the end to encourage children to think about both endpoints, and that they need to see that each space, rather than a mark, represents a unit. "[C]hildren must deal directly with the fact that the length of an object is the number of units (spaces, not marks, on the ruler) between the beginning point and the endpoint" (p. 24).

In 1998, Nunes and Bryant also wrote on this topic, suggesting that "children can conceivably be taught to follow a procedure for reading measurements on a ruler and still have little understanding of the logic of measurement" (p. 86), a sentiment echoed by Lehrer and Schauble

(2000). Nunes and Bryant (1998) cite a study in Nunes et al. (1995) of 92 English first years looking at their understanding of rulers. They reported 89% of the children, when asked to place numbers on a ruler similar to the one used in class, started with the number 1, not 0. They suggest this happened as counting does not start from zero. They also raise the issues of what the children were counting – the centimetre gaps or the lines, and their understanding of the concept of a unit. The interviews with the children rarely provided an insight into children's thinking in relation to these questions.

In a second task reported by Nunes and Bryant, children were asked to look at rulers and see if they were drawn correctly. Two of the rulers did not have equal units; 67% on one and 56% on the second responded they were incorrect. A second difference involved the use of zero. With metric rulers, 71% said a ruler starting with zero was correct while 76% said a ruler starting at 1 was correct. Interviews indicated that 20 of the 22 children did not mind if zero was present or not. In an American longitudinal study over three years Lehrer, Jenkins, and Osana (1998) also found that most students in Grade 1 would choose an equal-interval ruler for a measuring task, but would use both the equal unit and the unequal units rulers and read from the end of the second ruler if they were measuring something longer than one of the rulers. However, Grade 3 students would typically use the correct ruler iteratively. They also found that 41% of the students in Grades 1, 2, and 3 counted from one when working out the length of a piece of wood aligned at the two inch mark of a ruler, and even by Grade 5 some 20% were still not able to use a zero point correctly.

Outhred and McPhail (2000) state that learning length concepts is particularly important in the development of an understanding of measurement "as length is usually the first measurement process that students learn about" (p. 487). They go on to identify a considerable body of evidence that shows many secondary students do not have a thorough knowledge of the concepts of length, area, and volume. This includes the 4th American NAEP report (Carpenter et al., 1988) and the children's mathematical frameworks of British researchers (Hart, Johnson, Brown, Dickson, & Clarkson, 1989). Outhred and McPhail also note that the Australian TIMSS report (Lokan, Ford, & Greenwood, 1996) identifies that linear measurement is an issue. Commonly reported errors involve inadequate understanding of:

- the attribute being measured;
- what to count when using informal units; and
- the use of measurement instruments.

For Outhred and McPhail (2000) to be able to measure successfully, students need an understanding of measurement attributes and processes, and to be aware of the key differences between number and measurement concepts, especially the notion of continuity.

In their conceptual framework Outhred and McPhail identify three important stages of development, reminiscent of those used by Piaget et al. (1960). At the simplest level (level 1) students first need to identify the attribute they are measuring before they can compare quantities or measure. At level 2 they need to know that the quantity is unchanged if it is rearranged (conservation); that lengths can be subdivided into equal parts (units); and that the number of units used gives a measure of quantity. At level 3 comes repetition (or iteration) of units – with no overlaps and no gaps, and the idea that units can be combined to make larger units. They identify that a key understanding of the measurement process is the repetition (or iteration) of the unit, citing research (e.g., Battista et al, 1998; Outhred & Mitchelmore, 2000) that suggests this concept is not often taught. The concepts of a unit, and its iteration, are particularly important for the current study into understanding of scale as the number line and all linear measurement scales are built upon these ideas.

Rulers and basic measurement concepts are also the focus of Bragg and Outhred (2000a, 2000b). Their Australian study investigated the understandings that 120 Grade 1 to 5 primary students from low socio-economic status (SES) schools had of these. They found that by Year 5, while most students could use a ruler to measure a line, their understanding of the nature of units was poor. Many cited procedures for measuring like counting or ruler alignment, while only the more able had an understanding of the concept of measurement. Interviews indicated that these students had developed this understanding for themselves and were not taught it. Lehrer et al. (1998) would concur: "conventional instruction is not fostering the development of critical ideas in measure" and that "children's use of tools like rulers, especially in the primary grades, often masked (and overestimated) their understanding of foundational ideas in measure, like those involving iteration, identical units, and space filling" (p. 164).

Bragg and Outhred (2000a, 2000b) identified that only 19 of 72 students in Grades 3 to 5 could indicate that the number 5 on a ruler referred to 5cm and show a line 1cm long for a unit. When prompted, 54% counted along the marks, with some unable to resolve the conflict arising when they counted six marks. Twelve students (17%) counted the spaces, not the marks. Some simply identified that 5 indicated 'where the line ends'. Bragg and Outhred (2000a) further indicate that students of different ability had different success rates. For example, low ability students, although aware of the features of a ruler, were unable to maintain the length of the unit. In another task, low ability students had more ruler placement errors (whether to start at the end of the ruler, or zero, or one) and these persisted into Grades 4 and 5. Bragg and Outhred (2000a) claim that their results imply that "teachers focus on the techniques for using a ruler rather than on an understanding of how a ruler is constructed" (p. 117), causing little understanding to be developed. Furthermore, "such techniques may not provide the basis for developing ideas of units and scales or a framework for later measurement concepts" (Bragg & Outhred, 2000b, p. 97). They identify several critical understandings for learning the concept of measurement:

- the linking of the process of measuring to the construction of a scale, along with the role of zero on this scale;
- the distinction between counting discrete objects and subdividing a continuous item into units; and
- the nature of the unit and the structure and precision of unit iteration.

All of these relate to common misunderstandings held by students.

Irwin and Ell (2002) look at the move from informal to formal linear measurement. Their New Zealand study of 43 eight and nine-year-olds identified that many students did not understand the unwritten rules of measurement, and that increasing age did not seem to improve understanding, although teaching did. When measuring with informal units some students were observed to overlap objects to get a whole number count, leave gaps between objects, omit the first laid object from the count, and were unsure about how to deal with objects that extended beyond the length being measured (fractional measures). Some students also visualised the ruler, that is imagined the ruler in their head (often very accurately) rather than use a physical unit. On a task involving the extension of a partially drawn ruler, less than half the students in each age group showed they understood that the first unit was the same size as other units and started their ruler at zero. Only half of the students in each age group "drew adequately regular units" (p. 362). However, seven of the 15 students who started the ruler at zero were from the same class (although others in this class still made common errors). Finally, Irwin and Ell report that accurate measurement was not a task regularly undertaken by these students outside the classroom, with only one mentioning a measurement situation in which precision was important. Linear measurement was a school task, rarely reinforced in real life. In a Nashville study of 16 Grade 1 students, Stephan, Cobb, Gravemeijer, & Estes (2001) also found that when using their feet to measure the length of a rug, some students could omit to count the first placed foot. Likewise the students had problems if the last placement did not involve the use of a whole foot, giving answers like '14 and a little bit', '14 (and by turning their foot sideways) 15', or '14 and a half'. However both of these issues were remedied through teaching.

In commenting on the results of the CSMS measurement study Hart and Johnson (1980) provide a further indication of a range of issues associated with learning to measure.

> Some errors appeared to be committed by large numbers of children independent of type of school and textbooks or teaching schemes. In addition many of these regularly occurring errors were not restricted to any particular age group and some (as seen from the longitudinal study) persisted from year to year. (p. 146)
>
> In general the inappropriate strategies i e those which were limited in generalisability or were incorrect were characterised by a desire to
> (a) work within the set of whole numbers
> (b) use counting or addition

> The errors often indicated a lack of appreciation of the nature of measurement or the essence of numbers other than whole numbers (fractions or decimals), so that inappropriate rules (not necessarily ones that had been taught) were applied. (p. 149)

However, this CSMS study only focused on the measurement of length, perimeter, area, and volume, all of which are based on linear measure. Other studies also seem only to focus on such attributes (e.g., Costello, 1991; Lehrer, et al., 2000; Outhred & McPhail, 2000). The final paragraph in this section therefore reports on issues related to using scale to measure other attributes, focusing on the assessment items used with Year 8 students in the NEMP studies.

Reading a weighing scale where the arrow pointed to a labelled whole number caused few problems (Crooks & Flockton, 2002). About three quarters of students could locate 28°C on the unit scale of a thermometer where only multiples of five were labelled (Flockton & Crooks, 1998). Success rates dropped to under 50% where units were labelled but sub-divided. Working between units, 40% were successful on a weighing scale marked in 0.2kg increments, and 32% were successful with a ruler marked in 0.2cm intervals (Crooks & Flockton, 2002). (Note that the NEMP studies merely provide results, rather than attempting to identify causality for these.)

3.7.1 Discussion

The research on linear measurement suggests that learning to measure is not easy, and needs time and deliberation. It brings to notice the following issues for the development of an understanding of scale through linear measurement:

- the nature of a unit – as an interval – may not be well understood;
- some students may not understand that units need to be the same length;
- some may not have the skills to create equal units;
- some may not understand the importance of unit iteration – no gaps and no overlaps;
- some may not have developed an understanding of what a scale is; and
- the importance of, and the role of zero on measurement scales may not be understood.

The concept of a *measurement scale* (first developed through learning to measure length) also seems not to be a common object of study. Thus the use of linear scales to measure attributes other than length (e.g., weighing scales for weight, thermometers and oven thermostats for temperature, and car speedometers for speed) has not been given much emphasis in this chapter. This is in spite of the fact that these scales are often more complex than those used for linear measurement, so could present students with additional challenges.

Finally, the research on linear measure also indicates that some students may confuse the processes of counting and measuring. The early measurement tasks from the research reported above indicate a focus on unit counting. Such a process starts at one, so in early measurement students may be simply applying their knowledge of counting discrete objects rather than developing an understanding of measurement, and the nature of a scale. Measuring and counting confusion was also identified by Carr and Katterns (1984) in relation to using number lines to represent additions and subtractions, and by Bright, et al., (1988), Siegal and Smith (1997), and Mitchell (2005) in relation to using the number line to represent fractions. That the possible existence of such confusion has been found in a number of mathematical domains in which scale is applied suggests it is important to clearly introduce students to the conventions of measurement when the concept of a scale is first introduced. Furthermore, the findings indicate that items to assess understanding of the conventions used on a scale are important to include in the diagnostic assessment developed for this study.

In the next section the emphasis moves from the measurement applications of scale to the graphical applications of scale. Graphs have different purposes. They can be used within mathematics to explore the nature of an algebraic function, in which case like number lines they are decontextualised. Graphs can also be used as a representation of observations and data collected from the real world, as happens in statistics and the sciences. In these cases the graphs can be considered to be context based or even context dependent.

3.8 Graphing

As noted by Leinhardt et al. (1990, p. 19) "[a] graph cannot be interpreted fully without taking into account its scales." However, the importance of understanding scale when graphing has not always been recognised. Goldenberg (1988) reports "the near-total lack of literature on scale issues" (p. 151) when data identified that he needed to focus attention on scale for a study on the graphical representation of functions. Leinhardt et al. (1990) also report "[s]cale is an issue when using graphs for scientific data analysis and in computer-based instruction", but is not found to be so when introducing graphing in mathematics classes, possibly because "the scale is often assumed or given in mathematics instruction" (p. 17). Leinhardt et al. also note that "[m]any studies take for granted students' ability to construct the axes of the coordinate system. … Yet there is evidence that the construction of axes requires a rather sophisticated set of knowledge and skills" (p. 43). The next sections therefore report research on scales in the context of graphing, and to be is consistent with Figure 2.2, separate the more abstract approach to the algebraic graphing of functions from the contextual approach used in statistics.

3.8.1 Algebraic graphing

The CSMS graphing study is reported in Kerslake (1981). Tests were given to 459 13-year-olds, 755 14-year-olds and 584 15-year olds. Students had little difficulty with block graphs and using rectangular coordinates to plot integers (around 90% were successful). However, when the task involved joining several points plotted on a unit scale to make a line, or writing the coordinates for more points on the line, the success rates of students dropped dramatically. In an item to gauge students' understanding of infinity, when asked how many points lay on the line only 6% of 13- and 14-year-olds and 20% of 15-year-olds could give answers like ‘an infinite number’ or ‘as many as you like’. Most students provided a finite number. When the CSMS team established item difficulty, items on continuity and infinity were excluded, as some students could answer them without being successful at items deemed easier (i.e., with higher success rates).

The CSMS graphing test also included the plotting of fractional and decimal coordinates on a graph with unit scales and tested students’ understanding of graph scales. Here the fractional item was deemed to be straightforward so was placed at facility level 1, with around 80% of each year group plotting (1½, 4). For decimals, only 36% of 13-year-olds were able to approximately plot the point (4.6, 10.2). Some students interpreted the question as requiring two points to be plotted – (4, 6) and (10, 2). Another item on understanding the impact of changing a scale used three graphs, each showing a line. Two lines appeared identical, but one had a unit scale on the vertical axis, the other a scale in twos. The third graph also had a scale in twos, but with the line drawn correctly. Students were asked which two graphs showed the same information; 46.5% of 13-year-olds were successful with the item, placing it at level 2, with a similar success rate to items on interpolation. Overall, graphs in contexts were no easier to interpret than graphs related to abstract equations.

Goldenberg (1988) studied the graphical representation of functions, particularly in relation to computer and calculator use. He argues that scale affects visual perception so, for example, on a computer window shape and slope are artefacts of scale. Therefore, to use such technology effectively, the interaction between scale and shape need to be understood. He concludes that it is important for students to control scale and learn to deal with notions of continuity, of discrete points, and the infinitely large and small much earlier than is typically done. Goldenberg also reports the repeat of part of an experiment by Shultz, Clement, and Mokros (1986) on drawing and interpreting graphs of familiar phenomena like a bicycle going over a hill. This showed the small number of errors fitted a pattern and confirmed that scale issues were confusing to students. For the horizontal axis of the non-conventional representations, an ordinal scale for time was used generally by the students, but without evenly spaced intervals. For the vertical axis the data were “often purely nominal, listing events that took place over time but not treating them even as an ordinal scale, let alone an interval scale” (Goldenberg (1988, p. 149).

Dunham and Osborne (1991) talk of graphing in the context of American high school students studying pre-calculus. They identify that scales cause significant difficulties for high school students. For example, when reading from a graph “[w]e were continually amazed by the difficulties of students in extracting from an ordered pair (X_i, Y_i) the corresponding *x*-value, *y*-value or *f*(*x*)-value while operating from a graph” (p. 38). “The understanding and interpretation of scale is a persistent problem identified in recent studies of graphing ability” (p. 42). They also identify that science teachers report that students need instruction using asymmetric scales (that is different scales on different axes) and in choosing appropriate scales to make good use of graphing space. The core of the article addresses understanding of graphing, reporting interviews with 10 students. As Kerslake (1981) found, the interviews revealed "scale is not something that students notice regularly; rather they assume that most graphs are scaled on single unit intervals with equal scaling on both axes" (Dunham & Osborne, p. 42). The students also showed little understanding of stretching or shrinking a scale.

Discussion

These studies suggest that many older students do not have the understanding to work successfully with scales in an algebraic context. They add to the body of evidence suggesting that scales in general are poorly understood. The next section will continue to examine understanding of scale, but in the context of statistics.

3.8.2 Statistical graphing

Costello (1991) writing in England, McGatha, Cobb, and McClain (1998) in the US, and Nisbet (2003) in Australia indicate there has been a shortage of research into students' statistical understanding. Costello identifies that the shortage in England was due to relative neglect of statistics until the English national curriculum included three statistics-related attainment targets. The neglect was reflected by the absence of written material on statistical understanding relevant to secondary schools and the paucity of articles in the journals *Mathematics Teaching* and *Mathematics in School*. McGatha, at al. write:

> (T)here are actually only a handful of studies available that focus on students' statistical understandings. These studies fall into two categories: (1) studies that examine students' understanding of the mean and (2) studies that examine students' understandings in the context of data analysis. (pp. 2-3)

Here the second set of studies typically "outline the process by which students analyzed and reasoned about data in more innovative instructional approaches" (p. 3). Nisbett (2003) cites Konold and Higgins (2000), outlining the reason for the dearth of research being that many countries have only recently begun to teach data analysis at pre-tertiary levels, so investigating the reasoning of students at this level has not been a priority.

McGatha et al. (1998) describe a pilot study to identify students' existing understanding and ways of reasoning about statistics. Involving 30 seventh graders working in groups, the study included one task which looked at understanding the graphical representation of data. In the task, students were asked to summarise the results of a survey on the TV watching habits of 30 students so that, when the summary was put on a bulletin board, it was quick for parents to understand. The students had two graphical approaches. One was to treat the data as a set of points, the other was to create categories. In both, the data were seen as individual points that could be ranked, ordered and/or categorised, rather than as a global whole distributed along a continuum. The response of one group, and the researchers' analysis, illustrate this. The students started by grouping the data 0-10, 10-20, 20-30. One student then noticed the biggest number was 23 (actually it was 23.5) so the end group was changed to 20-25. During graphing, the grouping was again changed, to 0-10, 11-20, 21-25.

> Even though the graph is reminiscent of a histogram, we believe that for this group it was a bar graph over intervals. Changing the upper bound from 30 to 25 indicates that this group of students did not view this data as positioned within a space of potential data values. The way the group created the categories as described above indicates that they were partitioning the data items rather than creating intervals that fell along a continuum of potential data values. The spaces between the bars also indicates that each category was a separate subset of the data rather than intervals in the space of possible data values. (p. 11)

Nisbet (2003) reports on an item used with 43 Year 8 Australian students. They were asked to organise and represent a set of 50 data points; 30% did not reorganise the data at all while only 21% created groups with convenient intervals, but often with duplicated end points. Students were comfortable with this. Only 14% created a valid histogram (considered the appropriate graph) and 26% did not draw a graph at all.

In New Zealand, NEMP provides evidence of student ability to read graph scales. Year 8 students have little difficulty with a unit scale, or, if multiples were involved, where the number could be read directly from the scale (Crooks & Flockton, 1996, 2000; Flockton, Crooks, & Gilmore, 2004). In many problems gridlines were shown on the graph, ensuring an accurate reading. Several items, potentially requiring students to interpolate, had generous marking schedules allowing for reading to the closest labelled number (Crooks & Flockton, 2000). Only one item required accurate reading and working between the gridlines and the marked scale. Here only 58% of students were successful (Flockton et al., 2004).

NEMP also provides some evidence on students' ability to draw graphs. In general, the results showed that success rates were lower when students were working between gridlines. Over 85% could draw a bar or a point on a graph when working to a gridline or a labelled number (Crooks & Flockton, 1996, 2000; Flockton et al., 2004). Slightly fewer (80% or less) were successful when placing a point or drawing a bar between gridlines, even when this was only half-way between them (Crooks & Flockton, 1996, 2000). In other items, around 50% of students found working with tasks involving choosing or providing a scale difficult, even when all the points were common multiples (Crooks & Flockton, 1996). They were more successful with similar tasks when the scale had been started, or was provided (Crooks & Flockton, 2000).

A CSMS statistical item is reported in Kerslake (1981). Here a scattergram was used to show height and waist measurements. Students were expected to read several points, and mark one point. As NEMP found, around 90% of 13 to 15-year-olds successfully worked to and from gridlines (marked in tens), but when working between the gridlines (interpolating) this dropped to around 60% for the 13 and 14-year-olds. This placed interpolation at a lower facility level (level 2) than the other items (level 1).

When seeking to identify critical factors influencing graph comprehension, Friel, Curcio, and Bright (2001) examined research from a number of fields. They defined graph comprehension as the readers' ability to derive meaning from graphs created by themselves and others, and concluded that research identifies three components common to the different disciplines. The first of these is relevant to this study: "to read information directly from a graph, one must understand the conventions of graph design" (p. 152). Here they identify:

- "the *framework* of a graph (the axes, scales, grids, reference markings) give information about the kinds of measurements being used and the data being measured" (p. 126);
- the simplest framework involves an L shape with one leg representing the data being measured and the other providing information about the measurements being used;
- visual dimensions called *specifiers* (e.g., the bars on a bar graph) are used to represent data on the framework;
- reading the data is the most elementary skill in extracting information from a graph, and students find this easier than interpreting the data or reading between the data;
- "[s]tudents experience few difficulties with "read the data" questions" (p. 131), but can make errors relating to mathematical knowledge, language and reading, scale, and reading the axes; and
- students have issues related to *data reduction* – the organisation and grouping of data for graphing (and may not be able to choose an appropriate scale even when they can draw and read a given scale).

They further identify that research highlights the importance of context, mathematical knowledge, and experience for graph comprehension. "Those studying graph comprehension may need to consider the development of number knowledge" (p. 144).

Friel et al. (2001) also provide guidelines for school instruction on graphs. They suggest several elements need to be considered: sequencing the types of graph; developing an understanding of data reduction; and developing aspects of graph sense. In Grade 2, object graphs, picture graphs, line plots, and bar graphs with gridlines and unit scales are suggested for introducing students to the concept of graphing. Physical objects should be used initially before moving to representations of these and the more abstract graphical representations found in line plots and bar graphs. At Grades 3-5, bar graphs should continue to be used, with stem plots and pie graphs added. Scales should be introduced and used, but with care as the move from graphs with no data reduction to those with reduced data "may be confusing for students if they are not given opportunities to explore the transition" (p. 148). Initially Friel et al. suggest only exploring scaling on the frequency axis.

Other research into the ability to read and draw graphs is quite specific in stating that students have little difficulty in the literal reading of graphs (e.g., Nisbet, 2002; Pereira-Mendoza & Mellor, 1991; Sharma, 2003). Sharma (2003, p. 615) states that "many students can read tables and graphs" and cites other research to support this claim. Nisbet (2002) identifies "[a]ll the children appeared to be very familiar with the mechanics and conventions of drawing simple graphs" (p. 525). Pereira-Mendoza and Mellor (1991) cite a number of earlier studies indicating "that while students had few difficulties with the literal reading of graphs, they were often unsuccessful in answering questions requiring higher level cognitive skills" (p. 152). Their own Canadian research confirms this: 94% of their Grade 4 and 98% of their Grade 6 students were successful with the literal reading of graphs.

Identifying the nature of the problems asked by Sharma (2003), Nisbett (2002), and Pereira-Mendoza and Mellor (1991) was not always possible. Sharma did not provide the graph(s) or information that identifies the nature of the scales used. Nisbet did so, but these only required the Year 7 students to work with a unit scale (which explains the ease with which the students handled the graphs). Pereira-Mendoza and Mellor provide an exemplar bar graph (one of 12 graphs used) showing a multiples scale and gridlines. To answer correctly, students had to read along the gridline to the labelled number. However, Pereira-Mendoza and Mellor appear to contradict themselves about the ease with which students find the graph reading tasks when later in their article they state that "students make errors when faced with variable scales within the context of different graphical problems" (p. 151). They also state that in 'interpretive' tasks scale errors were one reason for lower success rates (52% for Grade 4 and 78% for Grade 6). Overall, the evidence from these three papers suggests that the impact of scale, in particular of working between marks or gridlines, may not have been considered in some of this research.

Finally in this section it is important to note that some literature acknowledges that statistics is based on measurement. For example, Stevens (1946) identified that different types of measurements exist, and that the nature of these measurements influence what statistics are appropriate and meaningful to use when analysing data. Lehrer and Schauble (2000) identify that "measure, broadly conceived, is core to children's understanding of data" (p. 110). The studies reported earlier in the section also allude to this link. For example, McGatha et al. (1998) and Nisbet (2003) indicate that when graphing the students in their studies had a tendency to treat numerical data as categories rather than measurements. Friel et al. (2001) talk of graphing as if it is a measurement process. This alignment of scale with measurement in the context of statistics is consistent with how scale is applied in algebraic graphing.

3.8.3 Developing an understanding of scale in the context of graphing

In this section the work of Moritz (2003), as referenced in Chapter 10, is addressed. Moritz suggests coordinate graphs are built on the general graphic feature "position represents value" (p. 228). Coordinate graphs apply this feature in two dimensional space to represent values of two variables. Moritz worked with 133 students from Grades 3, 5, 7, and 9 in two private schools in Tasmania; he used the SOLO taxonomy (Biggs & Collis, 1982, cited in Moritz, 2003) to develop four response levels for his task. This involved drawing a graph for a table of temperature change data over time. (The Biggs and Collis taxonomy is described by the authors as an assessment framework, and a general theory of development.)

Table 3.1: SOLO taxonomy and Moritz' response levels

Level	SOLO taxonomy	Level	Response levels (Moritz)
1	Prestructural responses not including relevant elements.	0	Non-statistical responses, where a picture or narrative shows the context, or axes and lines are drawn with some irrelevant values.
2	Unistructural responses using one relevant element.	1	Single aspect responses, with a table showing corresponding values or where values showing a single variable were provided (as on a Centigrade thermometer).
3	Multistructural responses using multiple relevant elements without integration.	2	Inadequate coordinate responses, where graphs show both variables but lack spatial variation (e.g., the results form a diagonal or lack appropriate scales), or lack correspondence (e.g., the sets of data are both graphed separately).
4	Relational responses appropriately relating relevant elements for the task.	3	Appropriate coordinate responses, where the graph shows what it should.

Discussion

The research into statistical graphing suggests that the ability of students to work with statistical graphs depends on the task's sophistication. For example, it is likely that reading unit scales causes few problems, but when working with multiples, especially when creating the scale or working between the marks, fewer students seem to be successful. This is consistent with Swan (1983) and Crooks and Flockton (2002) who report lower success rates for assessment items in which the unit is partitioned.

The difficulty students had creating a scale is of interest. Evidence from science mentioned in Dunham and Osborne (1991) and the statistical graphs in Crooks and Flockton (1996) suggests a pattern in which creating an appropriate graph scale is generally problematic. Student difficulty in choosing a scale to accommodate a random set of whole numbers could be worthy of investigation.

Finally, the response levels of Moritz (2003) suggest that the quality of the end product of a graphing activity is closely related to the quality of the scales students can produce. This in turn suggests that it is important for the contextual items to include graphical situations for assessing scale understanding.

Overall, the research reviewed in this chapter makes it clear that learning about scale is far from straightforward. However, before summarising it is appropriate to recognise an important theme in the literature as this theme may be observable in students' responses to the assessment developed to measure scale understanding. This relates to the confusion some students seem to have between counting and measuring.

3.9 Counting and measuring

The research outlined in this chapter identifies that scale aligns with measurement, that there is a set of issues around understanding measurement, and that the use of counting approaches with scale is not uncommon. This theme suggests that it may be of value to identify clearly the distinctions between counting and measuring. For example, how do we go about the process of counting a collection of objects? What do we emphasise as important, and what do we ignore? In comparison, how do we go about measuring the width of a doorway? The research on the use of representations clearly indicates that students need to know the conventions or rules of a representation if they are to use it successfully. Table 3.2 therefore identifies a number of differences between the two tasks. Linear measure is used because other forms of measurement scale build upon the understanding developed within linear measure and little was found about developing an understanding of other sorts of scale.

Table 3.2: Differences between counting and measuring

Counting	**(Linear) measure**
• In 1:1 counting, each object is given a number from the number sequence. Paying attention to each object is an important part of the process.	• Objects are ignored. Rather, the size of the space the object takes up (the interval) is considered.
• Objects can be different (sizes, shapes etc).	• To measure, a unit must be chosen or created, then used repeatedly to make the count.
• Spatial separation irrelevant.	• Units must be laid end to end (no gaps and no overlaps).
• Highest number reached is the total of the count. (Absolute in nature).	• Total of the measure is the size of the gap between the start and the finish. (Comparative in nature).
• Starts at one.	• Starts at zero.

It is possible to propose other differences of note. For example, counting can be said to be discrete, while linear measure is continuous. However it is possible to count 2½ apples!

3.10 Summary of findings

The literature review has identified that as a number line a scale can be used to represent numbers. To use scale effectively, students therefore need to have an understanding of the number sets that are used with a scale. Scale is also applied in measurement and graphical contexts, so students also need to recognise that scales exist in a variety of situations and to consistently apply knowledge about how a scale works. Overall, this suggests complexity in that students need to develop an understanding of scale as a representational object as well as learn the processes involved when the object is used. However a recurring theme throughout the literature is the problems some students have in understanding and interpreting scales and the time needed to develop a full understanding of scale.

The review identified the following, and in doing so indicates some directions to be explored when seeking to investigate scale understanding.

1) Scales are based on measurement principles.
    - Some students seem not to understand measurement, instead bringing their understanding of counting to measurement tasks.
2) As scales are based on measurement principles they assume many properties of numbers that may need to be developed for novices.
    - The existence and role of zero may not be understood.
    - Where fractions reside on the scale in relation to whole numbers may not be understood.
    - The continuous nature of the scale may not be understood.
3) It is possible that some students are not performing well with scales as they do not understand the conventions of the scale construct, so do not understand how the marks and other features convey information.
4) There appear to be issues related to students' understanding of non-unit scales.
5) Simple exposure and increasing age are apparently insufficient to create an understanding of scale.
6) The value of contexts when teaching and assessing scale understanding is worthy of investigation.
    - In doing this it is important to use contexts that should be familiar to students.
    - As student knowledge of linear measurement, and the non-unit scales commonly used for other measurement scales and in graphing may be weak, these may not be a reliable starting point for attempting to improve understanding of scale.

To close this chapter, it is worth noting that no literature was found that covers in a single study all of the research on scale to develop a global picture of student understanding. Issues in understanding scale were addressed within separate mathematical domains rather than across domains. This suggests that the topic of scale and the research question chosen for this study are worthy of investigation and may develop new knowledge. In the next chapter the literature review continues, exploring what the curriculum says students should have learned about scale by Year 7, and how various mathematics resources go about imparting this.

Chapter 4

Document Analysis

4.1 Introduction

This chapter is in two parts. The first part reviews four New Zealand curriculum documents current in 2005: mathematics; science; technology; and social studies. When talking about student understanding of linear measurement, Outhred and McPhail (2000) suggest "students' misconceptions may simply reflect teachers' inadequate measurement concepts, combined with syllabus documents that give little direction in terms of key concepts" (p. 493). Consequently, this curriculum review seeks to identify the understanding of scale that students entering Year 7 should have, and the quality of direction provided to teachers by the curriculum documents. In doing this, an answer will be formed to the research sub-question "what should Year 7 students know about and be able to do with scales?"

The second part of the chapter reviews some common primary school resources for the teaching of mathematics. This resource review seeks to provide an indication of students' level of exposure to scales and the sorts of activity common in Years 1 to 6. This is not to suggest that students in the study schools have been taught using these resources. Rather, it is assumed that the students come from a number of feeder schools, which use different resources in different ways, so their prior experience is unknown. However, such a resource review can indicate the level of support and direction generally available to teachers when teaching scale, and can help form an answer to the research sub-question "How do commonly available mathematics resources present the learning of scale?" The results will also provide triangulation for the findings of the curriculum review.

4.2 Curriculum review

4.2.1 *Mathematics in the New Zealand Curriculum* (Ministry of Education, 1992)

This section focuses on the mathematics curriculum document current at the time of this study (MiNZC). MiNZC splits mathematics into six *strands*, each with eight *Levels*, and outlines for teachers what students are expected to be able to do as they move from Year 0 (age 5) to Year 13 (ages 17 to 18). In early to middle schooling (Levels 1-5) each level is designed to take an average student about two years to complete; Levels 6-8 correspond both with the senior secondary school Years 11 to 13 as well as Levels 1-3 of the National Qualifications Framework. This analysis of MiNZC only considers curriculum Levels 1-4, as the study focuses on students in Years 7 and 8. MiNZC identifies that most of these students are expected to be working on a programme based on curriculum Level 4 (CL4) and below. Some gifted students may be working at CL5 (p. 17).

Within MiNZC, *mathematical processes* is the first curriculum strand to be detailed. It is cross-strand in that it is intended to be integrated into the delivery of the other five strands: *number, measurement, geometry, algebra,* and *statistics.* Each strand has its own achievement aims which outline students' learning opportunities within the strand. Content is then divided into *achievement objectives* (AOs) and *suggested learning experiences* (SLEs). AOs state what a student should be able to do after appropriate learning, while SLEs expand upon the AOs, address concepts that take time to develop, and contribute to the achievement of the broader curriculum aims (Ministry of Education, 1992).

In the years prior to this study, curriculum implementation was influenced by ERO; school reviewers looking at planning expected to see the programme of learning developed from the AOs in the curriculum statement (ERO, 2000, 2002). This has meant the material covered in the SLEs, especially where it does not match the AOs, was emphasised less in teaching. Consequently most of this analysis will centre on the AOs, as these are more likely to have formed the basis of work undertaken by students prior to Year 7.

The following subsections deal respectively with number lines, measurement scales, algebraic graph scales, statistical graph scales, and map scales. A brief discussion follows each detailed section of what students should be able to do.

Number lines

One aim of the number strand is to "develop an understanding of numbers, the ways they are represented, and the quantities for which they stand" (MiNZC, pp. 9 & 31). However, only two possible references to number lines can be found in the strand. These are SLEs rather than AOs and relate to decimals and integers, not whole numbers; therefore it is likely that the number line representation may not have been explicitly developed in classroom programmes.

Discussion

The omission of learning based on number lines is consistent with the emphasis of learning in MiNZC being based "within a range of meaningful contexts" (p. 26), and in this thesis is reflected in Figure 2.2, where number lines are identified as abstract applications of scale.

Measurement scales

One aim of the measurement strand is for students to "develop confidence and competence in using instruments and measuring devices" (MiNZC, pp. 9 & 57). Eight AOs relating to scales and a large number of SLEs support this. By the end of Year 2 (CL1), students should be able to order and compare lengths, describe the comparisons using mathematical language, and measure by counting non-standard units. By the end of Year 4 (CL2), it is expected that students can "carry out practical measuring tasks, using appropriate metric units for length, mass, and capacity" (p. 62). By the end of Year 6 (CL3), students are expected to "perform measuring tasks, using a range of units and scales" (p. 66). By Year 8 (CL4), they should be able to:

- carry out measuring tasks involving reading scales to the nearest graduation;
- read and construct a variety of scales, timetables and charts;
- design and use a simple scale to measure qualitative data. (p. 70)

Discussion

A commonly used conceptual model for developing measurement (Zevenbergen, Dole, & Wright, 2004) includes most of the stages outlined above, and provides additional detail.

1) Identify the attribute (what it is being measured).
2) Compare and order a set of items (seriation) – initially by direct comparison, later by indirect comparison.
3) Informal measurement using non-standard units – to give the idea that different units can be used in different contexts, and to create the need for common units.
4) Formal measurement using standard units – including a sense of the basic unit and appropriate language and symbols to describe it, the size of familiar referents, and estimation in relation to the unit.
5) Application to problem-solving contexts – involving the development of formulae.

The mathematics exemplars also expand on the AOs. "A mathematics exemplar is an annotated sample of student work produced in response to a set task" (Ministry of Education, 2001g, p. 1). These exemplars illustrate the major content areas and 'big ideas' in MiNZC, exemplify the major cognitive advances as students learn about a topic, and provide material to help teachers assess the level of student achievement. They are based on "authentic examples of work chosen from thousands of student work samples collected from schools throughout New Zealand" (Ministry of Education, 2001g, p. 1). The linear measurement exemplar links the stages of development visible in student work to the curriculum levels as follows:

Level 1i	use direct comparison – can identify the attribute to be measured and can compare objects by placing them next to each other.
Level 1ii	use indirect comparison – can use a third object (like a piece of string or an armspan) to decide which is of two items is longer.
Level 1iii	use non-standard measurement units repeatedly – can choose an appropriately sized unit and measures by counting the total number of units required.
Level 2	use standard measurement units – are familiar with the size of standard metric units, can accurately use an appropriate measuring tool by lining up the end with the start of the object, can read the scale and state the measurement with the correct units.

Level 3 use reasoned measurement – identify the nature of the measurement task, produce an initial estimate and a strategy to check the estimate, then use appropriate tools and accuracy to complete the task.

In MiNZC, there is a strong emphasis on developing the ability to use measurement, beginning at an early age. Over time, measurement with greater accuracy is also required. When read alongside the number strand and the exemplar, MiNZC indicates that students entering Year 7 should be able to read measurement scales involving both whole numbers and decimals. However, MiNZC seems to require students to 'leap' from counting non-standard units to being able to use measuring instruments like rulers and sets of weighing scales without first developing the concept of a scale. This omission does not give students the opportunity to develop an understanding of measurement conventions – like the role of zero on a scale, or the concept of a unit as an interval. Yet these conventions were identified by research as important in developing an understanding of linear measurement (Bragg & Outhred, 2000a, 2000b; Nunes & Bryant, 1998; Piaget et al., 1960) so this omission could lead students to having problems with scales.

Algebraic graph scales

In the algebra strand, one aim is to "develop the ability to think abstractly and to use symbols, notation, and graphs and diagrams to represent and communicate mathematical relationships, concepts and generalisations" (MiNZC, pp. 10 & 129). For graphs, three AOs support this, one each at CLs 2, 3, and 4. At CL2, students are expected to "use graphs to illustrate relationships" (p. 134), while at CL3 they "use graphs to represent number, or informal, relations" (p. 138). By CL4 students are moving on to "sketch and interpret graphs on whole number grids which represent simple everyday situations" (p. 142). A number of SLEs support these AOs at each level.

Discussion

The above requirements suggest that by Year 7, students should be familiar with graphs in the algebra strand, and be able to both create and interpret graphs that show relationships. The examples of graphs at all three levels indicate the use of qualitative graphs on which time is commonly used for the horizontal axis (e.g., Figure 4.1). The traditional approach of plotting points on a scaled Cartesian coordinate system is avoided. This qualitative approach follows the research-based suggestion of the time outlined in Leinhardt et al. (1990). They identify that, according to Bell and Janvier (1981, cited in Leinhardt et al., 1990), Janvier was amongst the first to encourage the use of qualitative graphs of concrete situations that are interpreted globally, the aim being for students to attend to the entire graph as a relationship between two changing variables that are described using words rather than numbers. The approach was suggested as a way to improve students' understanding of the concept of function in which global features like slope are attended to, rather than developing a pointwise interpretation (which was seen to be problematic), and seemed to be encouraged by the traditional quantitative approaches related to point plotting. Given that MiNZC has been the gazetted curriculum document since 1993, it seems possible that students may not have developed the concept of scale defined for this study from their work with graphs in the algebra strand of MiNZC.

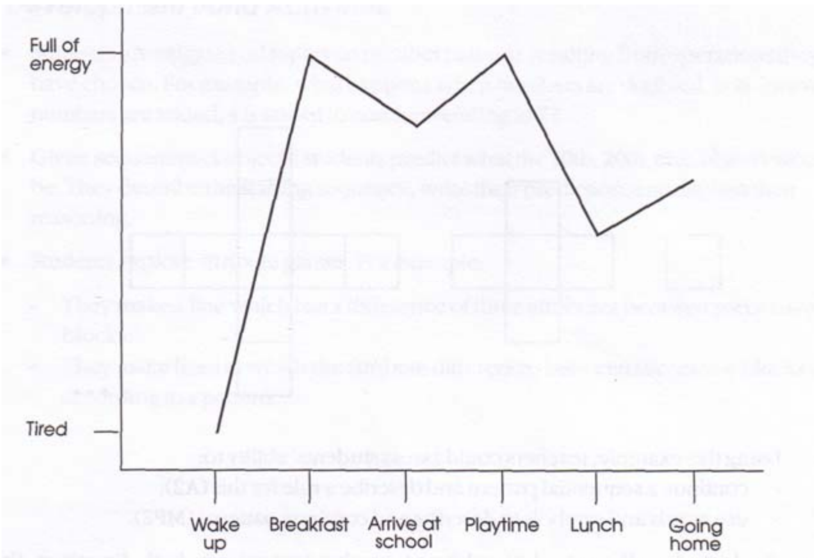


Figure 4.1. Example of a CL2 graph from the algebra strand (MiNZC, p. 135)

Statistical graph scales

In the statistics strand, two of the three aims relate to reading and drawing graphs.

- recognise appropriate statistical data for collection, and develop the skills of collecting, organising and analysing data, and presenting reports and summaries;
- interpret data presented in charts, tables, and graphs of various kinds. (MiNZC, pp. 10 & 169)

These are supported by seven AOs, three at CL2, one at CL3, and three at CL4. At CL2 students are expected to:

- collect and display category data and whole number data in pictograms, tally charts, and bar charts, as appropriate;
- talk about features of their own data displays;
- make sensible statements about the situation represented by a statistical data display drawn by others. (MiNZC, p. 174)

At CL3 students are expected to "collect and display discrete numeric data in stem-and-leaf graphs, dot plots, and strip graphs, as appropriate" (MiNZC, p. 178). At CL4 they:

- choose and construct quality data displays (frequency tables, bar charts, and histograms) to communicate significant features in measurement data;
- collect and display time-series data;
- estimate the relative frequencies of results and mark them on a scale. (MiNZC, p. 182)

SLEs support and expand on these AOs. For example, one CL2 SLE talks of students using pictograms, block graphs, bar graphs, stem-and-leaf graphs, and tally charts.

Discussion

At CL2, students are drawing graphs for both categoric and whole number data with pictograms, tally charts, and bar graphs. The SLE identified above extends this to other forms of display. What scales should be involved are not identified. The data display exemplar (Ministry of Education, 2001g) indicates that CL2 students "are able to group and organise category data and whole number data and present it in an appropriate format, such as a tally chart or bar graph, complying with standard conventions" (page not numbered). Student work shows unit scales with measurement conventions on both axes, so the development of non-unit scales is probably not intended.

Zevenbergen et al. (2004) provide a model for developing graphing that can be used to review the progression provided by MiNZC. They indicate that the contemporary approach in statistics is to use "a scaffolding approach where the need for various components emerges from problematising the process" (p. 293), rather than a more formal *x*-axis, *y*-axis process. The challenge for teachers is to progressively remove the concrete evidence and introduce the need for labelling. Zevenbergen et al. outline the following stages:

- Start with simple picture graphs that strongly relate to the students' personal experiences. For example, using square post-its on a baseboard, with students drawing on the post-its. Here a need for a common start line and even spacing is evident.
- Replace the pictures with squares, which creates a need for an *x*-axis label.
- Transfer to grid paper. Here the graph initially needs an *x*-axis label, but as data numbers increase, the need for a *y*-axis label emerges. (Use 1:1 scales initially, then move to 2s, 5s, and 10s, or whatever is appropriate.)
- Construct column graphs without grid paper. Students clearly label the *y*-axis, mark the scale and number it.

This progression from graphs based on physical objects, to ones involving representations of objects prior to introducing abstract graphs with an axis showing values is also found in the teaching suggestions of Friel et al. (2001).

Rangecroft (1994) also provides a progression for the development of graphing, but with more detail than Zevenbergen et al. (2004) or Friel et al. (2001). In doing this she makes more obvious the parallels between the development of graphical understanding and measurement understanding. Rangecroft's progression still addresses the importance of a gradual transition from real objects to the more abstract bar chart (sorting items into sets, developing pictograms where the objects are the same size, putting sets into lines to remove the need to count, stacking blocks and moving to situations where a baseline needs to be chosen, and using coloured squares to form a block graph) but specifies other important stages in the progression. These include:

- introducing ordinal data – for example, year groups in school – to lead to an ordered horizontal axis;
- using larger data sets to introduce the idea of "labelling spaces on the vertical axis as a more convenient way of finding the frequency than counting the individual squares" (p. 9);
- labelling the lines on the vertical axis instead of the squares – which "should arise naturally from the fact that it is the height of the column that is important" (p. 9). (In Rangecroft's experience "pupils find this change rather puzzling if it is introduced without explanation, but they accept it if the reason is explained to them" (pp. 9-10).)
- introducing the use of a non-unit scale when moving from early graph types to more varied display types – "one of the first problems which arises is the use of a scale on the vertical axis" (p. 10); and
- introducing bivariate data and the use of area instead of height to represent frequency (as in histograms and pie charts).

Rangecroft also identifies that scaling is a fundamental notion students must develop if they are to understand graphs. In statistics it involves *data reduction* (the sorting and grouping of data so a graphical display highlights the chosen aspect of the data).

Zevenbergen et al. (2004), Friel et al. (2001), and Rangecroft (1994) all outline a slower development of scale in graphs than is evident in MiNZC and the data display exemplar. These latter documents indicate that the concept develops early, with little specific attention being paid to its progressive development; students are expected to create a graph using appropriate conventions at CL2.

For CL3, students move on to displaying discrete numeric data, with two of the graph forms (dot plots and strip graphs) requiring a stronger understanding of measurement conventions and scale (see Figure 4.2). Bar graphs are still mentioned in the SLEs and the exemplar (Ministry of Education, 2001g), while the SLEs also include simple time series and measurement data. It seems that the expectation is for students entering Year 7 to have mastered the use of measurement conventions on scales, for example, they should number the marks not the spaces. Furthermore, they should have met non-unit scales, and have created scales then plotted real (random) data in relation to them.

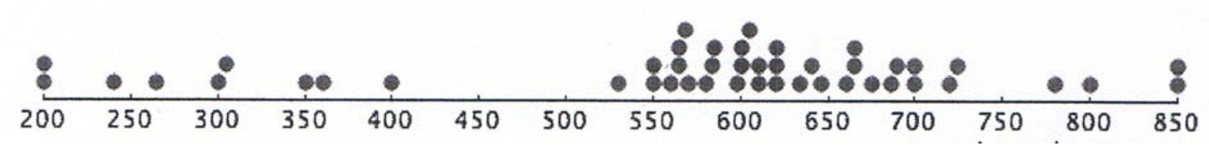


Figure 4.2. Dot plot showing measurement conventions on the horizontal axis

As with the development of the concept of scale in measurement, MiNZC gives little direction as to *how* this learning should take place. For example, the notion of scaling is not mentioned in MiNZC in relation to either statistical or algebraic graphs, so it is possible that scaling is not being properly developed by teachers. The expectation of a developing understanding of measurement also suggests that the teaching and learning of graphing could be hindered by a student's lack of understanding of measurement.

Map scales

Three AOs are relevant to maps, one at each of CLs 2, 3, and 4. These are supported by a number of SLEs. In CL2 students learn to "describe and interpret position using the language of direction and distance" (MiNZC, p. 96), and draw and read maps using the four compass points. At CL3, students are expected to "draw and interpret simple scale maps" (p. 100 & p. 101) and give and follow shortest/longest paths in mazes and maps. For CL4 the focus continues on investigating shortest and longest paths on a map but the AO emphasises specifying location using bearings or grid references.

'Interpreting scale maps' suggests that students have learned to read a map, so should be able to calculate real distances from the map. A scale-based strategy requires:

- reading a ruler accurately;
- accurately transferring that measurement to the scale; and
- reading between the marks on the scale (as unless both points are located on gridline intersections, and are placed along a horizontal or vertical line, the measurement will not be one that is marked).

This sophistication is consistent with CL3 measurement where students are expected to "perform measuring tasks, using a range of units and scales" (p. 66). However, an alternative (and simpler) strategy is to multiply the map distance by the numeric scale factor. This only involves the map measurement, not the ability to use the map scale.

For drawing a simple scale map of the school playground, students need to measure real distances and objects, and create a scale factor so the map fits on the available paper. Students may need to use the scale factor to calculate the dimensions of an object for the map, and measure accurately to place the object correctly in relation to other objects. Where accurate measurement of objects is not required, rather their relative location, this step may involve developing a coordinate system.

Discussion

The place that scale has in a developing an understanding of maps is not clear from reading MiNZC (Ministry of Education, 1992) particularly as the SLE on p. 101 is a repeat of the CL3 AO on p. 100. Bobis, Mulligan, and Lowrie (2004) suggest that for children's early years in

Australian schools the intention is for them to: use the language of position and orientation; place important things in the environment on maps; find paths on informal maps and mazes; and recognise a birds-eye view. In later years they should be: describing position in relation to directional language and coordinate systems; using formal measurement units to develop more accurate maps; and representing 3D space on scaled maps. Zevenbergen et al. (2004) suggest that in simple maps a different scale is frequently used on each axis. Typically this involves an alpha-numerical pairing with letters on one axis and numbers on the other. "Thus each square on the map is characterised by a distinct alpha-numerical pair" (p. 334). To read positions students should also be learning to read the *x*-axis before the *y*-axis. This suggests that simple maps use the convention of identifying and naming squares (rather than points) but may support the development of the '*x* before *y*' convention used in graphs. Overall, these references do little to clarify what the phrase *simple scale maps* implies. It is possible for students to have developed a sophisticated understanding of scale in the context of maps, and have learned about coordinate systems prior to Year 7. However, it is also possible that they are working towards a six figure grid reference system for CL4 (which in terms of scale understanding can be tied to the counting conventions used in measurement at CL1), and have not interacted with the scale aspect of the map.

Establishing what is intended by many of the AOs in MiNZC is not a straightforward task; the statements are general, and require interpretation. ERO (2002) support this finding, stating that "developing learning outcomes and relevant learning experiences from the achievement objectives in the curriculum statement was a difficult task" (p. 4). Based on the preceding analysis, students by Year 7 should have had experiences over a number of years working with 'real' scales to measure lengths, masses, areas, volumes, and temperatures. They should also have used scales in statistical graphs, but may not have used them with maps, or seen them in abstract number contexts – such as number lines or as part of a scaled Cartesian system. Creating coordinates and appropriate scales should have been met in statistics and possibly with maps.

In the next section the science, technology, and social studies documents are more briefly reviewed. Only Levels 1-3 are considered as these are the levels at which most students have met scales before reaching Year 7.

4.2.2 *Science in the New Zealand Curriculum* (Ministry of Education, 1993a)

Science in the New Zealand Curriculum focuses on exploring scientific concepts, using contexts meaningful to students as the medium. As such, measurement is an exploration tool rather than something to explore in its own right. For example, the 'information gathering' AO at CLs 1-2 is stated as "make observations and simple measurements" (p. 45). For Levels 3-4 it is "use appropriate instruments to enhance observation or to introduce quantification" (p. 45). With this approach, measurement scales are not created in science, the nature of scales is not explored, nor is the ability of students to work with them considered an important focus of their learning.

Graphs are likewise a tool to explore scientific concepts. A graph may be used to identify clearly the trend, pattern or relationship in data, or for reporting experimental results. Here the nature of the phenomenon being measured is not stated. It could be discrete or continuous. It could also be time related, like the growth of a plant, so students may meet such data in science before they study it in mathematics. It is therefore possible students' learning in science will form a body of understanding against which later mathematical explanations may be measured and interpreted.

4.2.3 *Technology in the New Zealand Curriculum* (Ministry of Education, 1995)

While the focus of the technology curriculum is on developing technological literacy "to enable students to participate fully in the technological society and economy in which they will live and work" (Ministry of Education, 1995, p. 5), the document recognises that students draw upon "knowledge and skills developed in other areas of the curriculum" (p. 20) and that undertaking technological activities contributes to the learning of these areas. The role of mathematics in the study of technology is explicitly recognised, with scales featuring prominently. Activities like surveying, graphing and describing trends, collating and interpreting statistical information, measuring, and making models and plans are all mentioned. "Calculating, measuring, and estimating skills can be practised and developed through technological activities, linking technology with mathematics. Graphs, tables, charts, and other visual presentations of data have a role in technological activities" (p. 18).

There is no assumption in the technology curriculum that students should already know a piece of mathematics or how to use a mathematically based tool. The above information also suggests that teachers of technology have an active role to play in students' learning of mathematics. As such, students may be getting some of their mathematics teaching, particularly in relation to scale, through the context of learning technology.

4.2.4 *Social Studies in the New Zealand Curriculum* (Ministry of Education, 1997)

The focus of the social studies curriculum is on exploring human society. It uses a variety of tools to bring meaning to this exploration. The scales met in graphs are part of a tool to explore social concepts, rather than something to learn about. This is reflected in the AOs, where graphing references are mostly linked to the inquiry process: "[t]he inquiry process involves students in collecting and analysing information about people, groups, communities, and societies" (Ministry of Education, 1997, p. 17). At CL1 and CL2 students are expected to use questions, collect, record and sort information, make a generalisation based on findings, and communicate their findings. At CL3 and CL4 they are to frame questions to focus inquiry, "gather and process information from a range of sources" and "analyse and respond to information in graphs, tables, charts, or percentages" (p. 19). Information is to be processed using appropriate conventions, a valid generalisation made (supported by evidence) and findings communicated using conventions appropriate to the mode of communication. In 'values' exploration, students are expected to "understand information which is presented in mathematical ways" (p. 19), while in social decision-making they need to "organise information to support logic and reasoning" and "present information clearly, logically, concisely and accurately" (p. 19).

The types of data students may encounter at different curriculum levels are not mentioned, but given the nature of human society it can reasonably be inferred they involve both qualitative and quantitative forms. Likewise, no mention is made of the graph forms for data display, but as the data come from real sources about human society, it is likely they could involve change over time (like population growth), or 'snapshots' of a community feature at a given moment – like the amount and types of waste produced by different classes in a school. The tasks could involve reading graphs created by others, as found on the internet, in newspapers and books, or creating an original display. Once again it seems likely that students are meeting graphs and types of data in social studies before they learn about them in mathematics.

4.2.5 Summary of findings

The curricula indicate students are expected to have the ability to work with measurement and graph scales in a variety of forms by Year 7. The curricula tend not to indicate that students should meet scales in abstract situations (as number lines or in Cartesian systems). Students are likely to have met scales involving multiples of whole numbers or fractions when measuring mass (depending on the chosen unit); and whole numbers, decimals or fractions, and integers when measuring length and temperature. Graph scales are likely to have involved multiples of whole numbers. As primary teachers cover all subjects, they possibly introduce students to contexts involving measurement scales and graph-related concepts while teaching science, technology, or social studies before they are introduced in mathematics. This suggests that students may be getting a 'double-dose', gaining exposure to a wider range of scales and data types than outlined in the mathematics curriculum. The learning about scale is also intended to be meaningful, with each meeting helping to build a richer understanding.

Here it is important to provide a word of caution. AOs are statements of intended outcomes, stating what students "should be able to achieve after appropriate learning experiences" (Ministry of Education, 1992, p. 16). MiNZC, and the other curricula give little guidance on *what* skills and understandings are required by students for them to reach these outcomes, or *how* these skills and understandings are to be developed. The development of the notion of, and an ability to use, scale is a case in point. While there are expectations of what students should be able to do in each strand, there is little to suggest how teachers get this to develop. There is no indication that teachers need to spend time building certain understandings, for example the role of zero on a measurement scale or the concept of scaling when creating and working with graphs, in spite of a sizeable body of research pointing to critical learning that needs to be addressed. This suggests these documents are relying on primary teachers having an 'expert understanding' of the content contained within the seven essential learning areas of the *New Zealand Curriculum Framework* (Ministry of Education, 1993b) as well as a sound knowledge of the pedagogy effective for imparting this content. It leaves individual teachers to construct a teaching programme ultimately based on their own understanding of, and experiences from, learning, a topic. This may mean, for example, that the essentials of scale identified in Chapter 3 are not being adequately addressed or that students are being introduced to what may be, for some, a bewildering series of data representations of which they have little understanding, and from which they learn little.

4.3 Resource review

In this section three sets of resources are reviewed: *Pearson Primary Maths; National Curriculum Mathematics;* and part of the *Figure It Out* series. The goal is to identify how these resources help students learn about scale, and the support they give to teachers in their teaching of scale.

4.3.1 *Pearson Primary Maths*

This section reviews the following resources:

- *Pearson primary maths 2a*, *2b, 2c* and *teachers guides 2a, 2b* (Pearson Level 2) (Wilkinson, 2000a, 2000b, 2001a, 2000c, 2001c);
- *Pearson primary maths 3a*, *3b, 3c* and *teachers guides 3a, 3b, 3c* (Pearson Level 3) (Wilkinson 2001b, 2002a, 2001e, 2001d, 2002b, 2001f).

These books are structured as a series of stages through which students work. Each book has a teacher's guide. Table 4.1 summarises the scale related activities identified, with a number indicating how many of the stages involve the use of scale in a particular branch of mathematics. The tick (✓) and cross (✗) indicate which types of number are used with the scales (a tick indicates this type of number is used).

Table 4.1: Summary of scale related material from the Pearson Primary Maths series

		Level 2	Level 3
Context	Number lines	4	7
	Measurement	2	2
	Graphing	8	0
	Algebra	0	4
	Probability	2	1
	Angles	0	1
	Statistics	7	0
	Maps	0	5
Type of number	Unit	✓	✓
	Multiples	✓	✓
	Decimals	×	✓
	Fractions	×	✓
	Integers	×	✓
Graph reading		To gridlines	No gridlines
Explanation of conventions used		×	×
Developmental progression		Some	Some

Discussion

These texts and guides give students practice with scales in a wide range of situations that includes number lines. The material in the Level 3 books is well in advance of the Level 2 books, but at both levels there seems to be a tendency for new representations of scales to be introduced with the assumption that they are already understood. For example, in book 2b (Wilkinson, 2001a) students are asked to estimate an amount of water, then measure it. No scaffolding is provided to help students to read a capacity scale and no mention is given to teachers to indicate that students may need to be supported as they learn to read such non-unit scales. Scales are also used to illustrate a new piece of learning, with the assumption that students understand what is drawn and recognise the drawing's relevance. For example, in teaching rounding, the decision about which number to round to is illustrated by locating the number on a number line.

Some developmental progression is found within curriculum strands, but this does not always cross strands, so students can be reading a time series graph early on, then later learn to work with points on a unit scale in algebra. In another example, a number line with a unit scale is used in book 2a (Wilkinson, 2000a), a scale marked in tens is introduced in book 2b (Wilkinson, 2000b), and measurement scales involving decimals are found in Level 3 (Wilkinson, 2001b, 2002a). However, students meet bar graphs in book 2a within the context of probability and as an illustration of doubling – before they meet them formally in statistics at the end of book 2b.

While children meet scales in a wide variety of representations there is no material in the workbooks that explains to students how to interpret and use those they meet. A few references in the teachers' guides alert teachers to issues students may need to address but there is little explanation of *how* to address these. As in the curriculum documents, students seem to be expected to pick up what they need to learn about scale from the exposure they are given, with teachers being expected to have the knowledge to impart what is needed.

One guide suggests that scales from meaningful contexts are better (Wilkinson, 2001d), but does not elaborate on what this term means. The illustrations in the workbook imply 'real-life' situations. For example, Wilkinson (2002a) shows a picture of a weighing scale in kilograms that is labelled every 200 grams, and marked each 20 grams (p. 67). On page 66 is a measuring jug with a scale in hundreds of millilitres, marked each 25 millilitres. However, in real-life measurement scales tend to be part of sophisticated instruments. For example, commercially available kitchen scales generally show two scales, one in pounds and ounces, the other in kilograms and grams. This suggests that using real measuring instruments may unnecessarily complicate early learning about scale.

4.3.2 *National Curriculum Mathematics* (NCM)

This section reviews the following resources:

- NCM Level 2, book 1, and Level 2, book 2 (Tipler & Douglas, 2000a, 2000b);
- NCM Level 3, book 1, and Level 3, book 2 (Tipler & Catley, 1998, 1999).

These texts are structured into the five content strands, in the order found in MiNZC (Ministry of Education, 1992). No guides are available for teachers to identify the key underlying issues for each exercise. Table 4.2 provides a summary of the analysis undertaken, using the conventions outlined for Table 4.1.

Table 4.2: Summary of scale related material from the NCM series

		Level 2 book 1	**Level 2 book 2**	**Level 3 book 1**	**Level 3 book 2**
Context	Number lines	4	0	3	4
	Measurement	10	0	5	0
	Graphing	4	0	0	0
	Algebra	0	1	2	6
	Probability	0	0	0	0
	Angles	0	0	0	0
	Statistics	4	3	3	0
	Maps	0	0	1	1
Type of number	Unit	✓	✓	✓	✓
	Multiples	✓	✓	✓	✓
	Decimals	✗	✗	✓	✓
	Fractions	✗	(half)	✓	(half)
	Integers	✗	✗	✗	✗
Graph reading		Unit scales	Working to marks and gridlines	Some work between the gridlines	Some work between the gridlines
Explanation of conventions used		✗	✗	✗	✗
Developmental progression		Some	Some	Some	Some

Within each book, the cognitive demands of the questions vary from strand to strand. For example, the emphasis of statistics in Level 3, book 1 is on graph types involving discrete data and counting conventions – pictograms, simple dot plots and block graphs (see Figure 4.3 for an example). Meanwhile in measurement, practice is given on reading scales involving multiples with unlabelled marks in-between.

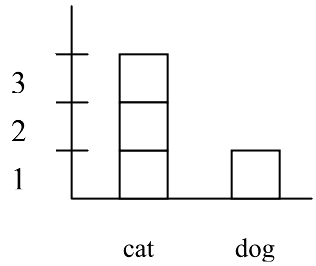


Figure 4.3. A block graph using counting conventions

In statistics the conventions used vary from question to question. For example, in Level 2, book 2, block graphs have numbers placed 'on the marks' of the vertical axis (Tipler & Douglas, 2000b), yet in Level 3, book 1, the numbers are 'between the marks' (Tipler & Catley, 1998). Different graphs in Level 3, book 1, also have different conventions for where to place categories on the horizontal axis – sometimes on the marks, sometimes between them.

Discussion

The NCM series gives students practice with scales in a wide range of situations that include number lines. Each book seems to be set at a higher level of sophistication and students are provided with practice at this new level. Since there are no teachers' guides, it is up to teachers to identify the development of understanding required to move students from the work in one book to the next, then include this as background and preparatory work for the exercises. Consequently it is possible for these developments to be overlooked or be taken as assumed knowledge.

In number, number lines act as a tool to help solve problems or support the learning of other concepts. The expectation is that students already understand the representation (e.g., see Tipler & Catley, 1998).

Overall, the approach taken to the teaching of scale mirrors that of MiNZC (Ministry of Education, 1992). Students are given practice at the form of task that they are expected to master at the particular curriculum level. In this approach it appears that the students are expected to pick up what they need to learn about scale from the exposure they are given, rather than from any specific teaching about scale.

4.3.3 The *Figure It Out* series

The Ministry of Education directly funds the development and publication of some resources for primary schools. In mathematics, the major publication is the *Figure It Out* (FIO) activity books. Each book focuses on an area of mathematics, introducing students to the important pieces of learning at a curriculum level. They have individual answer books with extensive teachers' notes explaining the key learning behind each activity, and identifying issues for teachers to address. Each book also links to mathematics from other strands. In some cases the notes suggest the use of activities from other books before a task is used.

As this series is extensive (over 50 individual titles) with many focused on number, a selection has been reviewed with reference to scale. These cover statistics, measurement, and algebra.

- Figure It Out Statistics, Levels 2-3, 3 and 3-4, and teachers' notes (Ministry of Education 1999a, 1999b, 2000a, 2000b, 2001a, 2001b);
- Figure It Out Measurement, Levels 2-3, 3 and 3-4, and teachers' notes (Ministry of Education 1999c, 1999d, 2000c, 2000d, 2001c, 2001d);
- Figure It Out Algebra, Levels 2-3, 3 and 3-4, and teachers' notes (Ministry of Education 1999e, 1999f, 2000e, 2000f, 2001e, 2001f).

Table 4.3 provides a summary of this analysis, using the conventions for Table 4.1.

Table 4.3: Summary of scale related material from the FIO series

		Statistics			**Measurement**			**Algebra**		
		L2-3	L3	L3-4	L2-3	L3	L3-4	L2-3	L3	L3-4
	Number lines	0	0	0	0	1	1	0	0	1
	Measurement	1	2	0	5	11	12	0	0	0
	Graphing	0	0	0	0	0	2	4	0	0
	Algebra	0	0	0	0	0	0	0	0	0
Context	Probability	0	0	0	0	0	0	0	0	0
	Angles	0	0	0	0	0	0	0	0	0
	Statistics	9	10	10	0	0	0	0	0	0
	Maps	0	1	0	0	1	0	0	0	0
	Unit	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Multiples	✓	✓	✓	✓	✓	✓	✓	✓	✓
Type of number	Decimals	×	×	✓	×	✓	×	×	×	×
	Fractions	×	×	×	×	✓	✓	×	×	✓
	Integers	×	×	×	×	×	×	×	×	×
Graph reading		B*	B*	B*	×	×	B*	B*	B*	B*
Explanation of conventions used		×	×	×	S*	S*	×	S*	S*	S*
Developmental progression		✓	✓	✓	✓	✓	✓	✓	✓	✓

B* reading between labelled marks is required
S* some explanation of conventions

Discussion

Table 4.3 shows that students are exposed to a variety of scales in the FIO statistics books, including those needed when 'real' data are collected. The books and notes assume students can choose an appropriate graph scale for their data, and can plot data on any chosen scale. It is also assumed students can read data from a variety of scales, and bring to statistics the ability to work accurately with measurement scales.

In the FIO measurement books, students meet scales while carrying out practical measurement tasks with real equipment. There is some emphasis on developing an understanding of the scales. For example, in the Level 2-3 book (Ministry of Education, 1999c) the first activity 'stretching it out' starts by using centimetre cubes to measure the length of an object. This is followed by a task requiring a cardboard strip to be marked in centimetres, for use as a ruler. In another activity students are expected to think of other ways of measuring if a ruler or the cubes are not available. The teachers' notes (Ministry of Education, 1999d) state that:

> [m]any of the difficulties students have with measuring length arise from their understanding of the ruler as a scale. Students also have difficulty with starting points or measuring from zero on the ruler. (p. 9)

The notes suggest giving practice with 'lining up' and the 'iteration of the unit' concept, along with matching students' marked strips to real rulers to help them understand how rulers were first made. The problems are based on a counting approach to measurement, with a 1:1 correspondence between the object (a centimetre cube) and a square on the home-made 'ruler'. The change in conventions from a 'counting-based' ruler to a 'real ruler' are not noted as important to address, so the purpose of zero on a measurement scale and the concept of a unit as an interval may still be missed.

Other activities, especially in the later titles, assume students have no difficulty reading an instrument scale to the nearest graduation, even those involving fractions or decimals. That students have the skills for choosing an appropriate instrument and measuring unit is also assumed. No activities address the process of reading a scale. Students are also expected to bring to measurement an ability to work appropriately with graphs.

In the FIO algebra books, students use graphs in a wide variety of situations. For example, in the Level 2-3 book (Ministry of Education, 1999e) some activities involve creating bar graphs, while others use line graphs. In all three books, the teachers' notes indicate students may have difficulty with some aspects of scale. For example, in the Level 2-3 activity 'a touch of sickness' students are stacking counters to create a graph. The teachers' notes for the activity address the difference between counting and measurement conventions "[i]t is conventional to show such relations as points on a number plane, … This means that a point on top of each column of counters is sufficient" (Ministry of Education, 1999f, p. 20). In the activity 'building graphs' (Ministry of Education, 1999e) students are building graphs with Cuisenaire rods acting as the legs of sheep. The teachers' notes suggest that "[t]hese points can be written as ordered pairs such as [(1, 4)(2, 8)(3, 12)(4, 16)…]" (Ministry of Education, 1999f, p. 20). For teachers using the FIO series, this may be the first activity in which students plot numeric data as this is introduced in the higher level measurement and statistics books. For the Level 3 book, the notes for 'possum poles' suggest that if students have trouble going straight from the series of pictures to the required graph, they should create a table first. Other notes mention that the last picture 'skips a minute', that is, a minute is omitted so the final data are not the next minute "as students may expect" (Ministry of Education, 2000f, p. 22). The Level 3-4 notes for 'stacks of money' suggest that when reading a point students need to draw a vertical line (parallel to the *y*-axis) to the point they want, then draw a horizontal line (parallel to the *x*-axis) across to the *y*-axis, where the number it hits is the required value; and that this strategy should be used when both extrapolating and interpolating (Ministry of Education, 2001f).

Other skills are assumed, for example, the abilities to create an appropriate graph scale and to translate between data representations, recognising their equivalence. Students learning algebra are also expected to be able to work accurately with measurement scales and recognise and use a variety of graph types.

4.3.4 Summary of findings

The reviewed resources seem to take their lead from the mathematics curriculum document. While they provide exposure to scale and different opportunities for learning about scale, in general it is not scale that is the focus of the lessons, rather scale tends to be a tool which is taken for granted as it is used 'in a range of meaningful contexts'.

4.4 Conclusions from the document analysis

The document analysis has identified that students should be familiar with using scale in the contexts of graphing and measurement (at school if not in the real world) so these are appropriate contexts to use when assessing scale. However, the four reviewed curriculum documents provide little indication of key learning to support the development of students' ability to use scale. The reviewed resources indicate that they generally assume scale learning 'just happens'; either students will pick it up while doing other things, or teachers know what issues to address, what misconceptions to confront, and how to do this. Learning about scale may therefore be haphazard and incidental. It seems that Outhred and McPhail's (2000) suggestion (related to the learning of linear measurement) that students' misconceptions may reflect teachers' inadequate concepts, combined with syllabus documents giving little direction for key concepts, may also be valid for the concept of scale in relation to the New Zealand mathematics, science, technology, and social studies curricula, and for the reviewed mathematics teaching resources. This suggests that identifying the key concepts of scale, and 'good practice' approaches for developing these concepts, could be useful for teachers and may improve students' understanding of scale.

Chapter 5

Methodology: General Research Design

5.1 Introduction

This chapter outlines the research methodology. Issues pertaining to the choice of research paradigm, method, procedures, data analysis, validity, and reliability are discussed. Details of the development and use of the assessment tool can be found in Chapter 6. The development of the assessment items is covered in Chapter 7.

For the convenience of the reader, the research question and its sub-questions are restated here.

What understanding of scale do Year 7 and 8 students in the case study schools have?

- What *should* Year 7 students know about and be able to do with scales?
- How do some commonly available mathematics resources present the learning of scale?
- What do students in the case study schools actually know about scale?
- What issues of understanding do the students have? and
- What does this research suggest as ways of addressing these issues?

The approach taken to answering these questions is explained during the rest of this chapter, and in Chapters 6 and 7.

Table 5.1 provides an overview of the research. It gives a timeline, shows which participants were assessed with what assessment, and where the results are reported.

Table 5.1: Timeline for the revised research design

	August 2005	September 2005	October 2005	November 2007
School 1	Interviews with Test 1 (Chapter 8 – Case 1)			
Teacher trials		Written test questions (Chapter 9)		
School 2		Written version of Test 1 (Chapter 10 – Case 2)		
School 4				Interviews with Test 2 (Chapter 12 – Case 4)

Note that when the thesis was upgraded from an MEd to a PhD, the research focus and design were changed (see Chapter 1). School 3 was removed from the design at this time as the nature of the findings indicated the need for a greater in-depth analysis of scale understanding. School 4 was added to probe further into student understanding. While the results from School 3 provided triangulation data for the findings from Schools 1 and 2, they added no additional insights related to scale understanding, so for the sake of brevity they are not reported here. (Also see Figures 5.1 and 5.5.)

5.2 Epistemological considerations: Ways of knowing

In developing the methodology for this study, consideration was given to the knowledge the research intends to generate, and the reliance to be placed on this. At the outset it is important therefore to focus on the broad approach to be taken. Often these approaches are defined as paradigms. In the post-modern world there are an increasing number of paradigms to choose from: For example, the scientific (positivistic), the naturalistic (interpretive), as well as approaches from critical theory, and feminist educational research (Cohen, Manion, & Morrison, 2007). Of these, the scientific and naturalistic paradigms seemed to best fit the intentions of the research so are considered in more detail below.

The scientific approach "is a way of comprehending the world, as a means of explanation and understanding, of prediction and control" (Cohen, Manion, & Morrison, 2000, p. 11). It arises from the natural sciences and focuses on uncovering the truth that lies behind what is observed and experienced. As far as possible the place of the researcher in this process is as an independent observer of the world, analysing relationships and identifying regularities between cause and effect. The goal is to discover the general laws that control all interactions in the universe. The positivistic researcher applies the methodology of the natural sciences to the social sciences. Theories and hypotheses derived from the observation of social phenomena are investigated empirically through the use of strict scientific method. The process involves observing controlled experiments and measuring the impact that manipulating variables has. Clarity of procedure is emphasised so fellow scientists can replicate the research. Quantitative data are produced and statistical theory and method are used to interpret the results (Cohen et al., 2000).

Naturalistic approaches focus on studying reality in its natural state as it is seen and experienced by participants; the "social world can only be understood from the standpoint of the individuals who are part of the ongoing action being investigated" (Cohen et al., 2000, p. 19). In this process the researcher must share the frame of reference of the subject and come to understand the individual's interpretation of the world if they are to make sense of their actions. Different perspectives can lead to other interpretations, hence the need for a rich description in which the role and impact of the researcher are acknowledged. Multiple data sources are also used to ensure the researcher's interpretation of the event is correct. Data are predominantly qualitative, with language playing a key role as an instrument through which the world is represented and constructed. Working hypotheses and theory emerge from the study of individual situations and are grounded in the data generated by the research act, so arise from research rather than precede it. Theory gives insights into, and understanding of, people's behaviour rather than describing universal truths (Cohen et al., 2000; Robson, 2002).

Criticisms have been made of both the positivist and the interpretive approaches. The application of the methodology of natural science to the study of human behaviour has been said to exclude the notions of choice, freedom and individuality, and omit the subjective and experiential perspective of the being. Scientific method also ignores the fact that unlike inanimate objects, human beings can react and learn from stimuli, so the process of observation is not neutral and can cause interaction effects among the objects of study. Observed situations also are affected by multiple rather than single variables, so efforts to control or limit the impact of variables can lead to artificial situations which have little relationship to reality. In turn, interpretive approaches have been criticised for using methods that are more open to error than those of the physical sciences. Merely interpreting reality in terms of the reports of its actors also neglects the impact of external forces that shape behaviour, and ignores the possibility of discovering useful generalisations about behaviour (Cohen et al., 2000).

In an attempt to use the strengths of each paradigm, while taking cognisance of their weaknesses, the researcher has chosen a mixed method incorporating aspects of both. The philosophical framework is fundamentally interpretive, because the research question seeks to clarify students' understanding. However, while it is important to identify what students can do, and try to make clear the thinking and understanding that led to their responses, this on its own is inadequate. In addition, it is important to develop an understanding of the magnitude of the issues students have when developing scale understanding. This requires a statistical approach to be used on a larger body of students than is commonly used in interpretive research, although it should be noted that even the numerical data tends to be used descriptively rather than inferentially. Table 5.2 provides a broad outline of the use of each paradigm in the research, while further details of the design are provided in the next section.

Table 5.2: Use of qualitative and quantitative paradigms in this thesis

Phase		**Form of analysis**
Design	• Literature review. • Curriculum review. • Resource review. • Assessment tool development.	• Rich description. • Rich description. • Rich description with some numerical analysis. • Rich description.
Case 1	• Trialling of the tool through interviews. • Analysis of what the tool has identified.	• Rich description. • Content analysis.
Teacher trial	• Triangulation study (tentative findings from Case 1).	• Content analysis.
Case 2	• Trialling of the tool as a written test.	• Content analysis, with some numerical and statistical analyses.
Design	• Further development of the assessment tool.	• Rich description.
Case 4	• Trialling of the revised tool through interviews. • Analysis of what the tool has identified.	• Content analysis, with some numerical and statistical analyses.

Now the philosophical approach to knowledge generation has been discussed, it is time to address the research methodology.

5.3 Methodology

5.3.1 An exploratory study

The research question this study set out to answer concerns the understanding that groups of Year 7 and 8 students had of scale. In the early chapters, a definition of the term *scale* was developed and the understanding of scale that students in Years 7 and 8 should have in terms of the curriculum was identified. The literature review indicated that there appear to be common issues when scales are used in different areas of mathematics, and also highlighted several starting points for exploration; for example, whether or not students have a better understanding of scales presented in meaningful contexts, and if students find working with marked lines easier than unmarked lines. Unexpectedly however, given the frequent use of scales in mathematics, no comprehensive studies of scale understanding were identified, suggesting that this research is fundamentally exploratory in nature. It was therefore important to adopt a methodology that enabled discoveries made during the course of the research to be further explored. The coming sections thus outline the adoption of a flexible design incorporating case studies within the framework of an evaluation methodology.

5.3.2 An evaluation methodology

"An evaluation is a study which has a distinctive purpose". It attempts to assess the worth or value of some innovation, intervention, policy, practice, or service (Robson, 2002, p. 202). It aims to judge "the worth or effectiveness of "something" in doing what it is supposed to do. Sometimes an evaluation may go further and focus on the unintended or side-effects of the "something"" (Hall, 2002, p. 1). Evaluations are also conducted to find areas of improvement (Davidson, 2005). Within an evaluation methodology, no single method of inquiry is privileged: "fixed or flexible designs can be used, with either qualitative or quantitative methods, or some combination of both types" (Robson, 2002, p. 202).

The concept of a *rolling design* for an evaluation is discussed by Davidson (2005). This is "basically an open-ended continuous improvement approach to doing evaluation" (p. 55). In this approach, the data collection is rolled out in phases. The initial phase is akin to a pilot, in that a small-scale data collection is undertaken, the findings of which are evaluated and used to inform the development of subsequent stages. This process of evaluation after each phase is continued, and opportunities to improve the data collection instruments and methods are taken. Additional pieces can be added to or deleted from the original design during the course of the evaluation to ensure that all important new information is captured and unanticipated findings are followed up. This form of design is appropriate for an exploratory study as it is not possible to predict the nature of the findings until data collection has been undertaken. Figure 5.1 shows the original MEd design of this study and highlights the use of evaluation to progressively explore student understanding of scale, and how those findings were to be used to develop the later research phases. As will be seen later, this design was modified to take account of contextual information and events that arose during the middle stages of the research.

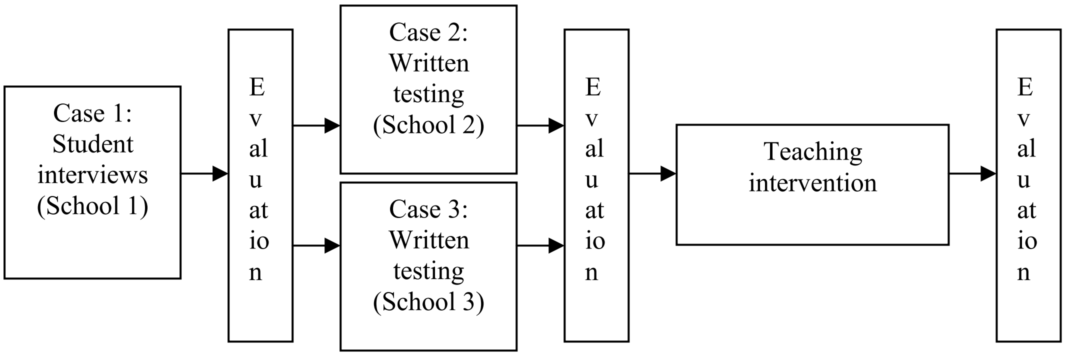


Figure 5.1. Original design highlighting the use of evaluation and case studies

Evaluation methodology was appropriate for the original study as it provided a framework within which the curriculum review could be undertaken. When no comprehensive studies of scale were identified the adoption of a rolling design provided the flexibility needed to explore in detail any discoveries made about scale understanding. Thus the methodology had the benefit of uniting the exploration of student understanding and the goal of informing the curriculum review under a common methodological framework. Since the goal of the curriculum evaluation was to evaluate the effectiveness of an existing curriculum document prior to the development of a new curriculum, the evaluation was summative in nature. A *summative evaluation:*

> asks questions such as: What impact, if any, does a policy, programme or some other type of government intervention have in terms of specific outcomes for different groups of people? It seeks to provide estimates of the effects of a policy either in terms of what was expected of it at the outset. (Government Social Research Unit, 2007, p. 1:4)

Typically a summative evaluation can have a formative effect on future developments and can carry with them strong implications for change (Robson, 2002). Thus the responses to the research sub-questions could not only be used to identify the effectiveness of the existing mathematics curriculum (Ministry of Education, 1992) for the teaching of scale, but also to suggest possible areas for improvement.

5.3.3 The curriculum evaluation

This section turns the focus from the overall research methodology to the specifics used in evaluating the effectiveness of MiNZC for the learning of scale (Ministry of Education, 1992). In doing this it focuses on *policy evaluation*.

> Policy evaluation is a family of research methods that are used to systematically investigate the effectiveness of policies, programmes, projects and other types of social intervention, with the aim of achieving improvement in the social, economic and everyday conditions of people's lives. (Government Social Research Unit, 2007, p. 1:10)

Curriculum as policy

A curriculum document can be identified as a *large-scale policy* as it "provides direction for interventions that are implemented across a system of providers under the same organisational umbrella" (Owen, 2007, p. 24). Policies are a general guide to action but do not necessarily specify the means by which the ends (goals) are attained. They provide an assertion of intents or goals, a guide to discretionary action, are a strategy to ameliorate or solve problems, and indicate sanctioned behaviour (Guba, 1984, cited in Owen, 2007). By evaluating the effectiveness of MiNZC (Ministry of Education, 1992) in providing direction for the teaching of scale, this thesis seeks to measure the effectiveness of a policy. As such it is not necessary to evaluate the programmes developed from the policy (text books, school schemes of work, and teachers' plans) as these can be expected to be discretionary and varied. What is important to consider are:

- What MiNZC was supposed to do, and whether it has done this;
- Possible side effects;
- The assumptions on which MiNZC is based (to see if they are the cause of any performance need); and
- The directions provided for action.

(based on Davidson, 2005). In particular, do students develop an understanding of scale from exposure to scale in a range of meaningful contexts?

5.3.4 The process of policy evaluation

> Policy evaluation and analysis requires a structured and organised approach to defining an answerable question, summoning appropriate and relevant evidence, critically appraising and analysing that evidence, identifying the risks and opportunities of a policy, programme or project, and determining the likely effects (positive and negative) of the project at hand. (Government Social Research Unit, 2007, p. 1:10)

This approach is explained in practical terms in Owen (2007), who identifies that the following issues need to be addressed when undertaking an evaluation.

- On what basis were the criteria for judging worth selected?
- What evidence was used, and what standard of evidence was required for judgement decisions?
- How were the conclusions presented and what decisions can be based on the evaluation?

These issues are addressed in the following paragraphs.

Criteria of worth and judgement of effectiveness

To judge the effectiveness of MiNZC (Ministry of Education, 1992) for teaching scale the question this research seeks to answer is '*does it fulfil its main purpose*?' That is, can students entering Year 7 answer the sorts of questions indicated as important by MiNZC? However, what proportion of students need to be successful for the policy direction to be considered a success? Here it is important to note that having a clear concept of the required achievement expectations is central to informing the curriculum review for the purpose of education has changed dramatically since MiNZC was written. In 1992, assessment was largely norm referenced in focus, particularly in the senior secondary school, but also at lower levels with assessments like the Progressive Achievement Tests from the New Zealand Council for Educational Research. By 2005 qualifications were gained from standards-based assessment, the goal being to make education productive for all students; educational practices should serve the purpose of equipping "all New Zealanders with the knowledge, skills and values to be successful citizens in the 21st-century" (Ministry of Education, 2008, p. 4). Multiple learning pathways were envisaged with different end points.

Given current expectations, the mathematics curriculum document needs to deliver 'success for all'. This is not inconsistent with MiNZC, for the achievement objectives were prefaced by the statement "*[w]ithin a range of meaningful contexts, students should be able to:"* (e.g., p. 130, emphasis in original). It is therefore appropriate to base the judgement decision upon this expectation. However, requiring 100% of students to be correct is unrealistic for the following reasons:

1) Educational measurement and assessment is never perfectly reliable. This means there must always be allowance for this lack of reliability;
2) In the case of this research the problem of reliability is compounded because the judgements are made on individual items, not on the whole test. This means that the allowance should be greater than what ideally would be chosen;
3) The literature on learning is clear that learning is uneven (e.g., Hall, 2000) so not all students can be expected to demonstrate the same understandings at a particular point in time; and
4) There are always a small number of students whose educational progress is limited. This is especially true in New Zealand which has an inclusive approach to the education of students with disabilities and learning problems.

Therefore the more conservative figure of 75% success was chosen in this study as the standard of evidence to measure whether or not students are meeting the curriculum expectations for scale understanding. This figure has been chosen as it is half-way between 50% (an important figure in norm-referenced assessment, and also indicates that there are as many students not reaching the standard as there are those who have achieved the standard) and 100% (success for all). The figure of 75% is an approximate guide which takes into account the four problems outlined above and is consistent with the notion of a 'standard' which most students should be able to reach, rather than a 'norm' which is focused on the notion of discriminating between students.

Sources of evidence

While there is a single type of evidence upon which the judgement decisions are made (student responses to the items developed to assess understanding of scale), this evidence is collected from multiple sources. For example, data are collected from several schools, and more than one test is developed. The process of data collection is also built upon two suppositions. The first is that the developed items can provide a measure of the intent of the curriculum. The scale construct, its relationship to the curriculum, and the items to measure understanding of the construct and curriculum were first explored in Chapter 2 and are further developed in Chapters 4, 5, 6, 7, and 11. This progressive development seeks to ensure that all reasonable care has been taken to establish the link between the items and MiNZC (Ministry of Education, 1992). The second supposition is that the students assessed at Schools 1, 2, and 4 have had exposure to MiNZC prior to Year 7. Given that MiNZC had been the official mathematics curriculum document since 1992, and that the changes in practice promoted by the NDP predominantly impacted on the teaching of number, assuming that the teaching of scale in measurement and statistical contexts was broadly in line with MiNZC is not unreasonable.

The presentation of conclusions and the decisions that can be based on these

How effective the scale-related learning of MiNZC is likely to be as a policy direction for the revised curriculum will be evaluated by considering under what conditions the students in the case study schools met the 75% success rate criteria. However, as the study is exploratory in nature, other findings cannot be anticipated in advance. Rather, implications for change will develop from a consideration of the identified student understandings and the relevant literature. To ensure the findings are robust so implications drawn from them are valid, all data will be subject to confirmation through the use of *triangulation*.

5.3.5 The use of triangulation

Within an evaluation methodology, the notion of triangulation is important as “evaluation is an intensely political undertaking”, so “solid, irrefutable evidence” must back up whatever conclusions are drawn (Davidson, 2005, p. 55).

There are various forms of triangulation, but the key idea is the use of two or more means of data collection when studying some aspect of human behaviour. For this study both *between methods* triangulation and *within methods* triangulation are used (Cohen & Manion, 1994). Between methods triangulation involves using more than one method of data collection for a given objective. This helps ensure that the interpretations, inferences, and conclusions drawn from the data are accurate. It does this by examining the issue from the perspectives of different information sources and data collection methods, seeing how well these support and point to a particular interpretation. *Within methods* triangulation uses the replication of a study as a check on reliability and theory confirmation (Cohen & Manion, 1994). The use of several cases is intended to provide opportunities for "verifying the repeatability of an observation or interpretation" and challenging explanations (Stake, 1998, p. 96).

Two other conceptions of triangulation, as evident in the way researchers investigate their data, are identified in Hall (2002). He identifies these as: *corroborative triangulation* and *coherence triangulation*. In corroborative triangulation, data from different sources are used to see if they corroborate each other (the normal interpretation of triangulation). In this study, one use of corroborative triangulation is the comparison of the findings of the research literature with those from the diagnostic assessment undertaken by students in this research (see Figure 5.2). In coherence triangulation different sources of information are used to identify if a coherent picture is emerging; coherence is then taken as a proxy for corroboration. "Corroboration is inferred through the overall coherence of the picture that is formed from the different data" (Hall, 2002, p. 3). In this study, one example of coherence triangulation is the use of the curriculum and resource reviews to build knowledge of how students develop their understanding of scale. Student data from the diagnostic assessment, and the findings of the literature review then add to the picture of the issues students have when developing that understanding. This is shown in Figure 5.3.

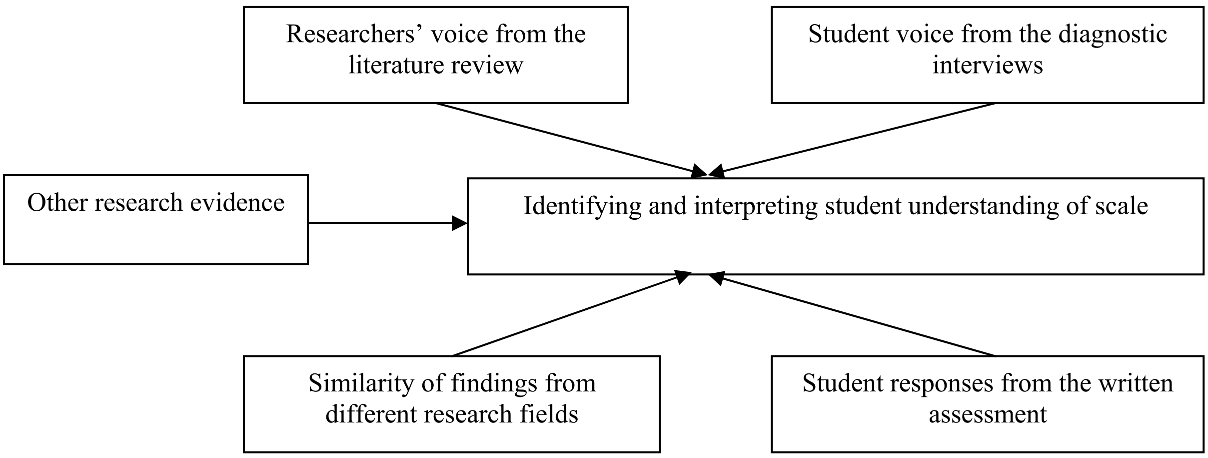


Figure 5.2. Corroborative triangulation used to establish student understanding of scale

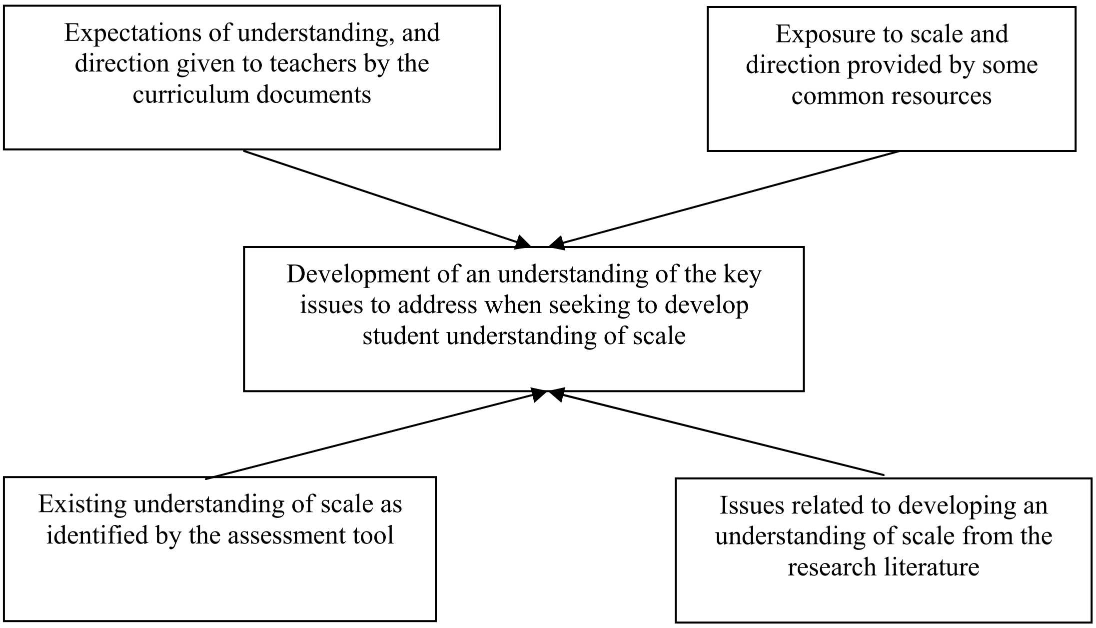


Figure 5.3. Coherence triangulation used to develop an understanding of how to address scale understanding

Overall, the use of varied data sources and collection methods aims to ensure that the interpretations made are as well informed as possible. Figure 5.4 shows a triangulation pyramid. This indicates how triangulation is used at different stages of the research.

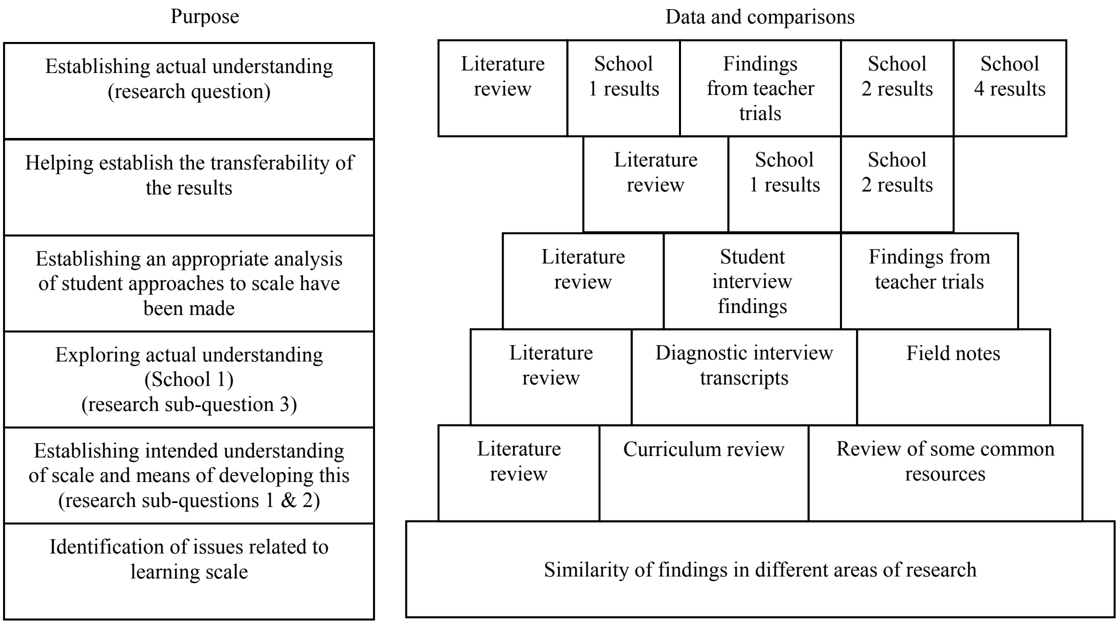


Figure 5.4. The triangulation pyramid: how triangulation is used in this study

5.3.6 Case studies and an evaluation methodology

To close the research methodology section, the role of case studies within evaluation is now considered; the next sub-section then discusses the use of case studies in this study. Robson (2002) notes that as most evaluations are concerned with the effectiveness and appropriateness of ‘something’ in a given setting, a case study strategy, rather than a sample, is often appropriate. Yin (2003) also identifies that case studies are appropriate for exploratory studies asking 'what' questions, and have a distinctive place in evaluation research. This suggests that case studies are likely to be a useful strategy for exploring understanding of scale. Figure 5.1 showed the use of case studies within the flexible design. Figure 5.5 shows how the nature of the findings quickly altered that design as discoveries (and further explorations of these) were made.

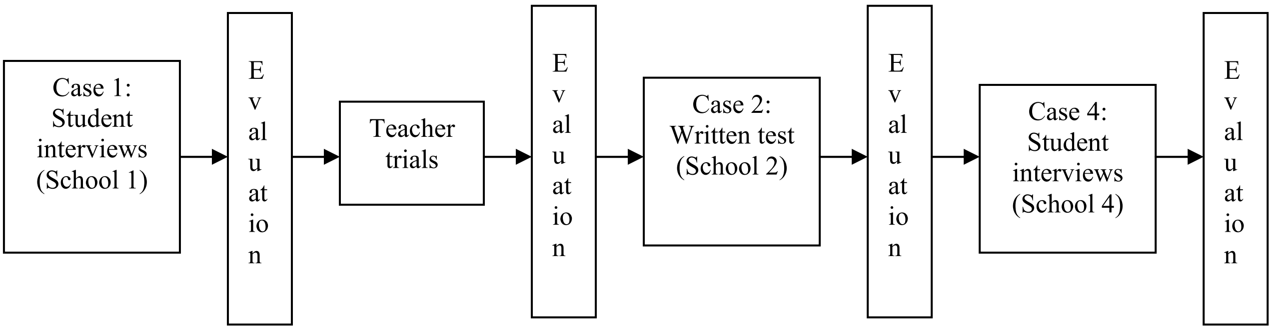


Figure 5.5. Final (flexible) design highlighting the use of case studies and evaluation

The use of case studies

"Case study is not a methodological choice, but a choice of object to be studied" (Stake, 1998). Case studies are used to "illuminate the general by looking at the particular" (Denscombe, 1998, p. 30). Case studies put the spotlight on one or a few instances of the phenomenon being investigated, looking in depth for things that would not have become apparent with a superficial process. The focus is a holistic approach to the natural setting, looking for events, connections and relationships, the 'why'. Case studies involve a rich narrative based on multiple data sources (Denscombe, 1998). They involve an empirical inquiry that investigates a contemporary phenomenon within its real-life context, allowing direct observation of the events being studied, and interviews with those involved in the events (Yin, 2003). The purpose is to "learn enough about the case to encapsulate complex meanings into a finite report but to describe the case in sufficient, descriptive narrative so that readers can vicariously experience these happenings, and draw their own conclusions" (Stake, 1998, p.100). To enable this to happen, the choice of case is important (Yin, 2003).

Choice of cases

When using a case study strategy, the cases are not 'sample units', rather they are chosen for a reason; a specific purpose within the overall inquiry. Case study designs can also involve a single case or multiple cases. Multiple case designs have the advantage of producing more compelling evidence, but can limit the possible choice of cases; for example the revelatory or rare case may not be possible choices (Yin, 2003).

For this study a multiple case design was chosen. In the original design, Case 1 was a small exploratory study, while Cases 2 and 3 were to involve two trial schools in a quasi-experimental teaching intervention based on the findings from Case 1. Each of these schools was initially chosen because they had been involved in the professional development (PD) for INP with the researcher. This meant:

1) The researcher had access to each school;
2) The researcher had developed a working relationship with the teachers and some students;
3) The students were familiar with being asked to explain their thinking;
4) The teachers were familiar with the intended style of PD and teaching, and were known to be part of a professional learning community; and
5) The students were familiar with the style of learning planned for the intervention.

Being currently involved with INP was considered particularly important for Schools 2 and 3 as it was identified that to have an impact, the (short) scale PD and teaching intervention should be attached to other long-term, in-depth PD with the characteristics the literature identifies as necessary for successful PD and learning.

In the event, the nature of the findings from Case 1 led to an early adaptation of the research design. A collection of teacher trials was added to test the reliability of these interim findings (see Chapter 9) while a fourth case study (School 4) was later added to explore further student understanding (see Chapter 12). While School 4 was also participating in INP, its participation in INP was not a factor in its choice. It was a larger, higher decile school (higher socio-economic status) so had the potential to produce different results from the earlier cases. As noted in Section 1.6, Case 3 was studied, but its findings are not reported in this thesis.

Case studies and the role of theory

In case study research, *theory development* is an essential part of the design phase. However, this is not possible in situations where the existing knowledge base is poor and the available literature provides no conceptual framework or hypotheses of note. *Statistical generalisation* is also not generally relevant as this process involves an inference being made about a population on the basis of empirical data collected from a random sample. Neither cases nor the individual studied in the case are ‘sample units’, having been chosen for a reason. As a consequence, *analytic generalisation* tends to be used instead. In this process, generalisation is made to theory rather than to another case (Yin, 2003). Other cases can then be used to establish the transferability of the findings.

As there was little existing theory to explain how students develop an understanding of scale, this thesis is exploratory. Where propositions or a conceptual framework were identified, these have been used as a starting point for investigation, but in general Case 1 was used to develop theory, and subsequent cases were used to test the transferability of that theory. For each case an assessment was used to identify student understanding. More information on these tests and how they were administered can be found in Section 5.5 while their development is addressed in Chapters 6, 7, and 11. The next section describes the participants in the study.

5.4 Participants

5.4.1 The case study schools

Students at three schools in the Wellington region participated in the study reported in this thesis. School 1 was used to trial the assessment tool, discover any improvements needed, and identify preliminary findings about student understanding of scale. At School 2 a larger body of students participated, sitting a written version of the assessment. This helped quantify the magnitude of the issues identified at School 1 and assessed the effectiveness of the written test format while also providing an opportunity to replicate the findings from School 1. Students at School 4 were then interviewed with a different assessment designed to build on what was learned and resolve outstanding issues (Figure 5.6).

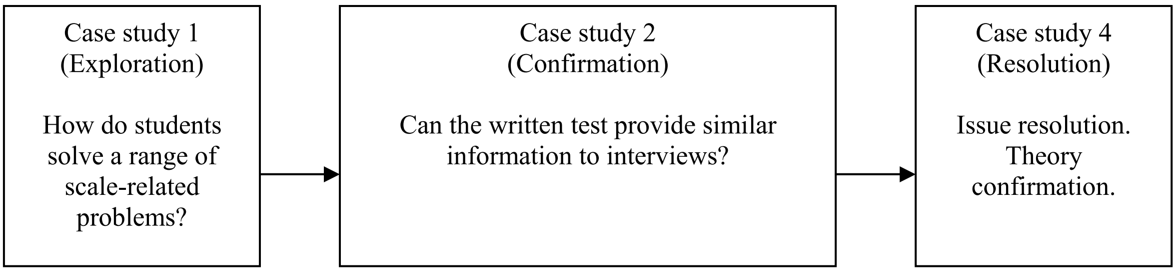


Figure 5.6. Simplified research design highlighting the role of case studies

School 1

School 1 is a suburban mid-decile intermediate of 300 students where classes mix Years 7 and 8. Current ethnicity data show that 69% of students are New Zealand European, 25% are Māori, and 6% Pasifika. Staff and students had previously participated in INP, so students had a mathematics programme in which the number strand of MiNZC (Ministry of Education, 1992) was re-interpreted through the Number Framework (Ministry of Education, 2005a).

Originally, 18 students were to be interviewed by the researcher, six from three different classes. Two 'more able', two 'average' and two 'weaker' students were requested from each class; a mix of boys and girls was also requested. Teachers were asked to choose students who would be appropriate to interview. Due to paternity leave for one teacher, students forgetting to bring the ethics forms, illness, and students being unexpectedly involved in a field trip only 13 interviews were possible over the available three days. At this stage it was deemed that there was sufficient information to show that the diagnostic items were worded appropriately for student understanding, and were fulfilling their purpose of providing useful information about a student's thinking. A range of student strategies for answering the items had also been identified.

School 2

School 2 is a suburban low-decile intermediate of 200 students where classes mix Years 7 and 8. Current ethnicity data show that 21% of students are New Zealand European, 72% are Māori, and 7% Pasifika. Eighty one students from the six classes and all six teachers volunteered to participate in the research. At the time staff were involved in the second year of INP, so again students would have had a programme of work in which the number strand was reinterpreted. No feeder schools had been previously involved in the NDP, so students entering Year 7 had a mathematics programme based on MiNZC (Ministry of Education, 1992).

School 4

School 4 is a suburban high-decile intermediate of 500 students in the second year of INP where classes mix Years 7 and 8. Current ethnicity data show that 69% of students are New Zealand European, 19% are Māori, and 12% Pasifika. Some of the feeder schools had previously been involved in the NDP so students may have been exposed for several years to a mathematics programme based on NDP resources as well as MiNZC. Thirty two students from two classes were selected by their teachers for an interview similar to those conducted at School 1. A new version of the assessment tool was developed for this purpose: Test 2. The items included several from the first test along with a collection of new items (see Chapter 11). Details relating to how these three cases were defined and analysed can be found in Section 5.4.3.

5.4.2 The teachers

The teacher trials involved Wellington primary and secondary teachers who were asked to locate a collection of numbers on different number lines, and explain in writing how they had done this. In total, 22 teachers from the first trial, eight teachers from the second trial, and three teachers from the third trial agreed for their work to be included in the study. As only 11 teachers were involved in the latter two trials, a great deal of caution is needed when interpreting the results, even though all of the teachers' responses provided positive evidence for the hypothesis being tested.

5.4.3 Defining the case and the unit of analysis

Yin (2003) identifies that when using a case study it is important to identify the case, its boundaries, and the unit of analysis. For this study, a case has been defined as a school, while the unit of analysis is the scale understanding of those students at the school who were involved in the data collection exercise. The school is an appropriate case as it provides a clear case boundary. Within each school the students are likely to have some common exposure to understanding due to sharing the classes in which they are taught. Teachers may also share a long-term plan for the structuring of the mathematics programme and could share planning as well. Defining the unit of analysis as a *cluster of individuals* operating within a school is a powerful tool for an exploratory study. It removes the need to work in detail on the understanding of each individual, which would restrict the use of the evidence. Working with the collective allows the different explanations of individuals to strengthen the picture of individual understanding. For example, the words of one student may not make clear what is understood, as that person may not be able to articulate their understanding clearly, whereas the words and explanations of a group of students with the same result for a question can shed multiple perspectives (some triangulation) on what is actually 'understood' and may have led to the response. How sense was made of the identified understandings is explored in Section 5.6; the procedures used in the study are considered next.

5.5 Procedures

5.5.1 The assessment for this study

A test, according to a joint publication of the American Education Research Association (AERA), American Psychological Association, and National Council on Measurement in Education (1999) is "an evaluative device or procedure in which a sample of an examinee's behaviour in a specified domain is obtained and subsequently evaluated and scored using a standardised process" (p. 3). The initial assessment developed for this study conforms to this definition. Furthermore, it is designed to identify student understanding of scale. Here, student misconceptions and lack of understanding need to be diagnosed if the tool is to be useful. For these reasons the assessment tool will be referred to as a *diagnostic test* or a *diagnostic* from now on.

5.5.2 Cognitive interviews

Cognitive interviews were used at Schools 1 and 4. These are interviews that focus on:

> … providing a view of the processes elicited by the questions. Concurrent and retrospective *think-alouds* and/or probes are used to produce reports of the thoughts that the respondents have either as they answer the survey questions or immediately after. The objective is to reveal the thought processes involved in interpreting a question and arriving at an answer. (Presser et al., 2004, p. 4)

In the context of mathematics education, such interviews are similar to those used in traditional Piagetian research.

When developing a test, it is important for the designer to verify that the items can be easily understood and provide the required data. This test of fitness for purpose was undertaken when conducting the one-on-one interviews at School 1. Interviews were appropriate as this phase of the research required detailed information and insights about student thinking which can only be collected through in-depth questioning.

The interviews were also used to develop an understanding of how students solve scale-related problems. For this, student responses were audiotaped and a transcript produced. An audiotape was appropriate as once started it is an unobtrusive method of collecting data. A transcript was appropriate as it allowed the "*interviewee's words to be treated as 'on the record' and 'for the record'*" (Denscombe, 1998, p. 109, emphasis in original). Written field notes were also taken, as non-verbal signals are not evident on an audiotape (Denscombe, 1998). Using the transcript and the field notes allowed the data to be triangulated to confirm the authenticity of the record. The field notes also provided a back-up to the audiotape in cases where there were problems with the tape or the audibility of the recording.

Along with interviews, written tests were also used to assess understanding of scale.

5.5.3 Written tests

Two forms of written test were used in this research. Firstly, the diagnostic from School 1 was used as a written test at School 2. This was to ascertain if it could provide information of a similar quality to the interview, and if students at this school had similar issues to those at School 1. A written test on its own provides a good deal of scope for different answers, hence has value as a diagnostic tool. Data from the interviews at School 1 also provided information to interpret student responses in terms of the strategies they may have used to reach their answer. Secondly, in the teacher trials, adults were given specifically designed items, with instructions to answer them, then explain how they had worked out their answer. These data were used to evaluate this form of testing, and as a form of theory confirmation; could the findings from School 1 be used to predict the form of response of such a different group, or would they refute those findings? The forms of data analysis are outlined below.

5.6 Analysis

The cognitive interviews provided rich qualitative data in the forms of audiotapes, examples of student work, field notes of student responses and observations of what the students were seen to do while answering and, from School 1, a follow-up semi-structured questionnaire. This questionnaire collected background information about the test and its difficulty that did not come to the fore as the items were answered. For example, one question asked students if they had worked before with each of fractions, decimals, and integers. All these data were then analysed to provide the preliminary 'map' of scale understanding.

The audiotapes were transcribed by a research assistant, with the researcher comparing the transcript to the field notes to confirm each record. For one interview (at School 1) the tape recorder was accidentally not switched on, but the field notes provided a good record of the student's answers and solution methods. To enhance the consistency of coding decisions and interpretations, the researcher 'marked', coded, and categorised all student responses, kept notes of decisions made when responses were interpreted, and used c*ontent analysis* to analyse the gathered data. "Content analysis is a careful, detailed, systematic examination and interpretation of a particular body of material in an effort to identify patterns, themes, biases, and meanings"

(Berg, 2007, pp. 303-304). Typically it is performed on existing records of human communications (books, documents, audiotapes, and video footage).

Analysis of the form of data provided by the interviews is not a neutral process, in that the process of coding, developing categories, and assigning data to these categories is subjective. Coding is a process whereby a person makes interpretations of the intent of the student. In doing this a person puts their own (adult) frame of reference onto the work, sometimes producing an inaccurate portrait of student thinking (Gould, Outhred, & Mitchelmore, 2006). The process of interpreting and deriving meaning is also value laden, while the selection of extracts involves a "*level of judgement and discretion on the part of the researcher*" (Denscombe, 1998, p. 135, italics in original). Categorising then allows similarities and differences in the responses to be considered for the purpose of identifying trends and patterns. Different categories may be created by different people who hold different knowledge, beliefs, and assumptions, so trends identified through this process may be more a product of the categories created than a reflection of real patterns in the data. However, this allows a situation to be viewed from a variety of perspectives, each of which can develop new understandings.

The overall approach taken to the data was interpretive and followed the process of *analytic induction*. This involved the verbal interview data being transcribed into text for analysis. The researcher then immersed himself in the data to identify themes using a constant comparison method of joint coding and analysis. The approach ensured the data analysis was grounded. It also allowed the information to be used to develop an understanding of the situation and to reduce the data by looking for patterns, trends, and meaning. Tentative hypotheses could also be formed and then tested against later data (Berg, 2007).

5.6.1 School 1

In the first phase of the analysis of School 1's data (the *open coding* phase) each item was considered independently. The responses for an item were put into natural 'categories' according to the answer given, and frequency tables of the different answers were drawn up. Next the responses within each category were analysed word-for-word to identify themes. Here similar phrases from different individuals were used to identify these themes, which were then coded. Coding was sometimes simplified by referring to the more elaborate responses of other students, and sometimes by referring to the research literature. For the chapter write-up, verbatim extracts from the transcripts were used to illustrate how students solved the items and to exemplify the range of thinking strategies used for each item. Illustrative examples that clearly articulated the exemplified thinking were chosen. Illustrative examples were provided as only a few students were interviewed at School 1. Where research identified a similar form of response this was also referenced, providing some triangulation for interpretations undertaken by the researcher.

Categorical labels were then developed for the identified themes. Where possible these labels summed up the strategy used. Sometimes these labels were based on the researcher's knowledge of the field of study and related to terms used in the literature (e.g., skip counting). At others they used indicative wording from the students' responses (e.g., a little bit less).

Responses for all questions were sorted into these categories, with both the interview transcript and the researcher's field notes being consulted as part of this process. At this stage trends across items were explored, in particular the responses for related items were compared. Commonalities between approaches to different items were used as a cross-check on the categories. This resulted in some categories being revised and the data resorted, so the categories could be used across questions while accounting for all the data. The sorted materials were then further analysed to isolate meaningful patterns. In undertaking this analysis the researcher's knowledge of the literature led to the proposal of a strategy hierarchy. The tentative findings from School 1 were then subjected to negative case testing, and later on, data from some items were recoded 'from scratch' to ensure consistency of interpretation over time and between schools.

5.6.2 The teacher trials

The teacher trials allowed the process of negative case testing to be used and provided a measure of the validity and fit for Test 1. Here groups of teachers were given written questions with the intention of interrogating the categories and the tentative hierarchy of strategies from School 1. The data from the trials were sorted directly into the categories from School 1, as the written explanations of method provided similar information to the interview transcripts from that school. Data were initially tabulated, although a mix of verbatim extracts and numeric totals were used for reporting. The extracts tend to be illustrative rather than random as they generally related to a particular point.

5.6.3 School 2

The data from School 2 also enabled the process of negative case testing to be used. Analysis began with a research assistant entering these data into an EXCEL spreadsheet. The researcher then checked atypical results against the original test paper to verify and clean the data. Several random papers were also cross-checked to ensure the reliability of the spreadsheet data. Tables identifying response frequencies were created. Such tabulated data allowed the researcher (and allows the reader) to identify item facilities and trends, as well as common errors. The *cause* of a response was then inferred by referring to the logic used by the students interviewed at School 1 and from similar responses found in the literature. Discrepancies between the data from Schools 1 and 2 were identified and have been reported, where possible using randomly selected examples of student work.

Data were later recoded and entered into SPSS, and this software was used to calculate correlation coefficients and a level of significance for each test item. The correlation coefficient provided a measure of the relationship between a correct response to an individual item and the total 'mark' a student received for the test. The level of significance gave an indication of how well the item measured this relationship. Both statistical procedures were appropriate given the form of the data, and the number of students involved. Note that further detail is provided in Section 10.3, and that this process was also used on the data from School 4.

5.6.4 School 4

School 4 was the final case. Responses from School 4 were treated similarly to those at School 1 to provide an independent check of the analysis undertaken at School 1, and where discrepancies were found the data at both schools were revisited, and categories adjusted where necessary. Data were also interrogated to answer specific questions that had arisen during the earlier stages of the research. In the write-up, verbatim extracts from the transcripts and examples of student work (randomly chosen where possible) support the discussion of the points raised.

5.7 Validity

5.7.1 Validity in relation to the research design

When designing research, it is important to adopt procedures that ensure its *validity*. In this sense, validity focuses on *fitness for purpose*. Does the research answer the questions or solve the problems it aimed to investigate? (Hall, 2007b). *Trustworthiness* is an equivalent notion to validity commonly used with qualitative research contexts. Research is trustworthy if the results can be relied upon because of:

- coherence of the overall design with respect to the research questions;
- the use of appropriate methods, strategies, and data analysis techniques;
- clarity and detail of the information about the research process;
- evidence of appropriate or reasonable interpretations; and clarity of reporting (Hall, 2007a).

Within an interpretive paradigm, the validity or trustworthiness of research is also considered in relation to the *plausibility* or *credibility* of the results. This aligns closely with the third and fourth bullet points above and requires a rich description of the research process and data to enable credibility to be inferred. Validity is also enhanced through consideration of the *transferability* of the results. Again a rich description is required so readers/users can decide if the results, or aspects of the results, fit their own context (Hall, 2007a).

Yin (2003) identifies that a criticism of case study research is that the "investigator fails to develop a sufficiently operational set of measures and that "subjective" judgments are used to collect the data" (p.35). Attempts have been made to address this criticism throughout the study. The development of the set of measures used to assess understanding of scale has been the subject of rich description. Multiple sources of evidence have been used at each stage of the research to ensure appropriate interpretations have been made. Attempts have been made to clearly present the collected evidence, the interpretations made of this evidence, and the links in the evidence chain. Important findings in the first case resulted in a design change in the study so these could be tested in subsequent cases. Procedures have also been adopted to maximise the possibility that valid tests have been developed for measuring scale understanding.

5.7.2 Validity and test design

During this study, a diagnostic is developed. The validity of a test, according to the AERA et al. (1999, p. 9) is "the degree to which evidence and theory support the interpretation of test scores entailed by proposed users of tests" and is the most fundamental consideration in the development and evaluation of tests. Evidence of validity needs to be accumulated to provide a sound basis for the intended interpretation of results. The process of validation then proceeds by developing empirical evidence, examining relevant literature, and/or conducting logical analyses to evaluate each of the propositions. Important contributions are made by other researchers' findings, and an analysis of the content of the test and the concept it is supposed to measure. The analysis of the responses of test takers provides validity evidence of the fit between what is being measured and what is intended to be measured (AERA et al., 1999).

Because two tests are used to measure student understanding of scale in this research, ensuring test validity is central to this study's credibility. Thus a range of procedures have been adopted to address test validity. They include:

- undertaking an analysis of the construct to map what is intended to be measured;
- reviewing relevant research to identify a need for the test;
- linking what is mapped in the construct to the proposed items to ensure coverage of the construct;
- using or adapting items recognised as useful in relevant research;
- detailing the procedures used to develop and administer the diagnostic, and interpret the results;
- outlining the intended interpretations and uses of the data;
- describing the characteristics of the examinees;
- analysing the trial interviews to ensure that what was intended to be measured was actually being measured;
- using evidence from research to support the interpretation of test results;
- using a teacher trial to test the researcher's interpretations of findings with a sample of adults;
- reviewing the concept map where the responses of test takers indicated a discrepancy with the map; and

- developing Test 2 on the basis of the responses of the test takers and the revised concept map.

It is hoped that the reader can determine for themselves whether or not appropriate efforts have been made to ensure test validity. In the next section, issues relating to reliability are addressed.

5.8 Reliability

5.8.1 Reliability in relation to the research design

Reliability focuses on the notion of accuracy, and requires attention to detail at all stages of the research in both forms of paradigm. In the interpretive paradigm, reliability is mainly assessed through an 'audit' analysis of the rich narrative, and is operationalised as *dependability*. If the research follows sound procedures, gives evidence of attention to detail, includes triangulation, and uses member checking, the results are interpreted as being dependable. In comparison, within the positivistic paradigm accuracy is inferred from evidence of 'repeatability', that is replication of the research produces similar findings. However, in qualitative studies replication studies are very often impractical (Hall, 2007a).

Yin (2003) identifies that in case study research, as every case is unique, it is not really possible to generalise from one case to another, or use the positivistic concept of reliability as this equates to another researcher repeating the same case study and getting the same results. Instead *replication logic* is used. Under replication logic the researcher is seeking to establish if the findings are robust and "worthy of continued investigation or interpretation" (p. 47). Either the exact conditions used in the case are replicated, or one or two conditions are altered to see if the findings can still be duplicated. In this study, the teacher trials used altered conditions to ascertain if contrasting results could be identified for predictable reasons, while Cases 2 and 4 sought result replication under altered conditions. How reliability was established for the diagnostics developed for the research is discussed in the next section.

5.8.2 Reliability and test design

Within the field of test construction, the concept of *test reliability* invokes the notions of *consistency*, *stability,* and *repeatability*. It focuses on the measurement process and whether a test provides an accurate or consistent measurement of performance when the testing process is repeated (Hall, 2007a). In addition to procedures identified in earlier sections, the following processes have been adopted to help ensure the reliability of the testing:

- stipulating procedures for test administration with consistent use of test materials to ensure fairness for all examinees;
- using the researcher as the interviewer and marker of all of the tests to control for marker variation;
- check marking to ensure consistent use of marking protocols; and

- monitoring the coding and the response patterns from later tests to look for consistency of response.

In this way, it is hoped that the responsibility of the researcher to investigate reliability within acceptable practical considerations has been met. The procedures to meet ethical guidelines are now explored.

5.9 Ethics

Educational research has an obligation to safeguard the rights and welfare of participants. This includes taking care to not disturb, or cause anxiety to, or harm participants and their communities, through either the details published in research reports, or participation in the research (NZARE, 1998). Procedures adopted for this research, and outlined here and in Chapter 6, have attempted to meet this obligation.

Informed consent is an important principle for research involving human participants. This requires potential participants to be given a clear description of what the research involves, how it will be reported, and the public availability of the reports. It allows prospective participants to make an informed choice about their involvement (NZARE, 1998). Such approval has been sought from, and granted by, all research participants, including the parents or guardians of the children. Approval was sought for the collection, use, and reporting of the data.

This research has been undertaken in a spirit of open enquiry, with the ideas generated being subjected to open scrutiny by teachers and peers. The knowledge generated is reported as objectively and clearly as possible, with limitations in techniques and the influence of particular theories noted.

This research has been conducted under the approval of the Human Ethics Committee of Victoria University of Wellington (College of Education subcommittee).

In the next chapter, issues pertaining to the test design and interviews are addressed in more detail.

Chapter 6

Methodology: Instrument Design and Interview Protocols

6.1 Introduction

Scales are such common mathematical objects that it is easy to take for granted that students understand how they operate. This research is designed to assess whether or not it is appropriate to assume that students develop an understanding of scale through exposure to them. To do this it is important to develop effective tools for measuring what the students in the case study schools understand about scale.

This chapter outlines the processes used for developing the tests. It starts by acknowledging the previous studies that have influenced the research and have helped shape its purpose before moving on to consider the development of the initial diagnostic (Test 1) and the protocols for its use. The development of the second diagnostic (Test 2) is then outlined. Note that a discussion of the development of the actual items for Test 1 can be found in the next chapter, while Chapter 11 contains a similar discussion of the development of the items in Test 2.

Table 6.1: Tools to measure scale understanding

Location of trial	Test version
School 1.	Test 1 (interview version).
Teacher trials.	Several specifically designed items (written format with explanation of solution method required).
School 2.	Test 1 (written version).
School 4.	Test 2 (interview version).

6.2 Background: The CSMS study and the NDP as contexts for this study

The CSMS research (e.g., Hart & Johnson, 1980) has influenced this study, with the rationale, methodology, and findings playing a role in determining the direction undertaken. Hart (1978) identifies that topics taught ad nauseam at school are often not well understood, and she asks why this happens, suggesting that if we can identify the causes of the errors, perhaps we can devise a cure for the students making them. In Hart (1981) she goes on to add that the way teachers present a topic is generally affected by what understanding they believe a child already has, and points out that:

> Very often secondary school teachers are re-teaching what was first presented in the primary school. In many cases it was not understood when first presented, or even when 'taught' the second time. After a history of failure at decimals or fractions it is little wonder that the next teaching of it in secondary school is greeted with dismay. Is there really any point in teaching something we know most children will not understand? One reason given for doing this is that the child will become familiar with the idea and understand it later. We have no proof of this, in fact our results show that the understanding does not

> ‘come’. Surely all that happens is that the child becomes familiar with a lack of success and that mathematics is something that you do but it makes no sense. (p. 217)

Have educators been guilty of making assumptions about what students understand? Have teaching resources reinforced this, not through what is taught, but through what is omitted and the way material is presented? The literature reviewed and the documents analysed earlier in this thesis indicate that these are strong possibilities.

The literature review also identified that students often develop their own strategies for solving problems, rather than learning those taught, and/or that algorithmic approaches to teaching causes problems (Behr et al., 1983; Bragg & Outhred, 2000a, 2000b; Hart, 1989; Mack, 1993; Nunes & Bryant, 1998). According to the stages of the Number Framework (Ministry of Education, 2005a) the strategies students use to solve number problems reflect their mathematical thinking; they interpret what is taught in relation to their own understanding of how numbers work. Over time and at their own pace, student understanding progresses from counting to an additive approach, and finally to multiplicative then proportional reasoning (Table 6.2).

Table 6.2: The Stages of the Number Framework

Numeracy Stage	Predominant form of thinking	Features of conceptual development
0 - 4	Counting.	Tend to use counting (rather than known facts) to solve problems. Counting in ones is common, but simple skip counts (e.g., 2, 4, 6, 8) may be used.
5 - 6	Adding.	Some knowledge of basic facts exists, especially in relation to addition and subtraction. Beginning to learn about multiplication. Tend to choose strategies based on adding or subtracting to solve problems.
7	Multiplicative.	A range of multiplication and division facts are known. Tend to apply strategies based on multiplication or division to solving problems, but can still use those based on adding or counting.
8	Proportional.	Can use a range of strategies to solve problems involving fractions, decimals, proportions, and ratios.

Diagnostic interviews with hundreds of thousands of students have identified that without intervention a sizable proportion of 11 and 12-year-olds still operate mentally at the counting stages. With teaching explicitly targeted at developing more sophisticated mental strategies this proportion decreases (Higgins, Irwin, Thomas, Trinick, & Young-Loveridge, 2005; Irwin, 2003, 2004; Irwin & Niederer, 2002). The research cited above suggests that, to assess scale understanding for this thesis, the developed diagnostics need to include items set at different levels of difficulty that allow students to use a method of their own choosing. However, how the students go about answering the items and possible links to number understanding need to be identified.

6.2.1 Links between scale understanding and number understanding

Friel et al. (2001) suggest that it may be important to consider number knowledge when considering graph comprehension, as mathematical knowledge and experience have been identified as characteristics of developing this. Within the NDP, the *numeracy project diagnostic interview* (NUMPA) (Ministry of Education, 2005b) is used to identify the mental strategies students have to solve number problems. In this one-on-one interview, the strategies students use are measured and mapped to the three strategy and four knowledge domains of the number framework (Ministry of Education, 2005a). The strategy domains are addition and subtraction, multiplication and division, and fractions, ratios, and proportions. As part of INP, teachers at School 2 were due to interview their students using the NUMPA close to the time the scale diagnostic was sat, so permission was sought to use those data as part of this research. The NUMPA is nationally recognised as a good measure of number understanding, so referencing scale responses to NUMPA results allows the impact of number understanding to be assessed. Also, as the NUMPA is in the form of a diagnostic interview, students who have participated in it are already familiar with the process of explaining their reasoning as part of an assessment.

In the next section the scale construct is revisited. The understanding gained from undertaking the literature review and the document analysis is used to refine and then map the content domain.

6.3 Mapping the content domain

Scales have been identified as existing in a variety of situations. Initially, all forms of scales and student experiences with them were considered for this study, with a definition of the construct being developed in Chapter 2. However, the reading undertaken has indicated that, as outlined below, further clarification of the construct is needed. In particular:

1) The full Cartesian plane has not been introduced.

   The curriculum review identified that students do not use the full Cartesian plane in Years 7 and 8, so graphs should be restricted to ‘the first quadrant’. Rational and real numbers are only explored as numbers greater than zero, so negative fractions and decimals should be excluded. The curriculum review also identified that students may not have met scales (that conform to the definition developed in Chapter 2) in the context of abstract graphs based on the Cartesian system. Therefore decontextualised graphs based on point plotting should also be excluded.

2) Map scales should be excluded from the construct.

   While the document analysis identified that simple scale maps are met at CL3 in MiNZC (Ministry of Education, 1992), and can involve important mathematical concepts, students may not have worked with the scale in this context. This is because it is possible to answer many map reading questions by multiplying a length read from the map by the map ratio, a process that is easier than using the map scale. Other map-based activities can involve the development and use of a map coordinate system which predominantly uses counting-based conventions. As graphing with coordinates is to be excluded, so too should maps and coordinates.

3) Scales on real-life measuring instruments are too complicated.
Real measurement instruments often involve two scales on a single line. For example, on a protractor, these scales are similar but reversed; weighing scales show units from different measurement systems; and a voltmeter uses the same unit but different sized intervals for a range of scales. It is assumed that to be successful in reading such instruments, students need to encounter simpler scales first. Test 1 therefore should use simplified versions of scales rather than real measuring instruments.

4) Large numbers can cause problems.
The NDP has identified that many students in Years 7 and 8 have issues reading large numbers, including saying the number one before or one after (Higgins et al., 2005; Irwin, 2003; Irwin & Niederer, 2002). Real-life scales on graphs often use large numbers, so a graph of New Zealand's population may have a scale in hundreds of thousands. As poor performance on a scale using large numbers may not indicate scale-related issues, problems in Test 1 should be restricted to small numbers.

In the next section the concepts within the scale construct are put into the form of a map. This seeks to make clear what is considered to be part of the construct, and how the different concepts are inter-related.

6.3.1 The concept map

Figure 6.1 indicates that there is a dynamic inter-relationship between the various aspects of scale. Firstly, there are a variety of applications in which scales are found, and a range of numbers that can be used within these applications. It also identifies that there are two main levels of understanding:

1) knowledge and use of scale conventions (that is understanding the object); and
2) being able to work with scales to read and locate numbers (understanding the processes applied to the object).

Lastly, it shows that several types of scale exist, each of which brings its own challenges when considered in relation to the application, the type of number used, and the level of understanding a student has of number.

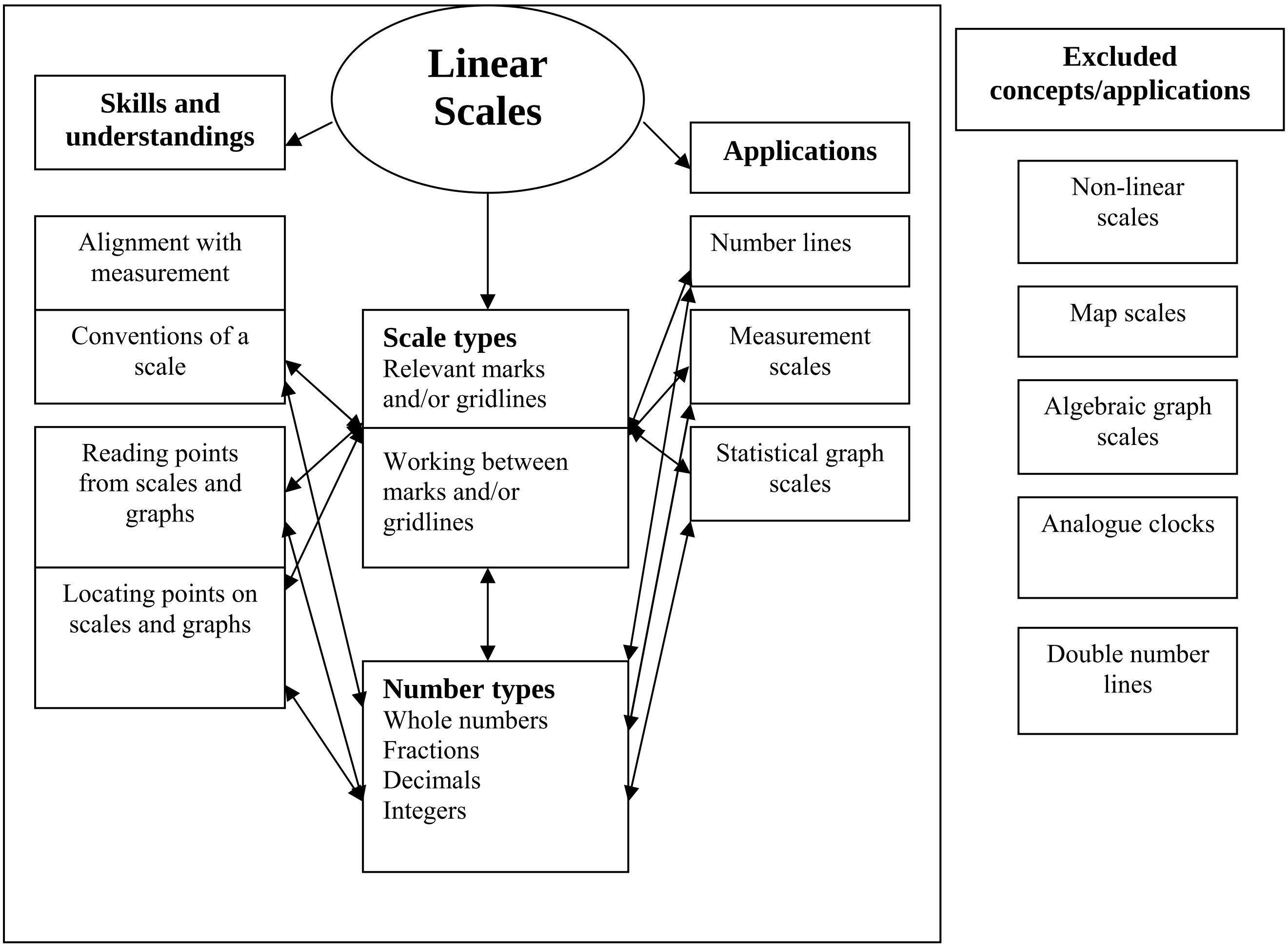


Figure 6.1. Concept map for scales included in this study

The concept map also shows that scale understanding 'aligns with measuring'. The literature identifies that students with a counting-based concept of measurement may measure from one, or read a broken ruler by looking at the highest number reached, rather than at the interval (see also Chapter 7).

Finally, the map identifies that working with and without marks may involve different understandings. When working with fractions, Behr et al. (1983) identified that the percentage of errors increased as the type of perceptual cue changed and that students used different strategies with different forms of cue. Bright et al. (1988), using the same four types of cue, noted that the relative difficulty presented by the cue types changed after a teaching intervention (see Section 3.6.2). This work raises a relevant question for the current study: do Year 7 and 8 students in the case study schools find scale intervals without marks harder to work with than those with marks? The question is relevant as this thesis seeks to identify student understanding of scale, so if different cue forms access different understandings it is important to undertake an exploration of understanding in relation to cue forms. Test 1 will use two of the four cue categories from Behr et al. to help answer this question (a 'complete set of cues' equates to working 'to and from the marks on a scale', while an 'incomplete set of cues' has similarities to working 'between the marks', see Figure 6.2). Such situations are commonly met when measuring or using graphs, so are important to include. The other categories (extraneous marks and irrelevant marks) are not commonly met in these contexts so were not considered for Test 1.

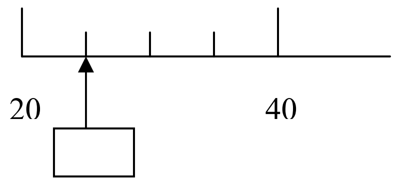


Figure 6.2a. A marked interval

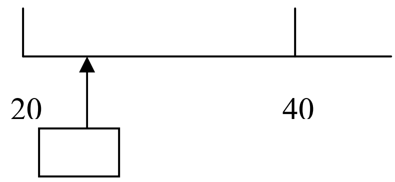


Figure 6.2b. An unmarked interval

Another question that follows from the one above is that if students do find problems with one type of cue harder, why? This question is relevant as it possible that the answer is related to student understanding. Here the central focus of exploration needs to be on how students approach the items. The interviews for Test 1 and Test 2 should help establish this, and whether or not the findings of Behr et al. (1983) apply to situations other than number lines for representing fractions. In the next sections, now that the construct has been mapped, the focus shifts to the processes used during the development of the diagnostics used for this thesis.

6.4 Fundamental assumptions

At this stage it is important to examine the assumptions on which the diagnostics are based, as these assumptions inevitably shape the tool by influencing both the items that are included and their nature. In turn this constrains the possible findings and thus can limit the conclusions that the research may reach.

Fundamental assumptions underlying this research:

- the diagnostics used in this thesis are being developed as tools to measure the understanding that students in Years 7 and 8 have of the defined scale construct;
- the ability to use measuring scales and graph axes effectively is part of understanding scale;
- the development of an understanding of where fractions, decimals, and integers can be found on a scale is an important part of the development of an understanding of scale;
- lack of understanding of scale is likely to adversely affect student progress in some aspects of mathematics; and
- identifying student understanding is the first step in enabling teachers to address misunderstanding or lack of understanding.

6.5 Concept and format development

The diagnostics are designed as a form of low-stakes assessment for which students need not prepare. The information gathered is solely for making judgements about student understanding. All items are short answer in format. This allows students to make their own response based on their understanding of the item, and allows an analysis of errors in this understanding.

A variety of assessment forms is experimented with (see Table 6.1). This is to ascertain the most appropriate form for using the items. As the items are designed to provide item-by-item information about the ability of students to work with a variety of scale-related concepts, they could be used individually to measure particular aspects of scale understanding. However, their validity and reliability as independent measures would need to be established.

6.5.1 The format of Test 1

In Test 1 there are two sections. Part A contains ‘bare’ number lines in a pure mathematics context. Part B uses parallel or similar items from contexts that students should have seen previously. These include measurement scales and statistical graphs identified by the document analysis as being familiar. For both parts, the item wording has been reviewed to simplify reading; however, problems linking to familiar contexts are prone to being ‘language rich’, often ‘telling a story’ to tap existing understanding. This additional use of language may create a barrier for some students that is not present with the number line problems.

It is hoped that the contextual items will identify whether or not the use of familiar, meaningful contexts results in higher achievement. If it were shown that the students treat number lines and contextual items with similar mathematical demands differently, the conclusion that might be drawn is that students do not have a coherent understanding of scale; that is they are influenced by the situation of the problem, suggesting an unstable or incomplete scale understanding.

6.6 The process for developing Test 1

The process for developing the initial test items has similarities to that used for the CSMS research outlined in Hart and Johnson (1980). For example, the present study aims to develop valid indicators of students’ level of understanding, with items suitable for both an interview and a written test. This process included the following steps:

- undertaking a concept analysis to help define the construct;
- investigating issues relating to student understanding identified in the literature;
- identifying examples of items used by other researchers, and students’ results with these;
- analysing curriculum documents to ascertain the understanding of the construct that students are expected to have;
- identifying and analysing relevant topics in some common resources;
- synthesising the information gained to identify appropriate contexts for items;
- selecting, modifying, and writing items that span the construct and are simple and easy to understand for the chosen year levels;
- consulting with colleagues to ensure that the items assess what the writer has intended;
- undertaking trial interviews to test the readability and general understanding students have of the items;
- modifying the items in light of the interviews; and

- creating a write-on version for use in schools.

In this process, efforts have been made to adhere to the standards outlined by AERA et al. (1999).

6.7 Interview protocols

The "methods used to conduct cognitive interviews shape the data they produce" (Presser et al., 2004, p. 8). Willis (2004) suggests that psychological research indicates that verbal reports are useful for tasks which:

- are at least somewhat interesting to the subject;
- involve verbal processing of information that enters conscious awareness – as opposed to being an automatic response; and
- emphasise problem-solving and decision-making.

In addition, Willis noted that:

- the environment for reporting should ensure that verbal reports are given during the course of the processing or very soon afterwards;
- social interaction between the experimenter and the subject should be minimised so as to focus the respondent's attention on the task; and
- the instructions should limit the degree of thinking about prior stimuli so as to avoid the possibility of these stimuli interfering with the respondent's thinking.

The tasks designed for Test 1 required students to process the problem, then use a mental strategy to solve it. While some of the students may not have found mathematics interesting per se, they were used to solving problems regularly in class, indicating that the first three suggestions for the use of verbal reports were likely to be satisfied as part of the students' normal classroom environment. To ensure the second set of suggestions was met it was decided to split the test into a series of individual response items that would be given to students one by one on paper. Students were given as much quiet time as they wanted to arrive at an answer, then immediately asked to explain their reasoning. This was to ensure they were given the best opportunity to remember what they had done.

Willis (2004) talks of the concept of *reactivity*. This is the process whereby the act of reporting verbally alters the course and structure of cognitive processing, creating results at variance from those of a silent control group. It happens as a respondent reflects on, and reacts to, what they are doing and saying as they report. Reactivity is significant for this study as the interview findings are used to provide a potential classification of the thinking processes of the silent experimental group at School 2; therefore the way reflective questions were asked needed to minimise the chance that students explained a method for solving the problem that was different to the method they actually used when first presented with the problem. For this reason, the main probe question was "how did you work that out?" The aim was to focus students immediately on what they had done when answering the item, while minimising comparison to, and reflection on, other items. It is also in line with Ericsson (cited in Willis, 2004) who states that the least reactive method to assess a participant's memory of the experience is to instruct them to give a retrospective report.

In addition to the probe question, minimal feedback on the correctness of a student's response was provided. This was to avoid a 'learning' process in which a student reflected on and revised responses identified by the researcher as incorrect. Non-verbal communication (smiling, head-nodding, etc.) by the researcher was used to encourage explanations, regardless of whether the logic supported a correct or an incorrect answer.

Willis also indicates that the cognitive demands of a task on each student need to be considered to see if follow-up questions are appropriate. Beatty (2004) suggests that some of the most useful findings from cognitive interviews are due to the adaptive probing of a skilled interviewer who improvises based on specific interview content. This implies that interviewing is an active process in which the interviewer needs to reflect on responses and non-verbal signals from the interviewee and make judgement calls about appropriate follow-up. This approach was taken and sometimes involved using unscripted clarifying questions. At others, it led to unclear explanations being accepted rather than explored further. The latter happened when, in the opinion of the researcher, the student was not sure why they had answered a problem in a certain way and was uncomfortable articulating their reasoning.

At School 1, after the main interview, a short retrospective interview took place. Students were again shown the items and asked to identify the sorts of item that they had not met before, those that they found hard to understand, and which items they found easy or hard. The aim was to establish which items needed revision, and to separate difficulties with task understanding from lack of content knowledge.

6.8 The interview process

To start each interview, students were asked about their understanding of the research, and the interview's purpose. Voluntary participation was stressed, and that while the interview involved a series of mathematics questions, it was not a test, nor would the results or what they said be reported to the school or impact on their school work. A 'paper ruler' marked in centimetre intervals (but not numbered) was laid on the desk. Students were told it was "in case you want it".

The first three interviews were used to trial the items and the interview process. While several small problems were found (e.g., the wording of the first item – see Chapter 8), the item structure was found to be sound and did not need revising. The retrospective interview identified that most students were comfortable with the items, although as expected some students had not met them all before.

6.9 The teacher trials

After the initial analysis of the data from School 1, a number of small groups of teachers were assessed and their results added to the study (see Chapter 5 for a discussion of the flexible design and Chapter 9 for details of the items used and the results of the trials). Interview analysis had indicated that students used a range of mental strategies to locate numbers on scales, and changed these strategies as the numbers and the scales changed. The adults in the first trial were given items to help identify how 'sophisticated' scale users answered such items, the goal being to validate or challenge the findings of the interview analysis. Teachers in the second and third trials were given items specifically designed to assess aspects of strategy transfer (e.g., whether adults used the same strategies on mathematically similar whole number and fraction problems).

6.10 The written version of Test 1

When adapting the interview items to a written test format, minimal changes were needed, although the 'paper ruler' was dropped as no interviewed student used it. Written instructions were developed for teachers, who acted as test administrators. These instructions were reviewed with teachers before the test was used. The question of absent students was raised and it was decided they should be given the paper if they returned within two days.

Discussions were also held about the information to be given to students prior to Test 1 being administered. The supplied ethics information had already given an outline of the research and the purpose of the diagnostic, although teachers decided this information should also be discussed verbally. It was agreed that students should not be prepared for the test, or be given information about its content beforehand. Papers were to be delivered after school on the day before the test was sat. These procedures were used to ensure that a comparable opportunity to demonstrate understanding was given to all students.

6.11 Test administration and assessment conditions

The written test was administered on a set date during class time. All teachers in a school were to use the test at the same time. Although a time limit was not imposed it was expected students would take 15 to 20 minutes on the test (roughly the same time that students at School 1 had taken for the interview). Students were expected to work in silence, not asking for help to answer the items. Initially they were only handed Part A. When this was completed, Part B was handed out. This procedure was appropriate as the two sections contained items that were either parallels or similar. It removed the possibility that students would go back and alter the number line items after 'recognising' the item in a familiar context.

6.12 Test 2

As Test 2 could already build on the work undertaken when developing Test 1, the process used for its development was greatly simplified, and was based on what had been learned. The stages of development involved:

- reviewing the student data and the findings from Test 1;
- revising the concept map;
- revisiting the literature to identify examples of items (and students' results with these) that span the revisions in the concept map;
- developing further the items already used, and designing the new items needed to explore understanding further; and
- redesigning the test structure.

The interview processes outlined in Sections 6.7 and 6.8 were unchanged.

Note that the development of the items for Test 2 is discussed in detail in Chapter 11, and that a copy of the test can be found in Appendix 2.

6.13 Conclusion

This chapter deals with methodological issues related to the development and use of the assessments developed for this study. The process of identifying the assumptions upon which Tests 1 and 2 are built and the use of these procedures are aimed at establishing that all reasonable steps have been taken to ensure the validity and reliability of the tests. In the next chapter, the development of the actual items for Test 1 is outlined.

Chapter 7

The Development of the Items for Test 1

7.1 Introduction

This chapter outlines the development of the items used in Test 1 to assess scale understandings. The corresponding information for Test 2 is in Chapter 11.

The scale understandings have been organised into four *aspects*. Aspect 1 deals with items related to student understanding of scale conventions. Aspects 2 to 4 address student understanding of scale involving multiples of whole numbers, fractions and decimals, and integers respectively; that is, their understanding of the conventions of a scale, and the processes students use when representing these sets of numbers on a scale.

Relevant items from the literature are used to inform the development of the items for each aspect. The items created for Test 1 are also shown, numbered as they are in that test (see Appendix 1) so they can be distinguished from the items that are drawn from other research. Note that items involving decontextualised number lines can be found in Part A of the test, and the problems set in what should be familiar contexts are in Part B. These parts are introduced separately in the chapter as the items in Part B are similar to those in Part A. For ease of reference, abbreviations for identifying the test items have been introduced; for example, item 2 in Part A is referred to as A2.

7.2 Aspect 1: Scale conventions

The concept analysis identified that to operate effectively with scale, students need a clear understanding of the conventions of a scale. Results in Nunes and Bryant (1998), Bragg and Outhred (2000a, 2000b) and Outhred and McPhail (2000) reinforce this as a line of enquiry. The item in Figure 7.1 was developed to measure this understanding.

1) On the line provided below, complete a number line to show where the numbers 2, 3, 5, 8 and 10 would go.

Figure 7.1. Item to test understanding of number line conventions (A1)

The 'paper ruler' marked in centimetre intervals was designed to support students with this item (see Section 6.8).

7.3 Aspect 2: Multiples scales

When units are aggregated for the scale, the concept analysis suggests questions should be asked to see if students understand that the number line conventions still hold. In particular, do they understand where the units are located?

NEMP results (Crooks & Flockton, 2002) indicate that many students are not proficient with such scales. In an item similar to Figure 7.2, Year 8 students produced the following 'correct' responses:

A	75 – 175	51%	
B	450 – 550	60%	
C	650 – 750	53%	(p. 22).

Even identifying the half-way number proved a major barrier for many, although only slightly more so than numbers appearing to be harder.

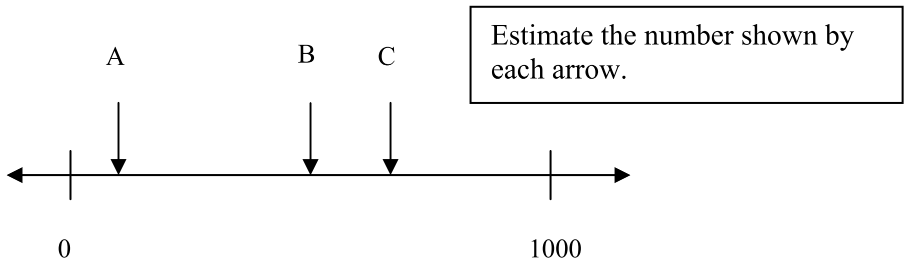


Figure 7.2. Item similar to NEMP (Crooks & Flockton, 2002, p. 22)

Ward (1979) used an item like that in Figure 7.3 in his research into the mathematics of ten-year-olds. He found that 13% of students did not attempt the item, while 42% answered incorrectly. He also found that 14% counted the notches rather than the intervals, thus gaining an incorrect result.

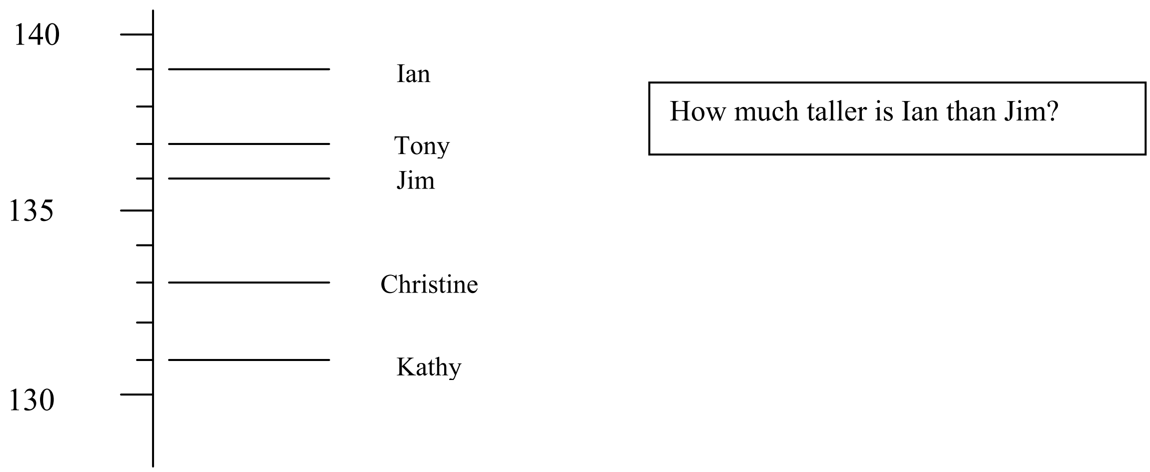


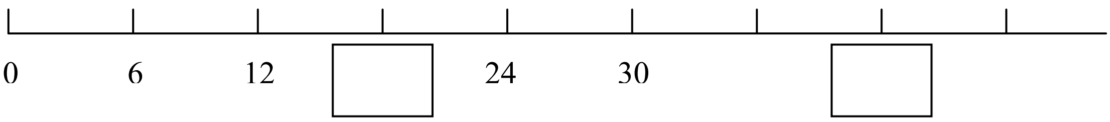
Figure 7.3. Item similar to Ward (1979, p. 100)

These items seem to assess quite different skills. The first uses a number line, while the second provides what should be a familiar context. In the NEMP item, students are working with an unmarked scale involving relatively large numbers, while in Ward's they are working with a unit scale involving smaller numbers. However, in both cases nearly half the students failed to answer correctly.

To assess understanding and identify misconceptions surrounding scales that involve multiples of whole numbers, the items in Figure 7.4 were developed. In A2 (the item labelled 2) students work with marks; the scale has been continued to see if they recognise a fixed amount of space always represents the same number of units. For A3 (item 3) the students work between marks.

A2 and A3 use a simpler scale than the NEMP example. Smaller numbers should negate the possible impact of number size. For A3, as the numbers are smaller, students are asked to identify a number other than the mid-point. The NEMP results suggests finding such a number is not an appreciably harder task.

2) The drawing below shows numbers on a number line. What missing numbers should go in the boxes?



3) On the number line above, put a cross where 11 should be.

Figure 7.4. Items developed for Test 1 (A2, A3)

One other item was developed to assess this aspect. Rangecroft (1994) cites the English Assessment of Performance Unit (APU), who in their *Review of Monitoring of Mathematics 1978-1982* (DES, 1982) found that "[f]or all types of graph the scale in use is a crucial factor, and scales without subdivisions of 1 or 10 are associated with lower success rates" (cited in Rangecroft, 1994, p. 10). Barcham (1996) comments that a frequent error is to assume each division on a scale is a number like 1, 10 or 0.1, and that students need to be reminded to check the scale. Hence an item with a simple scale in twos was developed to assess student strategies when working with scales that do not involve units or multiples of 10 (Figure 7.5).

4) What number is the arrow pointing to?

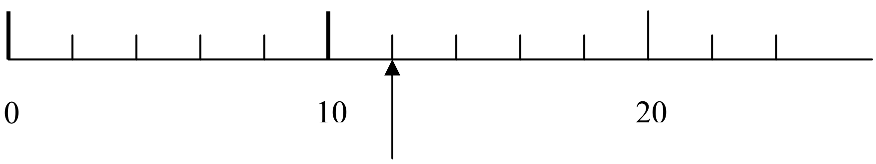


Figure 7.5. Whole number item with complete cues (A4)

7.4 Aspect 3: Fractions and decimals

When subdividing units, it is important for students to understand that the conventions of the number line still hold. In particular, that an equal amount of space always shows the same number of units.

For the item shown in Figure 7.6, Hart (1989) indicates that all of a sample of seven 10 to 11-year-olds "marked ¾ at a point greater than one, usually at the point 3 but sometimes between 3 and 4. When asked whether ¾ was more or less than one most said 'less' but did not mark the point thus" (p. 50).

Mark the number ¾ (with an ×) on the number line

0 1 2 3 4

Figure 7.6. Item similar to Hart (1989)

She noted that one student marked ¾ correctly on a number line from 0 to 1, but marked 6 when given a line from 0 to 8. (Other research has found students are more successful with number lines of length 1, e.g., Bright et al., 1988.) Hart suggests that the children "tended to view the number line as another region model and marked ¾ of the length of the line" (p. 51).

It is possible that students marking 3, or putting their answer between 3 and 4 had different issues. In 2001, the NUMPA identified that 42% of tested Year 7 to 9 students could not identify the symbols ½, 1/3 and ¼. When asked, some of Hart's students may have recognised the words, thus correctly identified that the number was less than one, but may not have recognised the symbol ¾, so marked a different point. Kieran (1976) also talks of an 'operator' interpretation of fractions in which ¾ of 4 is 3. Baturo and Cooper (1999) found that some students see the numerator in the fraction 6/3 as the whole number 6, with the denominator relating to how the whole is cut. Hart (1989) herself also identifies:

> The scrutiny of all the easy CSMS items (across topics) suggested that many secondary-aged children worked entirely within the set of whole numbers and any CSMS item which required the manipulation of non-integers was done incorrectly by over half the sample. (p. 46)

While the Hart item tests important understandings, it was modified to a number line from 0 to 8 for the present diagnostic (Figure 7.7). This allows error patterns for 'whole number thinkers' to be more obviously separated from 'regional' or 'operator thinkers'. Follow-up questioning can then focus on identifying the form and frequency of thinking behind specific errors. Marks have been added to identify the whole numbers. This should make it easier to accurately work out where ¾ is located.

6) Put a cross where ¾ goes on this number line

0 1 2 3 4 5 6 7 8

Figure 7.7. Fractional number line with incomplete cues (A6)

NEMP results (Crooks & Flockton, 2002) also indicate that many students are not successful working with fractions or decimals, although in relation to this study the usefulness of their items and reports is limited. The items often restrict student responses and thus their thinking (e.g., see Figure 7.9, students can only locate their answers in the interval zero to one) while the reports tend only to identify major errors. When using an item similar to a CSMS item introduced earlier (see Figure 3.4), but with a ruler as a context (Figure 7.8), the NEMP study found that only 32% of Year 8 students correctly answered 3.2 whereas 51% answered 3.1.

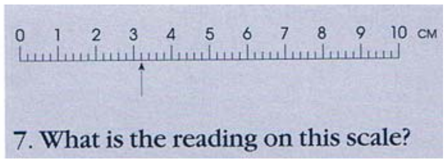


Figure 7.8. Item from NEMP (2002, p. 33)

In another item, students were asked to correctly place a series of number cards on a number line (Figure 7.9). Success was low on all items, with the proportion of students locating 0.1 correctly being similar to the error '3.1' for the item in Figure 7.8, which may indicate a response pattern for this kind of item.

0.1	54%	
$\frac{1}{10}$	38%	
0.25	26%	
0.5	36%	
50%	62%	
0.7	37%	
$\frac{4}{5}$	27%	
100%	66%	(p.23).

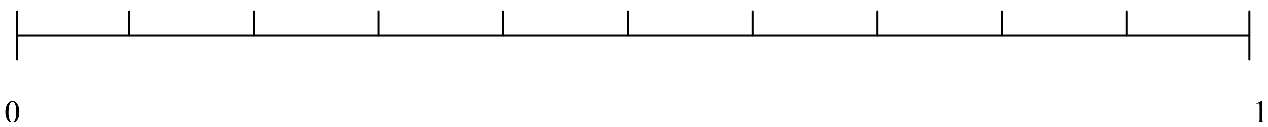


Figure 7.9. Similar item to NEMP (2002, p. 23)

These results signal that further exploration of student thinking in this area is required. Item A4 (see Figure 7.5 earlier) can be used to see if it was working between whole numbers that triggered such a response, or if a similar result happens with whole number scales. The absence of a context can also be used to see if the ruler and small intervals was responsible. In addition, A5 was developed to probe understanding further (Figure 7.10).

5) What missing numbers go in the boxes?

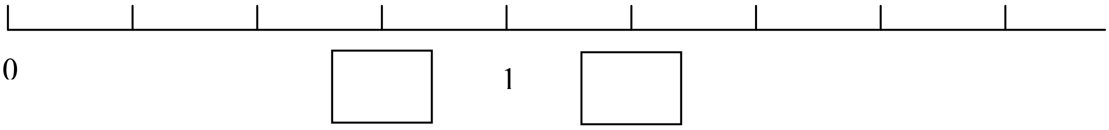


Figure 7.10. Complete cues (A5)

The item form provides more structure than Hart (1989), as marks are provided. When taken with A6 (Figure 7.7), the problems form a pair in which the same fraction needs to be identified on two different forms of number line. This allows for corroboration of, or challenge to, the findings of Behr et al. (1983) about the use of cues. Having one item in the form of reading the number shown and the other locating the fraction ensures that the items appear different.

In the final item for this aspect (A7, Figure 7.11) the focus is more firmly on decimals. In this item, student thinking is not restricted to numbers between zero and one so, should they desire to, students can place their cross at 4. Some simple scaffolding has been provided with marks indicating halves.

7) Using a cross, show where the number 0.4 goes on the number line below

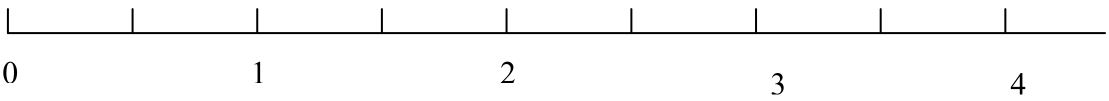


Figure 7.11. The decimal item (A7)

7.5 Aspect 4: Integers

The final item in Part A (A8, Figure 7.12) was suggested by the concept analysis and literature review. It examines if students understand that the number line conventions still hold when the line is extended to negative numbers. It should also identify if students recognise the role of zero on an integer scale. However, MiNZC (Ministry of Education, 1992) does not introduce integers until CL4, that is for students in Years 7 and 8. In the NDP, integers are also not introduced until Stage 7, also relatively late (Ministry of Education, 2005a). Students working below these levels may not have been formally introduced to integers and the integer number line, even though they may have met them informally through exposure to temperatures and thermometers. Lack of success on the item may therefore be related to lack of exposure, so only one simple integer item was developed. Note that the other number line referred to in the item is A1, shown earlier. Also note that the paper ruler mentioned previously could be useful when working on this item.

8) Please use the line below to draw another number line. On it show where –3 would go

Figure 7.12. The integer item (A8)

Now that all of the items in Part A of the test have been introduced, it is time to consider the contextual items in Part B. In this section only the graphical items are shown as the nature of the measurement scales is similar to Part A. Refer to Appendix 1 to see copies of these items.

7.6 The contextual items in Part B

Effective practice for the teaching and assessment of scale needs to be established. The literature review suggested that one important element of an exploration into student understanding of scale was to consider the impact of the use of contexts. With number lines, some context specific issues can be simplified or eliminated, while with contextual items students' informal knowledge and understanding may be tapped. Which would provide a better measure of understanding? It seems reasonable to assume that students with a strong understanding of scale should perform consistently on similar contextual and number line items. In the next section, a comparison of the number line and contextual items is provided to illustrate the degree of match between the items developed for this investigation.

7.7 Comparison of the number line and contextual items

Most of the number line and contextual items are closely matched. Table 7.1 identifies the skill each diagnostic item aims to assess and the differences between the item pairs.

Table 7.1: Comparison of the number line and contextual items*

Part A item	Part B equivalent
A1 Creating a number line to show a set of natural numbers. Unstructured number line.	**B5** Creating a scale on the horizontal axis of a graph to show a set of natural numbers. Similar numbers involved. Lines on graph paper make equal spacing easier. Issue of the role of zero on the axis.
A2 Identifying missing numbers on a scale with marks in multiples of six (interpolating and extrapolating).	**B2a** Identifying a missing number on a scale with marks in multiples of eight (extrapolating). The six and eight times tables are at the same level of the Number Framework (Ministry of Education, 2005a) so the change is not significant.
A3 Locating an unmarked number on the item 2 scale (interpolating).	**B2b** Reading between the marks on the graph scale (interpolating). Equivalent numbers used.
A4 Reading from a mark between the labelled numbers.	**B1** Reading from a mark between the numbers labelled on the thermometer. Similar numbers to A4 involved.
A5 (a & b) Identifying fractions on a marked scale, one less than one, one greater than one.	**B3 (a & b)** Identifying fractions on a marked weighing scale, one less than one, one greater than one. One number is identical, the other is similar.
A6 Locating a fraction on a scale – no marks. Scale goes from 0 to 8.	**B3c** Locating the same fraction on an unmarked weighing scale. Scale goes to 4 so the differentiation between students calculating ¾ of 4 and whole number thinkers responding '3' is not available.
A7 Locating a decimal on a scale. Large unit and halves marked.	**B4** Locating the same decimal on a ruler – no marks. As both have incomplete cues (Behr et al., 1983), performance should not change significantly.
A8 Creating a scale to locate an integer.	No parallel item. A contextual item using a thermometer has already been used.

* Note that the items in Part B are presented in a different order to those in Part A (e.g., item B5 is paired with item A1).

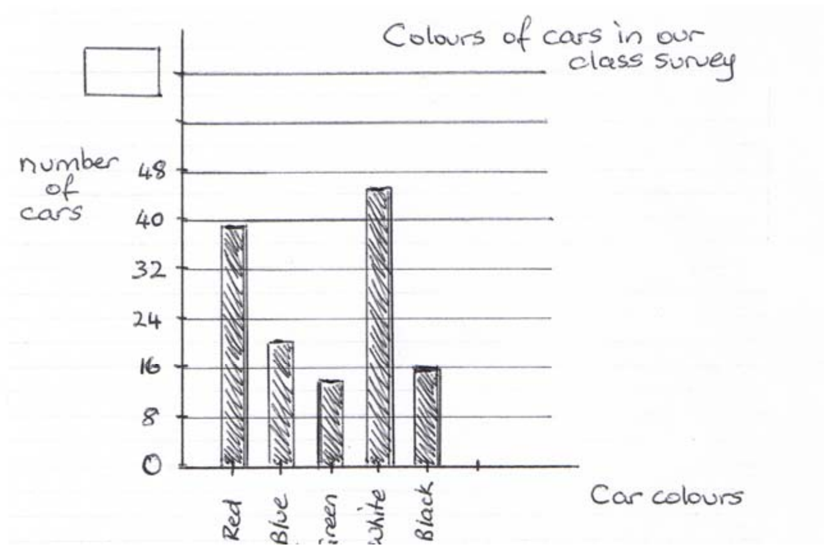
For the graphical items, attempts were made to restrict the assessed knowledge to scale. How this has been achieved is outlined below.

7.8 The graphical items

Bar graphs were chosen as a graphical context to measure understanding of scale since they are introduced in CL2 (Ministry of Education, 1992). They were also commonly used in the analysed resources (see Chapter 4). By providing students with a graph form that should be familiar, in a question that uses small numbers and in a context that students should be able to relate to, it is hoped that responses will allow issues around scale (rather than other factors) to be diagnosed (Friel et al., 2001).

B2 (item 2 in Figure 7.13) is a CL2 graph displaying data categories on the horizontal axis. In this graph, extending the scale and reading between the labelled values are the areas of interest (paralleling A2 and A3). Note that the graph has been reduced in size. This has been done with various figures in this thesis where doing so does not compromise the information contained, but does improve layout and reading flow. A full sized version of the item can be found in Appendix 1.

2) This graph shows the results of a survey a class did into favourite car colours.



(a) What missing number should go in the box?
(b) How many red cars did the class count?

Figure 7.13. Reduced version of B2

B5 (item 5 in Figure 7.14) involves drawing a CL3 graph in which students are expected to locate discrete numeric data on the horizontal axis. In this item it is the creation of a scale on the horizontal axis that is of interest (paralleling A1) so the numbers are closely related to those used in that item. The axis requires a unit scale and as unit scales are the simplest possible scale, the problems relating to developing an appropriate scale, as noted in Friel et al. (2001), should be minimised.

B5 can also assess scale-making. To construct the vertical axis, the gridlines can be used for multiples of two, although the size of each square would allow students to halve these to create a unit scale.

5) The teacher of a Year 7 class wants to find out about the families of her students. One question she asks is how many people live at home. Here are the results.

How many people live at home	2	3	4	5	7	9
Number of families	3	8	11	4	2	1

Finish the bar graph started below to show what she found out.

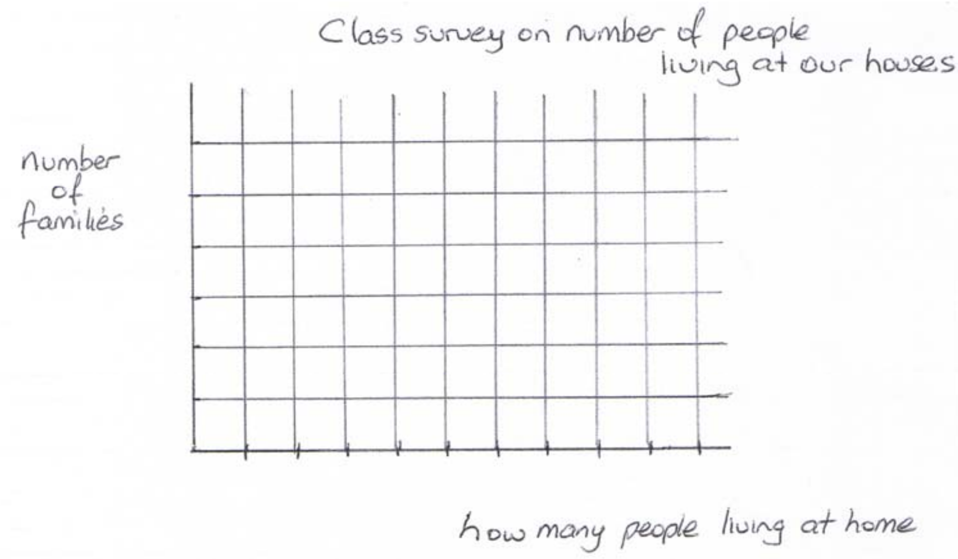


Figure 7.14. Reduced version of B5

Ward (1979) used a similar item to B5 on 569 English children. His problem required students to graph temperature data for five days, Monday to Friday. The numbers also suited a vertical scale of twos. Ward found that:

- 26% did not attempt the item;
- most of those who attempted the problem used a bar graph with temperature on the vertical axis;
- a few used a unit scale for the vertical axis and tried to extend the grid;
- 11% used a vertical scale in twos, although some others used a scale of fives;
- some just used the numbers in the table for the temperature scale;
- 7% numbered spaces on the vertical axis instead of marks. Ward identifies this as a common error amongst primary children. They:

> find it difficult to appreciate that on this kind of chart:
>
> a on the horizontal scale are discrete categories labelled at the centre of each block and representing a set of distinct results, and
>
> b on the vertical axis the numbered lines represent a measurement of temperature, frequency or whatever.
>
> (Ward, 1979, p. 121).

- interpolating between the numbers marked on the scale caused problems; and

- only 28% correctly drew positions for the temperatures on the scales they had chosen.

This suggests that B5 could also be used to find other data. However, as this requires students to apply skills outside the defined construct to correctly complete the graph, it was initially decided not to measure the additional data. The next section identifies where the items fit within the curriculum. The process was undertaken to ensure that the items are reasonable for students following a programme of study based on MiNZC (Ministry of Education, 1992).

7.9 Curriculum demands of the items

As the AOs in MiNZC do not mention the use of number lines to explore and develop insights into the number system (see Chapter 4), the identification of the curriculum demands of the number line items is not easy. One way is to analyse the demands of the numbers involved, and infer the demands of the items from this. Table 7.2 shows the results of this analysis in relation to both MiNZC and the Number Framework (Ministry of Education, 2005a).

Table 7.2: Curriculum demands of the numbers in the number line items

Part A: Items and number forms	**MiNZC reference**	**Number Framework reference**
A1 to A4: Whole numbers.	Number, level 1.	Stage 2 for the counting numbers. Stage 5 for skip counts in other than twos, fives and tens.
A5, A6: Fractions.	Halves, quarters, thirds, and fifths are met in story problems at level 2.	Stage 4 for proper fractions, stage 5 for improper fractions
A7: Decimals.	Number, level 3.	Stage 6.
A8: Integers.	Number, level 4.	Stage 7.

This analysis suggests that the numbers in the first part of the diagnostic cover the range of curriculum levels and framework stages to which students in Years 7 and 8 may have been exposed. This should allow a spread of results to be generated, with simpler items allowing those with lesser understanding some success, yet providing challenge for those with more understanding.

For the contextual items, students should have had considerable exposure to scales in meaningful situations by the time they reach Year 7 (see Chapter 4). As contexts tend to be more varied and sophisticated than number lines, Table 7.3 identifies where students ***may*** have met the scales during classroom mathematics. The table does not reference the number framework as this does not have measurement or statistics domains, nor does it reference other curriculum documents through which students may also have met many of these contexts. It is sufficient to show that students should have had exposure to these contexts in mathematics prior to Year 7.

Table 7.3: Curriculum demands of the contextual items

Part B: Item and context	Curriculum level and strand
B1 – thermometer.	CL3 measurement AO.
B2 – bar graph.	CL2 statistics AO.
B3 – weighing scale.	CL2 measurement AO.
B4 – centimetre ruler.	CL2 measurement AO.
B5 – bar graph.	CL3 statistics SLO.

7.10 Summary

Figure 7.15 shows aspects of the concept map covered by the diagnostic. These are coloured red.

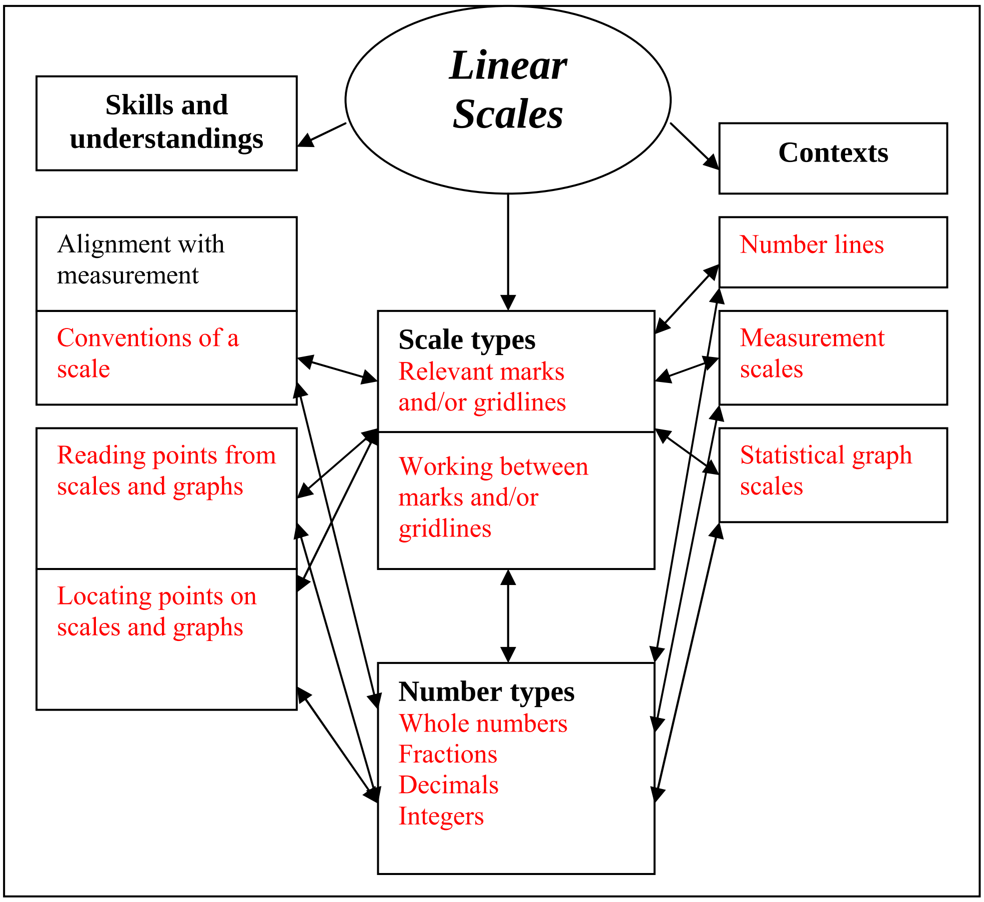


Figure 7.15. Coverage of the concept map by the items

The main features of the diagnostic are as follows.

1) Part A focuses on (abstract) number lines without a context; Part B focuses on scales in 'familiar' contexts.

2) The items can be sorted into four aspects:
   - Conventions of the number line (A1; B5);
   - Multiples of whole numbers (A2, A3, A4; B1, B2);
   - Fractions and decimals (A5, A6, A7; B3, B4);
   - Integers (A8).

3) In some cases, the items were developed from the literature, in others they were designed to fill gaps in the literature.

Chapter 8

Exploring student understanding: Case 1 – the cognitive interviews

8.1 Introduction

This chapter reports the initial exploratory study into scale understanding undertaken at School 1. In the early sections of the chapter each aspect of scale is explored in turn, item by item. After an outline of the results, a discussion section summarises trends and provides some interpretation, occasionally referencing the literature where this can help explain what has been observed. In keeping with the exploratory nature of this initial study, each aspect closes by considering the usefulness of the items as a measure of scale understanding.

The CSMS study identified significant error patterns in responses to test items. Hart (1978) talks of these in relation to the number operations test.

> The message that one receives from these errors and their high incidence is that there is a considerable amount of confusion in children's minds as to what they are doing in mathematics. Some errors will be symptomatic of the child being unable to grasp the level of abstraction being presented, others might arise because we never consolidated the teaching. Maybe we did not stress every detail of a process and were unaware of the gaps in the child's understanding. As maths teachers we perhaps tend to treat all errors the same way and feel that a demonstration of the correct method is sufficient for all to understand. (p. 39)

Specific attention is therefore paid to the errors and error patterns revealed by the items in Test 1. Student voice is used to identify the thinking and understanding behind particular solutions.

In the later sections of the chapter the emphasis changes from individual items to the test as a whole. The issues of item difficulty and whether the students are meeting curriculum expectations are addressed. Trends in the data are identified, with a response hierarchy based on thinking form being created and used to interpret student understanding.

Readers are asked to note that the layout of this chapter, like the other results chapters, is different from the rest of the thesis. The main findings are summarised at the outset to provide a structure to assist the reader and an indication of the varied findings. Details are then provided in the following sections. Readers are reminded that the development of the items for Test 1 is outlined in Chapter 7, while a copy of the test can be found in Appendix 1.

8.2 Main Findings

1) Most items in Test 1 provoked a range of responses and/or a range of explanations of thinking. The explanations indicated that the students were using *mental strategies* to answer the questions.
2) The existence or development of these strategies was not identified in the document analysis, so it is possible that some or most of these strategies lay outside students' formal learning about scale.
3) Each student showed that they had access to a range of strategies and chose one, or several, to solve a problem.
4) The strategies seem to link to particular forms of thinking. Some appear to link to processes based on counting, while others to additive or multiplicative processes. This enabled a hierarchy to be developed for the strategies.
5) The majority of the items in Test 1 appear to provide discrimination between students with low and high scores.
6) Students scored across the range of possible scores and the number correctly answering each item varied from 23% (indicating some items were 'hard') to 100%, indicating other items were 'easy'.
7) When assessing if the students had met curriculum expectations, the standard of a 75% success rate was only met for six items. All of these involved whole numbers and five of the six used number lines rather than items in context. This suggests that the group of students was not meeting the curriculum expectations outlined in Chapter 4.

In the next section, the aspect analysis begins, starting with items developed to measure understanding of scale conventions.

8.3 Aspect 1: Conventions of the number line (A1, B5)

8.3.1 Item A1: Creating a number line

The first interviews showed the wording of this item was problematic. Several alternatives were tried before adopting the wording "change the line below into a number line. On it show where the numbers 2, 3, 5, 8, and 10 go". Even with this wording, one student was unclear about the term "number line", writing a line of numbers "3 + 5 = 8 ÷ 4 = 2 + 8" (Student 7). This student did not realise his interpretation of 'number line' here was different to the one he used in other items. During the retrospective interview he indicated none of the items caused problems.

For marking, student responses were analysed in five different ways:

A1a In which direction do numbers get bigger?

A1b Did the students use marks?

A1c Did students place their numbers on the marks or in the spaces between them?

A1d Were the units used equal in size? In particular, where space for the numbers 4, 6, 7, and 9 was left, did this indicate the use of an iterated unit?

A1e Did the spacing between the numbers shown indicate an awareness of the need 'to leave more space for more units'?

Figure 8.1 exemplifies how this item was coded.

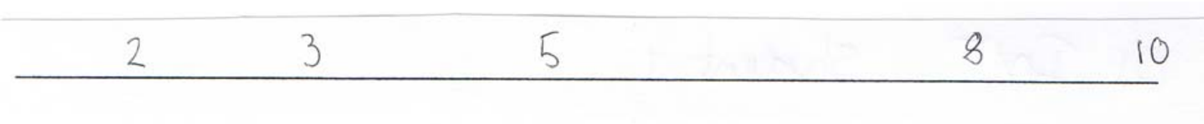


Figure 8.1. Work from Student 1

A1a Bigger to the right.
A1b No marks.
A1c Numbers used as marks.
A1d Units not equal in size.
A1e Space left for 'missing numbers'.

Results

A1a) Twelve students used the convention 'whole numbers get bigger to the right'.

A1b) Eight students placed marks on the line (and located numbers on the marks). Four students (including Student 1 – Figure 8.1) did not use marks but treated their numbers as marks.

A1c) Twelve students either located numbers on their marks or used their numbers as marks.

A1d) Only three students could place the numbers accurately in relation to one another. In doing this, Student 11 created equal spaces between the numbers 2, 3, 5, 8, and 10, but left no room for 4, 6, 7, and 9. Six of the eight using marks and three of the four not using marks did not have units of equal size, suggesting that the use of marks did not improve these students' ability to iterate a unit.

Of the unsuccessful students, Student 6 (S6) knew to leave space, although his marking in Figure 8.2 does not reflect this knowledge. "…I knew 10 would be the highest one so I put it up there and 2 is the lowest so I put it there and 3 is a bit higher than 2 so I put it there and then I knew 3 + 2 = 5 so I put that 2 spaces and then I put that 3 spaces…"

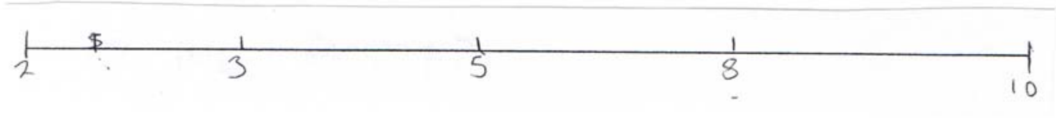


Figure 8.2. The number line of Student 6

Other unsuccessful students used a variety of strategies to locate the numbers. Two used halving:

I: OK, … you went 0 to 10 at the other end, you found half-way which is the middle, then you tried to split it up again here. How did that work, …?

S8: Well it's quite hard without a ruler cause if it was an even number it would be quite easy cause you'd find the half way and you'd make it half-way and then you'd just half that and half that.

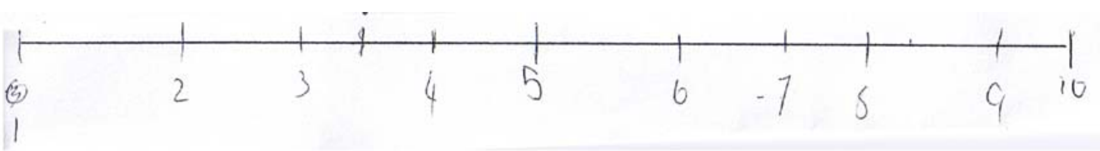


Figure 8.3. The number line of Student 8

A1e) Only nine students seemed to understand that a proportional amount of space needs to be left for numbers that are not shown. Conversely, understanding this did not guarantee that the finished line accurately showed the relative location of the numbers.

Discussion

All number lines were unique, although there were some trends.

- Students used a variety of processes (*strategies*) to answer the item. The strategy used sometimes had an impact on success. For example, students using halving adopted a strategy inappropriate to the numbers they were trying to mark.
- Some started their number line at zero, others at one or the lowest number given – two. Figure 8.3 shows that Student 8 changed the number marked at the start of her line from 0 to 1, but did not understand the impact this would have on the rest of the line.
- Students' use of marks did not seem to improve their success when spacing numbers. Intuitively, those using marks should find it easier to create equal units as they have a powerful visual check of their accuracy.
- What students produced on paper and what they thought they were producing did not always match. Some were not able to iterate a unit accurately, while others could not halve an interval accurately.
- Several students had difficulty choosing a unit size appropriate to the situation. This affected their ability to retain equal units as they 'ran out of space'. They tended to reduce the size of each unit rather than extend the line or start again.

8.3.2 Item B5: The bar graph

This item focused on students' ability to use their understanding of scale conventions when creating a graph axis. Tables 8.1 and 8.2 provide information about how students numbered the *x* and *y* axes respectively.

Results

Table 8.1: Horizontal (*x*) axis numbering

Omitted the item	234579 (Numbers provided for the *x* axis)	3811421 (Numbers intended for the *y* axis)	Unit scale 1234…	02468	No *x* axis scale provided	Total
2	4	4	1	1	1	13

- Three students placed their numbers on the gridlines/marks. Seven placed them between the marks.
- Four of the 10 students (4/10) who provided a scale swapped the *x* and *y* data, which has the benefit of creating a set of numbers in ascending order for the vertical axis. For the horizontal axis they placed the numbers in the order found in the table – instead of using the convention that larger numbers go to the right.
- Another four did not leave space for 'omitted numbers', seeming to treat the numbers as categories, not as part of a numeric scale.
- Only three students were consistent in their approach to this axis and their number line in A1.

Discussion

The results show that students did not see the task of creating a scale for a graph (B5) as equivalent to creating a number line (A1). McGatha et al. (1998) identified that students can treat the numbers on the horizontal axis of a bar graph as individual data points or categories. Responses to B5 indicate that few students recognised that the horizontal axis should show a scale, with spaces for 6 and 8. They seemed to treat the frequencies as belonging to categories, which markedly changes the visual picture provided by the graph. As a consequence it was decided to examine both the horizontal and vertical axes for knowledge of scale conventions. The vertical scale is now explored.

Results: the vertical axis

Table 8.2: Vertical (*y*) axis numbering

Omitted the item	**1234811 (ordered *y* axis numbers)**	**3811421 (*y* numbers in table order)**	**1234… Unit scale (no zero)**	**2468 (no space for zero)**	**02468 (2 had an implicit zero)**	**No scale provided**	**Total**
2	1	1	2	1	5	1	13

- Nine of the 10 students (9/10) who numbered their scale had larger numbers above smaller numbers.
- Seven students placed their numbers on their marks, two placed them in-between. One student did not label their unit scale.
- Five students did not leave space for zero, and two more, while leaving space, did not label it. If the issue of zero is considered separately, 8 of the 10 students (8/10) used an 'honest' scale, with spaces for 'omitted numbers'.
- Four students produced vertical scales inconsistent with the understanding shown in A1. In two cases the students had a better understanding of the vertical axis of the bar graph, in two it was poorer.

Discussion

As with A1, no two responses for B5 were alike although four students had incomplete graphs. (Two created at least one scale, but were unsure how to continue.) Students tended to treat their vertical axis more like a scale than their horizontal axis, but did not necessarily recognise it to be subject to the same conventions as their number line in A1. The role of zero on a graph axis seemed not to be commonly understood.

The ability to create an appropriate scale for one or other axis was no guarantee of a student being able to appropriately graph the data. In Figure 8.4, Student 4 used the same data for both axes. Other students ignored the axis labels, swapping the data sets. Many seemed not to understand the concept of a coordinate or understand how to create a coordinate from numeric data in table form, even though the form of graph in B5 was from CL3 – the standard programme for students in Years 5 and 6 (Ministry of Education, 1992). This may be a consequence of MiNZC promoting the use of qualitative graphs in algebra that are interpreted globally rather than using a formal *x*-axis, *y*-axis approach based on the Cartesian system (see Chapter 4).

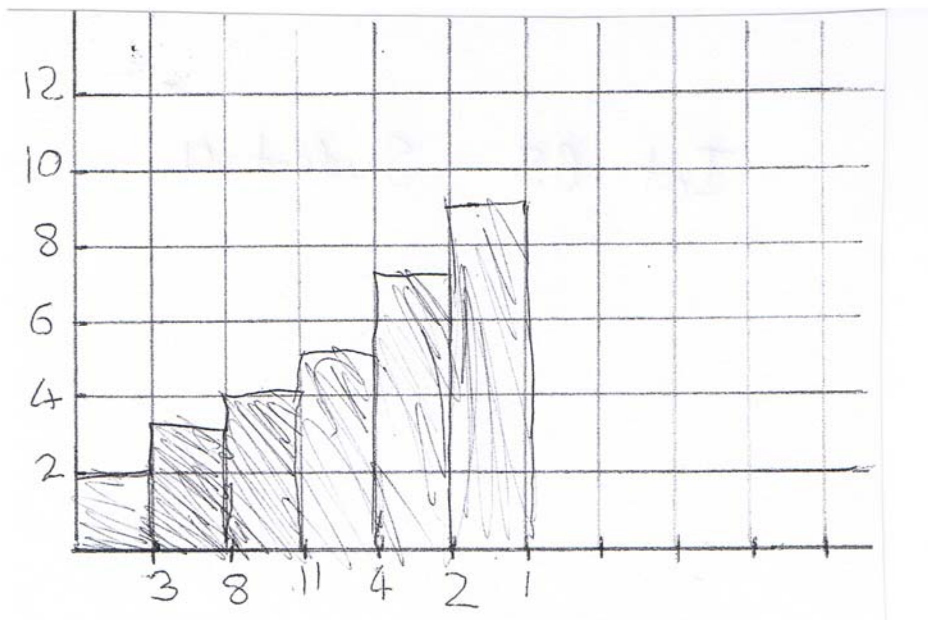


Figure 8.4. Bar graph from Student 4 (title and axis labels removed)

Item evaluation

Item A1 identified that some students have major gaps in their understanding of number line conventions, while others did not seem to have the skills to put them into practice. (If the item was marked simply right or wrong, only three students would have gained 'a tick'.) A1 therefore appears to measure important aspects of scale understanding.

While analysis of the written responses provided valuable information, the student's oral responses (and the visual cues provided by hand and lip movements) did at times help explain what the student was attempting to do. In several cases the oral response was not consistent with what was written. As such, this item may be best suited to an oral diagnostic.

Item B5 identified a number of issues that may appear to be beyond the scope of this study, but are intimately linked to an understanding of scale. In particular they are relevant to students' ability to identify scales in a variety of contexts, and bring to them their understanding of how a scale operates. In this instance, for students to be able to graph accurately, they should be able to create appropriate scales and plot points accurately in relation to them. As such, B5 measures important scale-related understanding and, as the oral responses and non-verbal signals during the interviews did not provide additional information to that obtained from the student's written response, the item is well suited to a written test.

8.4 Aspect 2: Using multiples scales (A2, A3, A4, B1, B2a, & B2b)

These items dealt with students' ability to work with whole numbers on the number line and in contexts. Each item pair (a number line and the similar item in a context) is introduced and discussed together.

8.4.1 Items A2 and B2a: Labelling marks

Results

For A2, all 13 students could locate 18 and 42 on the number line. When asked how they got their answer, some explanations involved skip counting: [its] "going up in 6s and then there's 12 so you have to put another 6 in there and then that's another 6 to make 36, and then another 6 to 42" (Student 5). Others used multiplication: "I just counted in sixes and what I did was, there was one, two and three and so I did three times six is eighteen and then for 42 I said seven times six is 42" (Student 2).

In comparison, only 9/13 were initially correct on B2a, an item involving the extension of a similar scale in eights. One student then self-corrected a skip counting error when asked to explain his thinking. Of those answering incorrectly, one did not seem to know her skip counts in eights: "making it into lots of eight, which was eight, sixteen, twenty four, thirty two, forty, forty eight and then fifty six, seventy two" (Student 2). Student 6 explained "I went 48 plus 8 equals 56" while Student 7 answered "cause its plus eight." These two students either did not notice the intervening mark or forgot to refer back to the question once they had worked out the size of the skip.

When comparing the responses for B2a and B5b, both of which relate to the vertical axis of a bar graph, the following trends are worth noting.

- If the issue of zero is again treated separately, eight students had 'honest' scales for both graphs (involving units or multiples).

- Two students answered inconsistently. Student 11 correctly answered B2a, but labelled her B5 axis 1, 2, 3, 4, 8, 11. Student 6 incorrectly identified B2a and used the vertical scale 3, 8, 11, 4, 2, 1 for B5.
- Two students correctly answering B2a were unsure how to create a bar graph for B5. One other student answering incorrectly was unsure how to create the vertical scale for B5.

Discussion

The evidence suggests that, until students are accomplished skip-counters, working with some multiples scales may introduce numerical errors that undermine the intended learning. It further suggests that ability to work with scale may link to understanding of multiplication, so investigating the possibility of a link between success in these items and students' understanding of multiplication (as identified by the NUMPA) is worthy of investigation.

Responses to A2 indicate that all students had an understanding of the convention that 'an equal amount of space shows the same number of units'. However, this understanding was not applied consistently across situations. For example, two students did not use an honest scale for the y axis on B5. Only three were correct with A1d (an unmarked number line). Here, while some students did not seem to have the skills to create intervals of approximately the same size, others did not seem to recognise the importance of doing so. Perhaps these students see each task in isolation, involving different understandings, rather than as related tasks involving common understandings about how scales operate.

Item evaluation

Locating 18 [A2a] appears to have been an unnecessary item as students tended to use the same explanation for this and locating 42 [A2b]. The other items provided useful information, especially when the responses were considered alongside those for the scale convention items.

8.4.2 Items A3 and B2b: Reading between the marks

Results for A3

- All students placed 11 correctly on the number line (A3), although in three cases the accuracy of the placement was marginal.
- Students used mental strategies to locate where to place their response. The strategy used by 10 students can be described as 'a little bit more, a little bit less'. The response of Student 4 is typical of this form of reasoning: "probably just before the 12, right here."

Three students identified there were six numbers in-between 6 and 12. For example, Student 5 was observed to create six spaces equal in size, and then counted along five of them and explained "you've got to get it in an even space." (In this context Student 5 used the term 'even' to mean equal.) The other two students using this approach were among the three who were less accurate with their placement. When the responses were compared to those of A1 and A8, it seems one knew that the spaces had to be an equal size, but found it hard to put this into practice. The other did not seem to realise all the intervals needed to be the same size.

Results for B2b

- Only seven students initially answered B2b correctly, although one student self-corrected when asked to explain their answer. Accuracy with the number line did not seem to transfer to the graphical context for all students.
- All five students with incorrect answers for B2b also had inaccurate placements in other questions. For example, three also had 'borderline' estimates for A3. One possible reason for lower accuracy in B2b is the size of the unit, in that the interval between marks was smaller than in A3. Another is that in A3 students could choose where to put their mark, while in B2b they needed to read from the bar to the scale, and did not have an accurate strategy for this, as the following quotes indicate.

  I: So how many red cars did the class count?
  S7: Forty
  I: … And why did you put forty?
  S7: Cause that's where the line is.

  I: And how did you get 38?
  S6: I just made a guess about 32 and 40. I just made 38.

  Student 7's strategy had the effect of simplifying the problem by reading to the nearest mark, rather than working between the marks.

- With the exception of Student 7, students used mental strategies to work out the answer. Six used 'a little bit more, a little bit less', two used an 'equal spaces' strategy and one guessed. One student identified she used multiple strategies. "It's a little under 40. I imagined eight little lines" (Student 12). Two others appeared to do something quite different to their verbal explanation. Both marked half-way, then seemed to count on in ones, while their verbal explanation was that the bar was a little under 40, so was 39.

Discussion

A3 involved locating a number on a number line, B2b reading from a graph axis. Ten of the thirteen students used the same strategy in both situations, providing evidence that reading and locating numbers on a scale involve similar skills. However, use of the same strategy in different situations does not imply a level of accuracy. Three students using the same strategy for both items produced incorrect answers for B2b. The other two students who incorrectly answered B2b used a different mental strategy for answering A3.

Item evaluation

The items provided useful information about scale understanding, although a student's solution strategy was not obvious from the written answer. Error patterns were also not distinctive. For example, three students used different strategies to arrive at the answer 38. One guessed, one used ‘a little bit less’, and one ‘counted on in ones’. Consequently these items may be best suited to an oral diagnostic where the logic of an answer can be explored. However, they will be retained for the written version of Test 1 so the responses of a larger group of students can be explored.

8.4.3 Items A4 and B1: scales in multiples of two

Results for A4

- All students correctly read from the number line in A4.
- Students used different strategies to obtain their answer. Four had counting-based strategies. For example:

  I: What number is the arrow pointing to?
  S2: Is it twelve?
  I: OK, sounds good. And how did you get twelve?
  S2: Well when I counted from ten, eleven, twelve, thirteen, fourteen, fifteen, it didn’t get up to twenty, so I counted in my twos which made it twelve, fourteen, sixteen, eighteen, and then twenty.

  S13: I just counted up in twos to get to ten and added another two to it.

  Four others had strategies that were additive or multiplicative in nature. For example:

  S7: I knew that ten plus ten is twenty and obviously that’s five cause it’s a half and just since, like, you can’t do ten plus ten so I decided it would be ten plus two.

  S1: Well, if you go one, two, three, four, five, five times two is ten, then six, seven, eight, nine, ten, ten times two is twenty. If you go to there it’s six and six times two is twelve.

  The remaining five students explained that the numbers were “going up in twos”. It was unclear whether their approach was counting, addition, or multiplication based.

- All students seemed to successfully count the marks for this item.
- No student assumed each mark represented a unit.

Results for B1

- In B1, four students provided incorrect answers. Three of these answered 12; one through a miscount, two through counting in ones.

  I: And how did you work out twelve degrees?
  S1: Cause the ten, zero, then goes one, two, three, four, five and five times two is ten and then it goes to six and six times two is twelve.

  S7: Cause there’s a little line, cause there’s a line there which is ten, and it goes up two more, that’s all.

The fourth student used a form of decimal logic to identify the solution.

> S13: Well, since there was only four lines between zero and ten. Each is like point two and it was on the second line above ten so I figured it would be ten point four.

Discussion

Ten of the 13 students used the same mental strategy when answering the two items. They can be said to have *consistent strategies*. However, once again some students correctly answered the number line item, but incorrectly answered the contextual item. Either the use of a context with this type of item may not be beneficial, or, as suggested earlier, interval size may influence a student's strategy. A4 had a large, easy to read scale, while the thermometer used small intervals. 'Smaller spaces' may trigger some students to use a unit scale or 'to use points', as on a ruler, suggesting they do not 'check what each mark stands for' (Barcham, 1996).

During the interviews, it became clear some students were having an issue with 'marks and spaces'. Some consistently counted marks, and were successful in doing so. Others consistently counted spaces, and were also successful. A third group was unsure what to count. As an example of one of the students in this group, Student 3 was unsure which marks to include in the count. When answering A4 he referred to the four marks between 10 and 20 "because it's going up in twos and it's four ones so I divided ten into four". For B1 he identified the interval was split into five: "because, well, you divide, again you divide ten into five and then like, count two, so that'd be fourteen – two twos." In this case Student 3 correctly answered A4 although it was unlikely he reached his answer by dividing 10 by 4. In other items he was not so successful. This suggests that a teaching activity to help students learn how to work with the marks and spaces between labelled numbers could help improve student success with scale related questions.

For the remainder of this thesis, students showing confusion with what to count will be said to have a *marks and spaces error*.

Item evaluation

While all 13 students correctly answered A4, the item provided useful information as it evoked different strategies from different students. B1 also used a scale in twos but provoked incorrect responses. This suggests the pair of items is useful to investigate responses to similar number line and contextual items.

8.5 Aspect 3: Fractional and decimal scales (A5a, A5b, A6, B3a, B3b, B3c, A7, & B4)

Three item pairs intended to examine understanding of fractional scales, and one item pair examined understanding of decimal scales. For A5a and A5b, students tended to give a single explanation for how they reached both answers. As this also occurred with B3a and B3b, the responses tend to be discussed together.

8.5.1 Item A5a: Reading ¾ from a marked number line

Results

Table 8.3: Response pattern for A5a

Answer	Omit	0	0.75	0.8	0.9	-2	Total
Frequency	5	1	3	2	1	1	13

- Decimal answers were frequently given for this item, with 3/13 correctly answering 0.75. Each of these three used some form of whole number thinking. For example, Student 12 identified there were four pieces, ¼ of 100 is 25, so ¾ of 100 is 75, so its 0.75. The other two students also worked with 100.

  S10: … I would divide it by 4 and it was 25 and I just did 3 25s is 75.

  S5: Well you can't get 4 into 10 so I worked to 100 and stuff.

- Student 1 seemed to use a counting approach, working in points: "cause it's zero there [points to zero] and zero point nine, one is after zero point nine." … "And one point one is after one."
- Neither of the students answering 0.8 could explain how they obtained this result.
- Student 7 seemed to think in ones, with each mark representing a unit: "I knew that one before zero is zero and one after one is two."
- Student 6 provided an integer solution. He seemed not to recognise zero as a number, locating -2 as the opposite of two: "because it's another line after 2 and so I thought it would be 2. So if that one is 2, then that one would be minus 2."
- No students used a purely fractional strategy.
- When moving beyond whole numbers, some students became visibly confused, especially those using counting-based strategies.

Discussion

Baturo and Cooper (1999) suggest that miscounting marks can be the reason for some errors. For A5a, if the students answering 0.8 counted zero as the first mark they may have divided the unit into five pieces.

Most students used a consistent strategy when answering A5a and A5b. For example, if they answered 0.8 for A5a, they answered 1.2 for A5b. The exceptions were Students 2 and 6. Student 2 was unsure what number went in the box before 1, but was comfortable putting 2 in the box after 1, which seems to suggest she believes that the numbers between 0 and 1 behave differently to those after 1.

Some paired results gave insights into student thinking. Students answering 0 and 2 seemed to think each mark represents a whole number. Those answering 0.9 and 1.1 seemed to think each mark was a tenth, counted from the closest whole number.

8.5.2 Item B3a: Reading ¾ from a weighing scale

Results

Table 8.4: Response pattern for B3a

Answer	**Omit**	**0.3**	**0.4**	**0.7**	**0.75**	**0.8**	**75g**	**Total**
Frequency	2	1	1	2	5	1	1	13

Students were more successful with this item than the equivalent number line item A5a. In the first instance more students attempted the item (11 compared to 8). Secondly, more were successful (five compared to three). The three students successful with A5a and A5b were also successful with B3a and B3b. Overall, the five successful students exhibited more varied strategies, and were not as dependent on whole number thinking.

> S10: It's divided into 4 parts, 0.25, 0.5 and 0.75 and so forth. And that one there 'cause it's pointing to 0.75.

> S13: Since that was a thicker line there than the other ones, I thought that would be point five, so, yeah, it would mean that was point seven five, so it would be zero point seven five and on this one it was one line past the two so it would be two point two five.

> S5: Well you can't put like a ½ between 5 and 10 so you have to do it into sort of 0.25 kilograms, 0.5 to 50 kilograms and 0.75 and that remains the same.

- Student 11 seemed to have a mental strategy to address the way the scale was marked, but was unsure of the conversion factor between grams and kilograms, obtaining the solution 75 grams: " 'cause each line is a quarter of the weight – a quarter of the number…"

- Two students used 'counting in points' to reach the answers 0.3 and 0.4:

  > S7: Because between a number and another number if it's like there's one and two for example, the numbers between would be zero point and zero point three, so zero point three [accompanied by pointing to various parts of the line].

  > S9: I think they're like zero point four … cause the one's on zero so it might be like zero point.

  In reaching this answer, Student 9 has demonstrated a marks and spaces error. She does not know that different conventions exist for measuring and counting, so counts four marks by including zero.

- Two students also worked in tenths, but seemed to fit these to the marks.

S6: Well I thought 0, that would be 0.3 and that would be 0.5 and that would be 0.6 or 7 and then it would be 1 [accompanied by pointing to the successive marks of the line].

S1: Well, there's 10 there [points to 1], so I thought 5, 6, **7,** [verbally emphasised] then 8, 9, 10.

Students 6 and 1 did not seem to realise that decimals can be longer than one decimal place, so fitted what they knew of tenths to the number of marks given. In doing this, Student 1 converted to whole numbers, while Student 6 worked directly with the decimals.

8.5.3 Response comparison for the item pairs

A5a and A5b formed an item pair that paralleled B3a and B3b. Thus results from these pairs can be compared with some confidence.

- The decimal-based strategies only used by some students on the weighing scale items suggest that the context may have provoked the use of knowledge not tapped by the equivalent number line item.
- Most students used different strategies when attempting the two pairs of items. For example, 10/13 were inconsistent in their approach to A5a and B3a. One student attempted neither item.

Discussion

The approach used by Student 10 to B3a suggests a 'cuts and pieces' approach; that is 'three cuts make four pieces' and 'four pieces means quarters'. This strategy seems simple enough to be more accessible to students than approaches like identifying ¼ of 100 is 25. Encouraging the use of this strategy could improve understanding of fractional scales.

A number of students used some form of whole number thinking, including those who correctly answered the number line problems. This not only supports the finding of Siegal and Smith (1997) and Mack (1993) that such thinking may hinder students' understanding of fractions, but extends this finding by demonstrating that whole number thinking can be used successfully by students.

Item evaluation

These items identified a range of strategies that students use when working between whole numbers. While not all responses could be explained by the students, analysis of the errors suggests that much of the thinking behind the written answers can be readily identified, so the items are likely to provide useful information about student thinking and understanding in a written test as well as an interview.

8.5.4 Item A6: Locating ¾ on an unmarked number line

A6 initially caused some misunderstanding. In the first interviews, students were not sure if they were expected to find ¾ of 8 or locate the number ¾ on the number line. After the interviewer discussed problems with item wording with the first students to be interviewed, the wording for this item was changed to "put a cross where *the number* ¾ goes on this number line". This wording was also used for the written test.

Results

Table 8.5: Response pattern for A6

Answer	**3/4**	**3 1/3**	**3 3/4**	**5 1/2**	**6**	**Total**
Frequency	6	1	1	1	4	13

- All 13 students attempted this item, while only eight attempted A5a. Six were successful – double the number who were successful with A5a.
- Students used a range of strategies to reach their solutions. Of those successful, four used versions of splitting the interval into four equal parts and counting along. For example: "cause from ¼ to a ½ is ¾ of 1" (Student 5). One other student identified that ¾ was ¾ of a number while the other was observed to locate half, then sub-divide the interval between ½ and 1 into half again.

S4: Well ¾ is a ¼ more than ½ and …
I: So you're splitting ½ and ½ again?
S4: Yeah.

- The two students who provided answers between 3 and 4 seemed to be confusing proper fractions with mixed numbers, reading the numerator of the fraction as the whole number 3.

[Student has marked 3 1/3.]
S2: Well what I did was I guessed where the half was and I knew that if three quarters was the one below one half, so that's it.

[Student has marked 3¾.]
S9: Cause like, the 3 then mark quarters, like a little bit away from four.

- Three of the students answering 6 seemed to be working out ¾ of 8, and using multiplicative strategies to find the fraction of an amount.

S13: Oh I just thought three quarters of eight, which is six.

S3: Cause you divide eight into four and then times it by three.

The fourth student treated the whole number line as the unit with quarters, implying four pieces. The student then tried a trial and error (skip counting) approach to find what goes into eight.

S8: I sort of thought that the 8 was the whole and I sort of divided it to what number can go into a quarter of it and I just sort of went, how many can you get 4, so I first went in 3s but I could only get 2 out of them and then I tried 4s but I realized that that was only the half so what I needed to do again was half again.

- Student 1 originally marked 6, but changed this to 5½ when asked to explain her answer. Her reasoning was as follows: "cause one, two, three, four and one, two, three, four, one, two, three, four, that's the third fourth." How this explanation relates to the location of her answer is unclear, but as noted in Chapter 6, further clarifying questions were not used when, in the opinion of the researcher, the student was uncomfortable articulating their reasoning.

8.5.5 Item B3c: Locating ¾kg on an unmarked weighing scale

Results

Table 8.6: Response pattern for B3c

Answer	0.4	3/4	3	3.75	Total
Frequency	1	7	3	1	13

- Seven of the 13 students used the same strategy to answer A6 and B3c. However, apparently consistent solutions were not necessarily an indication of consistent strategy use. For example, for B3c Student 7 (who answered 3) explained: "there's four, there's one quarter [points to one] cause zero is nothing, so there's one, two, three, four and three so one, two, three." [Pointing to each whole number in turn.] He seemed to treat the whole number line as a unit, then counted each unit as a quarter; in A6 he identified 6 is ¾ of 8.
- Five students correctly answered both items. Four used the same strategy. The fifth was the only student to identify that B3a and B3c were the same item, so copied the marks and arrow from the scale above, so could also be said to use the same strategy.
- Three other students answered one of the items correctly, but did not use the same strategy for the other. (Two answered '6' to A6, but correctly identified ¾ in B3c. One identified ¾ in A6, but marked '3' for B3c.) This suggests that these students did not see these items as different versions of the same mathematical problem. It also suggests that access to informal strategies supposedly made possible by a familiar context does not always result in a better demonstration of student understanding.
- The range of strategies used generally mirrored those used for A6.

Discussion

A number of the responses to these items are consistent with the literature. For example, Baturo and Cooper (1999) suggested that some students may confuse symbols for proper fractions and mixed numbers; Kieran (1976) identified that one way to interpret a fraction is as an operator which works on a number (as in ¾ of 8); and Bright et al. (1988) suggested that some students treat the whole number line as a unit (so if 8 is the whole, 6 is ¾ of it).

For A6, all students used a different strategy to that used on A5a, in spite of both items requiring students to work with the number ¾ on a number line. The three students who were successful with A5a swapped to cutting the unit into four equal parts, and counting on three of them. This indicates they had the strategy available, but did not see it as relevant to A5a and A5b. In addition, only one student identified that the only difference between B3a and B3c was the marking on the scale. Taken together, the data suggest that students see marked and unmarked number lines as different classes of problems that require quite different skills.

While there is some evidence that a context can invoke additional knowledge, it seems that other factors can also influence a student's choice of solution strategy.

Item evaluation: implications for subsequent testing

The fractional items gave insights into the thinking students brought to scales, and helped identify how some students interpret proper fractions. How students treat marked and unmarked number lines also warrants further investigation.

As a student's thinking strategies varied from item to item, and were confused by marks and spaces errors, these items, while useful in a written format, would probably be better within an oral assessment. However, it should be noted that even in the interview situation some students were unable to explain their reasoning. In some of these cases, the explanations of other students with the same response proved very useful in that they could be used to infer the logic that led to the response.

8.5.6 Item A7: Locating 0.4 on a partially marked number line

Results

Table 8.7: Response pattern for A7

Answer	-4	0.2	0.35	0.4	0.45	1.5	Total
Frequency	1	1	2	7	1	1	13

- A similar number of students were successful with this item as with A6, although different solution strategies were used for the two items.
- Eight students recognised the mark for ½, then used the strategy 'a little bit more, a little bit less' to locate their number. Issues arose with this strategy for three of these students. For two, it was their accuracy. For the third, the student initially located 0.6 instead of 0.4, then self-corrected when describing his strategy:

  S6: Cause that one would be 1.5 and 2 [points to relevant marks] so I thought there cause that's 0.5 [points to mark half-way between 0 and 1] so I just took away 1 from 0.5 and put it there.

  These are examples of why the strategy has been called 'a little bit more, a little bit less' – the piece taken seems an arbitrary amount, and can be used either side of the mark.

- Most were comfortable working with decimals. Only one student used 'whole number thinking' in their explanation: "cause that's five and four would go a little bit before five" (Student 9).
- Two other potentially useful strategies were used by students, although neither led to a correct placement. Student 10 attempted to cut the interval into five equal parts, but could not do so accurately. Student 3 seemed to be working from a remembered rule involving 10 parts (the points).

> [1.5 is marked]
> S3: Because with the points you divide it into ten so that you could put it where four is.
> I: Right.
> S3: So you kind of count four lines and that's the answer, you just go back one [from the half-way line].

- Student 12 seemed not to recognise that 0.5 had already been marked, dividing the interval between 0 and 0.5 into 10, and marking a little under half-way.
- Student 7 seemed to believe that decimals are located to the left of zero: "cause zero point four is before zero and it's four, so I put four on" [Figure 8.5].

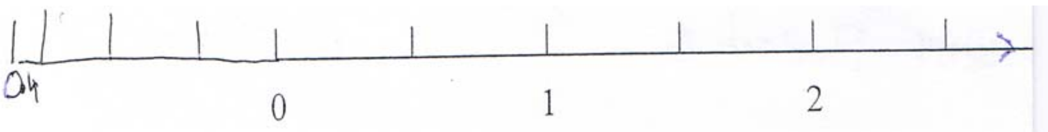


Figure 8.5. A7 response by Student 7: cropped version

8.5.7 Item B4: Locating 0.4cm on an unmarked ruler

Results

Table 8.8: Response pattern for B4

Answer	-0.4	0.3	0.4	0.5	0.6	Total
Frequency	1	3	7	1	1	13

- The same proportion of students correctly answered A7 and B4.
- Eleven students used the same strategy for both items, even though they themselves had to identify half-way in B4.
- Eleven students had potentially useful strategies for both items. Only four had both placements correct. For six students, accuracy was an issue on one item, and for one student it was an issue on both. For example, Student 7: "cause like zero point five would be about there [indicates where 0.8 would be], so the one before" [Figure 8.6].

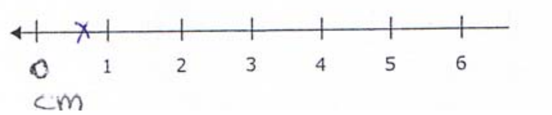


Figure 8.6. Student 7's response to B4, showing thinking inconsistency with A7

- One student, with an effective strategy for A7 identified that 0.4 would be found to the left of zero in B4: “I just drew some dots and then that was 3 dots, well from the arrow, 1, 2, 3, 4 so 0.4” (Student 6).

Discussion

The two students who did not use the same strategy on both problems did not seem to have a clear understanding of where decimals are found in relation to the whole numbers; exposing them to their discrepant thinking may be a useful way of addressing this issue.

While some students may have had an effective strategy to locate a number, they did not necessarily have the skills to place it accurately. This problem existed for both the larger interval in A7, and the smaller interval in B4.

Item evaluation

While these items were correctly answered by most students, they exposed a number of misconceptions that are worthwhile for teachers to identify. They also seem to identify where accuracy of placement is an issue.

As students used different strategies to locate their answer, and these were not necessarily obvious from looking at the written work, these items seem to be best suited to an oral interview.

8.6 Aspect 4: Integer scales (A8)

Ten students produced a scale, or enough to indicate where they believed ‘-3’ should be placed in relation to other numbers. As expected, some students were unsure how to proceed with this item.

While 6/13 (46%) provided scales that could be marked ‘correct’, information from this item was analysed in three different ways. This allowed information to be collected about student understanding of the place of zero on the number line and about how well units were iterated.

A8a Was -3 to the left of zero?

A8b Was zero shown?

A8c Were the numbers evenly spaced?

Results

A8a) Ten students showed some knowledge of negative numbers. Only one of these placed the negatives to the right of zero, suggesting most understood that negative numbers are smaller than zero, so are found to its left.

A8b) Zero was shown by 7/10 students producing a scale. One more marked zero, but did not label it. Two students had no space or mark for zero, but did show one, as in Figure 8.7.

This suggests that the students identified by Kuchemann (1981) as ‘skipping’ zero when counting from negative to positive may not recognise zero as a number.

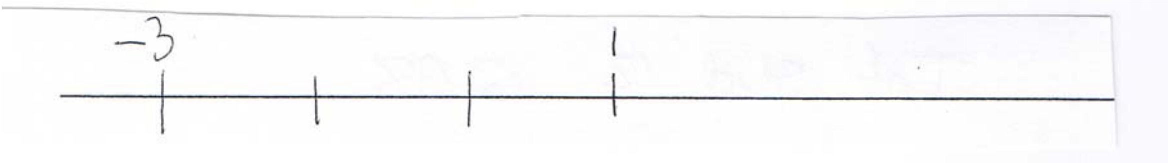


Figure 8.7. Student 12 omitting zero

A8c) Of the 11 students with marks on their page, only six had them evenly spaced. When compared to A1d, of the 10 students with a marked line in both items, five had uneven spaces in both items, two had even spaces in one, and three had even spaces in both. This seems to indicate there is some degree of consistency in students’ approaches to developing the spacing on a scale.

Item evaluation

The item was provided to test students operating at CL4, although it proved to be accessible to the majority of students. The results provide useful information about students’ understanding of the role of zero, and supporting information about students’ ability to space numbers evenly.

Most students who created a scale for this problem attempted to iterate a unit. Student explanations also provided little information that could not be read from the script, suggesting that the item would be suitable for a written diagnostic.

Now the results for each item have been discussed it is time to consider the bigger picture, and look for trends and patterns in the results. To start this process, item difficulty is considered.

8.7 Item difficulty

The process used to assess item difficulty did not follow the normal analysis pattern of psychometric testing as the analysis is only based on 13 students. Instead an inspection of the distribution was undertaken.

For each student a score was generated to make it possible to discriminate between students with different levels of success. For this purpose, 14 items were ‘marked’ correct or incorrect, generating a score out of 14 (A1, A8, and B5 were omitted, as they did not easily lend themselves to this form of marking). The students were then split into three groups on the basis of their total score; those with low marks (the bottom five students), middling marks (the middle three students), and high marks (the top five students). Note that as two students both obtained 11 marks, one was randomly assigned to the ‘top five’ group. Items were then analysed relative to the performance of the top and bottom groups, that is the items were analysed to see if they discriminated between these two groups. However, as the sample was small, and excluded consideration of the associated student comments, the results must be treated with some level of caution.

Results

The results of the analysis are outlined in Table 8.9. The level of discrimination considers the difference between scores for the groups. An item that generated a difference of three or more is identified as a strongly discriminating item, while those with a difference of two are classified as a discriminating item. Marginally discriminating items would produce a difference of one.

Table 8.9: Discrimination of the items between the low and high groups

Item	Number of students correct	Number in low mark group correct	Number in high mark group correct	Discrimination
A2b	13	5	5	No
A3	13	5	5	No
A4	13	5	5	No
A5a	3	0	3	Yes – strongly
A5b	3	0	3	Yes – strongly
A6	6	0	5	Yes – strongly
A7	7	2	2	No
B1	9	2	5	Yes – strongly
B2a	10	3	5	Yes
B2b	8	2	4	Yes
B3a	5	0	4	Yes – strongly
B3b	6	0	5	Yes – strongly
B3c	7	0	4	Yes – strongly
B4	7	2	4	Yes

The results show that the majority of items provide some level of discrimination between the two groups of students. When those items with no discrimination were reviewed, it was noted that all students correctly answered items A2b, A3, and A4, which accounts for the lack of discrimination and suggests that these items were ‘easy’. Only one item (A7, locating 0.4 on a partially marked number line) did not adequately discriminate between the groups. However, this item is part of a pair that explores the impact that a familiar context has on scales involving decimals so was retained for the written version of Test 1.

The next section reports student feedback sought during the short retrospective interview undertaken after the scale-related items had been answered (see Section 6.7). The questions focused on the students’ prior experience with fractions, decimals, and integers, how hard they found these number types (on a scale from 1 to 5) and how hard students found particular items.

8.8 Student feedback on items

All students indicated that they had worked with fractions before, while only 12 and 11 respectively indicated they had worked with decimals and integers. Table 8.10 shows the rating that students gave to how they found working with these numbers.

Table 8.10: Ratings for how students found working with different number types

	1 = hard	2	3	3.5	4	5 = very easy	Not done before	Mean
Fractions	0	1	6	2	3	1	0	3.4
Decimals	0	6	2	0	3	1	1	2.9
Integers	0	2	5	1	1	2	2	3.3

Students tended to rate fractions as slightly easier than decimals. This result is indicated by both the distribution of scores and the average scores. Overall, 6/12 (50%) rated fractions as easier than decimals, 3/12 (25%) rated them of equal difficulty and 3/12 (25%) rated decimals as being easier. One student indicated she had not worked with decimals before, and did not provide a rating, so was excluded from this comparison.

Students' ratings were cross referenced to the intended fractional and decimal items, to see if their perceptions of difficulty matched their performance. All six who rated fractions as easier used strategies based on decimals when responding to some of the items that asked them to read between whole numbers. None of these students used fractions. When scores were given for the fractional and decimal items, only 4/11 (36%) achieved marks consistent with their comparative ratings. This suggests that a student's perception of the difficulty of working with these number types was not a good indicator of their actual strategy use or achievement with the items.

The final check on item difficulty was to ask students which items they found difficult. In doing this, one student noted that rewording some items had made them easier. Six of the 13 students (46%) found A5a and A5b difficult, while 3/13 (23%) found the equivalent items B3a and B3b difficult. As these were designed as fractional items, this was not unexpected, with the identified difficulty level being reflected in the lower success rate for the items. Only two students identified that A8 (integers) was hard, while 10/13 (77%) indicated the bar graph in B5 was difficult. The issue here for most was the two rows of numbers in the table. It seems likely that discrete numeric data from CL3, and data tables in particular, may not have been met previously by these students.

One purpose of this research is to evaluate whether or not the curriculum expectations in relation to the learning of scale have been met. This is now addressed in relation to these students at School 1.

8.9 Curriculum expectations

The curriculum identifies expected outcomes for students in that the AOs are statements of what students should be able to achieve after learning. In Section 5.3.4 a standard of evidence was developed to identify whether or not the achievement expectations of the mathematics curriculum were being met in relation to scale. This was set at 75%, meaning that if 75% of the students in a group were successful with an assessment item based on the curriculum it could be said that the intended outcome had been met with respect to this item. That is, the group of students could be said to be achieving at expectation.

For only six items, all involving whole numbers (and all but one, number lines) could these students at School 1 be said to be meeting the curriculum expectations. These items were:

- A1a students knew that on a number line, the numbers should get bigger to the right;
- A2a students knew that 18 was half-way between 12 and 24 on a number line;
- A2b students could continue a scale in sixes;
- A3 student could locate 11 an appropriate distance from 12 on a number line;
- A4 students could identify the first mark past 10 on a scale in twos was 12; and
- B2a students could continue a scale in eights.

This result needs to be treated with a degree of caution due to the small number of students involved, but was a surprise in that the students were only meeting expectations for one contextual item, as the use of meaningful contexts is promoted by MiNZC (Ministry of Education, 1992).

In the next section, student responses to the items in Test 1 are explored further. To undertake this analysis, the strategies the students used were collated and summarised.

8.10 Strategies: a summary

When analysing trends in strategy use it became clear that a student tended to use a particular strategy across a range of problems. For example, ‘a little bit more, a little bit less’ could be used to locate 11 close to 12, and 0.4 close to 0.5. It also seemed that the strategies could be grouped into three broad categories related to particular forms of thinking. For example, ‘counting in points’ seems to link to counting processes. Others linked to additive or multiplicative thought. This suggests that the strategy facet of scale understanding may be related to number understanding, which provides a way to sort the strategies and develop a possible hierarchy (see Table 8.11). This first version of the hierarchy focused on the range of strategies that teachers participating in the teaching intervention for the original design of this research might develop with their students. (A second version can be found in Drake (2007) while the final version provided in Chapter 12 takes into account the later evidence from this research.)

Table 8.11: Initial classification of strategies for solving scale-related problems

	Strategy name	Main use	Description	Issues
Counting based.	Thinking in ones.	Scales with unlabelled marks.	Scales goes up in ones.	Only works for unit scales.
	Trial and error.	Scales with unlabelled marks.	A development of the units strategy. For example, students count along in ones. If that doesn't work try twos…	Only works for simple scales.
	Counting in points.	Decimal scales with unnumbered marks.	If there are marks between the (counting) numbers, count 0.1, 0.2, 0.3, …	Only works for number lines with 10 spaces (9 marks) between each unit.
	A little bit more, a little bit less.		Use to find numbers "just next to" other numbers.	The size of the "little bit more/less" can be arbitrary, causing low accuracy.
	Halving.		Eying up exactly where the middle is using the point of the pen as a marker.	Issues – some students are not able to work out where half is accurately.
Addition based.	Skip counting.	Scales with unlabelled marks.	A development of the trial and error strategy – using skip counts. For example – that's a big gap/number to fit on, lets try tens…	
	Bits and 'ths'.	Fractional number lines with marks.	For example, there are 5 ***bits*** (spaces), so each is a fif***th.***	
	Mixed methods.	Scales with unmarked intervals.	Combining the use of halving and "a little bit more, a little bit less".	
Multiplication based.	Marks and interval method.	Scales with unnumbered marks.	"There are 5 marks, the interval is 10, so each mark is 2".	
	Partitioning into 3, or 5 pieces.	Scales with unmarked intervals.	Students know how to accurately cut an interval into 2, 3, and 5 pieces, so use this information to read and locate numbers.	
	2, 3, 5, method.	Scales with unmarked intervals.	For example, locating 11 on a multiples scale of 6 by finding ½ way, and cutting the remaining interval into 3 equal pieces…	

It should be noted that the term *hierarchy* is used to describe the classification of strategies in Table 8.11. This term derives from the nature of the thinking processes involved – counting based, addition based, and multiplication based. It also needs to be stressed that while this hierarchy is based on the examination of student data, it is theoretical in nature, and suggests the hypothesis that students develop their understanding of scale over time, beginning with counting-based strategies before developing additive strategies and finally multiplicative strategies. This is consistent with the number framework, which suggests that when developing an understanding of number, students progress developmentally from counting-based strategies to additive, multiplicative, and proportional strategies (Ministry of Education, 2005a).

8.10.1 Developing the hierarchy

In creating the hierarchy, a number of strategies without obvious ties to a particular form of thought needed to be classified. The logic behind these decisions is now outlined.

Trial and error

This was classified as a counting strategy as students using it only seemed to recite certain counts, and did not access other number knowledge.

A little bit more, a little bit less

This was classified as a counting strategy as it can be used uncritically with little skill or understanding of number. It was used by all 13 students.

Halving

The decision on where to place halving was complex.

- Hunting (1989) details the progression students go through as they develop their ability to halve a unit – a process far from simple.
- Hart (1981) talks of one half as an honorary whole number, but in reference to students who are 11 to 16 years of age, and who are using what can now be identified as additive processes rather than the expected multiplicative ones.
- The number framework (Ministry of Education, 2005a) locates halving in the ratio and proportions domain, at stages 3 and 4 (within counting-based thinking), with students learning to answer problems like ½ of 8 (p. 15).
- The process of halving on a number line involves not only spatially partitioning an interval, but finding the midpoint of two numbers. If these numbers are small or simple, this process could be accessible to someone who is a counter.

In the interviews, halving was used by 12 of the students. The thirteenth got all items right without using this strategy. Here the frequency of use was an argument for considering this strategy to be a 'low-level' one.

Bits and 'ths'

As noted in Table 8.11, this strategy involves students accessing fraction knowledge relating to the number of partitions into which the whole has been divided. It was classified as additive as some students using it seemed to have no access to multiplication-based strategies, but were successful with their additive ones. Note that some students with multiplicative strategies also did not have access to this strategy.

2, 3, 5, method

Students using this strategy are showing multiplicative thinking. For example, when working on a scale in fifties, students locating the number 135 know they can split the interval into 10 pieces and know they can do this by halving the interval then cutting each half into five (so dividing by 10 is the same as dividing by two then five). This links to Stage 7 of the number framework (Ministry of Education, 2005b).

8.10.2 Student strategy profiles

Once the hierarchy of strategies was developed, its first test involved mapping student explanations to the hierarchy. The process of creating the strategy profiles helped identify omissions and inconsistencies in the hierarchy, which resulted in further development of the list of strategies and helped clarify how some of the strategies differed from one another. These developments are identified below.

Fitting tenths

This was classified as additive as it is more sophisticated than counting in points. It involves realising that the tenths and the intervals need to be aligned in some way. Conceivably, with a scale in fifths, this strategy could be successfully used by going 0.1 ***0.2***, 0.3, ***0.4*** …

Converting to whole numbers

Three students treated the whole of a fractional number line as a unit, then relabelled each mark as a fraction of that unit. This has the effect of converting an unmarked scale to a marked scale. This strategy was classified as additive as one feature of the counting strategies is that they seem reliant on knowledge of number sequences and little else, rather than having some understanding of how to operate on numbers. The strategy was also used by one student with low marks (but with a mix of strategies), one with middling marks and one with high marks.

n equal spaces

This strategy was separated from the '2, 3, 5 method' and classed as a simpler additive strategy. Students using this start with a subtraction to identify the width of the interval, then use this number to create that many equal pieces. For example '40 – 30 = 10, ten little lines'. However, as with many strategies its use can be compounded by marks and spaces problems as illustrated by this quote.

> Well if this is 32 to 40 and it's going up in 8s, it should just be ... you just like imagine that there are lines there, about 6 lines, and the lines should match up somewhere and is … a bit of an estimation without the lines there but … . (Student 8)

Other students seemed to know how many partitions to make, but found it hard to do this in practice. For example, when trying to locate 0.4 for A7, Student 12 initially tried to create 10 equal spaces, found this was not properly covering the interval, so decided 0.4 would be a little under half-way.

The 'n equal spaces' strategy is likely to be useful with a familiar scale like a ruler where the size of one tenth of a centimetre is known. However, it can be quite difficult to use as it involves accurately estimating $\frac{1}{n}$th of the interval (where *n* is the number of pieces that the interval is to be divided into), and the skill to repeatedly iterate the chosen unit. Multiple partitioning attempts were characteristic of students using this strategy.

Half and half again

It was noted earlier that halving may cause problems to students who are cognitively limited to counting processes if the numbers are larger. Repeated halving has been classified as an additive mixed method because of the need to operate on the numbers in a more sophisticated way. For example, quarter way between 20 and 40 cannot be found by simply looking at the numbers and identifying the middle number from knowledge of a number sequence; additional processes are needed, and the student needs to know that ¼ is ½ of a ½.

Students' strategy profiles

To create the student strategy profiles, the strategies used by each student were listed and the type of thinking they used in each item was identified. This analysis showed that those students with strategies classified as multiplicative also had access to lower level strategies, but not the converse. This evidence supports the hypothesis that a developmental progression exists for learning about scale.

The strategy a student chose to solve a problem seemed to depend on at least two factors (but there may have been others):

- the strategies a student had access to (not all students knew the same strategies);
- whether or not there were marks to work with.

Some students seemed to check their answer by using a different strategy from the one that they appeared to use when answering the item; possibly they explained their answer using a strategy that was easier to articulate than the one they were observed to use. In these cases students were credited with knowing and using both strategies.

Table 8.12 shows the range of strategies used by each student, but does not show where these were used in conjunction with another strategy to answer an item. Items A1, A8, and B5 were omitted, as they do not lend themselves to this form of analysis.

Table 8.12: Student strategy use

Score (out of 14)	Student number	Strategies	No of times used	Type of thinking
14	5	Little bit more… Halving. Skip counting. n equal spaces. Bits and 'ths'. Marks and interval. Whole number conversion.	3 2 2 1 2 2 4	Multiplicative.
14	10	Little bit more… . Skip counting. Bits and 'ths'. Marks and interval. Whole number conversion.	2 3 3 4 2	Multiplicative.
14	12	Little bit more… . Halving. Skip counting. n equal spaces. Bits and 'ths'. Whole number conversion.	2 3 4 4 2 3	Multiplicative.
12	4	Little bit more… . Halving. Skip counting. Marks and interval attempt. Conversion to whole numbers. Fraction as an operator.	4 3 2 2 1 1	Developing multiplicative strategies (but not transferring these to fractions). Has a 'marks and spaces' error.
11	11	Little bit more… . Halving. Skip counting. n equal spaces. Bits and 'ths'.	3 3 4 1 3	Additive (can work with fractions). Unclear how answer was reached for one item.

Table 8.12: Continued

11	13	Little bit more… . Halving. Skip counting. n equal spaces. Bits and 'ths'. Fraction as an operator.	3 5 4 1 1 1	Additive.
8	3	Little bit more… . Halving. Skip counting. n equal spaces. Bits and 'ths'. Marks and interval attempt. Fraction as an operator.	1 1 2 3 1 4 1	Developing multiplicative strategies (some errors in these – both applicability and execution).
8	6	Unit scale thinking. Halving. Little bit more… . Trial and error. Fitting tenths. Skip counting. Bits and 'ths'. Guess.	1 1 2 2 2 2 2 1	Additive (and can deal with fractional scales) No multiplicative strategies. Unclear how answer was reached for two items.
7	1	Little bit more… . Halving. Counting in tenths. Fitting tenths. Marks and interval. Fraction as an operator.	4 2 3 1 5 1	Multiplicative with whole numbers (but does not use this with fractions).
7	2	Unit thinking. Little bit more… . Halving. Trial and error. Skip counting. n equal spaces. Marks and interval attempt.	1 6 4 1 2 1 1	Additive.
7	9	Little bit more… . Halving. Counting in points (tenths). Skip counting. n equal spaces.	6 3 2 3 2	Additive (comfortably working within whole numbers and decimals). No multiplicative strategies.
6	8	Little bit more… . Halving. Trial and error. n equal spaces. Skip counting. Conversion to whole numbers. Marks and interval.	2 2 1 2 3 1 1	Additive. Developing multiplicative strategies (some errors in these strategies – both applicability and execution).
3	7	Unit scale thinking. Little bit more… . Halving. Trial and error. Counting in points (tenths). Conversion to whole numbers. Skip counting. Fraction as an operator.	2 2 1 1 3 1 2 1	Very little consistent thinking. Each problem discrete, unique, not linked. Has indications of early multiplicative thought.

Commentary

- The categories of thinking identified in Table 8.11 seem to be an effective way of looking at the strategies used by students. They appear to form a progression of difficulty. For example, students with high marks generally relied on strategies identified as additive or multiplicative.
- A student's total mark is not sufficient on its own to indicate their form of thinking. For example, one student might have used a multiplicative strategy, but was unsure of which marks to count. Another may have correctly answered the problem using a simpler addition-based strategy.

Now that the strategies used by students have been mapped, the consistency with which they have been used on both 'bare' number lines and contextual items will be considered.

8.10.3 Consistency of strategy use

The literature review identified that there is a common belief that relating mathematics to some aspect of the real world is an important approach to the teaching and assessment of the subject. Table 8.13 shows that there were significant consistencies between the strategies used for number line and contextual items when dealing with multiples and decimals on scales. There was significant inconsistency on some fractional items and between the conventions of a number line and a graph scale. On only one pair of fractional items did the context bring to light previously unidentified strategies. This may indicate that the use of familiar or real world contexts generally will not necessarily allow students to access additional knowledge, or provide a better picture of their understanding of scale.

Table 8.13: Consistency of strategy use when dealing with a number line and a similar item in a familiar context

Scale type	Items	Consistency rating	Percentage
Multiples	A2b, B2a	13/13	100%
	A3, B2b	10/13	77%
	A4, B1	10/13	77%
Decimal	A7, B4	11/13	85%
Fractional	A6, B3c	9/13	69%
	A5a/b, B3a,b	3/13	23%
Scale conventions	A1, B5	3/13	23%

It is also important to note that while there was some consistency in the way many students worked on 'bare' and 'contextual' problems, they did not seem to have a coherent concept of a scale and how it operates. Rather, some situations were seen to exist in isolation, to which a different set of 'rules' applied. For example, Figure 8.8 shows a response from Student 9, who, when asked how she got that answer, did not seem to imagine a thermometer could have anything other than a unit scale: "cause there's ten, that would be twelve". However, on a similar item involving a number line she correctly identified the scale went up in twos, suggesting her response to the thermometer item was influenced by the context (Drake, 2007).

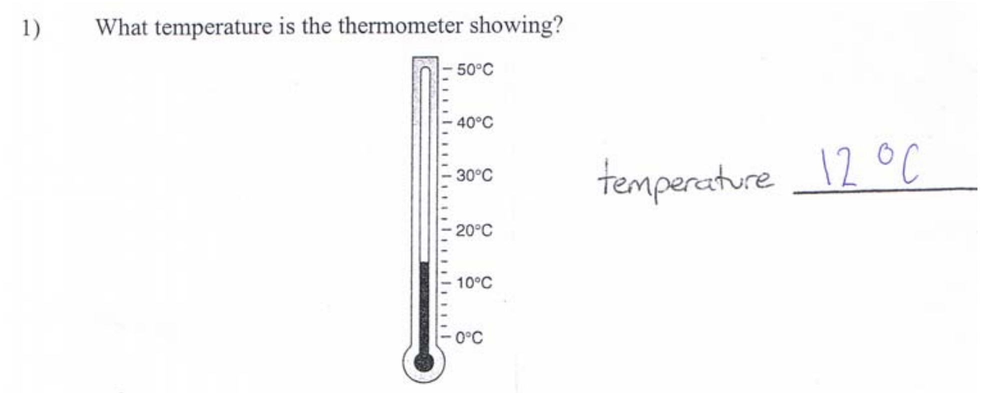


Figure 8.8. Unit scale thinking

Likewise in B5, some students seemed to have an unhelpful understanding of the structure of the bar graph. Here it can be said that using such a graph as a context for developing an understanding of scale introduced an element of *contextual pollution*. That is, it "introduced context situated knowledge that interfered with the intended learning about another topic" (Drake, 2007). In this case the contextual pollution was the misconception that the horizontal axis of a bar graph is not a scale so, for example, discrete whole number data recorded on this axis do not need to be placed in order of size.

8.10.4 Student strategies – where do they come from?

Because the strategies used by students are not mentioned in the curriculum document or reviewed resources, they can be considered to be a mix of *informal knowledge* and *prior learning* (Drake, 2007). Students may have developed some strategies for themselves while working with scales in classroom or real-life situations; others may have resulted from informal instruction while focusing on a learning task that happened to involve scales. In either case, the label 'informal knowledge' seems appropriate as the knowledge was not gained from formal teaching or instructional processes about scale. In addition, a few strategies can be identified as procedures taught to help students answer problems, and appear to be strategies that have been developed uncritically, then overgeneralised. For example, 'counting in points' was suggested as a way students can be taught to deal with decimals on a ruler. (We can almost hear a teacher say to the class 'each of these little marks between the numbers is a tenth, so its point one, point two, point three… .') Such strategies can be described as *prior learning.* Here prior learning can be defined as 'the body of skills, procedures and understandings a student has developed when working on classroom mathematics'. Note that this does not mean the skills and procedures are closely related to intended teaching. They are likely to include material that has been mislearned or incorrectly remembered, as well as material that has been reinterpreted in the context of what the student already knows.

To close this chapter, as a summary of findings has already been provided, some of the implications of the findings are now considered.

8.11 Implications

The following tentative observations can now be drawn from the analysis of this small sample of interviews.

1) Reading and locating numbers on a scale involves an active mental process. Students tend to use mental strategies rather than simply recall ‘known facts’ or recognisable ‘classroom procedures’ that fit the situation. Where such approaches (like counting in points) are used, they tend to lead to correct answers in only a few situations, ones the students are not able to adequately identify.
2) Using a strategy appropriate to the numbers and the type of number line does not guarantee a correct answer. Students also need to master how to count marks or spaces, to understand the conventions of a number line, and be able to partition accurately. The role of zero on a scale needs to be understood.
3) The contextual pollution observed in some students’ responses suggests that number lines may be a good vehicle for teaching ideas about scale.
4) Many students used decimals for problems more easily solved with fractions. They seemed unfamiliar with the measure sub-construct of fractions. Teaching this could lead to a general improvement in student understanding of scale.
5) Some students lacked knowledge of where fractions, decimals, and integers can be found on the number line. Work on fractions, decimals, and integers that utilises the number line could improve understanding of these number sets.

The nature of the findings in this chapter prompted a change to the research design (see Chapters 5 and 6). Chapter 9 outlines the first element of the changed design, a series of teacher trials to investigate how these adults answered scale-related items.

Chapter 9

Exploring understanding: The written tests for teachers

9.1 Introduction

The interviews at School 1 identified that the students used a range of mental strategies to answer scale-related problems. If these strategies are indeed common approaches to solving such problems they should be used by other populations, in particular adults. As detailed in Chapters 5 and 6, teacher trials were introduced to explore this conjecture. Teachers were chosen as a target population because they could also be used to test the concept of a strategy hierarchy based on the level of arithmetic thought; as people who have to teach mathematics, teachers are likely to be amongst the more successful mathematically, so they should be predominantly using the higher level strategies. They might also help identify other more sophisticated strategies not used by students.

9.2 Main Findings

1) The teacher trials showed that, like the students at School 1, teachers used mental strategies based on partitioning when working with unmarked number lines.
2) The teachers' strategies were among those identified at School 1, but tended to be from the higher levels of the hierarchy. This provides evidence to support the concept that the strategies for solving scale-related problems can be mapped to a hierarchy.
3) Written items asking the teachers to explain in writing the strategy for locating their answer did produce responses that, when considered alongside their number line, provided valuable information about the mental strategies used for answering the items.
4) Teachers tended to regularly use working marks. Students could benefit from being encouraged to do so more often.
5) The teachers' understanding of scale conventions was applied to an empty number line task.
6) Fractional language was commonly used to partition intervals on scales involving multiples of whole numbers. Responses also demonstrated that many of the teachers understood the relationship between multiplication, division, and fractions.

9.3 The first trial

9.3.1 The teachers

This trial involved small groups of senior primary, intermediate, and secondary teachers. A number of teachers from School 2 were involved, as well as small groups of teachers from a variety of

Wellington schools who were participating in short courses and PD workshops. In total, 22 teachers agreed for their work to be included in the study.

9.3.2 The items

These were specifically developed for the trial. As the sample involved adults, a different format was used for the items in that they were asked to locate a number on a number line, then explain in writing the strategy they used to get that answer. This format was introduced to assess whether or not strategies could be effectively identified when using a written test.

The number lines for the first trial involved whole numbers. The intention was to design items where the strategy (or strategies) used by the majority of teachers could be predicted, and the level of thinking identified from the strategy hierarchy. If that analysis was faulty, then the expected responses would not be evident, thus challenging the strategy-based interpretation. As only a small number of items were developed for the trial they will be introduced in the next section, alongside the results.

9.3.3 Results

The first item was designed to provoke the strategy of halving and halving again. Figure 9.1 shows the item, with the number in the box being the number to locate.

(a)

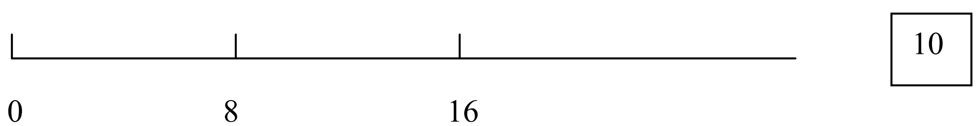


Figure 9.1. First teacher trial item

Seventeen of the 22 teachers identified that they used repeated halving. Three more divided the interval into four equal pieces. This latter strategy uses the knowledge that when finding 10, an interval of eight can be cut into the four equal pieces (quarters). Another teacher described cutting the interval into eight equal pieces, while the last halved then identified ‘10 was next to 12’. Note that quotes of teachers’ methods are not given as they were often abbreviated and notational. In some cases referring to the working marks on the diagram clarified these explanations.

The second item (Figure 9.2) involved locating a number that would be relatively easy to locate by partitioning the interval into five equal pieces. Most teachers therefore were expected to use this approach.

(b)

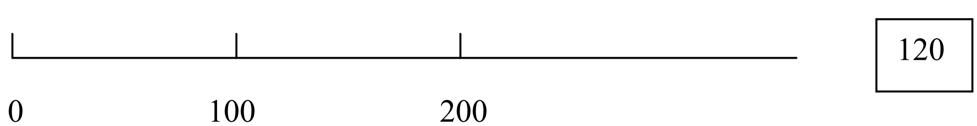


Figure 9.2. Second teacher trial item

In the event, only 12 teachers did this (for example, Figure 9.3). Of the others, three halved, then split the interval from 100 to 150 into five equal pieces (for example, Figure 9.4), which relies on the ‘2, 3, 5’ approach (knowing that dividing by 10 and dividing by two and then five produces the same answer). Six more halved and halved again, then took ‘a little bit less’. The final teacher split the interval into 10 equal pieces. According to the strategy hierarchy, all of these strategies can be identified as being additive or multiplicative in nature, so indicate the expected higher level strategy use.

Figure 9.3. Teacher work showing partitioning into five equal pieces (reduced)

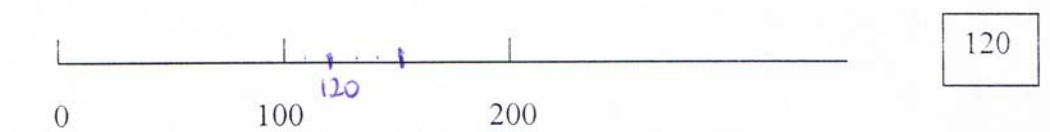


Figure 9.4. Teacher work showing halving then partitioning into five equal pieces

The third item (Figure 9.5) involved a problem many students would solve using ‘a little bit less’. However, the interview analysis suggests that this is an inefficient strategy, as the interval taken can be arbitrary, resulting in inaccurate placement. For the teacher trial, it was hypothesised that teachers would be aware of the drawback of the ‘guessing’ strategy, so would use a more sophisticated one – possibly one that had not been identified in the student interviews.

(c)

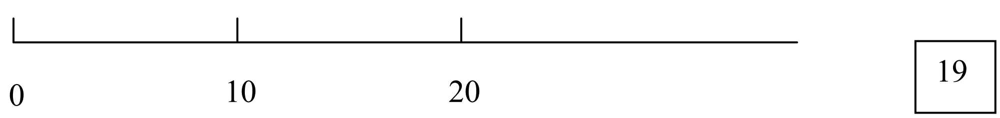


Figure 9.5. Third item

The strategies used for the item were varied, but showed that the teachers tended to use partitioning. Nine teachers used the 2, 3, 5 method of halving, then cutting the remaining interval into five equal pieces. One reversed this process, cutting the interval into five, then halving the gap between 18 and 20. Two repeatedly halved then located 19 a bit to the right of 18.75. Four claimed to use 10 equally spaced intervals (one wrote of placing nine equally spaced marks). The marks of two of these teachers suggested they initially tried ‘a little bit less’ (inaccurately). Of the last six, two halved to find 15, then tried ‘a little bit less’ while four used ‘a little bit less’. A number of this last group were as inaccurate as the students, suggesting that they were using an ‘arbitrary bit’ rather than a new and otherwise unidentified strategy.

For the fourth item, a more difficult number and scale was chosen, one that might force a change of approach to estimation (Figure 9.6). If partitioning continued to be used it was predicted that all would start by halving the interval.

(d)

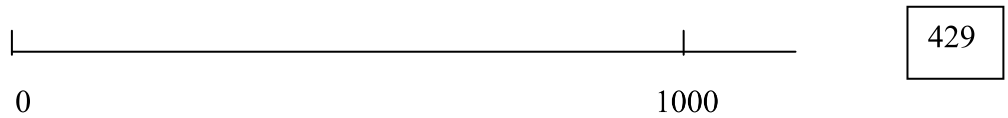


Figure 9.6. Fourth item

In the event, nineteen teachers started by halving. Unexpectedly three teachers either omitted or did not complete the item. Of those completing, most strategies were complex mixes of those already identified in the hierarchy. For example, five teachers halved to find 500, then split the interval from 0 to 500 into five equal pieces, then halved and halved again to find 425, and tried a little more than that. Four teachers halved, then divided the interval into five equal parts and estimated ‘a bit past’ 400. One halved to find 500, then split the interval from 0 to 500 into five equal pieces, and halved the interval from 400 to 500 to find 450 and split the remaining interval into five equal pieces to locate 430. Another reversed the last two stages. Only three teachers used halving then ‘a little bit less’, while another halved and halved again before taking ‘a little bit less’.

Some of the strategies showed a blend of whole number and fractional thinking. For example, one teacher halved then divided the interval into five equal pieces and cut the remaining interval into thirds. The only teacher who described cutting the interval into 10 equal pieces also talked of ‘moving a third past the last mark’.

9.3.4 Discussion

As the interviewed students had done, the teachers used mental strategies when working with these unmarked number lines. The common approach was to use partitioning, with a complicated problem tending to invoke more sophisticated partitioning strategies, rather than a transfer to estimation. For the first three items, only one teacher mentioned estimating: “1/5 of the way between 100-200. I estimated it” (Teacher 2). However, the number line showed this estimation really involved partitioning (Figure 9.7). Here it is also important to note that not all teachers have ‘mastered’ partitioning, for while the intervals are of equal size, there are six of them.

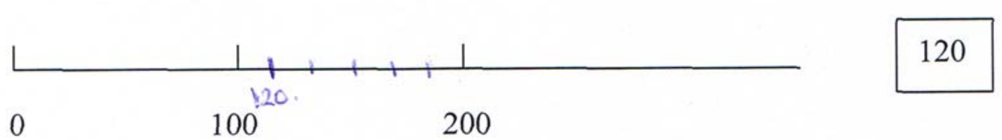


Figure 9.7. Teacher work showing a marks and spaces error

Even with the fourth item where it became very difficult to locate 429 on the scale without some form of estimation, only nine teachers wrote of ‘about’, ‘just behind’, ‘approximately’ or a ‘rough guess’. Three of these were the teachers identified to be using ‘halving then a bit less’ as their strategy. Overall, the responses to the four items therefore provide evidence to suggest that partitioning tends to be the common approach to working with unmarked number lines.

No new partitioning strategies were identified. As predicted, teachers tended to use ‘higher level’ strategies, with 19 of the 22 teachers using what was identified in Chapter 8 as a multiplicative strategy on at least one problem. The teachers also had a range of strategies to choose from. These findings suggest that a hierarchy is a useful way to categorise partitioning strategies. They also suggest that the developed list is a useful (but not necessarily exhaustive) summary of the strategies that can be used to locate numbers on an unmarked number line.

Most teachers used plentiful and obvious working marks on their drawings, again indicating the use of partitioning strategies. These marks generally coincided with the strategy described to locate their number. As they were so visibly used by the teachers, this suggests students should be encouraged to use more markings, even if they feel the need to erase incorrect ones.

Figure 9.8. Teacher working marks showing low quality unit iteration

The ability to iterate a unit was an issue for some teachers. Although stating the obvious, generally the teachers could halve an interval with accuracy, but when some tried to partition an interval into five equal parts less accuracy was noted (Figure 9.8). In some cases, this resulted in multiple attempts (Figure 9.9), in others the visible inaccuracy noted with the partitions was in sharp contrast with the detail in the explained strategy. Limited accuracy was most noticeable when a teacher was using ‘a little bit less’, or when trying to iterate a unit more than five times, both identified as lower-level strategies.

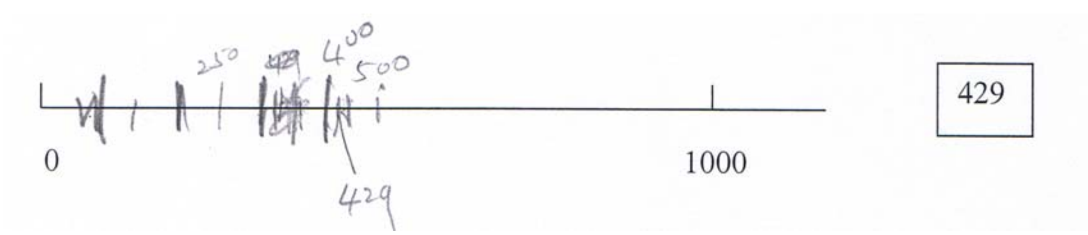


Figure 9.9. Teacher work showing multiple attempts at unit iteration

This last observation coincides with a classroom observation made by the researcher. Where a teacher draws a line on the board, then tries to split it into 10 equal pieces, marking sequentially, if the unit chosen is effectively iterated, the length chosen for the unit inevitably means the line is too short or too long, and needs to be changed. Combined, these two observations suggest many adults may find it difficult to accurately visualise ‘ten of’ or ‘a tenth’ in a measurement situation, which explains the development of strategies like the ‘2, 3, 5 method’ – it removes the need to repeatedly iterate.

9.3.5 The language of teachers

These four items all involved multiples of whole numbers. In answering them, the teachers commonly referred to fractions, or talked about division. Table 9.1 shows that all but one teacher wrote about halving, bisecting, or dividing the interval into two parts. (The other teacher used ‘*n* equal pieces’ for the first three problems and omitted the fourth.) Many teachers (14/22, 63.7%) also wrote of finding a fifth or dividing an interval into five equal pieces, and two teachers referred to using thirds. This suggests that fraction-based reasoning is an important aspect of the higher level partitioning strategies, as is understanding the link between multiplication, division, and fractions.

Table 9.1: Summary of teachers’ appeal to fractional or division understanding

Fractional language			
half	bisect	fifths	thirds
19/22 (86.4%)	1/22 (4.5%)	4/22 (18.2%)	2/22 (9.1%)
Appeal to division			
Divide into 2 equal parts		Divide into 5 equal parts	
1/22 (4.5%)		10/22 (45.5%)	

9.3.6 The empty number line item

The final item involved an empty number line (Figure 9.10). While this problem does not provide a scale according to the definition developed in Chapter 2, it was included to assess the understanding of scale conventions that the teachers applied. It was hypothesised that teachers would automatically treat the line as a scale so would label marks and include a proportional element to their response in which ‘larger skips’ would be used when showing the subtraction of a bigger number.

1) Use this empty number line to show how you solve the problem 92 - 39

Figure 9.10. The empty number line problem

Figure 9.11 provides two samples of teachers' work.

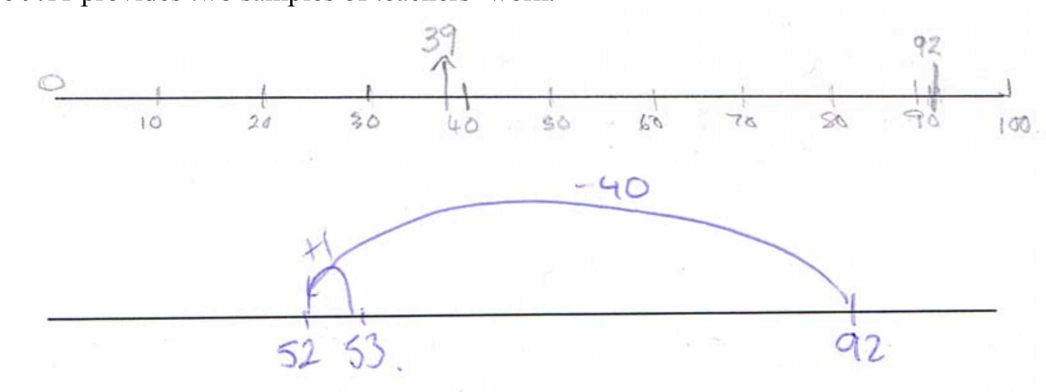


Figure 9.11. Teacher responses to the empty number line problem

Teachers' answers showed the expected response; although some illustrated they had represented the problem as a subtraction, others showed it as a difference (Figure 9.11): Generally, 'large skips' were shown when subtracting large numbers and 'small skips' were used to show the subtraction or addition of a smaller number. Skips of similar size tended to be used when showing skips of similar magnitude.

In the next section, the responses of a second group of teachers are explored. An additional group of eight teachers was given the items above plus two new ones. The additional items were developed to explore further the importance of fractional reasoning when partitioning.

9.4 The second trial

The responses for this group to the common items include no strategies not captured above, so will not be outlined here. The new items were added to test whether or not adults transferred strategies from whole number to fractional situations, something not observed amongst the interviewed students. Figure 9.12 illustrates the items and the response of one teacher. Of the seven teachers who answered the items, six clearly transferred their number line strategy from whole number multiples to the fractional number line. The seventh teacher incorrectly identified the number in 3(a) as 120. The eighth teacher did not complete these items, possibly as they were on the reverse side of the sheet so were not seen.

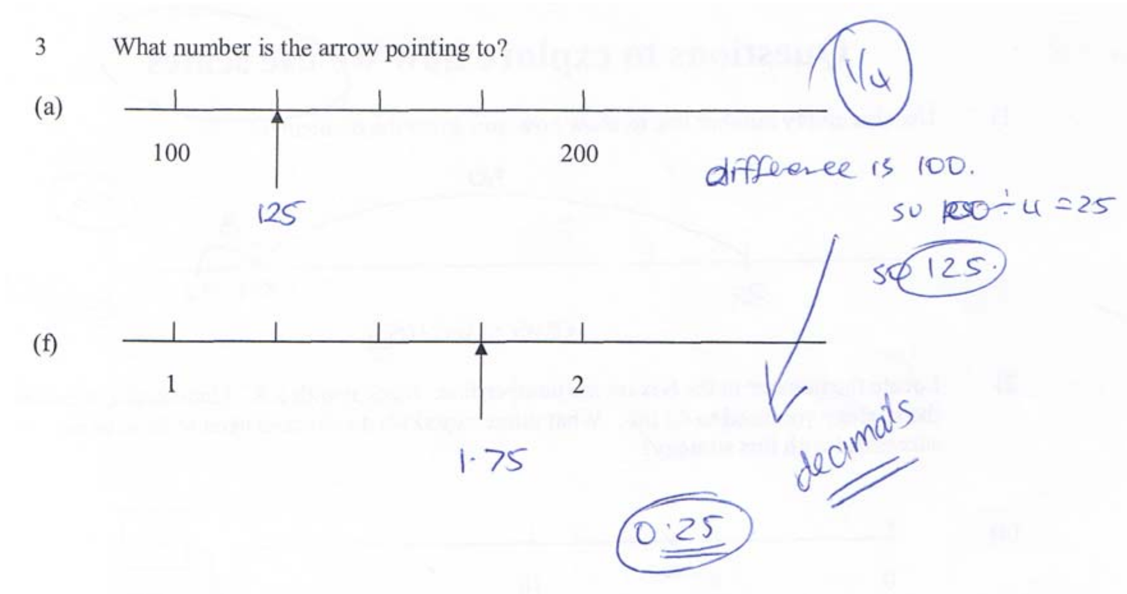


Figure 9.12. Teacher responses to additional items

9.4.1 Discussion

With only seven completed scripts it would be inappropriate to make general comments, even though the responses to the common items indicate that the teachers in this trial did not respond any differently from the teachers in the first trial to scale-related items. For this reason, another small trial was developed and is reported below.

9.5 The third trial

As noted in Chapter 6, a group of three teachers was given a new set of items to further assess aspects of strategy transfer. They were given items within the whole number domain then given similar items involving fractions. Based on the results of the second trial it was hypothesised that, as teachers of mathematics, they would transfer their whole number strategy to the fraction situation. To make what was being tested less obvious, the teachers first answered all the whole number items on one sheet of paper, then the fractional items on a second. The order of the items was also changed.

The first pair of items (1a & 2c) involved locating 13 on a scale marked in tens, and 3/10 on a unit scale. Two of the three teachers transferred their strategy of halving then dividing the interval into five, while the third used this strategy on the fractional item, but ‘halved and marked closer to 15’ on the whole number item.

The second pair of items (1b & 2a) involved locating 11 on a scale marked in sixes and locating 5/6 on a unit scale. Two of the three teachers used the same strategy of halving then dividing the remaining interval into three equal parts. The third (not the same teacher as above) did this on the whole number item, then reversed the process to find 5/6.

In the third pair (1c & 2b), teachers had to locate 42 on a scale marked in eights and 1¼ on a unit scale. Here the expected strategy involved halving and halving again– which all three responses showed.

The final problem (1d) also paired with 1c. In this item a scale of eights was still used – but this time each number was one more than in the item above, 1c. Two of the three teachers used the same strategies for each problem, while the third swapped from ‘halving and halving’ to ‘halving and dividing the interval into four equal pieces’. This meant all three had swapped strategies on one of the four paired items.

9.5.1 Discussion

As happened with six of the seven teachers in the second trial, the three adults from the third trial could transfer their whole number strategies to fractional problems. Some of the solution strategies for the scales involving whole numbers multiples show fractional thinking being used, similar to that shown in Figure 9.12. This suggests that the teachers in Trials 2 and 3 not only have an understanding of the relationship between multiplication, division, and fractions, but also recognise scale as a mathematical object to which their number knowledge can be applied, rather than seeing discrete problems that require different solution strategies.

While the teacher trials reported in this chapter cannot provide evidence of the understanding that students in Years 7 and 8 have of scale, they are an important source of triangulation. The results support the concept that a finite collection of mental strategies exists for partitioning unmarked intervals, and the idea that these strategies can be structured into a hierarchy. They also suggest that a written test requiring an explanation of the method used to reach the answer could be a useful tool for collecting information about the mental strategies used to solve scale-related problems. In the next phase of the research, the items in Test 1 were given to a larger group of students at a second school as a written test. Chapter 10 outlines the results and findings.

Chapter 10

Exploring Student Understanding: Case 2 – The Written Test

10.1 Introduction

The study at School 2 served a number of purposes. First and foremost it was to see if the responses of students from School 1 were similar to those of other groups of Year 7 and 8s. Second, the use of a larger sample meant the diagnostic was used as a written test, which allowed an exploration of whether a written form of the diagnostic could collect data on scale understanding as effectively as the diagnostic interview. Third, the larger sample also allowed the use of quantitative methods to assess how effectively the items discriminated between students of different abilities. Fourth, as student results for the NUMPA (Ministry of Education, 2005b) were available, links between scale understanding and number understanding could be investigated. Finally, it allowed the implications of the results for the teaching of these students to be explored.

Following the outline of the main findings, the chapter starts with an item analysis. As part of this, facilities (the percentage of students getting an item correct) are used to evaluate whether or not the students can be said to have met the curriculum expectations for learning scale. Both facilities and point-biserial correlation coefficients then help to identify items that discriminate effectively between students of different levels of understanding. Next, the items are grouped according to the four aspects of scale being considered (number line conventions, whole number multiples, fractions and decimals, and integers), and the results are explored in more detail. In particular, error patterns and similarities with the results from School 1, or from the literature, are noted. The existence of a possible relationship between the development of scale and number understanding (as measured by NUMPA results) is explored. The section closes by considering the implications for teaching. To round off the chapter, the results for two students are analysed in depth to identify if the written form of the diagnostic can give useful information about what students know, and general trends are noted.

As a reminder to the reader, the items developed for the test are discussed in Chapter 7, while a copy of the test can be found in Appendix 1.

10.2 Main findings

1) Most of the items correlated well with overall test performance. The levels of correlation found for the majority of the items suggest that the construct contains a coherent body of knowledge.

2) The written test was able to provide some useful diagnostic information about a student's scale understanding. However, when a student answered correctly it was not always possible to ascertain their thinking or identify next learning steps. Some errors could also be derived in a number of ways, so as you would expect, these results support the use of oral interviews where practicable.

3) Facilities identified that very few items had more than 75% of students succeeding, only one of which involved a context. This suggests that the students at School 2 had not met the curriculum expectation of being able to work with a variety of scales in a range of contexts by Year 7, as measured by the items in Test 1.

4) Errors noted in the literature and in the interviews at School 1 were also found in the work of students at School 2. Several additional errors were noted.

5) There was little evidence to support the notion that the students found contextual items more accessible than number line items; the exception being the pair of fractional items A6 and B3c.

6) The comparison of Test 1 and NUMPA results supports the concept of a link between scale and number understanding.

7) The errors in students' bar graphs suggest that a student's understanding of graphing cannot be isolated from their scale understanding. Graphs based on a Cartesian system have therefore been added to the concept map.

10.3 Item analysis

Table 10.1 outlines the items used in Test 1. The first column provides the item number, and a brief description of the skill being assessed; for example, A1a reports on the use of the convention that 'numbers get bigger to the right' for the number line item A1. The second column outlines the facility (success rate) of each item as a percentage of the 81 students who sat the test. This gives an indication of how easy or hard the students found the items. The third column provides Pearson correlation coefficients (actually point-biserial correlation coefficients (r_{pb}) as one variable is a discrete dichotomy) and the level of significance of the result from a two-tailed test (both calculated by the computer software SPSS). The correlation coefficient measures the degree of relationship between being successful in a particular item and the student's overall test score. As "the correlation coefficient is a standardised measure of an observed effect, it is a commonly used measure of the size of an effect" (Field, 2005, p. 111). An r value of ±0.1 indicates a small effect, ±0.3 a medium effect and ±0.5 a large effect (Field, 2005). The level of significance gives the probability of this result. For this purpose the significance is set at $p \leq .05$ (the conventional probability level for data of this nature). A level of significance less than .05 indicates that the effect is very unlikely to have occurred by random chance. Values reaching this level are identified with an asterisk in Table 10.1.

Table 10.1: Item analysis [N = 81]

Item	**Facility index**	**Discrimination index** r_{pb}	**significance**
A1a: Numbers get bigger to the right.	93	0.12	0.271
A1b: Using marks.	77	0.13	0.250
A1c: Numbers placed on, or used as, marks.	93	0.20	0.077
A1d: Equal spacing.	42	0.01	0.948
A1e: Spaces for 4, 6, 7, 9.	72	0.49	0.000*
A2a: Locating 18 on a scale in sixes.	90	0.39	0.000*
A2b: Extending a scale in sixes.	77	0.48	0.000*
A3: Locating a number close to a mark.	65	0.44	0.000*
A4: Reading a mark between labelled numbers (scale in twos).	68	0.59	0.000*
A5a: Reading a fraction < 1 (quarters). Marked number line.	33	0.46	0.000*
A5b: Reading a fraction > 1 (quarters). Marked number line.	28	0.51	0.000*
A6: Locating a fraction < 1 (unmarked number line in quarters).	11	0.32	0.000*
A7: Locating a decimal < 1 (partially marked number line).	47	0.39	0.000*
A8: -3 to the left of zero.	16	0.42	0.000*
B1: Reading a mark between labelled numbers (scale in twos).	69	0.60	0.000*
B2a: Extending a scale in eights.	92	0.60	0.000*
B2b: Reading between the marks (a bit less).	67	0.46	0.000*
B3a: Reading a fraction < 1 (quarters). Marked number line.	40	0.62	0.000*
B3b: Reading a fraction > 1 (quarters). Marked number line.	35	0.66	0.000*
B3c: Locating a fraction < 1 (unmarked number line in quarters).	36	0.60	0.000*
B4: Locating a decimal < 1 on an unmarked ruler.	28	0.46	0.000*
B5a: Space for 6, 8.	26	0.37	0.000*
B5c: *Y* axis scale includes zero.	37	0.59	0.000*
B5d: *Y* axis numbers on marks.	58	0.55	0.000*
B5e: Scale on *x* axis (ignore starting number).	22	0.60	0.000*
B5: Numbers get bigger to the right (*x* axis).	51	0.49	0.000*
B5f: Scale on *y* axis (starting at zero).	32	0.62	0.000*
B5: Numbers get bigger to the top (*y* axis).	47	0.55	0.000*

10.3.1 Item facility

As can be noted from these data, the items students found easy were (in order of facility index, highest to lowest):

- A1a: Numbers get bigger to the right.
- A1c: Numbers placed on, or used as, marks.
- B2a: Extending a scale in eights.
- A2a: Locating 18 half-way between 12 and 24 on a scale in sixes.
- A1b: Using marks to locate numbers on a line.
- A2b: extending a scale in sixes.

Note that the students at School 1 also found four of these items easy (A1a, A2a, A2b, & B2a).

These six are the only items for which students have reached the 75% success rate set as an indicator of whether students were reaching curriculum expectations. Three involve knowledge of number line conventions, which are not addressed in MiNZC (Ministry of Education, 1992), but need to be understood if students are to be successful users of scale. The other three involve identifying which whole number multiples should be placed on unlabelled marks so are as much to do with the students' knowledge of the appropriate number sequence as they are to do with scale (two used number lines and the third a graph axis). Overall, only 13 of the 28 items shown in Table 10.1 obtained facility indices over 0.50. It seems that the curriculum expectations for scale have not been met by the students from School 2; they have not demonstrated the ability to work with measurement and graph scales in a variety of forms.

The items that received facilities below 0.50 (from lowest to highest) were:

- A6: Locating a fraction < 1 (unmarked number line in quarters).
- A8: Locating -3 to the left of zero.
- B5e: Scale on *x* axis of bar graph (starting number ignored).
- B5a: Leaving space for 6, 8 on the horizontal axis of a bar graph.
- B4: Locating a decimal < 1 on an unmarked ruler.
- A5b: Reading a fraction > 1 (quarters). Marked number line.
- B5f: Scale on *y* axis of a bar graph (starting at zero).
- A5a: Reading a fraction < 1 (quarters). Marked number line.
- B3b: Reading a fraction > 1 (quarters). Marked number line.
- B3c: Locating a fraction < 1 (unmarked number line in quarters).
- B5c: *Y* axis scale of a bar graph includes zero.
- B3a: Reading a fraction < 1 (quarters). Marked number line.
- A1d: Units of equal size on a number line.
- B5: Numbers get bigger to the top (*y* axis).
- A7: Locating a decimal < 1 (partially marked number line).

The list includes all of the items involving fractions, decimals, and integers, as well as most of the analyses of the CL3 bar graph. Six of the items were from Section A and nine were from Section B. This is discussed later in Section 10.6.3. In the next section the level of discrimination that the items provide between students of differing abilities is discussed.

10.3.2 Discrimination

The items with values less than 0.30 (generally considered to be indicative of not discriminating) were for the most part those with high facility indices. Since most students answered these correctly, there was not a lot of scope for demonstrating discrimination between people. The exception to this pattern was one item involving spatial reasoning (item A1d). The judgement decision for this item was based solely on whether a student could produce intervals that were visually of a similar size, so measured spatial awareness. The low discrimination for the item is likely to be a result of the low relevance of this skill to the overall construct being measured. The remaining items all correlated well with 'overall' performance on scale as measured by the written test. This provides evidence that most of the items are discriminating between students with different levels of scale understanding.

In the next section attention turns to the results of groups of similar items – the aspect analysis. Aspect 1 looks at those items measuring understanding of number line conventions; Aspect 2 looks at items involving multiples of whole numbers; Aspect 3 relates to fraction and decimal items; and Aspect 4, integers.

10.4 Aspect analysis

10.4.1 Introduction

This section reports each aspect of scale understanding separately, beginning with a data table. In the top section of these tables, the first column gives the item number and a brief description of the measured skill. The second column details the success rate for the item as a percentage. The final column gives the percentage incorrect (including those omitting the item), then the percentage who omitted the item (e.g., see Table 10.2a). The second part of each table reports the identifiable error patterns for each item as percentages (e.g., see Table 10.2b). These data were coded by grouping the responses according to the errors and strategies identified amongst students at School 1 (Table 8.11) and/or findings from other groups of students reported in the literature.

Inferring the cause of an error in this way can lead to an incorrect attribution of causality as some students may have made a simple error not connected to understanding, such as not reading the question carefully. Likewise a student may answer correctly without understanding the concept being measured. This is a risk associated with diagnostic written tests and therefore acts as a caution against treating interpretations as more than suggestive. Note that the potential strategies of students successfully answering, or omitting, an item cannot be inferred from their answer, so did not provide useful diagnostic information. This is a drawback when using a written test.

After the quantitative data, general comments are made about the trends found in the responses, and particular findings relevant to individual items are discussed. Links with number understanding are then explored. Each aspect closes by considering the implications the results may have for teaching these students.

10.4.2 Aspect 1: Understanding of scale conventions

A1 (drawing a number line) and B5 (completing a bar graph) were used to measure students' understanding of scale conventions, with their rich data being analysed in a variety of ways. Hence A1a refers to the measurement of a particular convention in item A1. The results are presented in Tables 10.2a and 10.2b. Table 10.2a presents the item facilities (e.g., for A1a, 93% of students gave a correct answer; 7% gave an incorrect answer of whom the majority – 4% of the total group – omitted the item).

Results

Table 10.2a: Aspect 1 responses

Item	Percentage correct	Percentage incorrect	(omit)
A1a: Numbers get bigger to the right.	93	7	(4)
A1b: Using marks.	77	23	(3)
A1c: Numbers placed on, or used as, marks.	93	7	(4)
A1d: Equal spacing.	42	58	(3)
A1e: Spaces for 4, 6, 7, 9.	72	28	(4)
B5a: Space for 6, 8.	26	74	(27)
B5c: *Y* axis scale includes zero.	37	63	(25)
B5d: *Y* axis numbers on marks.	58	42	(25)
B5e: Scale on *x* axis (ignore starting number).	22	78	(37)*
B5: Numbers get bigger to the right (*x* axis).	51	49	(37)*
B5f: Scale on *y* axis (starting at zero).	32	68	(29)*
B5: Numbers get bigger to the top (*y* axis).	47	53	(29)*

* These figures include those students who did not label their scales.

Table 10.2b shows an analysis of the error patterns relating to Aspect 1. The first column identifies the possible strategies that were used (e.g., 'numbers get bigger to the left') while the remaining columns identify the items involved, classified according to whether the item utilises a number line (column 2) or a context (columns 3 and 4). Each cell in columns 2-4 provides the item number and gives the percentage of the total group (N = 81) who appeared to use the associated strategy. For example, the error pattern (strategy) 'numbers get bigger to the left' appears to have been used by only one student (1%) in relation to items A1a and B5. In contrast, the error pattern 'scale omits zero' is evident in the responses of 59% of the group in relation to the *y*-axis of item B5.

Table 10.2b: Error patterns for Aspect 1 items

Possible strategy	Number line item (% of group)	Contextual item (% of group)	Context (%)
Numbers get bigger to the left.	A1a (1)	B5 (1)	
Numbers order is mixed.	A1a (1)	B5 (*x*) (11)	B5 (*y*) (12)
Locating numbers in spaces.	A1c (3)	B5d (*y*) (17)	
Scale omits zero.		B5 (*y*) (59)	

Commentary

The number line and bar graph items provided two very different sets of results. Students were far more likely to use the conventions of a scale with the number line than when drawing a bar graph. Student responses relating to the four major conventions are analysed in more detail below. Reasons for the differences between the number line and bar graph responses are suggested.

a) Numbers get bigger to the right

This analysis looked to see if students understood and could use the convention that numbers are placed in order of size on a scale, and if laid horizontally get bigger to the right. This convention was commonly followed for the number line, but not so frequently with the graph axes. For example, Table 10.2b shows a small proportion of students used a mixed number order as part of their scale (11% on the x-axis and 12% on the y-axis). Most of these students did this on one or other axis, which may be due to the data table showing the numbers intended for the horizontal axis in size order. (Only one student used a mixed order on both axes.) Many of the students with a mixed order on the horizontal axis had swapped the data sets. This suggests they were trying to ensure their vertical axis had numbers in ascending order, but were not concerned with the number order on the horizontal axis. (Some students at School 1 also swapped the data on each axis, but were not asked why they had done this.) Those with a mixed order for the vertical axis tended to use the numbers from the table in the order in which they were found (see later Figure 10.2 for examples). This was also noted at School 1.

b) Numbers are located on the marks in a scale

For the number line, most students either used the measurement convention of locating the numbers on the marks, or treated their numbers as marks (see Figure 10.1). Only a few used the counting convention of placing their numbers between marks. However, being able to use this convention did not mean that students applied it to the axes of their bar graph, particularly the horizontal axis, even though they were using grid paper. (This was also noted at School 1.) In addition, almost all students with graphs drew bars that touched, indicating they were not clear on the related conventions for how to display discrete (counted) and continuous (measured) data.

c) Each interval represents an equal number of units

When working with a number line, the majority of students realised that space needed to be left for the numbers that did not need to be marked or labelled. However, 19 students created equal intervals between the numbers 2, 3, 5, 8, and 10, leaving no space for 4, 6, 7, and 9. For the graph it was common to find responses that did not show the zero frequency of 6 and 8, which distorted the visual image provided by the graph.

d) Each interval is an equal size

On the number line, students needed to iterate their chosen unit if they were to locate the numbers with any degree of accuracy. This was a skill that the majority of students (58%) did not possess (e.g., see Figure 10.1). In some cases this may have been partly due to their using their numbers as marks, as visually not all numbers are the same length. In others, the inaccuracy was due to the student not showing all the numbers between 2 and 10. However, some students did not seem to be aware that equal spacing was needed, and some who did seem aware of this need did not have the level of skill required to create equal units.

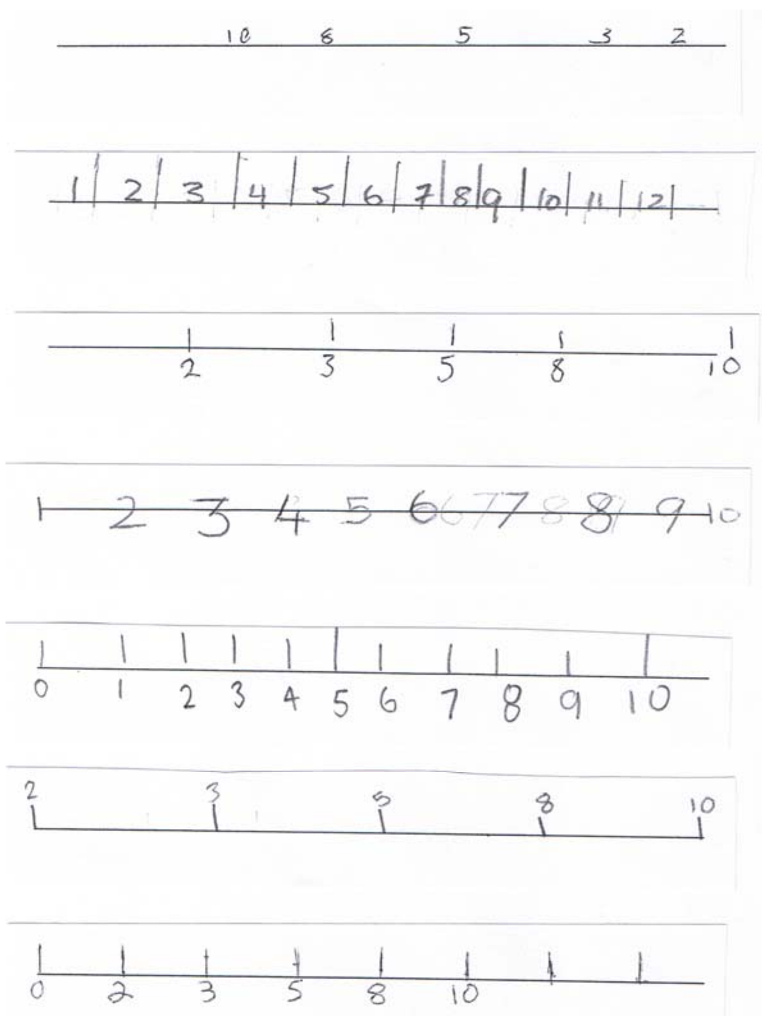


Figure 10.1. Student work samples for A1 (reduced)

The scales used by students

For A1, 32 students started their number line with 0, nine started with 1, and 35 started with 2 (the first number required). That students started at different points is consistent with the interviews. There was similar variation with the graph axes, both horizontal and vertical, which was unexpected. In fact, the low frequency of 'expected' responses, and the differences between responses to the number line and graph caused the researcher to look beyond the horizontal scale that was the original focus of the item and revisit the research on statistical graphing and to consider the graphs in more detail. The findings were relevant to this study, so are outlined below.

Students' graphs

At School 1, each graph was found to be unique. At School 2 it was rare to find identical graphs, even though there was a relatively large sample size. The following factors contributed to graph variety.

- Axis labelling was very varied.
- The role of zero on a graph axis was not clearly understood (e.g., 38% of students did not start their *y* axis at zero).
- Many students (41%) did not treat the horizontal axis of the bar graph as a scale similar to the one they created in A1, even though similar numbers needed to be labelled.
- The creation of one or more appropriate scales was no guarantee the information in the table was correctly converted to coordinates and plotted.
- Students' scale use was varied. For example, of the 32 students who located their numbers on the marks of the horizontal axis, 29 drew their bars in the spaces between the marks (28 had the bars touching). Of the 32, 18 drew their bar to the left of the labelled number, while five drew it to the right. Six more students either had no bars drawn, or the relationship between the bars and the axis numbering was unclear. Only two students centred their bars on the numbers without the bars touching. (Note that some examples of graphs can be found in Figure 10.2).

At least 28 students (35%) did not have an understanding of ordered pairs. These students made errors such as plotting the same number set on both axes, mixing the numbers so they were no longer in pairs, and not locating the numbers correctly in relation to the scales on the axes. This figure does not include those who swapped the axes for the data, maintaining the integrity of the ordered pairs. Nor does it include those situations where the lack of a scale made it unclear what had been plotted.

A number of studies indicate that the student responses to B5 are not unusual. For example, Ward (1979) identified that only 28% of the students in his study plotted points correctly in relation to the scales they had chosen for a graph (see Chapter 7). McGatha et al. (1998), in using measurement data 'that should be considered as a global whole distributed along a continuum', found some students treated the data as categories or discrete data points. In B5, discrete numeric data were used, but the results suggest some students treated the numbers as categories (e.g., the 24% of students from School 2 who used the numbers from the first row of the data table to label their horizontal axis, leaving no space for numbers with a zero frequency). As such this finding extends that of McGatha et al.

Developing an understanding of scale in the context of statistical graphing

Because B5 involved discrete numeric data, it effectively became a problem about graphing bivariate data (see Moritz, 2003). Students also needed to extract these data from a table. It seems they had several problems with this process. Firstly, while they may have been familiar with creating a bar graph for category data, a table of numeric data seemed less familiar. In particular, their understanding of the conventions for developing ordered pairs from a table was poor. Secondly, in a bar graph showing categories, the horizontal axis merely acts as a "common baseline" along which the bars are lined up (Australian Education Council, 1991, p. 165). With this baseline concept students are effectively displaying univariate data, so do not need to develop the idea that the horizontal axis is a scale. This concept only becomes relevant once students move on to data that can be considered bivariate. This helps explain why the students did not see creating a scale for the horizontal axis of a bar graph as a task similar to creating a number line. It is also clear that students' understanding of scale in a statistical graphing context cannot be isolated from their understandings of how graphs are constructed (thus providing another example of contextual pollution – see Chapter 8). It is therefore appropriate to analyse further the bar graphs students constructed, as this may help develop insights relevant to this study.

A 'framework for covariation?'

An attempt was made to apply to B5 the Biggs and Collis taxonomy and response levels used by Moritz (2003) (see Chapter 3). The area of interest was to see if these response levels could be adapted for the elements of scale upon which the development of an appropriate coordinate system relies. As this task only involved drawing a graph, the responses were initially classified under the following headings.

Level 1: Non-statistical responses / no answer provided.
Level 2: Single aspect responses.
Level 3: Inadequate coordinates.
Level 4: Adequate coordinates.

This framework allowed most graphs to be classified, but also helped identify several types of graph that were difficult to code. Some had no scales, and it was not obvious how the coordinates had been created. Others had scales and bars, suggesting some form of coordinate system had been used, but the values plotted showed little relationship to the numbers in the table. To separate these types from the non-statistical responses (e.g., see Figure 10.2) and those who had omitted the item, such graphs were given their own category, Level 1b. Graphs begun but not completed were also placed in this category. The original Level 1 was then named Level 1a. Three 'area graphs' (e.g., see Figure 10.2) were coded within this level. For these displays the data for each axis were counted then represented by shaded squares; so because the numbers for the horizontal axis added to 30, 30 squares on the grid were shaded – if the student had counted correctly. Twenty nine squares were shaded for the vertical axis data.

Another set of graphs looked like bar graphs, but were created from a single set of data – for example, by using the first row of data for both the horizontal axis and the height of the bars. In addition to these graphs, four students used an unlabelled horizontal baseline or a baseline numbered from 1 to 6, with bar heights from the first row of data. All such graphs can be classified as 'univariate responses' or 'Level 2'; that is, they only attended to one of the two variables in the data. These responses had serious flaws relating to understanding how to create a graph with non-categoric data.

Level 3 contained a third set of graphs that could be categorised as 'inadequate coordinate graphs'. In these, one common error was that the labelling of one or both of the axes did not create a scale, so the shape of the graph was not in accord to those drawn 'conventionally' (see Figure 10.2). (Moritz included some of the univariate responses outlined above in this level.) This level also included two graphs with the horizontal and vertical data reversed and three graphs where the bars were drawn horizontally. Here it is relevant to note that for these students the order in which the data appeared in the table seems to have been important; the data intended for the x-axis were in ascending order, while those for the y-axis were not. Reversing the data sets or drawing the bars horizontally allowed students to have the frequency data in ascending order, thus the graph conformed to the conventions for displaying categoric data on bar graphs while maintaining the integrity of the coordinates. Other graphs coded at this level include those in which both data sets were used (so there were different numbers used for 'x and y', but not necessarily paired values) and three with 'mixed order coordinates'. The 'mixed order coordinate' graphs generally had appropriate coordinates plotted, but had one or more pairs reversed, often where one number to be plotted was 'too big' to fit one of the scales (e.g., Figure 10.2).

Five graphs had appropriate scales and coordinates so were classed as Level 4 – but most of these still had inconsistencies. For example, the Level 4 graph in Figure 10.2 has the bars touching, which is generally considered to be inappropriate for discrete data. The first bar is also a different width, spanning the interval zero to two. However, this graph had the required scale and ordered pair development.

The process of assigning the graphs to levels basically maintained the four levels, but did not create a clear progression for the development of ordered pair and scale understanding. For example, some graphs coded as ‘Level 2’ had similar quality scales to those in ‘Level 3’ (so the scales of the Level 3 graphs were not necessarily more sophisticated than those coded as Level 2). In these cases the difference was a more sophisticated understanding of ordered pair construction in that the student used both data sets to create a bivariate display. Overall, the lack of a clear developmental progression led to the adoption of the term *category*, instead of *level* for Table 10.3 and Figure 10.2. (Note that Table 10.3 shows the number of students creating the forms of graph outlined above, while Figure 10.2 provides examples of student work to illustrate some of the categories and forms.)

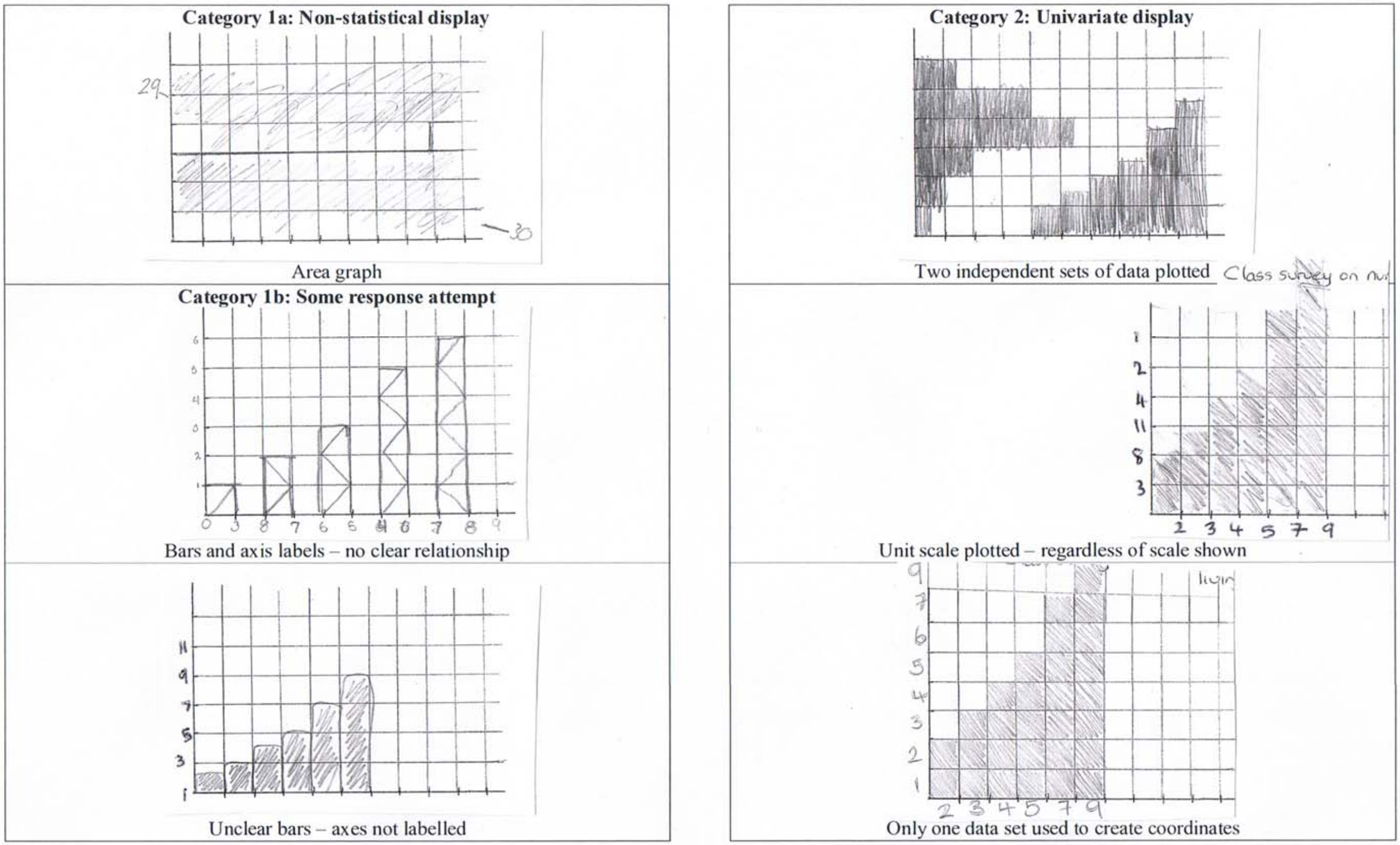


Figure 10.2. Student responses to B5 by category (reduced)

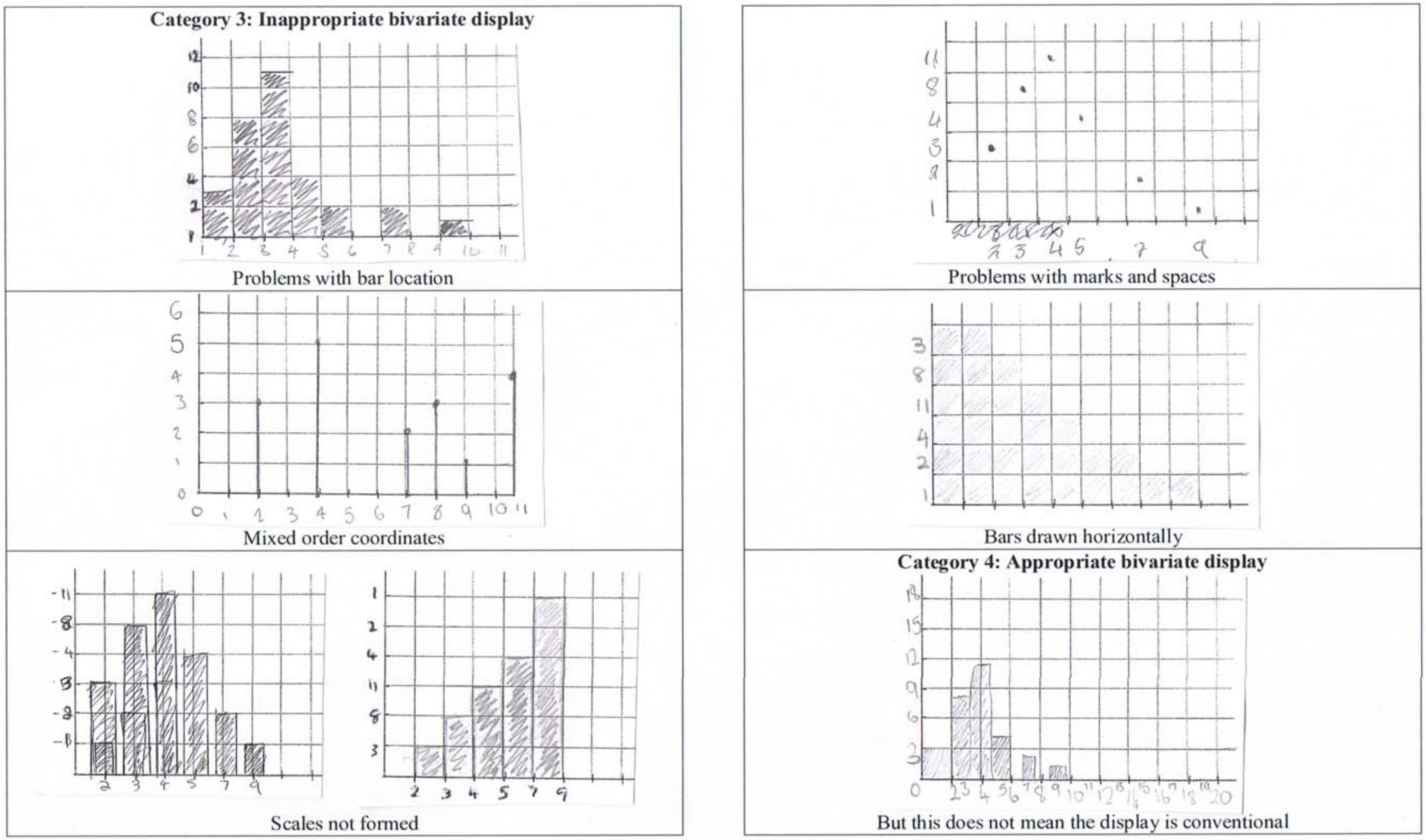


Figure 10.2 (continued). Student responses to B5 by category (reduced)

Table 10.3: Response categories for B5

Category	Type of response	Frequency
1a (non-statistical)	No response.	11
	Area response.	3
1b (some attempt at response)	Some graph form started.	12
	Unclear bars, axes not labelled.	3
	Bars and labelled axes – no clear relationship.	3
2 (univariate display)	One set of data reused to make ordered pairs.	12
	Each set of data plotted independently.	3
	Bar heights from first data set, *x*-axis not labelled.	3
	Bar heights from first data set, six categories for *x*.	1
3 (inappropriate bivariate display)	Scale not formed on either axis.	7
	Scales not formed for one axis.	5
	Point integrity not maintained.	3
	Unclear relationship between bars and scales.	1
	Bars drawn horizontally.	3
	Pairs reversed.	2
	Mixed order coordinates .	3
	Scales formed but plotting error.	1
4 (appropriate bivariate display)	Appropriate scales and ordered pairs.	5
Total		81

Discussion

While the process undertaken was of value to help sort the graphs and classify the errors found within them (so has been summarised here), the levels Moritz (2003) used could not be adapted to consider both scale and coordinate creation within the graphs. Rather, it appears that the process creates something that is task dependent, instead of illustrating general levels of understanding. This is not an isolated finding in that the CSMS team also found that Piagetian stages could not be compared across tasks when trying to describe stages of development in particular content areas (Hart & Johnson, 1980).

Marks and spaces

Another analysis attempted to find a link between the response categories and the placement of numbers either on or between the gridlines meeting the horizontal axis. This showed no clear trends. Even students with appropriate graphs could still locate their numbers ‘on the marks’ or ‘in the spaces’. This finding, when considered alongside the finding that the students who numbered the gridlines were not clear about where to place their bars, further highlights the confusion the students seemed to have of the process of drawing such a graph.

Implications for teaching

As at School 1, A1 provides evidence that not all of the students had a good understanding of the conventions of the number line, and that having an understanding of these conventions does not imply they have the skill to put them into practice. This suggests activities to improve understanding of these conventions would be valuable for some students.

The results for B5 also highlight the lack of scale understanding that the students brought to statistical graphing. The findings under the heading ‘marks and spaces’ provide further evidence to suggest that it is important to recognise that statistical graphs are scaled representations and to address scale understanding in student learning about these data displays.

Overall the results indicate that many of the students did not have a consistent understanding of scale, applying different understandings to the different situations. Generally, the students were more successful when working with the number line. This suggests some could benefit from working with a number line rather than graphs when learning about scale. The results further imply that some students may benefit from exploring the concept of an ordered pair in an algebraic context and being given practice plotting these on unit scales before dealing with graphing discrete numeric data in a context like statistics. This would be a change to current practice (see Chapter 4). It would allow students to come to terms with some of the conventions they need to master before meeting the distinctions between different data forms. Results from Ward (1979) suggest that students would find the topic easy, so with a little teaching around the convention of ‘across first and up second’, students at School 2 may be able to develop skills that greatly improve their ability to graph the data used in B5. However, it should not be assumed that teaching about scale in mathematical situations would be automatically transferred by students to situations involving statistical graphs. Explicit links would also need to be made.

10.4.3 Aspect 2: Scales involving multiples

This aspect covers scales that are labelled with whole numbers, and only require whole numbers to be read or located. Some items required students to work with unlabelled marks, and others to work between marks. The structure of the section generally follows that used for Aspect 1, while with one exception Tables 10.4a and 10.4b follow the conventions described for Tables 10.2a and 10.2b. The exception is that in Table 10.4b the items involving marked (but not labelled) intervals are separated from unmarked intervals.

Results

Table 10.4a: Analysis of Aspect 2 responses

Item	Percentage correct	Percentage incorrect	(omit)
A2a: Locating 18 on a scale in sixes.	90	10	(1)
A2b: Extending a scale in sixes.	77	23	(1)
A3: Locating a number close to a mark.	65	35	(14)
A4: Reading a mark between labelled numbers (scale in twos).	68	32	(0)
B1: Reading a mark between labelled numbers (scale in twos).	69	31	(1)
B2a: Extending a scale in eights.	63	37	(4)
B2b:Reading between the marks (a bit less).	67	33	(4)

As a guide for interpreting Table 10.4b, readers are invited to remember that as the students sat a written test, their responses to the items were grouped and coded according to the error patterns identified through the interviews at School 1. An in-depth discussion of particular errors can be found in Chapter 8, and a summary of many of the strategies students used can be found in Table 8.11

Table 10.4b: Error patterns for Aspect 2 items

Possible strategy for incorrect responses with marked intervals	**Number line item (% of group)**	**Contextual item (% of group)**
Intervening mark not noted.	A2b (15)	B2a (9)
Thinking in ones.	A4 (22)	B1 (22)
Counting in points.	A4 (3)	B1 (1)
Possible strategies for incorrect responses with unmarked intervals		
Reading to the nearest mark.✪		B2b (15)
Accuracy of placement (A3). Accuracy of reading (B2b).	A3 (7)	B2b* (0)

* Given the smaller scale in B2b, more leeway was given when coding correct answers, hence the smaller numbers coded with low accuracy. Without this leeway the proportion with low accuracy for B2b was 11%.

✪ This strategy is not based on partitioning. Reading to the nearest labelled number has the effect of making the problem a simpler one involving a marked scale.

Commentary

This and the next two sub-sections provide comments of a general nature relating to Aspect 2. Comments specific to trends with marked and unmarked intervals are reported under those headings.

Table 10.4a shows that over half of the students correctly answered each of the items involving multiples. However, only 22 students correctly answered all items, while 19 students correctly answered fewer than half of the items (including four students with no items correct). Table 10.4a also shows that students were generally less successful with the contextual items than with the number lines. Individual item analyses suggest even minor changes to a problem (that an adult would ignore) can mean some students see the problem as a new one, requiring different skills.

Most of the common answers provided by students at School 2 were identified during the interviews at School 1. This suggests that the strategies described in those interviews are common solutions strategies for such problems.

Marked intervals

The error patterns for A4 and B1 shows that about one quarter of the students provided answers that can be attributed to counting strategies. These students appear to treat each intervening mark as if it were either one or point one, apparently not realising they must work out what each interval represents before counting. Barcham (1996) also found students frequently assume the scale is in ones.

When extending a scale a smaller proportion of students seemed not to recognise that each new mark should be labelled with the next term of the relevant number sequence.

Unmarked intervals

As was found amongst the students at School 1, both the number line and the contextual items show that some students had difficulty locating and reading numbers accurately when working within an interval. The document analysis suggested students are often only required to read along a gridline or to a mark. Such an approach was illustrated in Chapter 8, when Student 7 from School 1 seemed to assume the answer could be read from the closest mark. At School 2, 15% of students were coded as 'reading to the nearest mark'; these students also may not have learned that reading between the marks is expected at times.

Individual item comments

In this section, findings relevant to particular items are noted. Readers are reminded that these relate to the data in Table 10.4b.

Item A2b: Labelling a marked scale

Of the six students with unusual solutions, three seemed to have problems with skip counting. Not unexpectedly this included the two students who scored below NUMPA Stage 4 for multiplication (Ministry of Education, 2005b). The others provided idiosyncratic (one-off) answers for which the logic was unclear.

Item B2a: Labelling a marked scale

There were more idiosyncratic answers for B2a than for A2b. Some of these could relate to students' ability to skip count in eights, as eight students gave even numbers between 52 and 72 that were not multiples of eight. However, there were also answers that seemed not to fit the question. For example, six students, with a variety of multiplicative stages on the NUMPA provided answers between 5 and 15. Note that the lower overall success rate with the contextual item was also observed at School 1.

Item A3: Locating a number between labelled marks

Eleven students omitted this item, a number of whom may not have noticed it. In the interview, the question was read out. In the written diagnostic it was a 'one liner' requiring students to add their answer to the number line for the item above (see Appendix 1). Of the 70 who answered, 53 (76%) were successful. Because this percentage is not much different to that for A2b, it suggests that there is little evidence to conclude that the students are better at either extending a marked scale or locating a number between those marked.

Item B2b: Locating a number between labelled marks

In the written test, both 38 and 39 were judged correct answers, as the top of the bar on the photocopy did not appear quite horizontal. Both responses indicate that the student noticed the bar was a 'bit below 40'. If both answers had been accepted during the interviews, 11/13 (85%) would have been considered correct.

Five students had an answer that was '10 out' (either 48 or 49). Four of these students may not have been able to correctly subtract from 40. One was at NUMPA Stage 4 for addition/subtraction so predominantly solved addition-based problems with counting strategies, while three were at Stage 5. Students at Stage 5 are able to cope with simple addition and subtraction problems using partitioning; however, subtractions from 'tidy' numbers like 40 are not always correctly undertaken early on. They can correctly work on the units column, but may fail to reduce the digit in the tens column by one.

This item was rarely omitted. Since it was intended to be of a similar difficulty to A3, this information provides evidence to support the argument that the number line item A3 may have been overlooked by some students.

Item A4: Reading between the labelled numbers

Students at School 2 were less successful than those at School 1, where all students answered correctly. Students with incorrect strategies tended to 'think in ones' or 'in points'.

Item B1: Reading between the labelled numbers

Twenty percent of students at NUMPA Stages 6 and 7 for multiplication used 'thinking in ones' with the thermometer, a higher figure than with the number line. This result is suggestive again of the concept of contextual pollution, as introduced in Chapter 8, but caution is needed in making such an interpretation without interview data.

Forty two students correctly answered both A4 and B1. Six of the students who correctly answered A4 swapped to 'unit thinking' for B1, while three gave a decimal answer. Such a strategy change was noted in the interviews, and interval size was suggested as a possible cause.

While the same percentage of students used 'thinking in ones' for both items, they were not always the same students. Only nine used this approach on both items. Many of the others answering 11 for A4 had idiosyncratic answers for B1. Using the strategy of 'counting in points' can account for only one of the decimal responses. Another, '10.4' could be explained by adapting the 'skip counting in twos' approach identified during the interviews at School 1. "Well, I counted 10.1, 10.2, 10.3, 10.4, 10.5, but that wasn't enough so I went up in twos."

Some of the other errors could have been obtained through incorrectly identifying the number of pieces into which the interval 10-20 was partitioned. This and the relatively low success rate supports the findings of Ward (1979) – that some students have difficulty working between the numbers of a marked scale, and some are unclear whether to count marks or spaces, and if counting marks, which marks to count.

Links to number understanding

Table 10.5 shows the results of analysing particular item responses in terms of students' NUMPA results. For example, for A2a, the incorrect responses to the item were compared to the student results from the multiplication/division numeracy domain. This comparison showed that neither of the two students who were below Stage 4 for their understanding of multiplication and division correctly answered the problem. The comment underneath the stage by stage breakdowns then provides a summary and/or additional information to help interpret the results. For example, the fact that according to the number framework students begin to learn skip counting in Stage 4 (Ministry of Education, 2005a) is pertinent information, so is provided to assist the reader. Readers are reminded that more information about numeracy stages can be found in Chapter 6, while a thorough explanation can be found in Ministry of Education (2005a, 2005b).

Table 10.5: Results of referencing the multiples items to NUMPA results

Item	**Area of Interest**	**Domain**	**Stage**					
			Below 4	4	5	6	7	8
A2a	Answer incorrectly	×/÷	2/2 (100%)	1/8 (13%)	3/16 (19%)	1/35 (3%)		
Comment	Students below Stage 4 generally have insufficient understanding of multiplication to skip count. No other trend evident.							
A2b	Answer correctly	×/÷	0	5/8 (63%)	11/16 (69%)	28/35 (80%)	18/20 (90%)	
Comment	There is some link between multiplicative stage and success. The task was accessible to many students with low multiplicative knowledge.							
	Answering 36	×/÷	0	2/8 (25%)	2/16 (13%)	7/35 (20%)	1/20 (5%)	
B2a	Answer correctly	×/÷	0	4/8 (50%)	7/16 (44%)	23/35 (66%)	17/20 (85%)	
Comment	Similar trend linking success to multiplicative strategy stage as seen for A2b.							
A3	Answer correctly	Fractions	5/13 (38%)		14/20 (70%)	14/26 (54%)	14/15 (93%)	6/6 (100%)
Comment	It could be hypothesised there is some link between success in this item and fraction understanding, as it involves partitioning an interval.							
A4	Answer correctly	×/÷	0	2/8 (25%)	8/16 (50%)	28/35 (80%)	17/20 (85%)	
Comment	Some relationship between correctly answering the problem and multiplicative understanding. The reverse trend was found for the use of counting strategies. This suggests a possible link between inappropriate use of counting strategies and understanding of multiplication.							
	Unit thinking	×/÷	2/2 (100%)	4/8 (50%)	6/16 (38%)	5/35 (14%)	1/20 (5%)	
Comment	There appears to be a link between a student's multiplicative understanding and 'unit thinking'. A similar trend was visible for fractional stages.							
B1	Answer correctly	×/÷	0	1/8 (13%)	7/16 (44%)	21/35 (60%)	15/20 (75%)	
Comment	Similar pattern to A4.							

Commentary

These results confirm a link between number understanding and success with scale problems involving multiples of whole numbers. While a particular stage at which students are able to answer such items cannot be identified, there tends to be greater success as multiplication or fraction strategy stage increases. In general, over 50% of students at NUMPA Stage 5 or above correctly answered these items. Students at Stage 5 for the multiplication/division domain can generally skip count, and know some of their simpler multiplication tables (Ministry of Education, 2005a).

Implication for teaching

Scales involving multiples of whole numbers are not only amongst the simpler scales that students meet, but are also commonly used in measurement and graphing situations. Developing an understanding of them should be an early stage in developing an understanding of scale. However, even when dealing with a marked scale in twos, roughly one third of students answered the number line and the contextual item incorrectly. The strategies identified at School 1 (and those used by teachers in the teacher trials) also showed that even when working with a multiples scale, fractional understanding tended to be invoked. For example, 15 is *half-way* between 10 and 20. This suggests that learning to work with multiples on a scale is no simple task, and involves understanding a variety of concepts. In the next section, attention is therefore turned to results for scales involving fractions and decimals.

10.4.4 Aspect 3: Scales involving fractions and decimals

Aspect 3 considers scales in which the partitioning of the unit is the focus. The section is structured similarly to previous aspects, starting with results tables (Tables 10.6a & 10.6b). The format of these tables follows that used for reporting the earlier scale aspects.

Results

Table 10.6a: Analysis of Aspect 3 responses

Item	Percentage correct	Percentage incorrect	(omit)
A5a: Reading a fraction < 1 (quarters). Marked number line.	33	67	(17)
A5b: Reading a fraction > 1 (quarters). Marked number line.	28	72	(17)
A6: Locating a fraction < 1 on an unmarked number line (quarters).	11	89	(5)
A7: Locating a decimal < 1 on a partially marked scale.	47	53	(7)
B3a: Reading a fraction < 1 (quarters). Marked number line.	40	60	(10)
B3b: Reading a fraction > 1 (quarters). Marked number line.	35	65	(3)
B3c: Locating a fraction < 1 on an unmarked number line (quarters).	36	64	(9)
B4: Locating a decimal < 1 on an unmarked ruler.	28	72	(5)

Table 10.6b: Error strategies for the fractional and decimal items

Possible strategy for incorrect responses with marked intervals	**Number line item (%)**		**Contextual item (%)**	
Whole number response.	A5a (19)	A5b (32)	B3a (10)	B3b (15)
Decimal response.	A5a (31)	A5b (27)	B3a (43)	B3b (51)
Fractional response.	A5a (27)	A5b (24)	B3a (27)	B3b (31)
Integers to the left of one.	A5a (6)		B3a (0)	
Whole number thinking (each mark is one).	A5a (7)	A5b (20)	B3a (3)	B3b (1)
Counting in points.	A5a (10)	A5b (7)	B3a (24)	B3b (19)
Not counting the first tile laid.	A5a (3)	A5b (2)	B3a (0)	B3b (0)
Fitting tenths.	A5a (1)	A5b (3)	B3a (1)	B3b (5)
Converting to whole numbers.	A5a (3)	A5b (3)	B3a (4)	B3b (3)
Marks and spaces errors.	A5a (4)	A5b (5)	B3a (16)	B3b (6)
Forgetting to record the closest whole number.	A5b (4)		B3b (3)	
Fractions or decimals only exist between 0 and 1.	A5a/b (6)		B3a/b (1)	
Possible strategies for incorrect responses when working with unmarked intervals				
Accuracy of placement.	A6 (4)	A7 (10)	B3c (4)	B4 (19)
Operator construct or treating the whole line as the unit.	A6 (20)		B3c (5)	
Fraction located in an incorrect interval.	A6 (46)		B3c (44)	
a/b located half-way between whole numbers *a* and *b*.	A6 (25)		B3c (19)	
a/b as *a* and the fraction 1/*b*.	A6 (9)		B3c (12)	
a/b as *a* and the fraction *a/b*.	A6 (5)		B3c (6)	

Commentary: marked intervals

A number of trends can be identified in the data for marked scales. For example, more students used whole numbers with the number lines than they did with measurement-based problems. More students used decimals with measurement-based problems than with number lines. This may have been due to the context tapping informal knowledge (Mack, 1993) or because the students expected to use decimals in a measurement context. It suggests that the number line and the measurement scale were seen as different constructs by some students. In a measurement context they recognise that decimals are found between the whole numbers, but treat the number line differently.

When working with both number line and contextual items, some students seem to have used the decimal counting strategies identified in the interviews. For example, for A5a 10% of students produced answers like 0.3 (counting in tenths from 0) and 0.9 (counting back a tenth from 1).

For each of the fraction reading items (A5a, A5b, B3a, & B3b) between one and three students provided answers that implied they had converted to whole numbers, possibly to use a form of thinking similar to that used by Students 5 and 12 at School 1, then failed to relate their answer back to the original situation.

In some instances a response did not fit into the categories identified at School 1. This may have been because such a response was not used or explained by the students at School 1, or because there were several potential explanations. For example, when answering A5a, the students at School 1 answering 0.8 were unable to explain their strategy, although Baturo and Cooper (1999) suggest this could arise from incorrectly including the marks at 0 and 1, giving five pieces instead of four. Further exploration of such responses is warranted.

Commentary: unmarked intervals

When locating a fraction on a number line, a large number of students placed their fraction in the *incorrect* interval (46% for A6 and 44% for B3c); some reasons for such placements were given during the interviews, but the number of students involved suggest this form of error is worthy of further exploration. This result contradicts Hart (1989), who found that for an item like A6 the majority of students seemed to treat the whole number line as the unit, placing their mark ¾ of the way along. (In this case, their answer would be 6.) The result is, however, consistent with Baturo and Cooper (1999) and Mack (1995). Mack found students with little prior fraction knowledge "confounded whole-number and fraction concepts and thought the numerator represented the number of wholes, or units, being considered and the denominator represented the number of parts in each whole" (p. 430), (so ¾ might be described as 3 pizzas each cut into 4 pieces). For A6 only 15/81 (19%) located ¾ between 0 and 1; poor placement in this interval was only a minor issue.

A second common error was for students to access an inappropriate fraction understanding. The interviews at School 1 found a student who treated the whole number line as a unit (as in Hart, 1989) while others were accessing the operator sub-construct, calculating ¾ of 8. Both approaches produce the same answer, so which strategy a student used could not be identified by the written test.

The success rate for the decimal items was not too different to those for the fractional items. This was unexpected as one item involved a form of ruler, something students should have been working with for many years. A higher number of accuracy errors may have been a contributing factor.

Item A5a: Reading ¾ from a marked number line

- Of those correct, 18/81 (22%) gave a fractional answer and 9/81 (11%) gave a decimal.
- As predicted by the interviews, students found this item much more difficult than those involving whole numbers.
- A few students can be classed as whole number or unit thinkers, for whom each mark represented one (see Table 10.6b). Students noting the 1, located 0 and 2 on either side. Students working from 0 could answer either 2 or 3, with 2 resulting from what Irwin and Ell (2002) called 'not counting the first tile laid'.
- Other than those who correctly answered ¾, only four students provided fractional answers. Of these, two answered ½ and one answered ¼. Figure 10.3 gives an example of how ½ can be reached; ¼. can be obtained by placing the same fractions in reverse order, and would indicate a common misconception about ordering fractions. It should be noted that such solutions also show a lack of understanding of the relative size of unit fractions.

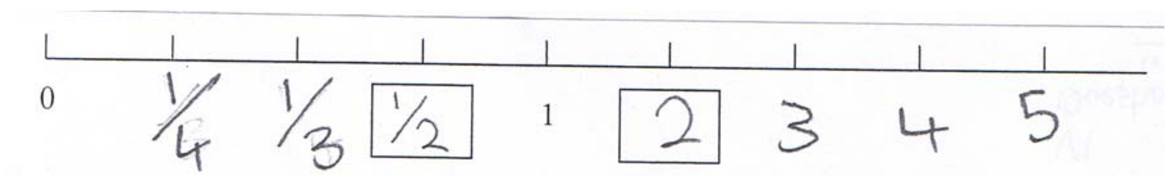


Figure 10.3. Student work for A5a and A5b

  - Some of the five students using integers may not recognise zero as a number, placing -1 to the left of 1 (e.g., Figure 10.4). These negative responses were only given by students at NUMPA Stages 6 and 7. These were amongst the more able.

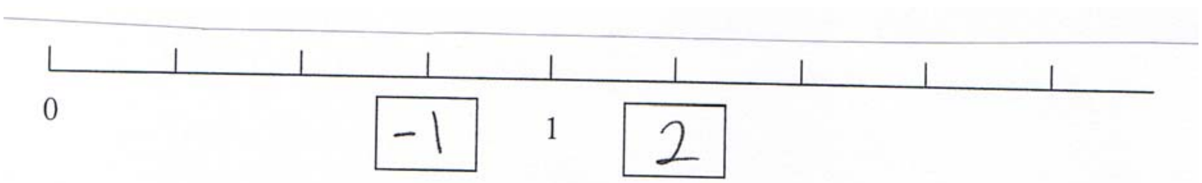


Figure 10.4. Student work showing integers to the left of one

Item B3a: Reading ¾ from a measurement scale

- When using the measurement scale, 13 students may have incorrectly started their count, producing the answers 4, 0.4, 4/5, 0.8, 8, and 80 (e.g., Figure 10.5). Most of these errors were not seen with A5a. The research into measurement scales, as well as Siegal and Smith (1997) who discuss fractions, indicates that some students have difficulty working out 'where to start' or whether to count marks or spaces. This form of error affected students at all NUMPA stages in both Years 7 and 8, regardless of which strategy domain was reviewed, suggesting that what was identified as the 'marks and spaces' error is an important one that needs to be addressed in teaching.

- Error pattern changes between A5a and this item (see Table 10.6b) suggest that some students see the measurement item as a different sort of problem to which they need to bring different skills. Strategy changes were also noted in the interviews.
- Twenty seven percent of students provided answers that can be attributed to counting strategies (e.g., 3, 4, 0.3, 0.4 and 0.9). As 0.8 could be reached by 'counting in point twos', this figure could be higher. Counting strategies were first noted in the interviews.

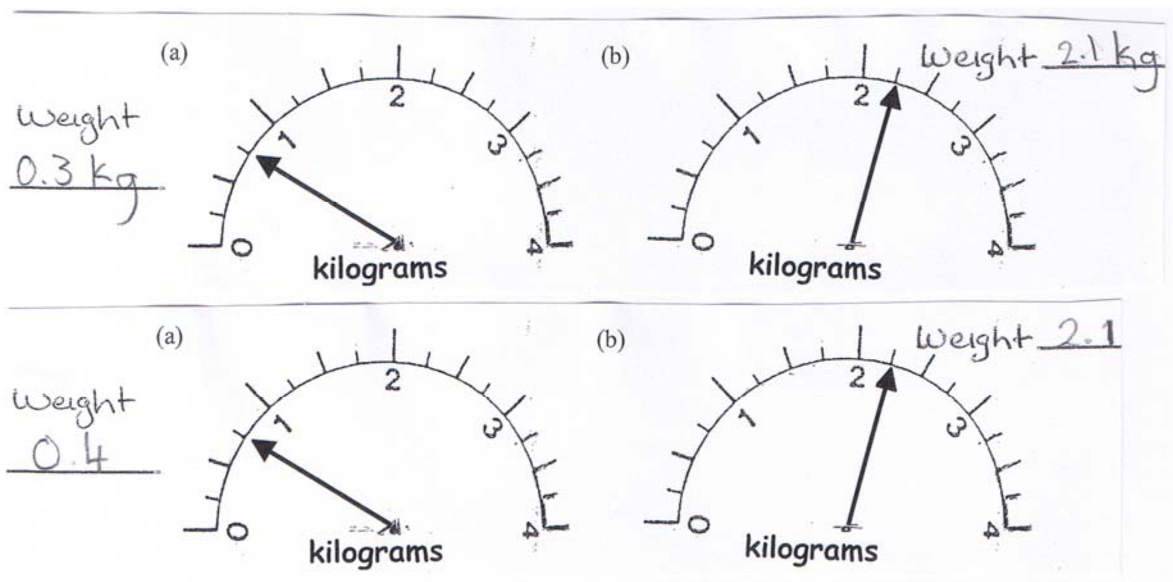


Figure 10.5. Student work for B3 showing counting strategies and 'marks and spaces' inconsistency

Item A5b: Reading 1¼ from a marked number line

- Fourteen students answered '2'. In four cases this was consistent with their response of '0' for A5a. For four of the others, it seemed to indicate that they believed fractions or decimals are found between 0 and 1, with whole numbers beyond this point (see Figure 10.6 and Figure 10.3 earlier). In such cases the conventions of scales were not consistently applied.

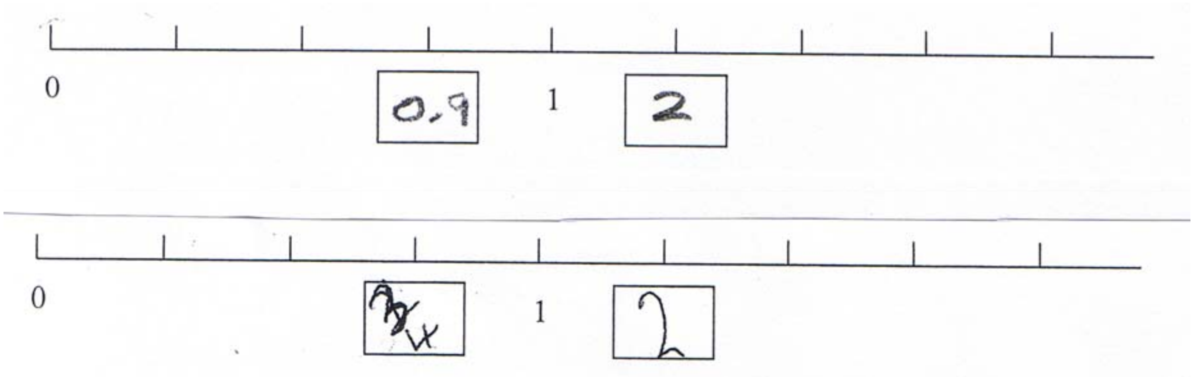


Figure 10.6. Student work for A5 showing fractions or decimals between zero and one

- One student answered ‘5’ and another ‘6’. Five can be reached by counting in ones from zero. For 6, zero is the first number in the count.

Item A6: Locating ¾ on an unmarked measurement scale

The interviews reported in Chapter 8 indicated the logic used to derive the answers 3 1/3 and 3¾. Other solutions between 3 and 4 were not met; however, Baturo and Cooper (1999) noted a student mark 6/3 as 6 and then considered thirds. Mack (1995) noted many of her students became confused when introduced to mixed numbers, returning to an early misconception. For example, a student could describe ‘a third’ as ‘one part out of three’, while ‘one third’ was described ‘this is the one and this is the third’, confusing the one for the whole number one. The solution 3¼ could therefore be derived from a similar logic to that of Student 9; ‘well it’s the three [pointing to the numerator] and there’s the quarter [pointing to the denominator], so it’s a bit more than 3’. Accuracy of placement may also account for some of the other less common responses but, as the interviews showed, this cannot be taken for granted; some placements that were slightly different to others were derived from deliberate and quite different strategies. Further work with such items is warranted.

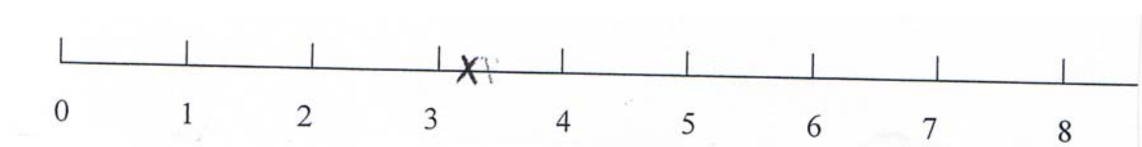


Figure 10.7. Student work for A6 showing deliberate placement of a mark at 3¼

While 33% of students were successful with A5a, only 11% were successful with A6, providing evidence to support Behr et al. (1983) who found that fractional number lines with a complete set of cues were easier than a number line with either a partial set or none.

Item B3c: Locating ¾kg on an unmarked weighing scale

Thirty five of the 81 students (43%) located ¾ between 0 and 1 while 31/81 (38%) located it between 3 and 4; the proportion locating ¾ between 3 and 4 being similar for A6. However, B3c had a higher success rate. One possible explanation for this is that some students who calculated '¾ of 8 is 6' for A6 misinterpreted the question, but corrected their error for the contextual question. This is consistent with the response change made by Student 3 at School 1.

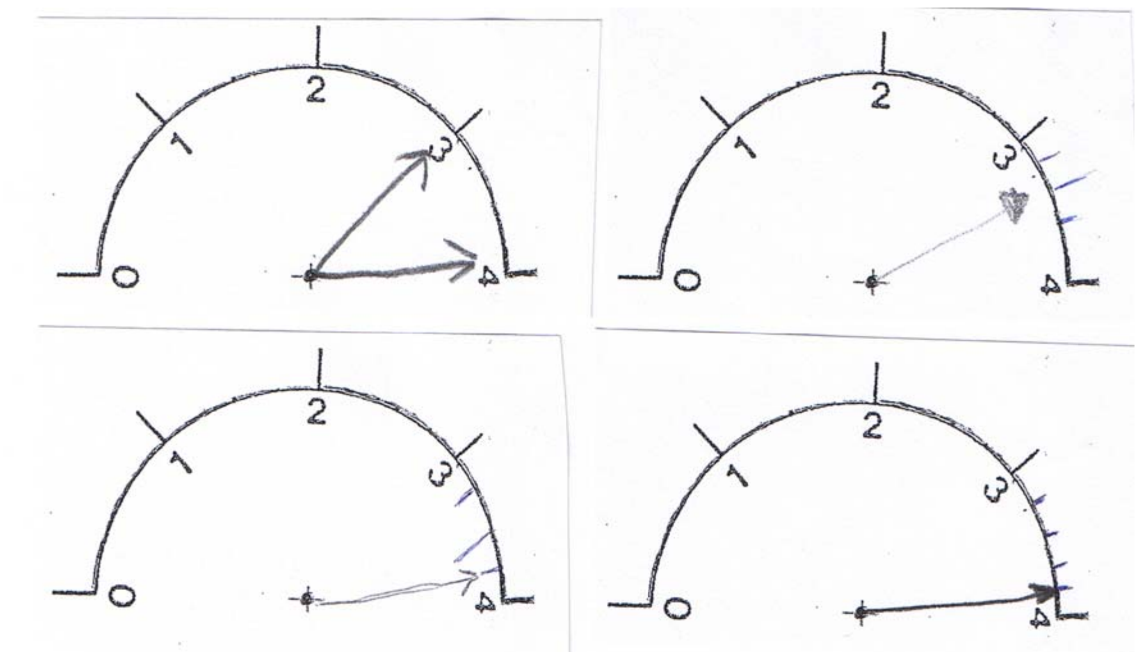


Figure 10.8. Student work showing deliberate location of ¾kg between three and four

Item A7: Locating 0.4 on a partially marked line

Overall, a much higher proportion of students were successful with A7 than any fractional item. More students located the decimal in the correct interval; 72% [58/81] located 0.4 between 0 and 1, while only 15/81 (19%) located ¾ between 0 and 1 in A6. This supports Siegal and Smith (1997) who suggest that some students do not view fractions as numbers located between the whole.

Item B4: Locating 0.4cm on an unmarked ruler

One student seemed to believe decimals are found to the left of zero on the ruler, but not when presented with a number line. Such inconsistency was also noted at School 1.

Fourteen students (17%) placed their response between 3 and 5; of these, 10 students also placed 0.4 between 3 and 5 in A7, but only one had the two answers in equivalent locations. When responses were compared to those from A6, only two did not locate ¾ between 3 and 4. This group seems to have little understanding of where fractions or decimals are located on a scale. When NUMPA strategy stage information was used to further analyse these results, they showed that while no student was beyond Stage 6 for fractions, they were not simply those students with weaker number understanding.

Table 10.7 shows the number of students who answered the number line and the contextual items correctly or incorrectly. It shows that many more students were correct with the number line than the ruler. A possible reason for this is that half-way was marked on the number line, but not on the ruler. It also shows that most students tended to either answer both items correctly or incorrectly, rather than answering only one of the items correctly.

Table 10.7: Answer consistency for A7 and B4

		B4	
		Correct	Incorrect
A7	Correct	17	21
	Incorrect	6	37

By Year 7, given that students have been working with rulers since Year 3 (Ministry of Education, 1992), it seems reasonable to expect that they have developed some understanding of the instrument. These results, along with those of Piaget et al. (1960), Nunes and Bryant (1998), and Bragg and Outhred (2000a, 2000b), show that this understanding cannot be assumed.

Neither A7 nor B4 provided fully marked intervals, but no student partitioned the (unmarked) centimetre in B4 into 10 equal pieces to make it look more like the ruler they recognised. It is possible the decimal itself caused problems for some. They may have learned to reunitise when using a ruler, measuring to the nearest centimetre, then counting on in millimetres to avoid the need to use or understand decimals.

Links to number understanding

Table 10.8 shows the results of analysing particular item responses in terms of students' NUMPA results. (The conventions used in this table are those used for Table 10.5.) Overall it shows that when reading from a marked scale, while there was no particular relationship between providing a fractional answer and improved understanding of fractions or multiplication, only those with an understanding of multiplication were able to respond correctly with a decimal. For the decimal items, accuracy of placement was an issue for students at all levels of understanding.

Table 10.8: Results of referencing the fractional and decimal items to NUMPA results

Item	Area of Interest	Domain	Stage					
			Below 2	2-4	5	6	7	8
A5a	Correct (fractional answer)	fractions	0	3/13 (23%)	3/20 (15%)	9/26 (35%)	2/15 (13%)	1/6 (17%)
Comment	Students with better fraction understanding were not more likely to provide a fractional answer. A similar result is found when multiplicative stages were cross-referenced. This suggests strategies leading to a fractional answer are not restricted to those with access to multiplicative thinking. It supports the classification of 'bits and ths' as an additive strategy (see Table 8.11).							
A5a	Correct (decimal answer)	fractions	0	0	2/20 (10%)	3/26 (12%)	2/15 (13%)	2/6 (33%)
Comment	Of the nine students correctly using a decimal, four were at Stage 6 for multiplication; five were at Stage 7. This indicates they had mastered their multiplication tables (Ministry of Education, 2005a), and suggests decimal solutions are tied to an understanding of multiplication and division. This observation is supported by student responses from the interviews at School 1.							
The results imply fractions may be easier to use than decimals on such items so, if taught how to do this, students with lower levels of number understanding may be able to access such items successfully.								
Item	Attribute	Domain	Below 4	4	5	6	7	8
B3a	Correct (decimal answer)	×/÷	0	0	0	4/35 (11%)	12/20 (60%)	
Comment	Only those with higher multiplicative stages produced a decimal answer. This supports the findings for A5a.							
A5b	Whole number thinking	fractions	0	3/13 (23%)	3/20 (15%)	2/26 (8%)	5/15 (33%)	1/6 (17%)
Comment	Some students with better fraction understanding still seemed to use 'whole number thinking' past 1, answering '2'.							
A5b	Correct (decimals)	×/÷	0	0	0	3/35 (9%)	6/20 (30%)	
Comment	Only students at Stages 6 and 7 were successful with decimals. This is similar to the findings for A5a and B3a.							
The proportion of correct answers was low for any fractional stage. Figures were similar to those for A5a, and only exceeded 50% at Stage 8.								
B3b	Correct (fractions)	×/÷	0	0	3/16 (19%)	7/35 (20%)	4/20 (20%)	
Comment	Once again, only students at Stages 6 and 7 successfully provided a decimal answer. This provides some additional evidence that fractional approaches may be more accessible to students of a wider ability range.							
A6	Answer correct	fractions	0	0	1/20 (5%)	4/26 (15%)	3/15 (20%)	1/6 (17%)
Comment	Students at a variety of stages correctly answered the item. Those at higher stages were not more likely to be correct.							
A6	Solution between 3 & 4	fractions	0	8/13 (62%)	10/20 (50%)	8/26 (31%)	4/15 (27%)	1/6 (17%)
Comment	As the frequency of solutions between 3 and 4 decreases as strategy stage increases, this suggests such thinking is less common among those with better fraction understanding.							
A6	Operator approach	fractions	0	0	2/20 (10%)	5/26 (19%)	5/15 (33%)	4/6 (67%)
Comment	Treating the whole number line as the unit, or using the operator sub-construct seems to increase with fractional stage. Students with better fractional understanding were more likely to use these forms of approach.							

Table 10.8 (Continued): Results of referencing the fractional and decimal items to NUMPA results

B3c	Answer correct	fractions	0	1/13 (8%)	3/20 (15%)	12/26 (46%)	8/15 (53%)	5/6 (83%)
Comment	There appears to be some link between fraction understanding and success with this item, which is a different pattern than for the number line item A6. However, as with B3a and B3b, only students with higher multiplicative stages answered correctly. (Stage 5, 1/16 [6%]; Stage 6, 15/35 [43%]; and Stage 7, 13/20 [65%].)							
A7	Answer correct	fractions	0	3/13 (23%)	9/20 (45%)	12/26 (46%)	9/15 (60%)	5/6 (83%)
A7	Answer between 0 & 1	fractions	0	8/13 (62%)	12/20 (60%)	18/26 (69%)	14/15 (93%)	6/6 (100%)
Comment	The proportion correct, and locating 0.4 in the correct interval, rises steadily with multiplicative stage. Accuracy of placement is an issue for some students at all ability levels.							
B4	Answer correct	fractions	0	2/13 (15%)	5/20 (25%)	6/26 (23%)	8/15 (53%)	2/6 (33%)
B4	Close to correct	fractions	0	5/13 (38%)	9/20 (45%)	12/26 (46%)	10/15 (67%)	6/6 (100%)
Comment	Students across a range of stages were successful, regardless of strategy domain considered. However, when close to correct answers are considered a pattern emerges, suggesting accuracy of placement is a significant error pattern for students of all abilities.							

Implications for teaching

One trend identified in the data is that some students with little fraction understanding successfully answered a number of the items using fractions, while only students with higher fraction understanding were successful with decimals. This suggests that developing teaching activities to strengthen students' fractional strategies could improve performance. These activities could be appropriate for most ability groups and should include a focus on developing an understanding of marks and spaces; for example, by giving them practice at folding and cutting paper into equal sized pieces (noting how many cuts are needed for how many pieces). Repeating such work with a line (which cannot be folded) could be important for introducing strategies that ensure the correct number of pieces are created, and checking that the pieces are equal. Working with partitioned lines to identify how many pieces into which an interval has been cut could also be beneficial.

Some students seemed not to understand that fractions and decimals can be found between all the whole numbers, some may have been confused between the different fraction sub-constructs, and some misinterpreted proper fractions. This indicates that it is important to begin work on the measure construct of fractions by linking fractions symbols (and what they mean in relation to the part-whole model) to locations on the number line. Here the number stick (Figure 10.9) seems a valuable piece of equipment, as it can mirror a marked number line, but is easy to change from one denominator to another.

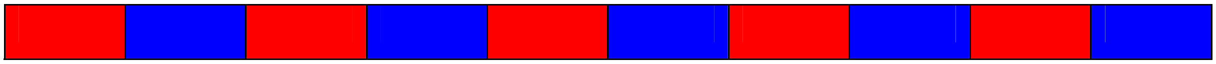


Figure 10.9. A number stick

Because the number framework (Ministry of Education, 2005a) indicates that students should be introduced to improper fractions or mixed numbers as they work towards Stage 5, it seems sensible not to introduce students to the measure construct of fractions until they are also working towards this stage.

In the next section results for the item assessing the fourth scale aspect are considered.

10.4.5 Aspect 4: Scales involving integers

The main focus of this aspect was to see if students understood that the conventions of the number line continued when the line was extended from the positive to the negative and understood the significance of zero on the number line. However, the results were difficult to code as the item did not gather the intended information. As happened with some other items, the student responses led to the item being adapted for School 4.

Results

Table 10.9: Analysis of Aspect 4 responses

Item	**Percentage correct**	**Percentage incorrect**	**(omit)**
V1-A8a: Locating -3 to the left of 0.	16	84	(5)
Possible strategies for incorrect responses	**Number line item (%)**		
Only one (negative) number marked and labelled.	A8 (12)		
No negative numbers shown.	A8 (54)		
Negative number(s) to the right of zero.	A8 (27)		

Item A8: creating an integer number line

This item was originally intended for students with greater scale understanding, being pitched at CL4. However, the 13 students who answered correctly were not solely those with more sophisticated strategies in the NUMPA +/- domain. One student had counting as their most advanced additive strategy, while six could not work successfully with three digit numbers (Ministry of Education, 2005a). The interviews also suggested that the concept of a negative number may be more familiar to students of this age than originally expected. Students may therefore be able to develop an understanding relatively early on of where negative numbers can be found in relation to whole numbers.

In contrast, 44/81 (54%) of students labelling their mark did not recognise the significance of the negative sign, ignoring it and marking 3 instead. In addition, 32/81 (40%) of students simply put a mark where they thought the number should be, and showed no other numbers. Four of these students did not label their number, so it was not clear if they were marking -3, or did not recognise the negative sign as a symbol with meaning (so were marking 3). Such responses were not met during the interviews where, as requested, most created a number line and marked their location for -3. The nature of these students' responses meant that little meaningful analysis could be completed.

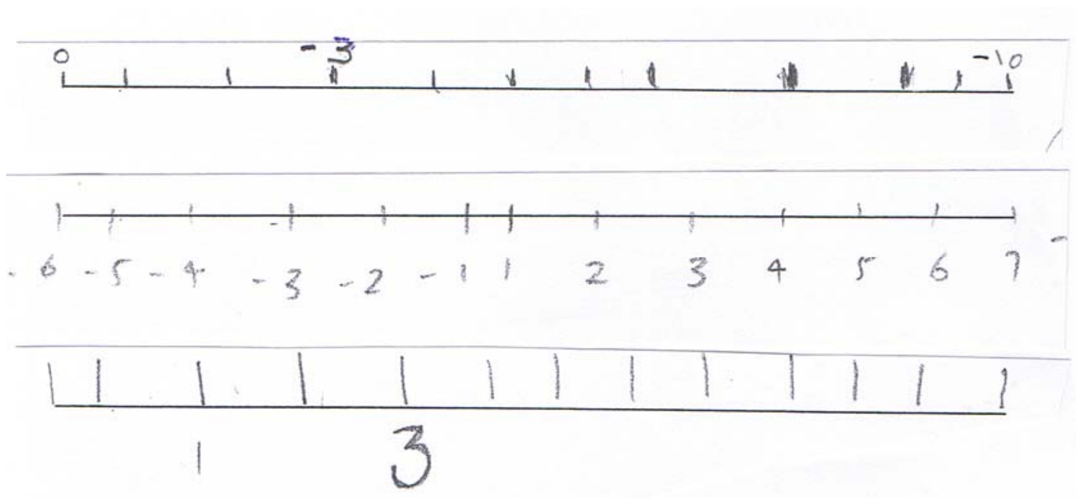


Figure 10.10. Student work samples for A8

Implications for teaching

Of those marking and labelling a single number for A8, 19/32 (59%), also completed A1 by marking only the numbers requested. This suggests that some students may not know the convention of showing at least two numbers so the scale can be determined. They may also not be aware that zero tends to be used as the second referent. This is another convention of scale that needs to be addressed in teaching.

It seems sensible that the concept of an integer could be introduced through practical work on temperature, which is a context that students may meet regularly if watching the weather. The use of the thermometer (with a unit scale) as part of this work would highlight both the importance of zero and where the negatives can be found in relation to the natural numbers. The integer number line could be introduced after this work.

In the next section, the effectiveness of the written test for measuring understanding of scale is explored by example. This is done by presenting in depth the results of two randomly selected students. The goal is to identify whether or not the written test can provide useful diagnostic information about what the student understands and needs to learn.

10.5 Effectiveness of the written test as a diagnostic

10.5.1 Student 41

Aspect 1: Conventions of the number line

> The student's response to A1 shows numbers getting bigger to the right, and being used as marks. Spacing between numbers is not even but some space is left for those not shown.

This suggests that Student 41 needs to improve her understanding of number line conventions. Using marks to locate numbers would probably improve her spacing. Practice should be given on dividing up an interval or a line into equal parts – along with discussing strategies for ensuring those parts are equal. This includes using marks to show where any 'missing' numbers should go.

> For the Bar graph (B5) no scale was shown on the horizontal axis, and no spaces left for the 'missing numbers'. For the vertical scale, 1, 2, 3, 4, 8, 11 were spaced proportionately, with space left for zero.

Student 41 seems to understand that the vertical axis of a graph is a scale, responding consistently with her A1 number line. However, the horizontal axis seems to display categories. The integrity of the coordinates was also not maintained, with the coordinates (2, 2)(3, 3)(4, 4)(5, 5)(7, 7)(9, 9) plotted. This suggests that work on algebraic point plotting using ordered pairs and tables of values, and on creating scales for axes within that context, could be beneficial before returning to statistical applications of scales.

Aspect 2: Multiples scales:

> A2, A3, A4: all correct. Contextual items: two out of three correct, with the idiosyncratic answer '5' for B2a. (This item could be revisited with the student as it is inconsistent with her number line response on a similar item.)

Student 41 seems to have mastered scales involving multiples, but the strategies she used to answer these items cannot be ascertained. According to the table of strategies (Table 8.11), it is possible she used low-level strategies that only have limited effectiveness.

Aspect 3: Fractional and decimal scales

> A5a, A5b, A6 answered 3, 3, and 3¼ respectively. B3a, B3b, B3c answered 0.8, 2.2, and 3.5. All incorrect.

Responses for Section A suggest that Student 41 has little understanding of the measure sub-construct of fractions, and confuses the symbols for proper fractions and mixed numbers. This analysis is supported by her NUMPA results. Her fraction stage of 2-4 suggests that one of her next learning steps is that fractions can be found between the whole numbers, and beyond 1. Responses from Section B are all inconsistent with Section A, although the use of decimals for B3a and B3b indicates that she may believe that decimals are found between the whole numbers on measurement scales. The numbers 0.8 and 2.2 can be arrived at through a marks and spaces error, counting the marks at both ends of a unit, or from ‘skip counting in 0.2s’.

Decimal problems A7 and B4. Both correct.

These results support the idea that Student 41 has some understanding about decimals being located between whole numbers.

Aspect 4: Integer scales

For A7 a single mark was put on the line.

As this item is intended for students with greater number knowledge, the lack of success was expected; however, the convention of showing at least two numbers on a scale would be worth introducing, initially in the context of whole numbers.

10.5.2 Student 61

Aspect 1: Conventions of the number line

The student’s response to A1 shows numbers getting bigger to the right. Marks are used, with numbers placed on these. The spacing used is even with an appropriate amount of space left for the ‘missing numbers’.

This evidence suggests Student 61 has mastered the conventions of the number line.

The bar graph for B5 was not completed. No horizontal scale was provided. However, a vertical scale of 10, 20, 30, with space for zero, was provided. Numbers were placed on the marks.

Student 61 seems to have some understanding that the vertical axis of a bar graph should be a scale, even though he produced an inappropriate one as the numbers to locate did not exceed 11. However, he seems unsure of how to deal with discrete numeric data for the horizontal axis. This suggests that Student 61 may benefit from working on ordered pairs and tables of values in the context of algebraic graphs, but could also benefit from direct work on data tables within a statistical context.

Aspect 2: Multiples scales:

All items correct.

The evidence suggests that Student 61 has mastery of multiples scales, although the strategies used to answer these items cannot be ascertained.

Aspect 3: Fractional and decimal scales

A5a, A5b, A6: all correct – the first two being answered with fractions. B3a, B3b, B3c answered ¾, 4/9 and 3½, only the first being correct.

The evidence suggests that Student 61 has some understanding of fractions, including improper fractions, but this is not consistent from situation to situation. The solution 4/9 can be obtained by inverting the correct answer of 9 quarters, while the response 3½ for B3c suggests that the distinction between proper and improper fractions is not fully understood. Again the strategies and understanding used to create these answers cannot be ascertained, so a teacher would need to explore these before identifying whether the student's strength with number lines could be used to improve measurement scale understanding.

Decimal scales A7 and B4 answered 0.9 and 0.5 respectively. Both incorrect.

The answer 0.9 can be obtained by using the strategy 'a little bit less' to locate 0.4 – but working from one instead of 0.5. This suggests that some work identifying (and labelling) the value of a mark in a partitioned interval may be worthwhile. The response 0.5 could relate to accuracy of placement. Again, identifying what the student has actually done is important before identifying what action is needed to address the issues.

Aspect 4: Integer scales

For A7 the number 3 was shown and marked, located to the right of 0.

This suggests that Student 61 was not familiar with negative numbers, but understood the need for some form of comparison (or interval) to be shown on the number line to give a sense of the scale.

10.5.3 Discussion

In their written form, A1, A8, and B5 allow rich analyses to be undertaken, whereas the interviews at School 1 did not provide much insight into what students were doing. For each question, a number of aspects can be analysed, with the responses indicating an awareness of, and an ability to use, scale conventions. Conventions can be classed as knowledge – things we either know and can use, or do not know or cannot use, so written evidence can identify whether we have this knowledge or skill.

For the other items where students had incorrect answers, the analysis of errors undertaken in Chapter 8 suggests possible causes and next learning steps, but these are not definitive in that in some cases different errors, or a combination of errors, may lead to a particular result. Where students were correct, no indication of how the answers were reached was gained. The interviews suggested working with scale is strategy based and could be developmental in nature. They identified that a variety of strategies of varying degrees of sophistication can be used to successfully answer many items. Thus the fact a student was correct does not provide adequate information to the teacher looking for the next learning step to develop student thinking. In this sense, for diagnostic purposes the written format is inferior to the interview format for the majority of the items.

10.6 Summary of findings

10.6.1 The diagnostic items

Most of the diagnostic items provided good correlation between the results for an individual item and the total for the test. The notable exception was the item where the ability to iterate a unit accurately was being measured, which solely relied on spatial reasoning. The overall results suggest that most items discriminated appropriately between students of different levels of scale understanding. In general, the contextual items also tended to have higher correlations than similar number line items; however, in only one case could it be said that a contextual item had a higher success rate. Since the contextual items generally had lower facility indices (closer to 0.5), the potential for discrimination was greater.

The analysis of two student test papers indicated that the written test could provide some useful diagnostic information about an individual's scale understanding. The items related to the use of scale conventions were the most effective in this format, but many of the other items required the use of mental strategies which the provided form of question did not probe. However, the interviews at School 1 did allow access to student thinking, and a different form of written question may be able to do the same, as indicated by the teacher trials in Chapter 9.

10.6.2 Student results

Most of the students at School 2 did not meet the (implied) curriculum expectation of having the skills to work with scales in a variety of forms by Year 7. If a score of 75% was taken as an indication that a student understood the scale concepts measured by the test, only 14/81 (17%) of students reached this level of attainment. Up to 35% did not correctly answer problems involving whole numbers on scales, and between 7 and 78% showed they did not know, or could not apply the various scale conventions to a number line or bar graph (see Tables 10.2a & 10.2b earlier). For the fractional items, no more than 22% correctly answered any item, while less than 50% correctly answered the decimal items – which included a ruler. No student was able to produce an appropriate graph for discrete numeric data; only five students used appropriate scales and coordinates and all five had their bars touching – a convention appropriate for continuous or grouped data. The scales in graphs could also not be considered in isolation from the students' apparent understanding of graphing.

Errors noted in the literature and the interviews were also found in the work of the students. This suggests that the errors identified at School 1 could be the result of common misunderstandings, so the explanations given during the interviews may be indicative of the form of understanding that leads to such errors. A few additional errors were identified, but the logic behind these could not always be ascertained.

10.6.3 Comparison of results for number line and contextual items

Table 10.10 shows the facilities for the pairs of similar items. For five of the pairs (shown in blue), students had a higher success rate for the number line item. In one of these cases (A1a & B5) the difference may be due to the impact of the context. For A2b and B2a, it may be due to having one scale in sixes and the other in eights; while for A7 and B4 the difference may be due to the interval being partially marked in A7. In only one case (shown in red), when locating ¾ on an unmarked measurement scale, did students have a notably higher success rate for the contextual item. Overall, there is little evidence to suggest that the use of contexts for scale allowed the students to access informal learning or prior knowledge.

Table 10.10: Success rates for number line and comparable contextual items

Number line item	Success rate	Contextual item	Success rate
A1a: numbers get bigger to the right.	93	B5: numbers get bigger to the right.	51
A1c: numbers placed on the marks.	93	B5d: numbers placed on the marks.	58
A1e: space left for 4, 6, 7, 9.	72	B5: space left for 6 & 8.	26
A2b: extending a scale in sixes	77	B2a: extending a scale in eights.	63
A3: reading between the marks	65	B2b: reading between the marks.	67
A4: reading a number between labelled marks (number line marked in twos).	68	B1 reading a number between labelled marks (thermometer marked in twos).	69
A5a: reading ¾ from a marked number line.	33	B3a: reading ¾kg from a marked weighing scale.	40
A5b: reading 1¼ from a marked number line.	28	B3b: reading 2¼kg from a marked weighing scale.	35
A6: locating ¾ on an unmarked number line.	11	B3c: locating ¾kg on an unmarked weighing scale.	36
A7: locating 0.4 on a partially marked scale.	47	B4: locating 0.4 on an unmarked ruler.	28

Pairs of items were also analysed for patterns amongst correct and incorrect responses (e.g., Table 10.7). In general (as in that table) most students either correctly or incorrectly answered both items. However, with the whole number and decimal items there was a tendency for more students to answer the number line rather than the contextual item correctly. This tendency was reversed for the fractional items.

10.6.4 Links to number understanding

The comparison of students' scale responses to NUMPA results indicated a link between scale understanding and number understanding. With one exception, the proportion of students successful with scale-related problems tended to increase as multiplicative or fractional stages increased. The exception related to the intended fractional items. Students with a range of stages were able to successfully answer these items using fractions (although students with better fraction strategies were not more likely to use fractions). However, only students with more sophisticated multiplicative strategies tended to be able to correctly answer with decimals.

10.6.5 Implications for teaching

The results detailed in this chapter show that there are particular gaps in scale understanding amongst the students at School 2. To address these, the development and use of a number of learning experiences are suggested.

1) Activities to help develop students' understanding of the scale conventions. These could focus on the conventions of a number line, then compare measurement scales and most graph axes to what has been learned about the number line. This approach would allow an understanding of the conventions to develop before students need to apply that understanding in measurement and graphing situations.
2) Activities that focus students on identifying how many pieces each interval has been divided into when working with a marked interval.
3) Activities that provide practice with subdividing an interval into *n* equal parts, and in iterating a unit along the length of a line. Developing these skills should help improve accuracy of placement.
4) Activities that provide practice working with scales that show multiples, and that develop the strategies needed to read and locate numbers on such scales.
5) Activities that develop an understanding of 'the continuity of the number line'; that fractions and decimals can be found between the whole numbers, and which fractions and decimals are found between which whole numbers.
6) Activities that focus on the concepts underlying graphing. For example, that graphing often involves plotting pairs of numbers, that the order in which these numbers are placed on the graph is important, and that these number pairs cannot be split. Getting students to initially work with graphs in the domain of algebra would allow these understandings to become cemented before the particular issues of statistical contexts are introduced.
7) Activities that build an understanding that the axes of most common statistical graphs are scales, and conform to scale conventions.

In the next chapter, the development of the items for Test 2 is discussed.

Chapter 11

Further item development for the cognitive interviews

11.1 Introduction

The results for Test 1 indicate that in spite of careful design, the items did not fully cover the construct. This was partly due to unexpected results; for example, the integer item did not provide the expected information. The test results also raised a number of questions. For example, some students' use of decimals suggested a procedural response based on little understanding, so how would they cope when given a decimal for which this approach was problematic? This chapter outlines the development of a second test that builds on what has been learned to date. In doing this some of the items move beyond what is currently found in the literature. Also, because the items in Test 1 identified that students were generally not more successful with problems set in a context, this issue does not need to be explored further in depth (although some item pairs are used). This allows greater flexibility for developing items and enables other issues to be explored. For example, some items in Test 2 have been designed to test the hierarchy of strategies in Chapter 8. Item sets have also been developed to consider:

- Student strategies with fractional scales involving thirds, quarters, and fifths;
- Differences in responses when working with marked and unmarked intervals;
- The impact of changing number size;
- Whether students transfer strategies used with whole numbers to similar problems using other number sets;
- Whether students use the same strategies when reading and locating a number;
- Student understanding of the conventions for plotting a point on a Cartesian system; and
- Issues when non-unit scales need to be created for a graph.

Test 2 uses a new test structure. Number line and contextual items are no longer separated into A and B sections, instead they are mixed throughout the test. To ensure that the item numbering conveys similar information to Test 1 (but cannot be confused with Test 1 numbering) number line items are labelled with an N (so N4 is item 4, and is based on a number line) while contextual items are labelled with a C. Items relating to an aspect were also not located together. One finding from Test 1 was that small changes to a problem that an effective user of scale would ignore may cause some students to change their solution strategy. To reduce the impact that grouping similar items may have on how some students attempt to solve each question, items relating to an aspect are now placed throughout the test. The following sections introduce the developed items, aspect by aspect.

11.2 Aspect 1: Scale conventions

Table 11.1 summarises the items for this aspect. The first column gives the item number, the second column identifies the relationship of the items to those in Test 1, the third described any changes that have been made to an existing item and the final column provides a brief comment.

Table 11.1: Aspect 1 items

Item number	Relationship to Test 1 items	Changes	Comments
N1	Same item (A1).	None.	
Not used	B5.	NA.	The bar graph (B5) is not included in Test 2.

Since both the number line and bar graph items provided rich information about student understanding of scale (in both the cognitive interviews at School 1 and the written test at School 2) these items needed no further development or exploration of student responses. However, item N1 has been retained to allow some baseline comparisons between schools.

11.3 Aspect 2: Scales involving multiples

At School 2, approximately one third of students were unable to identify 12 on a number line marked in twos, an unexpectedly high figure given the age of the students and the results from School 1. In addition, the comparison of results to NUMPA stages (Ministry of Education, 2005a) suggested some link exists between multiplication understanding and scale understanding. Some unexpected factors may also impact on student approaches, for example the size of the interval.

Table 11.2 shows the equivalent information for Aspect 2 to that presented in Table 11.1 for Aspect 1. While a number of items from Test 1 have been retained or adapted, new items have been developed to further explore scale understanding related to whole number multiples. The new items are discussed in more detail below.

Table 11.2: Aspect 2 items

Item number	Relationship to Test 1 items	Changes	Comments
N2	Similar to A2b.	The number to be read has been changed.	Should not make a difference to responses.
N3	Similar to A3.	Part b added to item.	b should require a different strategy.
C4	New item.		Links to A4 and B1 – smaller interval.
N5	New item.		Unmarked interval.
N6	New item.		Marked interval (partitioned into four).
N10	Item A4.	Unchanged.	Provides baseline data for comparison.
N17	New item.		Unmarked interval.
C18	New item.		Identifying half-way.
C20	New item.	Variation of C4.	Locating the number instead of reading a number.
N22	New item.		Unmarked interval.
N25	Similar to A4.	Size of numbers increased.	

11.3.1 Marked scales

Items C4 and C20: Reading and locating a number on a scale in twos

On its own, item 4 (Figure 11.1) is likely to add little new knowledge about student understanding of scale as an item in Test 1 also had scales marked in twos. However, C4 and C20 (Appendix 1) are an item pair developed to explore the impact that a minor change (one which would be ignored by an experienced user of scale) has on student responses. In C4, students are asked to read a number from a scale. In C20 they are asked to locate that same number on the identical scale. The items are separated within the test to reduce cross-over of knowledge due to proximity. What impact will this change have? C4 also pairs with N10, which is the number line item A4 from Test 1 (see Appendix 1). Here it can be used to examine the impact that changing the interval size may have on students' responses. A third pairing for C4 is with C24 (see Figure 11.10 later). Both of these items depict weighing scales with small intervals and five partitions; one involving whole numbers and the other decimals. This pair can be used to consider strategy transfer from whole numbers to decimals.

4) Tom is weighing himself on a set of bathroom scales. How heavy is Tom?
(The scales measure people's weight in kilograms.)

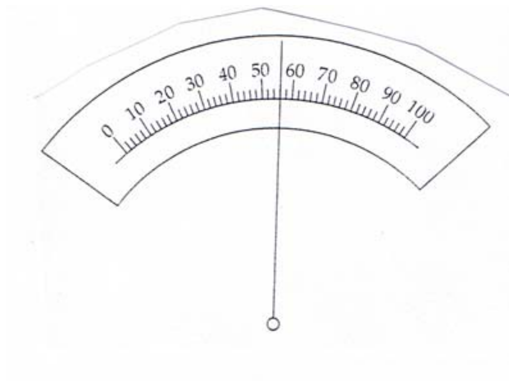


Figure 11.1. Test 2, C4

Item N25: a line labelled in hundreds with marks for twenties

This item is a development of A4. Instead of a scale in tens with marks for twos, a zero was added to the numbers on the line (Figure 11.2). A second change was to ask students to locate a number rather than identify the location of an arrow. Would these changes increase the level of difficulty or change students' error patterns? The original item was retained in the test to provide a comparison (N10).

25) Draw an arrow to show where 120 is

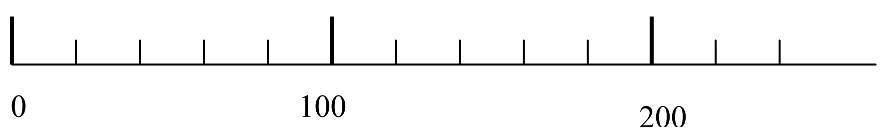


Figure 11.2. Test 2, N25

Item N6: Read an arrow on a line marked in 25s

Item N6 (Figure 11.3) provides a further exploration of the impact of number size, using the same numbers as N25, but with four intervals not five. Would this have an impact on student strategies or results?

6) What number is the arrow pointing to?

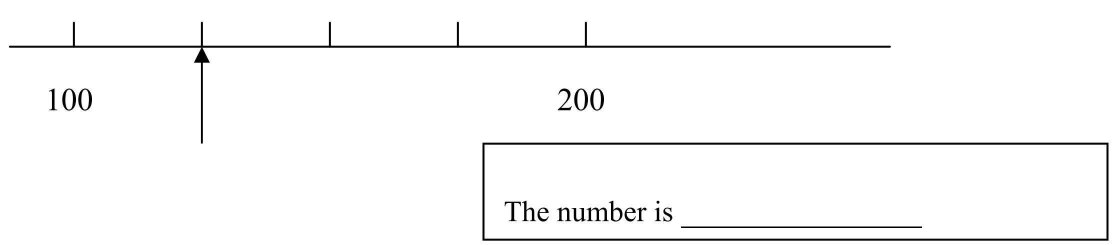


Figure 11.3. Test 2, N6

11.3.2 Unmarked scales

Item N3: working between the marks

In Test 1, A3 allowed students to use the strategy ‘a little bit less’ to locate a number on an unmarked number line. What would students do if the number was not “right next to” a marked number? For Test 2, an extra item was added to N3 so students had to locate not only 23 on a scale in sixes, but also 14.

Item C18: Half-way

This item was introduced to verify the categorising of halving as a counting strategy. If indeed halving is a low-level strategy, almost all students should be able to use it to answer this question, in spite of the number size (Figure 11.4).

18) On the dashboard of a car, there are two large dials, a speedo and a rev counter. Here is a rev counter. How many revs is it showing?

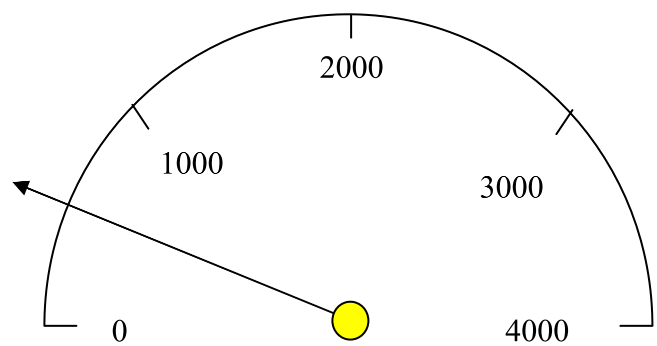


Figure 11.4. Test 2, C18

Item N5: locating a number on a fragment of an unmarked line

This item was developed to further investigate students’ responses to unmarked intervals, and aligns with A3. A3 can be answered using ‘a little bit less’, while this item, according to the teacher trial, can be answered using a variety of strategies (for example, partitioning the interval into five equal pieces, or halving then partitioning the remaining interval into five). All of the strategies used by teachers on an item similar to this were more sophisticated than ‘a little bit less’. The item aims to provide evidence to support or further develop the hierarchy of strategies.

5) Put a cross where 52 is on this number line

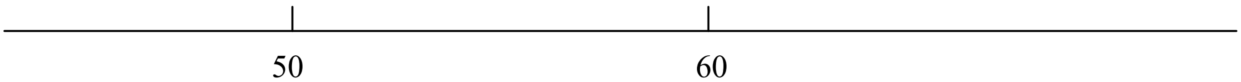


Figure 11.5. Test 2, N5

Item N17: locating a number on an unmarked line

N17 was also developed to help verify or challenge the hierarchy of strategies. It is similar to N5 so could also be used to consider whether or not students retain strategies when working on mathematically similar problems.

17) Put a cross to show where 130 is on this number line

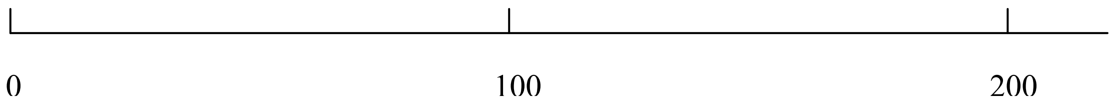


Figure 11.6. Test 2, N17

Item N22: Halve and halve again

This is the most complex whole number item developed (Figure 11.7). The simplest strategy to locate 625 is to find half way between 500 and 1000 (750), and then half way between 500 and 750. However, only students with a good understanding of number are likely to recognise that 625 is a quarter of the way from 500 to 1000. What approaches will students who do not have this understanding of number take?

22) Draw an arrow to show where 625 can be found on this number line

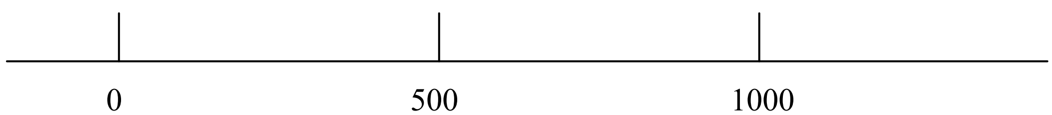


Figure 11.7. Test 2, N22

11.4 Aspect 3: Scales involving fractions and decimals

In Test 1 a number of items involved scales with fractions and decimals. However, these items had limited variability, and generally had low success rates. Results from Schools 1 and 2 also indicated that some students appeared to use decimals in a procedural manner. In Test 2 it is therefore appropriate to use a broader set of items to explore how students approach fractions other than quarters, and decimals that do not involve tenths. Table 11.3 is similar to Tables 11.1 and 11.2, providing details of the items for this aspect; subsequent sub-sections introduce each new or adapted item in greater depth.

Table.11.3: Aspect 3 items

Item number	Relationship to Test 1 items	Changes	Comments
C9	New item.		Decimal (scale in tenths).
C11	New item.		Scale marked in fifths.
N12	Similar to A7.	Decimal to be located changed. Part B added.	Partially marked scale. Changes may invoke different strategies.
N13	Item A5.	Unchanged.	Provides comparison for items in thirds and fifths. Pairs with N21 and N26.
C14	Item B3c.	Unchanged.	Provides baseline data for comparison. Unmarked interval.
N15	Similar to A6.	Fraction changed to thirds.	Marked interval.
C19	Similar to B4.	Decimal to be located changed.	Same strategy should be invoked. Unmarked interval.
N21	Similar to A5.	Fractions changed to fifths.	Marked interval.
N23	New item.		Unmarked interval, reading between units.
C24	New item.		Scale marked in fifths.
N26.	Similar to A5.	Fractions changed to thirds.	Marked interval.
C28	Similar to B3c	Reading rather than locating.	Unmarked interval. Pairs with C14.
N29	New item.		Scale marked in twentieths.

11.4.1 Marked intervals

Items N13, N21 and N26: number lines with quarters, fifths and thirds

For Test 2, A5 from Test 1 was developed into three similar items, the difference being the number of intervals into which the unit was partitioned (e.g., see Figure 11.8). Would students treat these similar items with a ‘bits and ths’ approach, which can successfully be used to answer all three, or would they do something different?

21) What missing numbers go in the boxes?

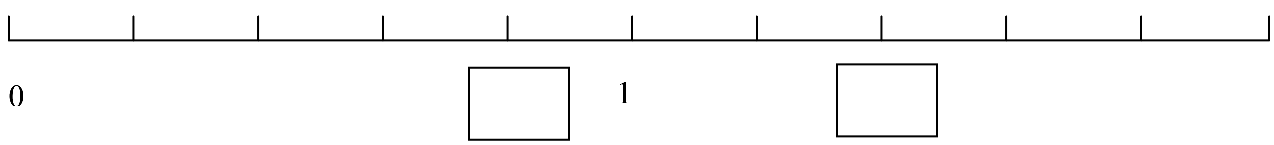


Figure 11.8. Test 2, Item N21

Item C11: the thermometer

This thermometer was introduced to explore further student understanding and use of decimals and fractions in a context where their use was appropriate. The context was chosen as it was felt that many students should be familiar with having their temperature taken when ill.

11) Here is a thermometer. What temperature does it show?

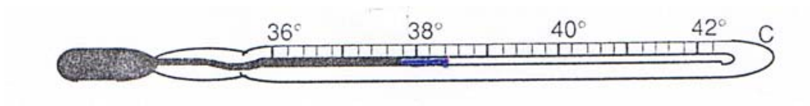


Figure 11.9. Test 2, Item C11

Item C24: a weighing scale marked in fifths

Similar to C11, this item uses a context where the use of fractions or decimals is appropriate. Here the interval size is slightly smaller, but every whole number is labelled. What impact will these minor changes have on student strategies and results?

24) Jane is weighing a stone for a science project. How heavy is the stone?
(The scales are in kilograms.)

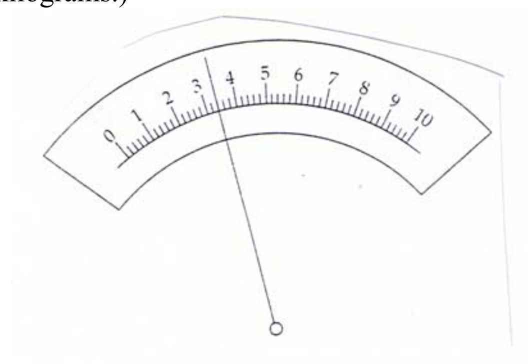


Figure 11.10. Test 2, Item C24

Item N29: the number line in twentieths

This item was added to explore further the responses of students when working between whole numbers. The style of the item is similar to one used by Swan (1983) who found that when answering such items, some students started at the labelled number to the left of the arrow, then counted along the scale ignoring other labelled numbers and the size of the interval. Swan identified that, if given three similar number lines, marked in fifths, tenths, and twentieths, such students were comfortable labelling the same marked point with a variety of different numbers.

29) What number is the arrow pointing to?

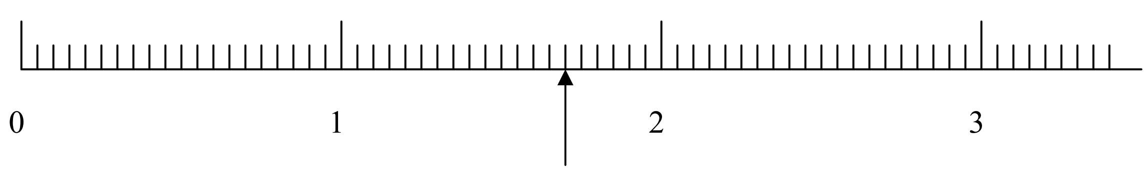
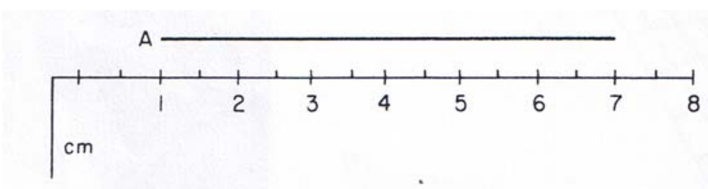


Figure 11.11. Test 2, Item N29

Item C9: the broken ruler problem

The broken ruler problem is an item which has been used often to assess understanding of linear measurement. It was not used in Test 1 as a simpler ruler-based task was given. However, Test 1 results suggested that some students bring their understanding of counting to scales. A ‘broken ruler’ item was therefore added as it provides an opportunity to measure the proportion of students for whom this was an issue. It also provides an indication of whether or not students understand that measurement involves intervals.

The broken ruler problem occurs in two basic forms. In one, someone is given a broken ruler that does not start from zero, and is asked to measure something. In the other, a drawing of a ruler is lined up against a rod that does not start at zero, and the person is asked how long it is. Crooks and Flockton (2002) have two versions of the task, one for an item presumably shorter than the piece of ruler, another for one presumably longer. On the first, 64% of Year 8s were successful (within half a centimetre) while on the second 55% were successful. The CSMS team also had a version (Figure 11.12).



How long is the line in centimetres (cm)?

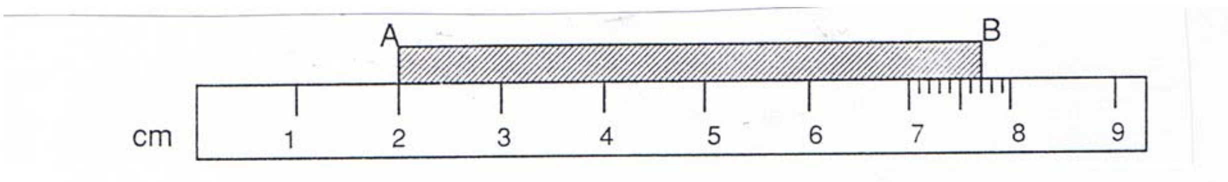
Figure 11.12. CSMS measurement item (Hart, 1978, p. 40)

Hart (1978) reported that 46% of Year 1 secondary students responded that the line was 7cm long; this figure dropped to 23% by Year 3.

Bragg and Outhred (2000b) and Irwin and Ell (2002) also used versions of the problem. In Bragg and Outhred’s task a ruler broken off at the 3.5cm mark was given to students to measure a line 9cm long. The scale went from 4cm to 20cm. In Grade 5 only 58% of students were successful. Students who could align a ruler correctly commonly started their measurement from 1.

Irwin and Ell photocopied part of a classroom ruler showing 3 to 7.5cm on one side and 236 to 287mm on the other. Students measured four lengths, two longer and two shorter than the segment. There were no marked differences in the results of 8 and 9-year-olds. For a length of 5.6cm, many students visualised the start of the ruler, while for an 8cm length more students iterated the ruler segment, causing “more confusion in counting lines or spaces” (p. 363). On the shorter lengths, students often simply used the numbers on the ruler “rather than what was missing” (p. 363). Some were observed changing their method from one length to another, which to the authors suggested the students were unsure what to do. Only when measuring the 5.6cm line were more than 50% of a group of students successful, and this was the 8-year-olds. Figure 11.13 shows the 'broken ruler' item for Test 2. Similar to the item used by Lehrer, Jenkins, and Osana (1998), the rod starts at 2. This helps distinguish between students who simply read the end point, and those who count from one.

9) Use the ruler shown below to work out the length of the shaded rod



The length of the rod is ________________

Figure 11.13. Test 2, Item C9

11.4.2 Unmarked intervals

Item N12: the decimal number line

For this test, A7 was altered to see how students would cope when two decimals, 0.3 and 0.21, were to be located, and one of these did not appear to fit the 'counting in points' strategy. It was expected that locating 0.3 could provide a little more challenge than 0.4, for which the strategy 'a little bit less' could be used. This strategy seemed likely to be less accurate for 0.3. The other decimal, 0.21, could not be exactly located on the number line, given the scale. Would students recognise this? Would 0.21 also be recognised as a smaller decimal than 0.3?

12) (a) On the number line below use a cross to show where the number 0.3 is.
(b) Also on the number line below draw an arrow to show where the number 0.21 is

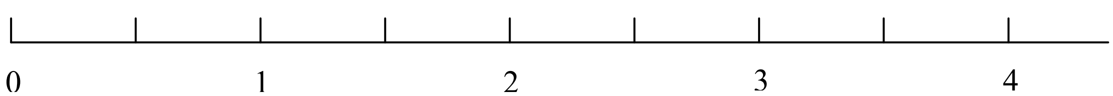


Figure 11.14. Test 2, Item N12

Items C14 and C28: reading and locating

For this pair of items, B3c from Test 1 was used as Item C14. It was then written as a 'scale reading' task for C28 (Figure 11.15). The context of the problem was also changed slightly. The intention was to focus on what impact the change from 'locating' to 'reading' had on responses.

28) Pua is weighing a book as part of a mathematics test. How heavy is the book?

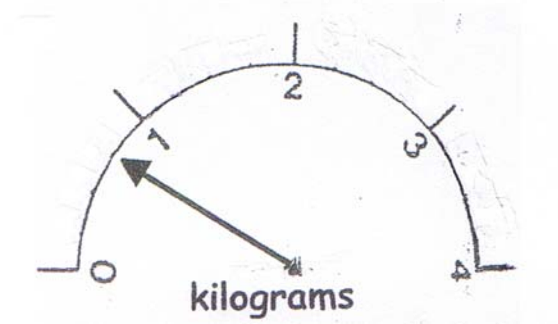


Figure 11.15. Test 2, Item C28

Item N23: what's my number?

This item was developed to explore further how students went about answering problems on unmarked intervals. The location of the arrow was measured to be 1.6, which students locating 1.5 as a working mark could quickly identify. How many would do so? The problem also pairs with C19, in which students need to locate 1.6cm on a ruler.

23) What number is the arrow pointing to?

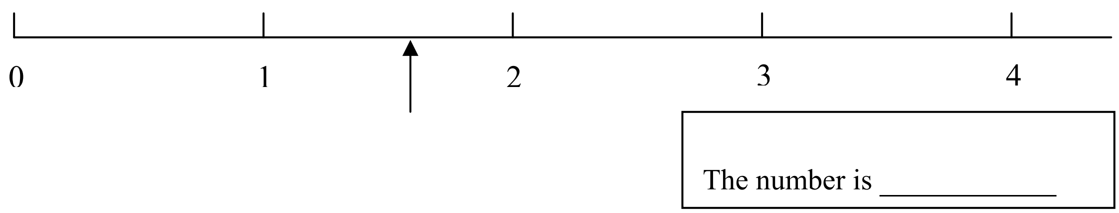


Figure 11.16. Test 2, Item N23

11.5 Aspect 4: Scales involving integers

At School 2, 54% of the students did not recognise the significance of the negative sign when answering A8. Forty percent of students (including some of those who omitted the negative sign) only marked a single number. Item A8 was therefore deemed to be quite difficult. However, changes were made to see if the prompt of a second negative number would result in a clearer indication of the understanding a student had of the relative size of integers. Table 11.4 indicates that again only the one item was used for Aspect 4.

Table 11.4: Aspect 4 items

Item number	Relationship to Test 1 items	Changes	Comments
N16	Similar to A8	Two integers to locate.	Change made to ensure at least two numbers are marked.

Item N16: the integer number line

16) Use the line below to make a number line. On it show where –1 and –4 would go

Figure 11.17. Test 2, Item N16

11.6 Aspect 5: Scale creation and coordinate plotting

While algebraic graphs, creating graph scales, and plotting coordinates were not originally considered to be part of the construct being assessed, the analysis of Test 1 results suggested that students' understanding of scale in graphs could not be separated from their understanding of graphing. A fifth aspect, scales and coordinates, therefore has been added to Test 2. Table 11.5 identifies these items, while more detail of their development is given in the subsequent sub-sections.

Table 11.5: Aspect 5 items

Item number	Relationship to Test 1 items	Changes	Comments
N7	New item.		Plotting a point (unit scale appropriate).
C8	New item.		Creating a scale to show a set of numbers.
N27	New item.		Plotting a point (non-unit scales required).

Item N7: plotting a coordinate on a unit grid

In MiNZC, algebraic graphs were not identified as a common context for scale, particularly at CL3 or below (Ministry of Education, 1992). Consequently items based on algebraic graphs were not considered for Test 1, even though it was clearly a context in which an understanding of scale is important. Given the responses for the bar graph in Test 1, Item N7 was developed to 'fill the gap'. It represents the simplest case in which students create a scale and plot a point. How will they go about answering such a problem? Ward (1979), in his study of the mathematics of ten-year-olds, provides an item involving a simple map laid on a grid. An example of how to locate a place was provided using the coordinate (3, 4), then the locations of two other items were asked. Ward found 45% of students incorrectly answered the first item and 44% the second; 16% and 10% omitted the items respectively. The majority of errors were from swapping the number order in the pair. While the overall performance of children was low in relation to the other test items, these questions were rated the best liked. This suggests that items such as N7 could be easily understood by students in Years 7 and 8, and introducing them even earlier could develop student understanding of the mechanics of using a coordinate grid prior to working with bar graphs and numeric data.

7) Plot the point (3, 4) on the set of axes below. Mark it with an X

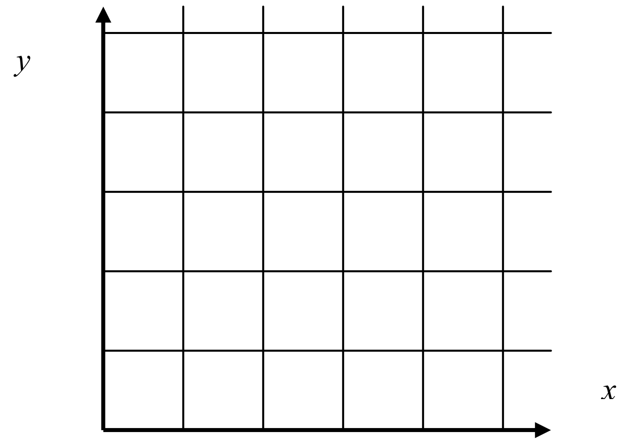
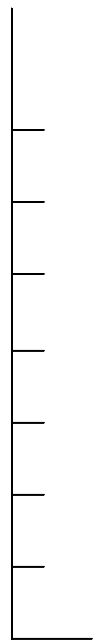


Figure 11.18. Test 2, Item N7

Item C8: creating an appropriate scale

Rangecroft (1994) suggests that the use of a non-unit scale when moving from early graph types to more varied display types is one of the first problems arising from the use of a scale on the vertical axis, using the term *scaling* to mean the development and use of appropriate scales. C8 seeks to explore student understanding of scaling as part of their understanding of scale, assessing whether or not students have come to grips with this concept in their work with graphs. The item structures the process by providing a marked line, and attempts to access existing knowledge by suggesting the line is the vertical axis of a graph. Students are then required to create a (non-unit) scale for the line and locate three numbers on that scale. Students attempting to use a scale in multiples of four will be able to comfortably plot 4 and 8, but not 37, as the line needs to be extended. Meanwhile a scale in fives would mean that 37 would be slightly above the top mark. The numbers have also been chosen so that, depending on the multiples used for the scale, at least one of the numbers needs to be located between those marked.

8) The line below is the vertical axis from a graph. On this axis, the numbers 4, 8, and 37 need to be plotted.



(a) Create a scale that allows you to show all of these numbers
(b) Put the numbers 4, 8, and 37 on your scale. Mark each number with a cross

Figure 11.19. Test 2, Item C8

Item N27, creating non-unit scales and plotting (8, 15)

This last item for the aspect asks students to plot a point on a small grid that will not accommodate unit scales. Furthermore, unless students extend or partition the grid the point cannot easily be located by using identical scales on both axes. Rather a scale in twos on the x axis and a scale in threes on the y axis are the most appropriate. Friel et al. (2001) suggest that students may not be able to choose an appropriate scale for a graph even when they can draw and read a given scale. However, while this item can be cross-referenced to N10 to see how many of the students who can read a scale in twos can successfully complete this question, realistically, because students may not have worked with plotting coordinates on a unit scale within an algebraic context, it is not expected that many students will be successful with this more complex item.

27) The point (8, 15) needs to be shown on this grid. Plot the point (8, 15) and mark it with a cross. Remember to label your scales.

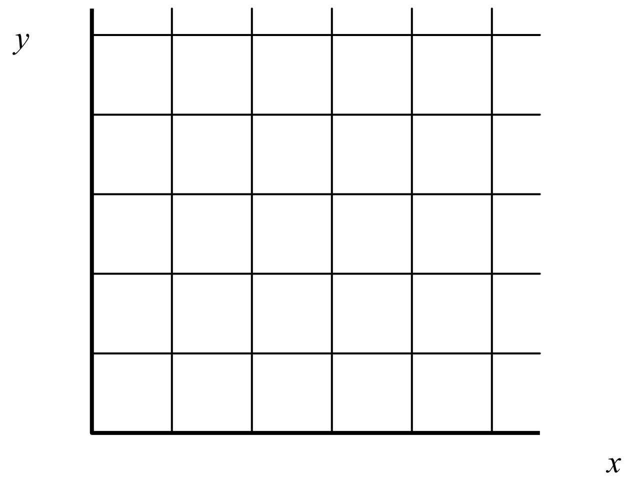


Figure 11.20. Test 2, Item N27

11.7 Summary

Now that the new items have been introduced, the contextual items will be mapped to MiNZC (Ministry of Education, 1992), the concept map developed in Chapter 7 will be revised, and the main features of Test 2 will be summarised.

Table 11.6 shows that most of the contextual questions in Test 2 focus on measurement scales based on CLs 2 and 3. The table requires no further explanation.

Table 11.6: Curriculum demands of the contextual items

Item and context	Curriculum level and strand
C4 – weighing scale.	CL2 measurement AO.
C8 – graph axis.	CL2 statistics AO.
C9 – ruler.	CL2 measurement AO.
C11 – thermometer.	CL3 measurement AO.
C14 – weighing scale.	CL2 measurement AO.
C18 – rev counter.	CL3 measurement AO.
C19 – ruler.	CL2 measurement AO.
C20 – weighing scale.	CL2 measurement AO.
C24 – weighing scale.	CL2 measurement AO.
C28 – weighing scale.	CL2 measurement AO.

Figure 11.21 shows the revised concept map to which the conventions for plotting a coordinate, algebraic graph scales, and the concept of scaling have been added. All boxes are now coloured red, showing that items have been developed to cover all aspect of the construct.

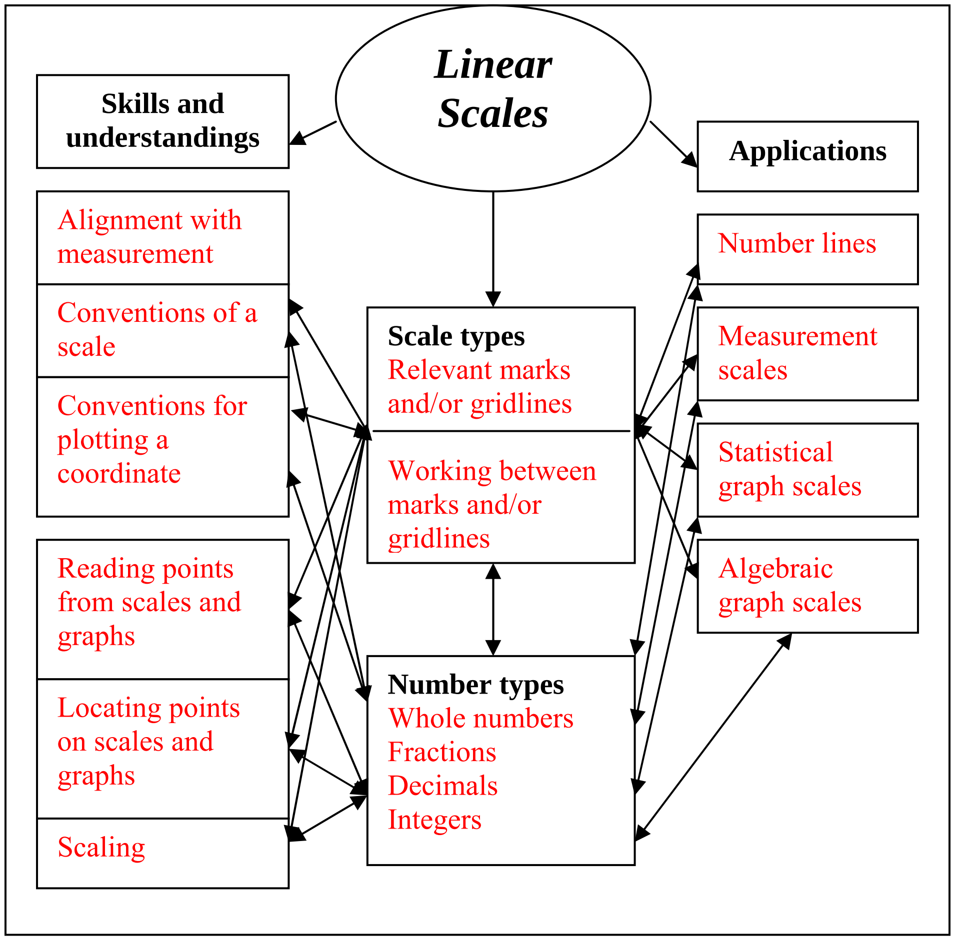


Figure 11.21. Coverage of the concept map by the diagnostic items

Finally, the main features of Test 2 can be summarised.

1) Test 2 is to be used as a cognitive interview; using the protocols developed for the School 1 interviews with Test 1 (see Chapter 6).

2) Number line and contextual items are mixed throughout the test.

3) The items can be sorted into five aspects:

- Conventions of the number line;
- Multiples of whole numbers;
- Fractions and decimals;
- Integers; and
- Scale creation and coordinate plotting.

4) In some cases, the new items were based upon the literature; in others they represent the development of an item from Test 1.

In the next chapter, the findings from the cognitive interview at School 4 are discussed.

Chapter 12

Exploring Student Understanding: Case 4 - The Cognitive Interviews

12.1 Introduction

This chapter reports the 32 interviews undertaken at School 4 to explore unresolved issues related to student understanding of scale. In general, the structure of the chapter follows that of Chapter 10, starting with a summary of the main findings. Section 12.3 then outlines the results of the item analysis while Section 12.4 revisits the strategy hierarchy from Chapter 8, using the responses of the students to the new items to challenge and further develop that hierarchy. The aspect analysis is reported in Section 12.5. This draws upon the revised strategy hierarchy to categorise and report student error patterns, which in turn confirms that the revised hierarchy can explain student responses. The chapter finishes by considering trends across items.

Readers are reminded that the development of the items for Test 2 is outlined in Chapter 11, while a copy of the test can be found in Appendix 2.

12.2 Main findings

1) The majority of the items correlated well with overall test performance. This indicates that, in general, the items in Test 2 discriminate effectively between students of different levels of understanding. The exceptions were those items requiring greater spatial awareness.
2) The levels of correlation found for the majority of the items suggest that the construct contains a coherent body of knowledge, the exception being the spatial awareness required to accurately iterate a unit.
3) The interviews at School 4 provide further evidence for the existence of a range of mental strategies for answering scale-related problems. Responses to the new items led to the reclassification of some strategies and the addition of several previously unidentified strategies (see Section 12.4 and Tables 12.2 and 12.3).
4) The stages in the strategy hierarchy (counting, additive, and multiplicative) can be identified as being based on students' growing understanding of scale. The counting strategies tend to be uncritical in that they may involve the use of a number sequence by students but its relationship to the interval or the marks is not understood. In comparison, the additive strategies involve the critical use of a number sequence, in which the student understands that the sequence must fit the interval. Finally, the multiplicative strategies involve an awareness of the relationship between the number of pieces and the size of each

piece. In this both the size of the interval and the partitioning (potential or provided) is understood.

5) Many of the students interviewed at School 4 seemed not have adequate language to explain what they meant when talking about scale. It follows that an interview which allows follow-up or probing questions to be asked is currently the best way to identify student understanding of scale.

6) Results from the aspect analysis confirm the impact that number understanding has on responses to scale-related problems. Indeed, number understanding can be considered as a compounding factor when attempting to develop student understanding of scale.

7) A number of the items provided evidence of a student's understanding of the symbols used for decimals, fractions, and integers.

8) Interviews at Schools 1 and 4 indicate that many students solved equivalent problems in a context and on a number line by using the same mental strategy. Those successful with scale also tended to use the same strategy to solve items that involved similar scales, even if they involved different number sets.

9) Minor changes to a problem that are not mathematically significant can have a major impact on how some students solve them. This is particularly true of those students using strategies based on counting.

10) Items based on asking students to read a number from a scale can be more effective in identifying student understanding than those asking a student to locate a number.

11) Student understanding is not always easy to ascertain, even in an interview. In particular correct answers can mask 'low-level' thinking, and responses to item pairs may be needed to identify this.

12) The scale interval is fundamentally multiplicative, although the ruler based on a unit scale is additive in nature.

13) The students generally had weak understanding of coordinates and scale creation. Many of the errors with coordinates related to a lack of understanding of the conventions used, while errors related to scale creation may be the result of lack of exposure to such tasks.

14) The evidence from item facilities (and from the detailed analysis of student responses) indicates that the interviewed students cannot be said to have met the curriculum expectations in relation to the items in Test 2; the students met the required standard of evidence (a facility index of 75%) for only three of the 12 contextual items. Analysis of student responses also demonstrates that a correct response to an item does not necessarily indicate an understanding of the scale used in the item.

12.3 Item analysis

This section reports the analysis undertaken on the items in Test 2. Note that as the sample size at School 4 is much smaller than at School 2, the significance of the correlations tend to be lower than those identified in Chapter 10.

Table 12.1: Item analysis, Test 2 (N= 32)

Item	**Facility index**	**Discrimination index (all)** r_{pb}	**significance**
N1a: Numbers get bigger to the right.	97	0.20	0.264
N1b: Using marks.	41	0.18	0.322
N1c: Numbers placed on, or used as, marks.	84	0.41	0.019*
N1d: Equal spacing.	41	0.00	0.999
N1e: Spaces for 4, 6, 7, 9.	72	0.64	0.000*
N2: Extending a scale in sixes.	81	0.49	0.004*
N3a: Unmarked interval. Locating a number close to a mark.	75	0.31	0.080
N3b: Unmarked interval. Locating a number a third of the way.	78	0.14	0.441
C4: Reading a mark between labelled numbers (scale in twos).	59	0.61	0.000*
N5: Locating a number a fifth of the way between marks.	59	0.28	0.127
N6: Reading an unlabelled mark (quarter-way).	69	0.46	0.008*
N7: Plotting a coordinate.	19	0.40	0.025*
C8a: Locating 4 appropriately.	38	0.10	0.598
C8b: Locating 8 appropriately.	34	0.00	0.999
C8c: Locating 37 appropriately.	22	0.32	0.076
C9: Broken ruler problem.	44	0.59	0.000*
N10: Reading a mark between labelled numbers (scale in twos).	84	0.46	0.011*
C11: Reading from a thermometer marked in fifths.	28	0.75	0.000*
N12a: Locating a decimal < 1. Partially marked number line (1 d.p.).	41	0.52	0.002*
N12b: Locating a decimal < 1. Partially marked number line (2 d.p.).	38	0.75	0.000*
N13a: Reading a fraction < 1 (quarters). Marked number line.	41	0.68	0.000*
N13b: Reading a fraction > 1 (quarters). Marked number line.	41	0.65	0.000*
C14: Locating a fraction < 1 (quarters). Unmarked scale.	41	0.52	0.002*
N15: Locating a fraction < 1 (thirds). Unmarked number line.	28	0.58	0.000*
N16: -4 to the left of -1.	41	0.68	0.000*
N17: Locating a number. Unmarked scale – more sophisticated strategy?	56	0.20	0.263
C18: Reading a number from an unmarked scale – half-way.	88	0.55	0.001*
C19: Locating a decimal > 1. Unmarked ruler.	78	0.57	0.001*
C20: Locating a number. Scale marked in twos.	78	0.48	0.006*
N21a: Reading a fraction < 1 (fifths). Marked number line.	41	0.68	0.000*
N21b: Reading a fraction > 1 (fifths). Marked number line.	41	0.68	0.000*
N22: Locating a number quarter way between marks (large number).	48	0.15	0.399
N23: Reading a decimal from an unmarked number line.	69	0.53	0.002*
C24: Reading a decimal from a marked weighing scale (fifths).	63	0.77	0.000*
N25: Locating a number between labelled marks (scale marked in twenties).	72	0.65	0.000*
N26a: Reading a fraction < 1 (thirds). Marked number line.	19	0.65	0.000*
N26b: Reading a fraction > 1 (thirds). Marked number line.	9	0.52	0.002*
N27: Plotting a coordinate appropriately.	13	0.39	0.027*
C28: Reading between whole numbers (unmarked scale).	25	0.50	0.004*
N29: Reading between whole numbers (scale marked in twentieths).	22	0.47	0.007*

* Items with probabilities ≤ 0.05.

The probability of getting a good or even a high discrimination index is related to: the relevance of the item to the overall construct being measured; the quality of the item wording (i.e., it is free of ambiguity); and the item having a facility approaching 50% (because the value of 50% provides the maximum potential for discrimination between students as it breaks the students into two equal groups and potentially maximises the standard deviation). In Test 1, only one item (A1d: the ability to iterate a unit) provided a low discrimination index. It was suggested that this occurred because the item solely measured spatial awareness. For Test 2 a number of items were developed with the intention of further testing the strategy hierarchy. Some of these items involved students locating numbers in unmarked intervals in positions that were not close to either a mark or half-way. Analysis of responses showed that for some students these items did not invoke a 'higher level' strategy, rather the repeated use of a 'low-level' one. Other than N1d (the repeat of A1d), four items (N3b, N5, N17, & N22) thus unintentionally became measures of these students' spatial ability to choose an appropriate 'unit' for a particular interval and iterate it. None of these five items provided high discrimination indices and thus provide further evidence to suggest that these items were assessing something different from the majority of the other items in the test.

12.3.1 Item facilities

In Chapter 5 a standard of evidence of 75% was set to measure whether or not the assessed students had met the curriculum expectation of being able to work with scale in a variety of meaningful contexts. Below are listed the nine items from Test 2 for which this standard of evidence was met (in order of facility index, highest to lowest).

- N1a: Numbers get bigger to the right.
- C18: Reading a number from an unmarked scale – large number (half-way).
- N1c: Numbers placed on, or used as, marks.
- N10: Reading a mark between labelled numbers (scale in twos).
- N2: Extending a scale in sixes.
- N3b: Locating a number a third of the way. Unmarked interval.
- C19: Locating a decimal > 1. Unmarked ruler.
- C20: Locating a number. Scale marked in twos.
- N3a: Locating a number close to a mark. Unmarked interval.

It should be noted that the standard was met for only three contextual items. For one of these items (C20) students met the standard when locating 56kg on the scale, but did not meet the standard when reading 56g from the same scale (C4).

Students at all three schools could: identify that the numbers on a number line should get bigger to the right; identify a number that was 'half-way', whether it was found on a number line or a contextual scale; and continue a number line in sixes. With number lines, students at School 1 could also locate a number close to a labelled number and read from an interval marked in twos. Also with number lines, students at School 2 either used their numbers as marks, or placed numbers on the marks, and could extend the labelling of a scale.

Below is a list of the 23 items which students found difficult, having a facility index of less than 50%. They are placed in order of facility, lowest to highest.

- N26b: Reading a fraction > 1 (thirds). Marked number line.
- N27: Plotting a coordinate appropriately.
- N7: Plotting a coordinate.
- N26a: Reading a fraction < 1 (thirds). Marked number line.
- C8c: Locating 37 appropriately.
- N29: Reading between whole numbers (scale marked in twentieths).
- C28: Reading between whole numbers (unmarked scale).
- N15: Locating a fraction < 1 (thirds). Unmarked number line.
- C8b: Locating 8 appropriately.
- C8a: Locating 4 appropriately.
- N12b: Locating a decimal < 1. Partially marked number line (2 d.p.).
- N1b: Using marks.
- N1d: Equal spacing.
- N12a: Locating a decimal < 1 on a partially marked number line (1 d.p.).
- N13a: Reading a fraction < 1 (quarters). Marked number line.
- N13b: Reading a fraction > 1 (quarters). Marked number line.
- C14: Locating a fraction < 1 (quarters). Unmarked scale.
- N16: –4 to the left of –1.
- N21a: Reading a fraction < 1 (fifths). Marked number line.
- N21b: Reading a fraction > 1 (fifths). Marked number line.
- C9: Broken ruler problem.
- N22: Locating a number quarter way between marks (large number).

This list includes the integer problem, all three of the items written for scale creation and coordinate plotting, all but three of the items developed around fraction and decimal understanding, and six of the 12 analyses relating to contextual items. However, due to the comparatively small number of contextual items used, and the spread of results relating to them, the in-depth analysis undertaken to ascertain what students actually understand and are able to do with scale will be considered before comment is made on whether or not these students can be said to be meeting curriculum expectations in relation to the items asked in Test 2.

12.4 Student strategies: A review

In a test, when different items are used it is common to receive different answers. As new items were asked in Test 2, the hierarchy of strategies was revisited as part of the analysis of the School 4 interviews. This led to the clarification of the differences between some strategies and the identification of some new ones (see Tables 12.2 and 12.3). Two tables are provided as the interviews confirm that students tend to use different strategies on marked and unmarked intervals.

An example of a change made is the redefinition of the strategy 'trial and error' as 'reciting a count'. The interviews showed that a number of students chose a count but did not check it 'filled the interval'. Sometimes the choice of the count for the item was puzzling. For example, when asked to identify the numbers ¾ and 1¼ on a number line (N13), Student B said the numbers were 15 and 25, "because I added these, a 5, 10, 15 and then 20 and 25". When answering N26 Student M did something similar (Figure 12.1). Student responses from School 2 were reconsidered in the light of this new understanding. It seems likely that a number of the errors likewise involved students reciting counts they knew, such as the 2s, 5, or 10s, rather than having a strategy to match the count to the interval. Students who are consciously trying to match the count to the interval, which is usually evidenced by their using a skip count to check the interval is properly filled, or using multiple attempts, are now more clearly classified as 'skip counting'.

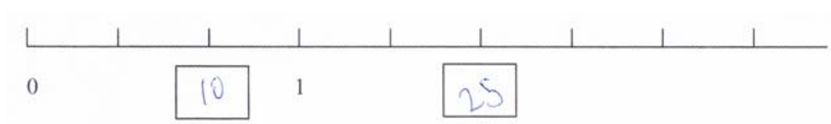


Figure 12.1. Item N26: Student M's response, classified as 'reciting a count'

The revised hierarchies are provided overleaf in Table 12.2, then a summary is made of the new findings as they relate to the classification of strategies.

Table 12.2: Typical student strategies for numbering a marked scale

Thinking type.	**Strategy name.**	**Example of strategy.**	**Useful with…**	**Comments.**
Counting based.	Counting in ones.	Each mark shows one more, as all scales go up in ones…	Unit scales.	Only works for unit scales.
	Counting the space.	Counting on in ones (with fingers) to establish the size of the interval.	Scales in whole numbers multiples.	Some students are not able to maintain the count using their fingers.
	Between.	The number is between 50 and 60 so it goes there.	Not particularly useful.	
	Halving.	The one in the middle is half (or half-way).	Unit scale or scales in multiples with half-way marked.	Being able to halve the numbers being used can limit usefulness.
	Reciting a count.	Students may use trial and error; counting in ones and if that does not work, assume that twos do…	Scales marked in multiples of a number. Can be adapted for decimals.	Only works for simple scales. Tends to be no checking strategy for ensuring the correct count is used. For example, the recited count might not fit the interval – like 30, 60, 90, 1. With integers, negatives may be to the left of one.
	Counting in points.	If there are marks between the (counting) numbers, count 0.1, 0.2, 0.3, … Some students also count back in tenths from the nearest whole number.	Scales marked in tenths.	Only works for number lines with 10 spaces (9 marks) between units. Also need to know 'where to start counting' that is, to place the first number in the count on the first mark past the whole number (marks and spaces errors).
Addition based.	Skip counting.	A development of the 'reciting a count' strategy – using skip counts: 'that's a big gap, lets try tens…'. The recited count fits the interval: 'I tried 20s, 20, 40, 60, but it didn't fit, so I tried 25s and it did' or '3 lines between 100 and 200, 25+25+25+25=100.'	For interpolating and creating a scale. For extrapolating, this only requires a continuation of the scale with the correct 'skip'. Some students can skip count with fractions – thirds, quarters, and fifths.	Being able to skip count in the correct multiple is essential. Some students try a number of counts (trial and error) until they find one that fits. Explanations can show additive reasoning.
	Fitting tenths.	For a scale marked in quarters: 'that would be 0.3, that 0.5 then 0.6 or 0.7 then 1'.	Scales marked in tenths.	Does not work if the number of pieces does not divide neatly into ten, but can give approximate answers.

Table 12.2: Continued

	Cuts and pieces.	Its cut into five, so each is a fifth.	Fractional and decimal scales.	Marks and spaces errors possible.
	Mixed methods.	Halving and halving: '150 is half-way between 100 and 200, halve again to get 125 and 175'. '56 is half-way and a bit more, the third mark'. 'Half-way is five, 6, 7, so its 5.7cm', (using the fact 5+5=10).	Useful for intervals partitioned into four pieces.	Some students ignore existing marks when identifying half-way.
	Converting to whole numbers.	Treating the entire number line as the whole: '¼ is 1, ½ is 2, ¾ is 3 and 1 is 4'. Reunitising tenths as whole numbers: 'I counted in tens and it didn't work so I counted in fives and it did.'	Not useful for fractional scales. Decimal version can work on scales in tenths.	Fractional version simplifies an unmarked scale by changing it to a marked scale.
Multiplication based.	Marks and interval method.	There are 5 marks, the interval is 10, so each mark is 2. 4 pieces, $100 \div 4 = 25$, so first mark is 125. 4 pieces, ¼ of 100 is 25 so ¾ is 75. There are 20 marks, arrow is on the 15th, 15/20 is ¾ (or 0.75).	Any non-unit scale. Also useful for decimals.	Need to know reliably how to count the marks (or intervals).
	Whole number conversion.	An application of the 'marks and interval' method: '4 pieces, ¼ of 100 is 25 so ¾ is 75, so its 0.75'.	Decimal and fractional scales.	Need to be able to find a whole number division that equates to the fraction being converted. Marks and spaces errors possible.
	Treating the fraction as an operator.	Recognise the highest number on the line as the number being partitioned: 6 is ¾ of 8 or 'last number is 6, I want 2/3. As 4/6 is 2/3 the mark goes on 4.'	Not particularly useful.	Misapplication of operator sub-construct to locating a fraction on an unmarked line.
	Times table.	It's the six times table, so the 7th mark is $7 \times 6 = 42$.	Good for labelling a known scale.	

Table 12.3: Typical student strategies for partitioning unmarked intervals

Thinking type.	**Strategy name.**	**Example of strategy.**	**Useful with…**	**Issues and limitations**
Counting based.	A little bit more, a little bit less.	Make a mark 'a bit' to the left or right, so 23 is 'a bit to the left' of 24.	Locating numbers 'just next to' other numbers.	The size of the 'bit' can be arbitrary, causing low accuracy.
	Halving.	'Eying up' where the middle of an interval is using the point of the pen as a marker.	Finding midway between two other points.	Some students are not able to accurately work out where half-way is.
	Between.	The number is between 50 and 60.	Not particularly useful.	
	It's in the middle.	¾ is between 3 and 4, so it goes in the middle.	Not particularly useful.	
	Using bits.	14 is 2 (bits) to the right of 12 on a scale in sixes. 600 is a bit to the right of 500, so 625 is a little bit to the right of 600.	Scales involving multiples. In larger scales the 'bit' can be 10s or 100s, and 'larger and smaller' bits used. Can be used with decimals and fractions.	The bit tends to be an arbitrary amount, and is not used to 'fill' the interval, so there is no checking mechanism for its accuracy.
	Ten in-between.	Between each number it goes in ten.	Locating tenths on unit scales.	The size of the pieces can be arbitrary.
Addition based.	*n* equal spaces.	Draw in marks for each unit, counting along in ones 'to fill in the gap'.	Scales involving whole number multiples. Can be successfully transferred to decimals or fractions.	Reliant on the choice of an appropriately sized interval for the 'unit' and successful iteration of that unit. Often used repeatedly to obtain the correct sized intervals.
	Mixed methods.	½ and ½ again, or combining the use of halving and 'a little bit more, a little bit less'.	Works well dividing an interval into four (quarters). Can be used on most intervals.	
Multiplication based.	Partitioning into 3 or 5 pieces.	Students know how to accurately partition an interval into 2, 3, and 5 pieces.		Requires students to identify how to cut the interval into the correct number of pieces.
	2, 3, 5 method.	Repeated halving for eight (eighths). Locating 11 on a scale using multiples of 6 by finding ½ way, and cutting the remaining interval into 3 equal pieces…	Scales that can be divided into 2, 3, 4, 5, 6, 8, 9, 10, 12…	Relies on understanding division; such as ÷ 2 then ÷ 5 gets the same answer as ÷ 10.

12.4.1 New findings about student strategies

Ten in-between

This strategy can be considered the equivalent of 'counting in points' for the unmarked number line, or as an embryonic version of '*n* equal spaces'. In this approach, students understand that decimals are found between the whole numbers, and when the whole number is divided up you get 'ten of them in-between.' Students at School 4 who predominantly solved items with counting strategies, or a mix of counting with some adding were found to use this, hence its classification as 'counting-based' (although some students with multiplicative strategies used this approach when locating 0.21). These quotes illustrate this form of understanding.

C28 (answering 0.8)
S(L): Because it's ten in-between here and it's not enough to be 9.

S(M): because you know 10 again and then 90 would probably be about there then 80.

S(DD): I think oh zero point like seven and because it's right there and it's got a space towards the one so it would be 0.7.
When revisiting the item at the end of interview
S(DD): Because I thought it was like 1, 2, 3, 4, 5, 6, 7 and there was space like 9, 8, 7.

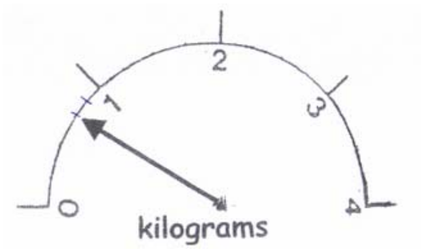


Figure 12.2. Item C28 (reduced): Student M's response

Note that this form of thinking could not be easily identified by looking at the students' written responses. Only the answer was provided in the majority of cases and it was rare to find working marks (Figure 12.2).

At the end of the interview, 27 students were asked why they used a decimal to answer N23 (labelling an arrow between one and two). Six students used explanations showing a procedural understanding of decimals that can be linked to the 'ten in-between' strategy.

S(B): because 1 and then 1.2, 1.3, 1.4…

S(F): … I reckoned it would be in the lines to go up to ten and then go over to the two.

S(M): … because it's one and then you put a decimal dot, and you put another number after it.

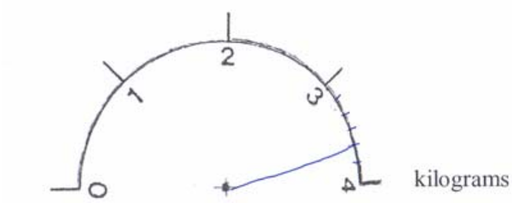
S(FF): Like you only have 1 and as you go up you add 1 at the end of it without putting a dot … it can't be more than 10.

Another 11 answered that it was because the arrow was between 1 and 2. Only five mentioned that the answer could also have been a fraction. These students tended to have the highest NUMPA results.

Some students used 'ten in-between' thinking to answer fraction problems, and in doing so showed they did not really understand this form of number.

N30 (locating ¾ on a number line)
S(W): um three one there, three, two there, that would be three four.
I: Right.
S(W): It's going in a line. [Finger moving to show points going up]
I: Three ones…three twos…three threes…and three fours. So it keeps on going up to three tenths?
S(W): Yep.

Student M shows how 'ten in-between thinking' relates to student understanding of the ruler, and can be complicated by marks and spaces errors.



S(M): … it's always got like, five little things and in between that its got heaps of little milli ones. And so those are the main ones and then **…** so it should be between three and four.

Figure 12.3. Student M locating ¾kg on a weighing scale (C13)

Student L shows how 'ten in-between' thinking can develop into an additive strategy. He was able to solve problems with five, ten, and twenty partitions using doubling and halving, but not others. For example:

C4: Because there's 5 things instead of 10 so one must equal two.
C20: Because you get to the 50 and then there's only 5 places so its half of the six.
N21: There's five here [points to marks between 50 and 60], and there's 10 to the …[60], so I just take a number here [points to 4th line] and double it.

In N6, he identified the interval had been partitioned into four, but was unsure how to answer the problem, saying "five would be easier".

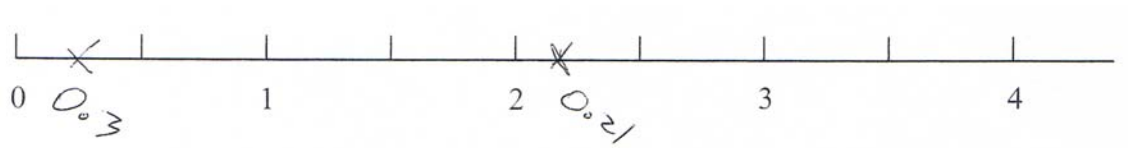


Figure 12.4. Student response to N12 showing ten in-between thinking

Closer to

On first impressions, Student I used a ‘closeness’ strategy not based on partitioning when working with an unmarked number line. For example, locating 52 in N5: “because there’s room for the 51 and enough for the others to go after.” In later items it became clear that Student I could find half-way, and referenced this point. For example locating 1.6 in N23 “1.50 would be round here [points to middle of interval] so that will be just a bit over from it.” This helped interpret her response to C14, where she placed her mark for ¾ a little before 3½: “Because there’d still be room for the 3/5 and the rest of the numbers behind and there’s a gap for the … rest of the numbers before the 3/5.” It seems Student I thinks ten fractions, 3/1, 3/2… exist between each whole number, as happens with decimals. Based on an analysis of all of her responses, it was decided Student I was actually using a form of partitioning based on halving and ‘ten in-between’, so her ‘closer to’ explanation was coded as an additive mixed method.

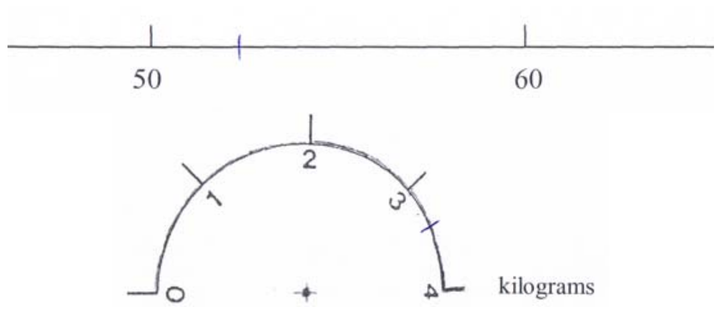


Figure 12.5. Student I’s responses to N5 and C14 (reduced)

Student M, who also appeared to use a ‘closer to’ explanation, likewise showed he was referencing ½, and ‘ten in-between’. Students M (and E) also used a similar approach with an interval of six in N3b (see Section 12.5.2, item N3b).

Bigger to the right

This strategy is neither a partitioning or estimation strategy. It is based on identifying one number, and saying another number is bigger than it, so it goes to the right (by an arbitrary amount). Smaller numbers can also be located to the left. The strategy can be used on marked or unmarked scales, involving multiples or decimals.

N25: (locating 120 on a scale marked in 20s.)
S(U): Because it's 100 and that's one hundred and two, [points to the mark for 120] so one hundred and twenty it's…it's probably about there [mark drawn just before half-way].

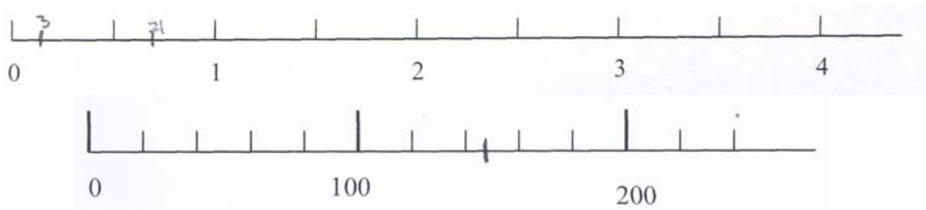


Figure 12.6. Student U's responses to N12 and N25 (reduced) showing 'bigger to the right' thinking

Between

This is a partitioning strategy related to 'bigger to the right' which can be used on both marked and unmarked intervals; it is not open-ended in that the response identifies an interval within which the number lies. In a written problem involving a marked interval this strategy can appear to be the result of a marks and spaces error, and with an unmarked interval, an inaccurate placement. Use of the strategy does not exclude the possibility of obtaining a correct answer. For example, Student U explained "it's 50 and 6 is somewhere in there" when describing how he answered C20.

Figure 12.7. Student U's responses to C20 (reduced) showing 'between' thinking

This strategy was also used by students locating fractions on the number line. For example, when as a supplementary question Student BB was asked to locate ¾ on a number line (Test 1, item A6) he located it at 3½, explaining "so there is the three and there is the four" [points to the numbers 3 and 4 on the number line].

It's in the middle

In this variation of 'between' thinking, students identify that the arrow or the number is between the labelled numbers, and give 'half-way' as their answer. A small number of students at School 2 provided answers that can be attributed to 'in the middle' thinking on eight of the items in Test 1, although there was no apparent logic behind these responses at the time. In Test 2 Student J provides an answer that suggests this form of thinking when answering C4 (see Figure 12.8): "because half of 10 is 5 and half of 5 is 2.5 and it was on the five in the middle." Here the strategy explained bears no relation to the partitioning of the interval, but does suggest the answer is 55 because the arrow is in the middle.

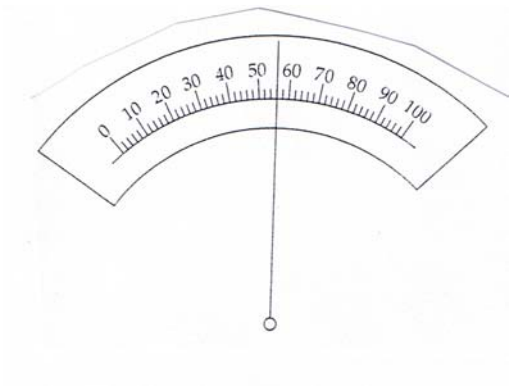


Figure 12.8. The scale from Item C4

Using bits

This strategy involves the repeated use of ‘a little bit more’. It is an inefficient strategy. For example, seven students used this on both N5 and N22. Two incorrectly located both numbers and five incorrectly located one. The strategy was also more commonly used by the mathematically weaker students. It is classified as a counting strategy as limited number knowledge is used, and those using it take an arbitrary amount and make no effort to see if it iterates to fill the interval.

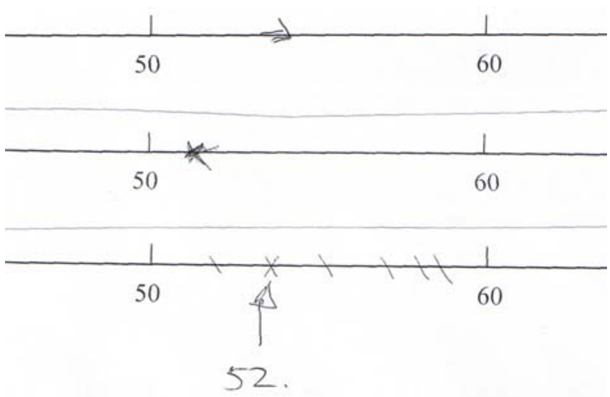


Figure 12.9. ‘Using bits’ for N5 (reduced): the strategy is not always visible in a written response

Whole number conversion

A consistent theme throughout this study has been the use of ‘whole number conversion’ strategies by students. By converting problems to whole numbers, they can answer problems without recourse to fractions or decimals. Students at all three schools have made occasional errors using this strategy, forgetting to reconvert their answer to the situation from which the problem was drawn.

Visualising marks

When dealing with an unmarked interval, rather than using working marks some students *visualised* the interval with the marks; that is, they saw a picture of the interval in their head. This approach was more common amongst students using ‘ten in-between’ thinking, presumably as these students had a mental picture of an interval divided into ten pieces from using a ruler. However, it was clear that some students who visualised did not have an understanding of ‘how many pieces’. For example, when asked to explain his answer for C28, Student II responded that “there’s lots of little lines in between each of those, so that’s about ten of them”. Irwin and Ell (2002) also found some students working with a broken ruler visualised the missing start of the ruler. Visualising has not been categorised as a partitioning strategy as it does not seem to fit this form of thinking.

Imaging marks

Another strategy noted amongst the students at School 4 was to *image* marks; that is mentally partition the interval rather than physically do so. This sometimes involved taking an arbitrary piece and mentally iterating it (as in ‘using bits’) or mentally partitioning the entire interval. Some students were seen to move their finger, nod, or move their eyes when doing this. An example of this was Student CC in finding 625 between 500 and 1000 for N22:

> I: I notice you were sort of hovering your pen around the middle and around this way as well. What was going on there in your head?
> S(CC): It was just 500, 6, 7, 8, 900.

Another example is Student W, explaining his location of ¾ [just before 3½] for C14: “because I tried to figure out where half-way was between the three and the four and um, I sort of put little dashes in my head, like where …1, 2, 3,...”.

12.4.2 Discussion

The strategies classified at the lower levels of the hierarchy have proved to be accessible to more students, and can give some success. In some cases, ‘harder-looking’ problems had high success rates because of this, such as students who could iterate an arbitrary bit to locate 130 on N17. However, accuracy was found to affect such strategies (when dealing with unmarked intervals), while marks and spaces errors impacted on the results for marked intervals (see Section 12.5). A second confounding factor was the use of visualising and imaging strategies.

Visualising and imaging marks caused problems when trying to identify the strategy a student used because there was not necessarily any visible evidence of their use. This is particularly problematic in a written situation, but even in an interview the interviewer needs to be aware of their existence. For example, for N3b, a student could say that they located 14 by finding a little bit less than half-way, but actually they may have identified half-way then imagined the remaining half split into three equal pieces. This could explain the accuracy with which some students could use apparently low-level strategies.

12.5 Aspect analysis

The first section of the aspect analysis takes a similar format to that shown for School 2 in Chapter 10 in that a table reports the results for each item. However, the second part of each table is subtly different because student responses have been ascertained in an interview. This means that the possible strategies used by students do not need to be inferred; instead the incidence of particular errors can be reported.

Comments relating to the results of specific items are provided after this initial section. However, Test 2 was designed to collect data from sets of similar items so trends across items could be further investigated. In some cases these comparisons are reported as part of an individual aspect, but in others the item sets cross aspect boundaries. In these cases the reports of these investigations follow the aspect analysis but often refer back to, or build on, what has been said in this section. Where significant additional comments are made about an item, the section number is noted for the reader.

12.5.1 Aspect 1: Scale conventions

Item A1 was the only item for this aspect, and was retained to provide some baseline data on understanding of scale conventions. The results show great similarity to those of School 2, with the exception of N1b, for which far fewer students at School 4 used marks.

Table 12.4: Aspect 1 responses (N = 32)

Item	**Percentage correct**	**Percentage incorrect**	**(omit)**
N1a: Numbers get bigger to the right.	97	3	(3)
N1b: Using marks.	41	59	(3)
N1c: Numbers placed on, or used as, marks.	84	16	(3)
N1d: Equal spacing.	41	59	(3)
N1e: Spaces for 4, 6, 7, 9.	72	28	(3)
Error pattern		**Percentage**	
Not using marks.		N1b (56)	
Locating numbers in spaces.		N1c (6)	
Interval for the unit not reasonably equal in size.		N1d (56)	
Space not left for the ‘missing numbers’ 4, 6, 7, 9.		N1e (25)	

12.5.2 Aspect 2: Scales involving multiples

For Aspect 2 a wide range of items was used to investigate student understanding, results for which are reported in Table 12.5. Responses to many of these items have also been mapped to the revised strategy hierarchy to provide readers with an indication of the variety in the strategies that the group of students used to solve each problem. The resulting strategy tables (e.g., Tables 12.7 & 12.8) indicate the difficulty a teacher may have when identifying student understanding as they clearly show that students may successfully answer some items using strategies based on little understanding.

Table 12.5: Aspect 2 responses

Item	Percentage correct	Percentage incorrect	(omit)
N2: Extending a scale in sixes.	81	19	(0)
N3a: Locating a number on an unmarked scale (close to a mark).	75	25	(0)
N3b: Locating a number on an unmarked scale (near half-way).	78	22	(0)
C4: Reading a marked scale (twos).	59	41	(0)
N5: Locating a number a fifth of the way between marks.	59	41	(0)
N6: Reading a marked scale (twenty fives).	69	31	(6)
N10: Reading a mark between labelled numbers (scale in twos).	84	16	(6)
N17: Locating a number on an unmarked scale (part-way along).	56	44	(0)
C18: Reading an unmarked scale (half-way).	88	12	(0)
C20: Locating a number on a marked scale (twos).	78	22	(0)
N22: Locating a large number quarter way between marks.	48	52	(3)
N25: Locating a number on a marked scale in twenties.	72	28	(0)

Error pattern	Item number and percentage			
Marks and spaces.	C4 (3)	N6 (6)	N25 (3)	
Forgetting to reference the closest labelled number.	N6 (3)	N25 (0)		
Counting in ones.	C4 (25)	N10 (6)	C20 (3)	N22 (3)
Counting in points.	C4 (3)			
Accuracy (inability to iterate accurately).	N3a (19)	N5 (13)	N17 (13)	N22 (9)
Its in the middle.	N3b (9)	C4 (6)		
Bigger to the right.	C20 (3)	N17 (6)	N22 (3)	
Reciting a count.	C4 (6)	C20 (16)		

Commentary

No more than two students omitted any particular item. Responses to the items have also been useful to confirm the impact that number understanding has on scale. The responses also show that minor changes to a problem – changes that an experienced user of scale would ignore – may lead to some students changing their approach.

Item N2: Extending a scale in sixes

This item provides useful information about student strategies. For example, 26/31 (84%) used skip counting. Three of these students did not know their skips in sixes, so used finger counting to help add on six each time, and one was successful with this. Six of the students identified that the numbers were the six times tables before skip counting in sixes. Of the other five students, two ‘counted in ones’ to identify the size of the gap between numbers, then tried this again to add on each six (one doing so correctly), and three identified and used the six times tables and did not apply a skip count. The strategy of the thirty-second student was not explained.

Overall the success rate for this item is similar to that of School 2; approximately one student in five did not answer correctly. The main source of error was not related to understanding the conventions for continuing the numbering of a line, rather the students’ understanding of number. The results thus provide evidence that number understanding has an impact on success with scale. However, it needs to be noted that N2 cannot be used to reliably predict errors with other marked lines (e.g., see Table 12.6). It shows that of the seven students answering N2 incorrectly, only three answered N6 incorrectly and five answered N25 incorrectly.

Table 12.6: N2 as a predictor of other incorrect responses

N2 incorrect	Also with N6 incorrect	Also with N25 incorrect
6	3	4
1 (through counting in ones).	0	1

Item N3a: Locating a number 'close to' a labelled mark

This is another item from Test 1. The main difference between the results from Schools 2 and 4 is the number of students omitting the item, supporting the theory that some students at School 2 missed the item as it did not have its own number line. The other result to note is that 25 students (78%) at School 4 used the strategy 'a little bit less' instead of more accurate (but more complicated) strategies. Such strategies could be challenged by asking students "how do you know you are right?" This follow-up probe may be useful to ascertain if a student has higher-level strategies that could be used for confirming the accuracy of their response.

Table 12.7: Strategy use for N3a

Strategy	Classification of strategy	Number using the strategy	Number correct
Halve then third or halve then 3 equal pieces.	Multiplicative.	2	2
Sixth.	*n* equal pieces. Additive.	1	1
n equal pieces.	Additive.	4	2
A little bit less.	Counting.	25	19

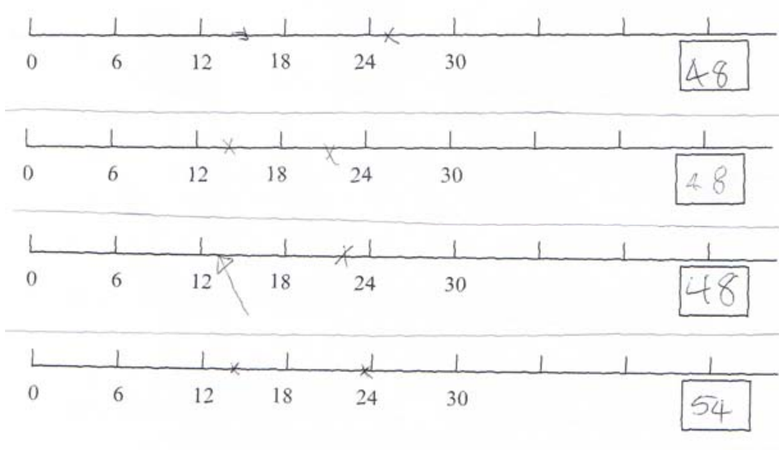


Figure 12.10. Examples of responses to N2 and N3 (reduced)

Item N3b: Locating a number near half-way

When items requiring students to identify numbers part-way through an interval were developed, it was not anticipated that a common method of solution would be to choose an arbitrary unit and iterate it, without checking the unit 'fills the gap'. Such a strategy was certainly not identified amongst those participating in the teacher trials. Many of the students using this strategy were successful with this item (Table 12.8).

Table 12.8: Strategy use for N3b

Strategy	Classification of strategy	Number using the strategy	Number correct
Partitioning into 3 pieces/thirds.	Multiplicative.	3	3
Bit before ½.	Mixed method, additive.	3	3
n equal pieces.	Additive.	4	4
Closer to.	Mixed method, additive.	2	1
Using bits.	Counting.	16	14
It's in the middle.	Counting.	3	0

In Table 12.8, two students have also been identified using a 'closer to' strategy. When probe questions were used to explore the reasoning of Student M, he identified 14 was sort of in the middle, but closer to 12. Student E talked of 14 being two after 12 and four before 18, applying addition and subtraction in his explanation. Both students referenced half-way in their responses to other items. For example, in the broken ruler problem both counted the millimetres by starting at 5 and adding two more on. These responses are part of the evidence base for classifying 'closer to' as a mixed method additive strategy.

Items C4 and C20: Reading and locating

C4 and C20 used the same weighing scale labelled in tens with marks for twos. In C4 students were asked to read 56 from the scale. In C20 they were to locate 56 on it. Student responses to this pair of items were revealing in that nearly half changed strategies when answering the items. Of these, nine were forced to change strategy, as they had used a counting strategy to answer C4 with 53 or 50.3. (Five still incorrectly answered C20, two were accidentally correct, choosing an inappropriate strategy which happened to give the right answer and two correctly recited a count.) Overall, 19/32 (59%) answered C4 correctly, while 25/32 (78%) answered C20 correctly.

Table 12.9: Strategies for answering C4 and C20

Strategy	Classification of strategy	Number using the strategy		Number correct	
		C4	C20	C4	C20
Division (100 ÷ 4).	Marks and interval, multiplicative.	2	4	2	4
5 lines, going up in twos.	Marks and interval, multiplicative.	4	4	4	4
Half and a bit more.	Mixed method, additive.	4	5	4	4
Going up in twos.	Skip counting, additive.	8	5	7	4
Going up in threes.	Skip counting, additive.	0	1	0	1
Made 10 intervals.	*n* equal pieces.	0	1	0	1
There's not ten so its going up in twos.	Reciting a count. Counting.	2	5	1	5
4 lines.	One multiplicative (marks and interval).	0	2	0	1
Counting in ones.	Counting.	8	1	0	0
Counting in points.	Counting.	1	0	0	
It's in the middle.	Counting.	2	0	0	
Bigger to the right.	Counting.	0	1	0	1
A little bit more.	Counting.	0	1	0	0
Unclear.	Unclear.	0	1	0	0

Table 12.9 shows it was the students who were using lower-level strategies that tended to answer incorrectly and to change their approach for the items. The results also suggest that an item asking students to read a number from a scale may be more effective in identifying student understanding than one asking the student to locate a number.

Item N5: Locating 52 on an unmarked scale fragment

This item shows that it is possible for students to successfully use a variety of strategies correctly, although students using strategies classified at the higher stages were more likely to be correct.

Table 12.10: Strategy use for N5

Strategy	Classification of strategy	Number using the strategy	Number correct
½ then 5 equal pieces	2, 3, 5 method. Multiplicative	4	4
Fifth of the way	Multiplicative	2	1
½ and ½, bit to the left	Mixed method, additive.	4	3
½ and ½	Mixed method, additive.	1	1
n equal pieces	Additive	5	3
½ and a little bit less	Mixed method, additive	1	1
Closer to	Mixed method, additive	1	1
Using bits	Counting	12	4
Little bit more	Counting	2	1

Item N6: Reading 125 from a marked scale

Table 12.11 shows the strategies used by students to answer this item, and the way they have been classified in the strategy hierarchy. It shows that students can interchangeably use fractions, division, or multiplication when using a 'marks and interval' approach, as this relates to their number use, not their approach to the scale. It also shows that students with a multiplicative or additive strategy tended to answer the problem correctly. Of those incorrectly skip counting, Student D correctly skipped in 25s, but forgot to reference the closest labelled number (100) answering 25, while Student II made a marks and spaces error, answering 120. During the course of answering this item four students swapped strategies when they identified that their initial strategy was not working. For example, Student X swapped from skip counting in twenties to using an incorrect multiplication as he did not count the marks. Student AA first tried skip counting in twenties, then swapped to 'halve and halve again'. Where students have changed strategies, both have been included in the count in Table 12.11, which is why the total number of strategies listed exceeds 32.

Table 12.11: Strategy use for N6

Strategy	Classification of strategy	Number using the strategy	Number correct
Fractional (quarter of 100 is 25).	Marks and interval, multiplicative.	4	4
Division (100 ÷ 4).	Marks and interval, multiplicative.	2	2
4 gaps/lines so 25s.	Marks and interval, multiplicative.	2	2
Multiplication (4 × 25 = 100, 5 × 20 = 100).	Marks and interval, multiplicative.	1	0
Halve and halve.	Mixed method, additive.	7	7
Skip counting.	Additive.	11	8
Addition (e.g. 25 + 25 = 50, 50 + 50 = 100).	Skip counting, additive.	2	1
Reciting a count.	Counting.	2	0
Guess.	Not applicable.	3	0

Item N10: Reading a scale marked in twos

This item was a repeat of Item A4 from Test 1, and provided both a strategy comparison and baseline data. Eighty four percent of students answered this item correctly at School 4, compared to 68% at School 2. Of the students with incorrect answers, one counted in ones from zero, while another counted in ones from ten. The strategies used mirror those outlined for A4 at School 1 and for C4 from this test, so will not be discussed further here. Twenty five percent of students at School 2 produced answers that could be attributed to counting-based strategies.

By the time students reached this item they had already worked on another item with a marked interval in twos. For more effective users of scale, this tended to have an impact on their strategy explanation, in that it was often shorter, with more understanding assumed. (This was a trend that continued as the students encountered more items to which similar strategies could be applied.) To code such explanations the individual response could not be taken in isolation; rather, responses for earlier items that the student deemed to be similar needed to be taken into account.

Item N17: Locating 130 on an unmarked scale

This item and N10 produced very similar results, so only one of these items is necessary for any diagnostic test on scale understanding.

Table 12.12: Strategy use for N5

Strategy	Classification of strategy	Number using the strategy	Number correct
½ then 5 equal pieces.	2, 3, 5 method. Multiplicative.	6	4
Nearly a third.	Multiplicative.	1	1
½ and ½, bit more.	Mixed method, additive.	3	3
½ and ½.	Mixed method, additive.	1	0
n equal pieces.	Additive.	4	3
½ and using bits.	Mixed method, additive.	5	3
½ and a bit less.	Mixed method, additive.	2	0
½ and closer to.	Mixed method, additive.	1	0
Closer to.	Mixed method, additive.	1	0
Using bits.	Counting.	6	4
Bigger to the right.	Counting.	2	0

Item C18: Half-way

Twenty eight students correctly answered this problem, supporting the classification of finding 'half-way' as a low-level strategy. Four students (with the weakest number knowledge of the 32) answered incorrectly. Three of these students did not use the concept of finding half-way (or a number in the middle) on any other problem. Students BB and HH identified that the number was 100, BB because it was between 0 and 1000, HH "because it's on the middle". (Student HH used the concept of finding half-way in C19). Students B and II talked of making skips in 100 and 50 respectively.

This item identified that some of the students did not understand the concept of half-way. In identifying these students C18 has provided valuable diagnostic information.

Item N22: Locating 625 on an unmarked number line

This item was designed as a test of the impact that number size had on student strategies. The response of Student J illustrates the impact that a larger number like 625 could have. She earlier used 'halve and halve again' on both marked and unmarked intervals, using this strategy on N6 to identify an arrow pointed to 125. For N22 her initial attempt located half-way as 750 (but only after self-correction). Her second attempt involved splitting the interval into five equal pieces (hundreds), then halving and halving the interval between 600 and 700. Clearly Student J did not realise that 625 was half way between 500 and 750.

The item also proved to be useful when identifying the strategy level some students had access to. For example, this was the only item upon which Student T used a multiplicative strategy. For four other students it was one of two items on which they used multiplicative approaches.

Note that while some students may have had a strategy for locating 625, not all were able to use this successfully, as the 'units' they were trying to create were not of a similar size. Also note that no students were successful with this item using counting strategies, although some attempted to differentiate the size of their 'bits', for example using a big bit for 100 and a small bit for 25.

Table 12.13: Strategy use for N22

Strategy	Classification of strategy	Number using the strategy	Number correct
5 equal pieces then ½ & ½.	Mixed method, multiplicative.	2	1
5 equal pieces (twice) bit more.	Mixed method, multiplicative.	2	1
½, five equal pieces, 4 equal pieces.	Mixed method, multiplicative.	2	1
½ then 5 equal pieces, bit more.	2, 3, 5 method, multiplicative.	1	1
5 equal pieces and closer to.	Multiplicative.	4	3
½ and ½, bit less.	Mixed method, additive.	1	0
½ and ½.	Mixed method, additive.	3	3
4 equal pieces.	Additive.	1	1
½, bit less for 600 & a tiny bit more.	Mixed method, additive.	4	1
Closer to.	Mixed method, additive.	3	1
Bit more than 500, little bit more than that.	Using bits. Counting.	7	0
Bigger to the right.	Counting.	1	0
Counting in small skips for ones.	Counting.	1	0

Item N25: Locating 120 on a scale marked in twenties

Table 12.14 shows a similar strategy picture to Table 12.11 for N6. Two students identified that there were four lines, so the scale had to be in twenties. This response has been classified as an additive strategy because an interval of 100, four partitions and a partition size of 20 cannot be linked multiplicatively.

Occasionally, student explanations of their strategies are unclear. This can be the case particularly when a response like "it's going up in twenties" is given. Generally, further probe questions are needed to clarify the strategy, but in some cases they do not do so because, for example, the student has no logic to explain why they chose twenties.

Table 12.14: Strategy use for N25

Strategy	Classification of strategy	Number using the strategy	Number correct
Division (100 ÷ 5).	Marks and interval, multiplicative.	2	2
5 gaps/lines so 20s.	Marks and interval, multiplicative.	2	2
4 gaps/lines so 25s.	Marks and interval, multiplicative.	1	0
Multiplication (5 × 20 = 100).	Marks and interval, multiplicative.	1	1
Skip counting in twos or twenties.	Additive.	12	11
Skip counting in tens.	Additive.	2	2
Going up in 20s.	Unclear.	1	1
4 gaps/lines so 20s.	Skip counting, additive.	2	2
Reciting a count.	Counting.	2	0
Counting in ones.	Counting.	1	0
Bigger to the right.	Counting.	1	0
Guess.	Not applicable.	1	1

Items N6 and N25: Intervals partitioned into four and five

In N6 students were reading from a scale partitioned into 25s, while in N25 they were locating 120 on an interval of 100 partitioned into five pieces. Nineteen students (59%) correctly answered both items, while seven (22%) correctly answered one. That such a percentage only answered one correctly suggests that the pair of items is useful to establish which students have not mastered the use of such scales.

With this item pair it was common for students to use different strategies on the two problems. For example (as was expected) the seven students who used halve and halve on N6 all used a different strategy on N25. Five switched to skip counting in twos or twenties, one to a skip count in tens and the seventh student to “four marks so twenties”. All of these strategies are classified at the same level in the strategy hierarchy, providing more evidence to support that hierarchy.

Items N10 and N25: The impact of number size

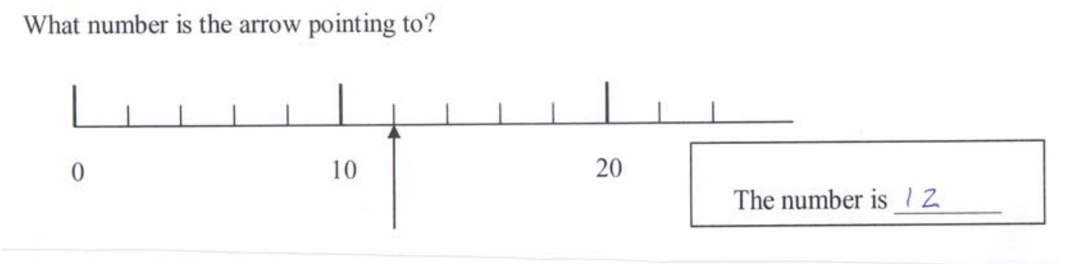
Item N10 required students to read a scale marked in twos, while for N25 they needed to locate 120 on a scale marked in twenties. Overall, the results suggest that the change of number size had an impact on a small number of students. Table 12.15 shows that of the 24 students who correctly answered both items, 18 used the same strategy on both items, and that students who changed strategies were less likely to be successful. Three of the students who were able to skip count in twos were unable to do so with twenties; two swapping to reciting a count in tens, and one to ‘bigger to the right’. In doing this they changed from a strategy that is classified as additive in nature, to ones that are classed as counting strategies. Four other students with both items correct swapped from a multiplicative strategy on N10 to skip counting in twenties. In seems when number size was an issue these students changed to a lower level strategy. This provides evidence that the categorisation of these strategies is appropriate.

Table 12.15: Consistency of strategy use

Number correct	Consistent strategy use	Inconsistent strategy use	Total
Both correct.	18	6	24
One correct.	0	3	3
Neither correct.	1	4	5
Total	19	13	32

The multiplicative nature of the interval

The strategies used by students answering the items for this aspect illustrate that the scale interval is fundamentally multiplicative. For example, for students to work out the value of each mark in N6 requires the calculation of a difference (the interval is 200 – 100 = 100), the identification that it is cut into four pieces, and a calculation based on multiplication, division, or fractions (each successive mark shows 25 as 100 ÷ 4 = 25 or 4 × 25 = 100 or ¼ of 100 is 25). Students who incorrectly count marks are unlikely to connect multiplicatively the width of the interval, the number of pieces and the size of each piece, even if they have good number understanding. Figure 12.11 shows they are likely to be restricted to using lower level strategies.



- S(O): I like saw that they were 10 apart but there was only 4 in between so it had to be twos.
- S(EE): … there was 4 marks so I did 12, 14, 16, 18, 20.
- S(JJ): … there's four lines there. And two times four is eight but … so like with two five decimals together makes ten.

Figure 12.11. Student responses to N10

Such multiplicative reasoning is generally beyond the reach of students until they are operating at least at Stage 6 of the number framework (Ministry of Education, 2005a). This equates to Level 3 in the new curriculum (Ministry of Education, 2007).

For unmarked number lines, students also benefit from using higher level number skills. To locate 130 on N17 (Figure 12.12) they have several options. Some of those used by students at School 4 are shown in Table 12.16. All three rely on some understanding of multiplication.

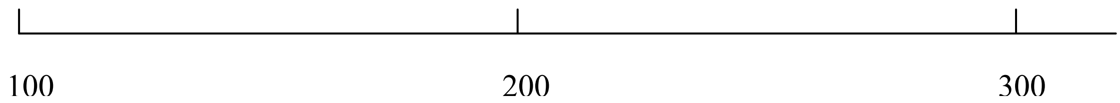


Figure 12.12. N17

Table 12.16: Some strategies for locating 130

Option 1	**Option 2**	**Option 3**
Cut the interval into 10 little pieces (as 10 × 10 = 100). 130 is the third.	Half-way between 100 and 200 is 150. Half-way between 100 and 150 is 125. 130 is a bit to the right.	130 means the interval is cut into tens. I can get ten if I halve first, then divide the bit between 0 and 150 into five, and count along three.
3/4 correct.	3/3 correct.	4/6 correct.

Other students correctly answered the problem by 'using bits' and by 'halving' then 'using bits'.

12.5.3 Aspect 3: Scales involving fractions and decimals

Aspect 3 also utilises a much broader range of items in Test 2 than Test 1.

Table 12.17a: Aspect 3 responses (facilities)

Item	Percentage correct	Percentage incorrect	(omit)
C9: Broken ruler problem.	44	56	(0)
C11: Reading from a thermometer marked in fifths.	28	72	(16)
N12a: Locating a decimal < 1 on a partially marked scale (1 d.p.).	41	59	(3)
N12b: Locating a decimal < 1 on a partially marked scale (2 d.p.).	38	62	(16)
N13a: Reading three quarters from a marked number line.	41	59	(6)
N13b: Reading a fraction > 1 (quarters). Marked number line.	41	59	(6)
C14: Locating three quarters on an unmarked weighing scale.	41	59	(3)
N15: Locating two thirds on an unmarked number line.	28	72	(13)
C19: Locating a decimal > 1 on an unmarked ruler.	78	28	(0)
N21a: Reading a fraction < 1 (fifths). Marked number line.	41	59	(9)
N21b: Reading a fraction > 1 (fifths). Marked number line.	41	59	(9)
N23: Reading a decimal from an unmarked number line.	69	31	(3)
C24: Reading a decimal from a marked weighing scale (fifths).	63	37	(6)
N26a: Reading a fraction < 1 (thirds). Marked number line.	19	81	(22)
N26b: Reading a fraction > 1 (thirds). Marked number line.	9	91	(22)
C28: Reading between whole numbers (unmarked scale).	25	75	(6)
N29: Reading between whole numbers (scale marked in twentieths).	22	78	(9)

Table 12.17b: Error patterns for Aspect 3 responses

Error pattern	Item number and percentage						
Counting from 1 (no place for zero). Marks and spaces error.	C9 (28)	N13a (9)	N21a (9)	N26a (3)	N15 (6)	N29 (6)	N24 (3)
Reading from the end of the ruler.	C9 (13)						
Counting in points.	C11 (38)	N12b (3)	N13a (12)	N13b (9)	N21a/b (9)	N26a/b (3)	N24 (19)
Between.	C14 (6)	N15 (6)					
It's in the middle.	C11 (3)	C14 (3)	N24 (3)				
Counting in ones.	N12a (3)	N12b (3)	N13a/b (9)	N21a/b (6)	N26a/b (9)		
Bigger to the right.	N12b (16)						
Ten in-between thinking.	N12b (22)	N21a/b (6)	N26a/b (3)	C14 (13)	C28 (35)	N15 (9)	
Hundredths in-between.	N12b (9)						
Reciting a count.	N13a/b (3)	N21a/b (6)	N26a/b (3)	N19 (3)	N24 (3)		
Negatives to the left of one.	N13a (6)	N21a (13)	N26a (6)				
Not referencing the closest unit.	N13b (3)	N21b (13)	N26b (3)				
Fitting tenths.	N26a/b (3)						
Treating the whole line as the unit.	C14 (22)	N15 (13)					
Accuracy.	C14 (6)	N15 (3)	C19 (16)				
Confusing mixed numbers and proper fractions.	C14 (19)	N15 (3)					
Treating the fraction as an operator.	C14 (6)	N15 (6)					

Item C9: The broken ruler problem

Of the 14 students who answered the broken ruler problem correctly, ten were asked to explain their solution method. In doing so nine indicated they had used addition or subtraction. For example, Student AA answered "because 2 is there and 7 is there, so that would be 5 … and then I counted these little lines, that's a five there [pointing to half-way] … and then you've got to add on two …" In contrast to this were the nine students providing an answer between 6 and 7. They explained a 'double count', with the start of the rod counted one, and the first millimetre counted one again. These students are unlikely to learn the 'additive' strategy by themselves as the numbers 2 and 7.7 are not additively related to their answer. Something must first challenge their conceptual understanding.

Four students provided answers between 7 and 8, reading 7 from the ruler, then counting the millimetres. One student used a single count, answering 14, while another answered 7/14. That over 40% of students produced answers between 6 and 8, thus showing limited understanding of the nature of measurement and how to use a ruler after seven or more years of schooling, is a cause for concern. However, this is not an isolated finding (see Chapter 11). It suggests that the way linear measurement has been taught to these (and other) students needs review.

Item C11: Reading from a thermometer marked in fifths

Student responses to this item indicated that those who were successful tended to attend to the interval. For example, seven of the nine explicitly mentioned the middle of the interval (39) and six stated the number of partitions in the unit. In comparison, only five of the other students noted that 39 was the middle of the interval, and this did not have an impact on their subsequent working. (It seems the visual cue of the ten pieces took precedence over the size of the interval.) These results, along with those of C9, suggest that effective users of scale attend to both the interval and any partitioning.

Table 12.18: Strategy use for C11

Strategy	Classification of strategy	Number using the strategy	Number correct
Division (100 ÷ 5).	Marks and interval, multiplicative.	2	2
5 gaps/lines so 0.2s/twos.	Marks and interval, multiplicative.	3	3
4 gaps/lines so twos.	Marks and interval, multiplicative.	1	1
20 per space.	Unclear.	1	1
Skip counting in twos.	Additive.	1	0
Trial and error.	Skip counting. Additive.	1	0
Counting in points.	Counting.	13	0
Counting in ones.	Counting.	2	0
Guess.	Unclear.	1	0
Tenths	Unclear.	2	2

Item N12: Locating 0.3 and 0.21 on a partially marked number line

This item was a development of A7 from Test 1. A similar proportion of students at both Schools 2 and 4 successfully located a one-digit decimal on the line. The success rate for N12b at School 4 was also similar, suggesting the two-digit decimal was no harder.

Table 12.19 shows the strategies students used on N12a. It shows students using counting-based strategies were unlikely to answer correctly. Further discussion of the results for this item is found in Section 12.8.2.

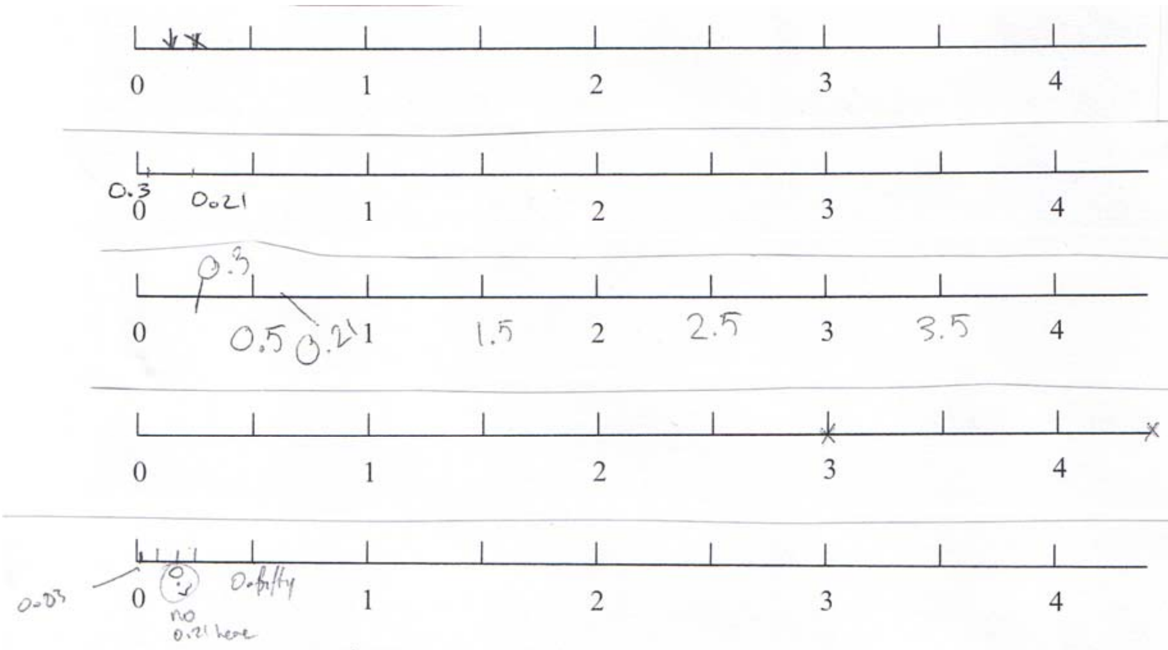


Figure 12.13. Exemplar student responses to N12

Table 12.19: Strategy use for N12a

Strategy	Classification of strategy	Number using the strategy	Number correct
2, 3, 5 method.	Marks and interval, multiplicative.	4	4
½ and ½, bit more.	Mixed method, additive.	6	5
½ and around half-way.	Mixed method, additive.	3	2
½ and a bit less.	Mixed method, additive.	1	0
Counting in points.	Counting.	9	1
3 is close to zero.	A little bit more. Counting.	4	0
Counting in ones.	Counting.	1	0

Items N13, N21 and N26: Quarters, fifths, and thirds

In Test 2, three variations of A5 from Test 1 were used to explore response patterns with different fractions. These visibly similar items assessed students' ability to work with quarters, fifths, and thirds (respectively) with numbers under and over one. The comparative success rates for these items are shown in Table 12.20.

Table 12.20: Proportion correctly answering each item

N13a	N13b	N21a	N21b	N26a	N26b
13/32 (41%)	13/32 (41%)	13/32 (41%)	13/32 (41%)	6/32 (19%)	4/32 (13%)

The equal figures for the first four items do not represent the results of a particular group of students who were successful on all four problems. Nearly half of these students only answered one of the pairs correctly, and only 4/32 had all six items correct.

Twelve students used the same strategy on each item, with only one of these students having all six correct. Only four of the other 11 students had one or more of the items correct. Some students transferring strategies indicated that they were responding to the visible similarity of the items, and did not examine their mathematical content.

The only student to transfer strategies and correctly answer all six items used 'bits and ths' then converted her responses to decimals. The responses to this set of items showed that only seven of the students at School 4 knew of, and used, the 'bits and ths' strategy (which has been given the easier to remember name 'cuts and pieces' in Table 12.2). These students were amongst the most mathematically able, both in terms of their results for Test 2 and for the NUMPA (Ministry of Education, 2005b). For N26a, all six who answered correctly started out by identifying the interval was cut into thirds.

Table 12.21 shows the strategies of students who correctly answered an item pair. It shows that successful students tended to change their strategy from item to item, and that an interval partitioned into either four or five can be successfully answered using a skip count. Those answering N26 correctly first identified the interval had been cut into thirds then went on to convert this information to decimals for their answer. While other students could have been successful had they understood the 'cuts and pieces' strategy (three pieces make thirds), students limited to decimal-based skip counts were unlikely to succeed. This suggests that a single fractional item involving either quarters or fifths is not sufficient to assess fractional (or decimal) understanding, replicating the finding for intervals partitioned into 20s and 25s (Section 12.5.2).

Table 12.21: Strategies of students correctly answering an item pair

	N13 (quarters)		**N21 (fifths)**		**N26 (thirds)**	
	Number using strategy	Number correct	Number using strategy	Number correct	Number using strategy	Number correct
Division (e.g. 1 ÷ 5 = 0.2).	0	0	4	4	0	0
½ and ½ .	6	6	0	0	0	0
Cuts and pieces.	3	3	1	0	4	4
Skip counting.	2 in 25s 4 in 2s 2 in ¼s	2 0 1	4 in 20s 6 in 2s 1 in 0.2s	2 4 1	0	0

Student understanding can be effectively masked by apparently correct responses

Correct responses have already been shown to be problematic when using a written question, but they can also be misleading in an interview. For example, five students used a strategy involving twos or twenties to answer N13 (a scale in quarters), then used a similar strategy on

N21 (a scale in fifths). Three of these answered N21 successfully, the fourth answered 0.8 and 0.4 (forgetting to reference the nearest whole number), and the fifth answered 8 and 4 (a whole number conversion error). Student X was one of these five students and answered the three similar fraction items as follows.

N13: [explaining 0.8 and 1.2]
S(X): Its going up in twos, twenties [changes answers to 0.80 and 1.20 and explains] so people don't get confused, some people think one half when they see 1.2

N21: [0.80 and 1.40]
S(X): It's going up in 0.2s

N26: [0.75 and 1.75]
S(X): its ¾ of the way there, I looked at the amount of things between each number

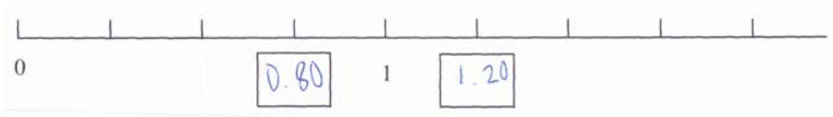


Figure 12.14. Student X's response to N13

Following the interview, probe questions identified that Student X did not consistently count either the number of marks or spaces (things were marks), and did not realise it was always helpful to do so. Instead he often looked at the item and decided that (in his own words) "he knew the system."

Such results show students can sometimes be successful with 'inappropriate' strategies and that a written response cannot reliably be used to provide evidence of student understanding. Items involving five marked intervals were found to be particularly prone to giving such 'false positive' results. Ten and 0.1 would do likewise.

Responses to individual items can provide false information

The above example also shows that certain items, on their own, are inadequate to accurately measure student understanding. While working with the same items, for N13 Student DD explained there were three things so it went 2, 4, 6, 8; appearing to make a marks and spaces error. For N21 he explained it was 2, 4, 6, 8, 10. It seems that in answering these problems DD is reciting a count in twos, but is not showing an understanding of the relationship between the count and the filling of the interval. As another example, Student W used a skip count in 25s for N13, and when that did not work for N21, used thirties. This suggests that it is important to use problems where the interval has been split into both four and five pieces when diagnosing student understanding of scale.

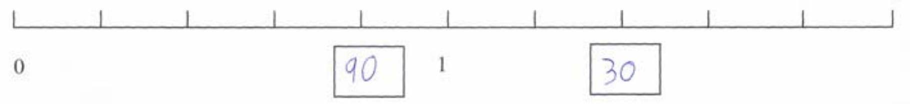


Figure 12.15. Student W's response to N21

Further discussion of the results for these items can be found in Section 12.7.2, 'beliefs about number.'

Item C14 and C28: Reading and locating ¾kg on the same unmarked weighing scale

C14 was a repeat of B3c from Test 1 (locating ¾kg on the scale) so comparative achievement data are available. These show that at all three schools a sizeable proportion of students answered incorrectly (School 1, 46%, School 2, 65%, School 4, 56%). Many of the strategies identified amongst the students at School 1 were also used by students at School 4. For example, *n* equal pieces, ½ and ½, treating the whole number line as a unit. More in-depth questioning led to the clarification of some students' thinking. For example, Student U used '10 in-between' thinking: "because that's three and then I go 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. So I'm thinking it'll be four." The strategies of 'between' and 'it's in the middle' were also identified (see Tables 12.2 & 12.3, and Sections 12.4.4 & 12.4.5). These strategies help explain a number of the responses to this item by students at School 2.

Some of the unsuccessful strategies used by students again seem to relate to a lack of understanding of fractions symbols. For example, for C14 Student G located the three then "just went in quarters", placing her mark a bit above 3. Student II located the three then went "one four", again locating her mark a bit above 3. In both of these cases the lack of understanding of the symbol was accompanied by a recognisable partitioning strategy. This was also noted in the responses of other students.

Table 12.22: Strategies for answering C14 and C28

Strategy	**Classification of strategy**	**Number using the strategy**		**Number correct**	
		C14	C28	C14	C28
n equal pieces.	Additive.	8	3	7	3
½ and ½.	Mixed method, additive.	7	6	6	5
After ½ and before 1.	Between, counting.	2	0	1	0
Always 5 then 10 little milli ones.	Explanation unclear.	1	0	0	0
½ then closer to 1.	Explanation unclear.	0	2	0	0
Partitioning into 5.	Multiplicative.	0	1	0	0
2, 3, 5 method.	Multiplicative.	0	1	0	0
½ then 10 equal pieces.	Multiplicative.	0	1	0	0
Halving and using bits.	Mixed method, additive.	0	2	0	0
Closer to.	Mixed method, additive.	0	2	0	0
Using bits.	Counting.	0	6	0	0
A little bit less.	Counting.	0	1	0	0
Lots of little lines in-between.	Visualising.	0	1	0	0
Goes in tens so about 0.8.	Visualising.	0	1	0	0
5/6, not a smaller fraction like ¾ or 2/3.	Unclear.				

Decimals or fractions, or decimals and fractions?

C28 brought to the fore differing understandings about where fractions and decimals are found in relation to the whole numbers. For this item (supposedly reading ¾ from an unmarked weighing scale), 25/32 chose to answer with a decimal, 5/32 a fraction. When reading an unmarked scale, how should students perceive the line? In this item, starting by halving can lead to the answer ¾ (or 0.75), as the arrow is clearly in the middle of one of the halves. Starting from the belief that decimals live between the whole numbers is likely to lead to an expectation of a one decimal place answer. In this case 0.8 or 0.80 (given by 14/32) is a good approximation of ¾, so should both be coded correct? The initial distinction of 0.75 being correct and 0.8 incorrect has been retained as it was symptomatic of differences in student thinking. For example, 11 of the students answering 0.8 gave explanations that indicated they understood that ten pieces exist between whole numbers. By considering the answers to other items (particularly N23) it became clear that only the most mathematically able students could consider both perspectives (fractions and decimals) and choose which they thought would give them the correct answer. Is this surprising? It should not be. Both fractions and decimals on their own 'fill the interval', so why should students think both fill the *same* interval? This concept relates to the equivalence of fractions and decimals and the continuity of the number system. A belief that the decimals lie between the whole numbers helps explain why some students located ¾ between 3 and 4, even though they could read the fraction as three quarters.

Reading and locating

C14 and C28 formed a second 'reading and locating' pair requiring ¾ to be located and read. For C14, 16/32 (50%) located their answer between 0 and 1 (13 answering correctly). Nine students (28%) located their answer between 3 and 4. For C28, having the arrow between 0 and 1 had a significant impact, as 28/32 (88%) provided an answer between these numbers.

When Test 1 was developed, it was not considered that some students might use different skills when reading and locating a number on the same scale. However, the example pairs (C4 & C20, C14 & C28) used to explore the impact of this change show that some students (those tending to have low-level strategies) answered some of these tasks differently. For example, in C20, students were forced to locate a number that would not exist if counting in ones or points. In C14, students who did not understand fractions were being asked to locate a number they did not understand. The results suggest that items asking students to read or locate a number on a scale need to be carefully thought through if they are to be diagnostically useful. They also suggest that items involving reading the scale are more likely to identify issues with student understanding than items requiring a number to be located; this possibility explains a result of Bright et al. (1988) who found that items where the fraction was given were easier than those where the representation was given.

Item N15: Locating 2/3 on an unmarked number line

The strategies students used for this problem have been identified in Table 12.17b, and are also discussed in Sections 12.7.1 and 12.8.1 so do not need to be repeated here. However, Student W provides an example of how student understanding of scale is affected by a collection of different factors, each of which is prone to error. In this case he did not understand fraction symbols so located 2/3 between 2 and 3, but also made a marks and spaces error by drawing three marks for thirds. It should be noted that such compounding of errors explains why the error total can exceed 32 in this table as well as in some of the later tables in the chapter.

S(W): …because I put in two thirds so I went like a line there, a line there, a line there. Oh no, that's wrong … Um, yeah and then I sort of figured out so that'll be one third, two thirds…ish. So it should be there.
I: So you went one third, two thirds, oops, that's wrong. So why does it move from there [the third mark] to there [the second]?
S(W): Um, because that would have been three thirds [points to third mark].

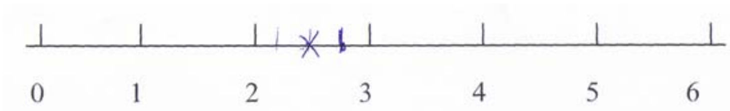


Figure 12.16. Student W's response to N15

Table 12.23: Strategy use leading to correct answers for N15

Strategy	Classification of strategy	Number using the strategy	Number correct
n equal pieces.	Additive.	8	6
Bit over ½.	Mixed method, additive.	4	3
Guess.	Unclear.	1	0

Note that many of the errors are summarised in Table 12.17b

Item C19: Locating 1.6cm on an unmarked ruler

This item allowed the use of the same strategy as B4 in Test 1, in which 0.4cm had to be located on the same ruler. Success rates on the items were very varied (School 1, 54%; School 2, 28%; School 4, 81%). Changing the item from a decimal between 0 and 1, to one between 1 and 2 removed certain error patterns so may explain in part the higher success rate at School 4. For example, no student located 1.6cm to the left of zero, as happened with 0.4cm at both Schools 1 and 2. Nor did any student locate 1.6cm between 16 and 17 and no-one located it at 16, as five and four students respectively had done with 0.4cm at School 2.

Table 12.24: Strategies for answering C19

Strategy	Classification of strategy	Number using the strategy	Number correct
2, 3, 5 method.	Multiplicative.	2	2
½ then 10 equal pieces.	Multiplicative.	1	1
½ and ½, bit over ½.	Mixed method, multiplicative.	2	2
Bit over ½.	Mixed method, additive.	21	19
n equal pieces.	Additive.	3	3
Closer to.	Mixed method, additive.	1	0
Using bits.	Counting.	2	0
Whole number thinking.	Counting.	1	0

Item N23: Reading 1.6 from an unmarked number line

Many of the findings for this item are reported elsewhere. For example, the item was used to identify 'ten in-between' thinking in Section 12.4.1. The results are also compared to C19 in Section 12.8.1, where responses to pairs of similar contextual and number line items are reported. Table 12.25 shows that no new strategies were identified by the item, and that the use of lower level strategies tends to be associated with higher error rates.

Table 12.25: Strategies for answering N23

Strategy	Classification of strategy	Number using the strategy	Number correct
2, 3, 5 method.	Multiplicative.	1	1
Bit under 2/3.	Mixed method, multiplicative.	1	1
½ then 10 equal pieces.	Multiplicative.	1	1
½ and ½, bit over ½.	Multiplicative.	2	2
bit over ½.	Mixed method, additive.	19	16
n equal pieces.	Additive.	2	1
Closer to.	Mixed method, additive.	1	0
Using bits.	Counting.	2	0
Half.	Counting.	1	0

Item C24: Reading a weighing scale marked in fifths

This item is another that was paired with other items to test the impact of certain changes on response patterns. Those results are reported in Section 12.8. Table 12.26 shows the impact of having a number of problems of a similar style in Test 2. Students who recognised this similarity progressively abbreviated their answers, adding comments like 'it's the same as the other one'. For Table 12.26 the responses for C24 have not been 'referenced back' to such earlier explanations to illustrate the point that some individual responses need to be considered in relation to similar items if an accurate understanding of student thinking is to be developed.

Table 12.26: Strategies for answering C24

Strategy	Classification of strategy	Number using the strategy	Number correct
4 lines, each is 0.25.	Unclear.	1	0
5 lines, each is 2 (or 0.2 or 20).	Unclear.	5	5
Just below ½.	Mixed method. Additive.	2	2
10 equal pieces.	Additive.	1	1
Going up in twos.	Unclear.	3	3
Skip counting in twos (or 0.2s or 20s).	Additive.	9	8
Trial and error.	Skip counting. Additive.	1	1
Reciting a count.	Counting.	1	0

Item N29: A line partitioned into twentieths

This item provides different information from the others assessing decimals on scales, so has merit for a scale diagnostic. In particular, analysis of student responses showed that of the ten students who counted in points, seven incorrectly read their decimal, giving answers like 'one point fourteen'. The two students counting in twos also did this. In comparison, when students were asked to read the decimal 0.21 for N12, the interview analysis showed no clear link between how they read the decimal and where they located it. The students incorrectly answering this item were often not those incorrectly answering other items with counting strategies (e.g., C4 & N13a).

Table 12.27: Strategy use for N29

Strategy	Classification of strategy	Number using the strategy	Number correct
Multiplication and division (100 ÷ 20 = 5, 5 × 14 = 70).	Marks and interval, multiplicative (whole number conversion).	2	1
Fractional (15/20).	Marks and interval, multiplicative.	1	0
20 lines so halved.	Marks and interval, multiplicative.	4	3
1 line makes 0.05 so skip counted.	Mixed method, multiplicative.	1	1
½ and a little bit more.	Mixed method, additive.	3	1
Skip counting in 10s then 5s.	Additive.	1	1
Skip counting in twos.	Additive.	2	0
Ten in-between.	Counting.	1	0
Counting in points.	Counting.	10	0
Counting in ones.	Counting.	2	0

12.5.4 Aspect 4: Scales involving integers

Only a single item was used to assess understanding of integers on the number line, as they were from CL4. As expected, a high proportion of the students (47%) did not recognise the mathematical importance of the negative sign.

Table 12.28: Aspect 4 responses

Item	Percentage correct	Percentage incorrect	(omit)
N16: -4 to the left of -1.	41	59	(3)
Error pattern	**Percentage**		
No negatives shown.	(47)		
No place for zero.	(6)		
Integers between 0 and 1.	(3)		
Negatives to the right of zero.	(6)		

The revised item format was more useful, in that it allowed students to show whether or not they understood the order convention in relation to integers, so is worth retaining in this form for future diagnostic testing. However, as found with other items in which students were required to create a scale, the number lines drawn tended to be unique, so each conveyed different information about the understanding of the student who drew it (see Figure 12.17). The results for this item will be discussed further in Section 12.7.

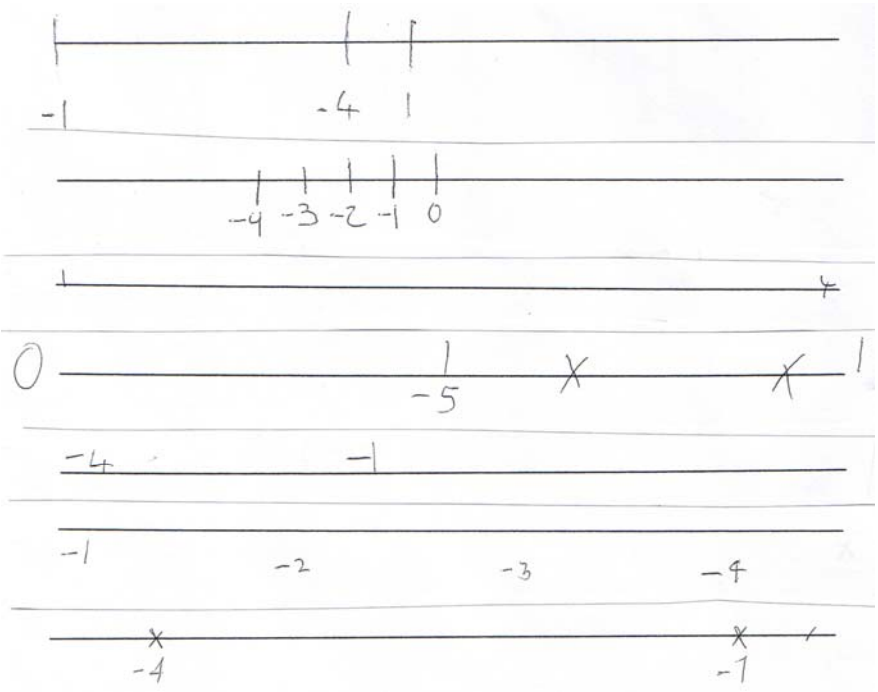


Figure 12.17. Examples of student responses to N16 (reduced)

12.5.5 Aspect 5: Scale creation and coordinate plotting

This new aspect reports the findings of the items introduced to explore student understanding (and error patterns) when creating scales and using coordinates. It was introduced as a consequence of the findings from School 2 for B5, which identified that students' understanding of scales in the bar graph could not be separated from their understanding of graphs and graphing.

Table 12.29: Aspect 5 responses

Item	Percentage correct	Percentage incorrect	(omit)
N7: Plotting a coordinate.	19	81	(25)
C8a: Locating 4 appropriately.	38	62	(41)
C8b: Locating 8 appropriately.	34	66	(41)
C8c: Locating 37 appropriately.	22	78	(47)
N27: Plotting coordinate appropriately.	13	87	(56)
Error pattern	**N7 (%)**	**C8 (%)**	**N27 (%)**
Coordinates reversed.	(28)	(NA*)	(3)
A square marked (so no zero).	(19)	(13)	(0)
2 squares marked (so no zero).	(6)	(NA*)	(0)
A line marked.	(3)	(NA*)	(0)
Starting from the top of the scale, not the origin.	(13)	(6)	(3)
Unit scale created or wanted.	(NA*)	(16)	(19)
No zero.	(0)	(3)	(0)

*NA Not applicable

The curriculum review identified that students in Years 7 and 8 may not have met the concept of coordinate plotting in an algebraic context, so the low success rates and the high proportion of students omitting items N7 and N27 were not unexpected. Responses to the individual items are discussed in more detail below.

Item N7: Plotting a point on a unit grid

For this item a number of error patterns were identified. For example, six students worked in squares, one on a line and two labelled two separate points. Four students did not know to start their count at the origin, instead beginning at the top left-hand corner as they would when reading text. Ten moved vertically before horizontally, illustrating they did not know the graphing convention of horizontal movement first. It was common for students not to label their scales, suggesting they were used to working with unit scales only.

As most of these errors relate to knowledge of scale conventions, they seem to be ones that could be easily rectified through teaching. This finding supports the suggestion that algebraic graphing could be a good context for developing an understanding of graphing conventions, and suggests that the form of this question was not something these students had experienced.

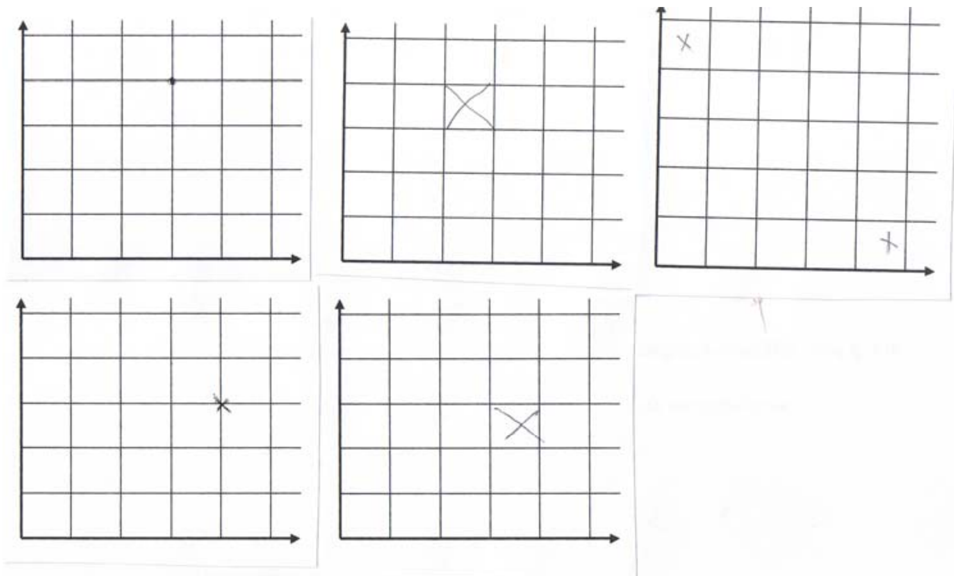


Figure 12.18. Examples of student responses to N7 (reduced)

Item C8: Scaling

Thirteen students were unsure how to approach this question (only one omitting N7) even though additional explanation was given and reference was made to the axis being part of a graph, like the question before, with the researcher pointing to the vertical axis. The wording was varied and simplified for the later interviews, but this had little impact. The results for N27 suggest that even if this had been a graph drawing exercise, many of these students would have been unsure how to proceed.

Students who successfully answered the item tended to use a scale in tens, even though this did not fully utilise the axis (two of these started by skip counting in fives, finding they only got to 35, then decided to work in tens). Two others started in fours, but were unsure where 37 would be located. Only one student attempted to choose a scale based on the number of partitions on the axis, and this student counted spaces not marks, then incorrectly located both 4 and 8 (see the second to last scale in Figure 12.18).

S(N): I thought that, um, 36, divided by 6 was 6 so there's a spare space so I put 37 there. And then I kind of worked out where, um, those two went.

Only one of the four successful students was amongst the more successful with scale (and number), while two scored poorly on the other scale items. The responses suggest three things. First, success with this item does not correlate with being an effective user of scale; second, that most students at School 4 are not familiar with creating non-unit scales; and third, most students do not know the convention of creating a scale that fully utilises the axis. It was also not common for students to label their scale, even though the item asked them to do this (see fourth to last example in Figure 12.19).

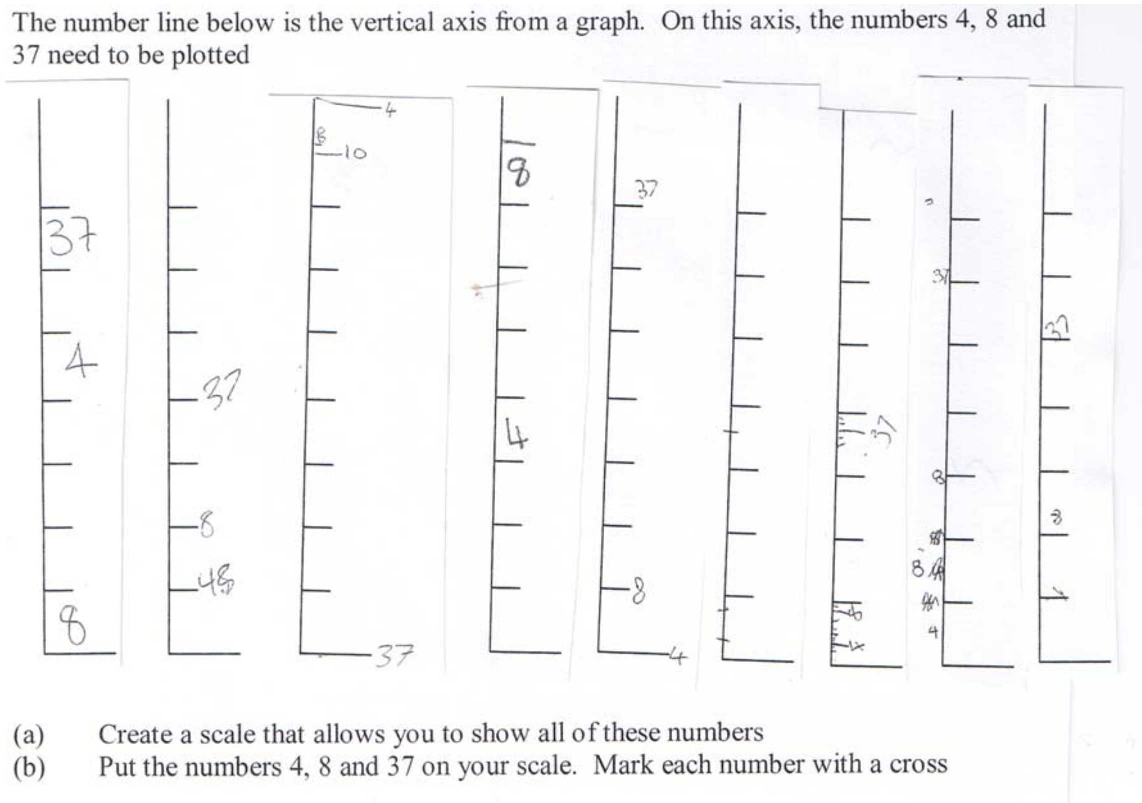


Figure 12.19. Examples of student responses to C8, showing original wording

Item N27: Plotting a point where two non-unit scales are required

As expected, N27 proved to be a very challenging item. For example, the six students who wanted a bigger grid (so they could use a unit scale) were amongst the most able mathematicians at the school. Given that skip counting could be used to identify an appropriate scale, this suggests that non-unit scales are something these students have rarely worked with. This possibility is reinforced by the fact that one of the four successful students was amongst the less successful with the other scale items.

Table 12.30: Plotting a point that requires creating two non-unit scales (N27)

Attempted item	Correct	Honest scale on *x* axis	No scale marked
10/32 (31%)	4/32 (13%)	3/10 (30%)	7/10 (70%)

Is creating a scale more difficult?

Crooks and Flockton, (1996) and Friel et al. (2001) suggested that students find creating a scale for a graph more difficult than plotting points on, or reading from, a graph. This study raises a question in relation to their finding. Students K and P showed they could successfully locate and read whole numbers, fractions, and decimals on scales, using multiplicative thinking when appropriate. However, when plotting points they reverted to using a unit scale, even though for N27 this meant there was insufficient space to complete the task. For N8 they were unsure what to do. For N7, Student P located the point (3, 4) in the square rather than on a gridline intersection. She sums up the problem nicely: "I haven't really learnt about graphs and all that sort of stuff." It seems these students are not aware that their number line and measurement skills can be applied to graph scales. Using something other than a unit scale for a graph was also not part of their experience. Results for the bar graph at Schools 1 and 2 also indicate some students did not understand that graph axes should be scales, although these same students could work effectively with whole numbers on number lines – and read from and locate numbers on the vertical axis of a graph. Is the lack of success with scale creation a function of lack of exposure and students learning about the use of scales in isolated contexts, each with their own sets of rules, rather than being a more difficult task? Further research is needed here.

This ends the discussion of the individual scale aspects. In the following sections, some trends identified across the aspects are identified. In the first of these, the language that students used when talking about scale is considered.

12.6 The language of scale

In Chapter 2 time was spent defining the construct and clarifying the language to be used when considering features of scale. This was important for clear communication of concepts and thinking. During the data analysis, a review of all of the interview responses from School 4 was undertaken when it became clear that little of the mathematical language described in Chapter 2 was being used by students. This involved considering the words students used when they were asked to describe how they worked out their answer to each of the questions; that is, the language they used for the retrospective questioning. The analysis showed that, when trying to refer to a particular feature of a scale, very few students had adequate mathematical language to identify that feature, which hampered their communication. For example, words like 'gaps', 'places', 'skips', 'jumps', and 'spaces' were used to refer to an interval. 'Notches', 'dashes', 'small ones', 'little things', 'things', 'bars', 'numbers' and 'arrows' were used to describe marks. Some students referred to the same feature using a variety of terms for different problems. Others simply avoided naming what they were talking about, as the following quotes show.

> S(A): Because there is…because there is 4 in between him and 60 so it goes up in twos. (C4)
>
> S(G): Because there's not enough to go in ones. (C20)
>
> S(L): Because there's five things instead of 10 so one must equal two. (C4)
>
> S(S): Um, because three and then two…um. (C24)

In several cases, students adopted the language that the interviewer used while trying to clarify their strategy, a feature that was also apparent when a similar retrospective analysis was undertaken of the interviews from School 1.

Chapter 4 identified that while students are expected to meet scale in a wide variety of situations, the reviewed curriculum documents provided little guidance on what skills and understandings students needed to develop for them to use scale successfully in those situations. The reviewed resources likewise provided exposure to scale, but did not focus on learning about scale. That many of the students at School 4 lacked an appropriate language to discuss the features of scale is consistent with the Chapter 4 analysis of the possible impact of such a curriculum approach. The lack of an appropriate language suggests that (as indicated by the student quotes) some students may find thinking and communicating about scale difficult. The finding further suggests that one of the first tasks of a teacher seeking to develop students' understanding of scale is to consciously build the language of scale. Many of the terms and definitions outlined in Chapter 2 would be appropriate.

12.7 The impact of number understanding

12.7.1 A real test of symbol understanding

While this research focussed on understanding scale, a number of items challenged students' understanding of symbol notation. N15 tested their understanding of fraction symbols, which many showed was not strong when they located two thirds between 2 and 3. Students F and U located their marks at $2\frac{1}{3}$, talking of the two, and placing their mark three numbers away from this. Further questioning identified they were seeing the interval between the whole numbers as containing ten pieces, so two thirds was really 2.3. Student G, with the same location talked of 'two and then thirds', confusing proper fractions and mixed numbers. Students HH and FF marked 2½, explaining that two thirds was between two and three, even though they were able to correctly read the fraction.

For all of these students there is obvious confusion between proper fractions and mixed numbers. Perhaps it is time to stop talking of fractions as being made from two numerals that need to be interpreted in relation to one another, and recognise that many are made of *three* numerals, with a whole number and a fractional part.

N12 (locating 0.21 and 0.3 on a partially marked number line) provided a similar test for understanding decimal notation. Here many students showed they did not understand the place value concept for decimals. For example, 16/32 at School 4 located 0.21 to the right of 0.3. Some doing this seemed to believe decimals only have one decimal place, applying their knowledge of rulers, or decimals in other situations.

I: Can you put where that one is …? [0.21]
S(A): Wouldn't that make it around there [extends line and marks -0.1] because won't there be ten in each gap?
I: What makes you think there's going to be ten in each gap?
S(A): Because in a ruler there's always 10mm in each centimetre.

S(R): Um, well, zero point two one, and so, um, first of all I thought it went higher than 0.3 which would be like that space over there [points to space after 0.5]. And then I thought that it was actually like, zero point twenty one, so it would be 2cm and one.

S(FF): Zero point twenty one… [marks 2.1]
I: Why did you think it goes there? …
S(FF): Because I remember when we used to have to do these things like 4.3 + thing and you couldn't go past the 10 so I chose there.

For others, decimals had two places, so treated 0.5 as nought point fifty, and confusing 0.3 for 0.03. Here the correct reading of the decimal did not seem to be a factor in this interpretation.

I: Could you put where you think those two numbers are on the number line?
S(J): There's 21 [marks line correctly].
I: Yep.
S(J) There's three [marks 0.05].
I: Okay. So how did you work out that that was 3 and that was 21?
S(J) Because from 50 that's 25 [pointing].
I: Oh, I get it. Half of 50 is 25 so that's worked out where 21 is going to be because it's a bit to the left isn't it?
S(J) Yeah, and then I did three is going to be pretty close to zero.

N16 tested understanding of integer notation. As expected, many students did not recognise these numbers. For example, 15/32 at School 4 did not recognise the negative sign as a symbol with meaning when asked to locate -1 and -4. They located 1 and 4 instead. Of the 16 using the signs, Students S and T placed the integers to the left of one, with no space for zero. Student W placed them between 0 and 1, while Students X and EE had -4 to the right of -1. While the item gave good information on some students, especially when used with follow-up probes, it would be useful to have a problem that would give better data about the place of zero on the number line. The item in Figure 12.20 is of the style of Kuchemann (1981) and is designed to do this.

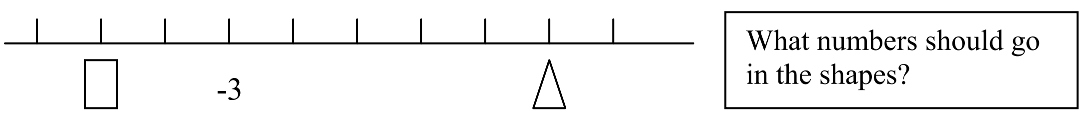


Figure 12.20. Integer item to explore the understanding of zero

Discussion

The results outlined in this section indicate that student ability to work with scale is dependent on their understanding of the number symbols with which they are working. This finding supports Dufour-Janvier et al. (1987) who identified that to use a representation, students first need to understand the representation. The finding however, uncovers a 'chicken and egg' scenario; a student cannot work with scale effectively if they do not understand these forms of number, but a good test of whether they understand the numbers is to get the student to locate them on a scale. This suggests that it is important to use number lines and scales when developing such number concepts, as the number line can visually show the relationship between the number symbols and the (relative) size of the number(s), which is not easy to convey in other contexts (a comment also made by Kieran (1976) in relation to considering order amongst fractions).

12.7.2 Beliefs about number

Negatives to the left of one

When asked to identify ¾ on a marked number line (A5a in Test 1, N13a in Test 2), approximately one in fifteen students at each of the three schools responded with negative numbers. Their answers indicate that they believe the negatives are found to the left of one. Student A indicated he believed there were 'ten in-between' each mark, hence his answer of -10. Student S seemed to choose a random number (-4), but used this consistently when answering N13a, N21a, and N26a. Student 6 from School 1 seemed to choose -2 as the opposite of 2. These responses are another illustration of how a student's understanding of number affects how they answer scale-related questions.

Decimals are found between the whole numbers

In reviewing responses, some students at School 4 showed they believed that only decimals are appropriate when labelling numbers on the number line. For example, Student C indicated an understanding of quarters, and that three pieces gives thirds, but was not able to use this information unless he knew the fraction to decimal conversion. For N21, he miscounted the marks so used 'counting in points', renaming one as 0.6.

N13: [explaining the answer 0.75.]
S(C): I just forgot [the] 0. bit, I just counted it up in quarters.
I: So you actually know that 0.25 is a quarter? [nods]

N21
S(C): It had six notches so that would be 0.6 for one.

N26
S(C): It went up in thirds.
I: So it goes up in thirds. So if it's going up in thirds how do you work out it's 0.75? So what went through your head there?

S(C): It's half in the middle there.
I: Yep.
S(C): It's about where 1.75 would be.

This issue was explored with Student AA, when she was having a similar problem.

[Counts 2 lines in-between 0 and 1]
S(AA): So they look like they're going up in thirds, but what's the third of one?
I: You're allowed to write it in as a fraction rather than a decimal if you want. If that's going to be easier.
S(AA): Oh, really?
I: Of course you are.
S(AA): Oh, so that's two thirds…
I: Yep.
S(AA): And that's one and two thirds.
I: Yep.
S(AA): There we go.

Student AA was the only student to provide a fractional answer for this problem.

Numbers below 1 are different to those above

Some students at Schools 2 and 4 provided responses showing that they believe the numbers below one work differently to those above one. For example, Students A and S seemed to believe the negative numbers are located to the left of one, while Figures 10.3 and 10.6 provide examples that suggest the students believe that decimals (or fractions) are only found between zero and one. Other students recognised that fractions and decimals are also found above one, but changed the interval size between units. For example, students H, P, and Q 'split' their number line on N26, switching from thirds to fifths.

I: … why point 66 recurring?
S(P): Un, because if you take one again and one hundred, a hundred, a third of the hundred is 33.33 recurring and two thirds its 66.66 recurring so if we just put, take out the point and put it behind the whole number.
I: Ok. And why is this one 1.4?
S(P): Because there's five spaces so it should be in 140.

Discussion

Students' beliefs about numbers may help explain some of the responses identified in this study. For example, students who believe that decimals are found between whole numbers may, when asked to locate a fraction like ¾, not place it between 0 and 1 as they know this space is already taken up by the decimals. This in turn suggests that a student's understanding of number has implications for any diagnostic test on scale.

Both Test 1 and Test 2 contain several sections. In Test 1 there is a clear partition between the number line and the contextual items. In addition there are questions relating to the different aspects of scale understanding: number line conventions, whole number multiples, fractions and decimals, integers. Test 2 adds scale creation and coordinates plotting. For many of the number types certain items seem to act as a test of understanding, so in each of these areas a 'ceiling effect' may exist within a scale test. For example, a student comfortable working within the realm of whole numbers may need to develop a new understanding of fractions before being able to work successfully on fractional items. This can require a paradigm shift during which a more sophisticated understanding of number develops. This suggests that some of the student problems with understanding scale are fundamental to their understanding of the number system so will not be easily overcome. It also suggests that only limited gains in understanding should be expected from short interventions to improve scale understanding.

12.8 Strategy transfer

The impact of context and the use of different number sets on similar scales were explored through the interview process at School 4. The goal was to gather more information about strategy transfer. The results are outlined below.

12.8.1 The impact of context

Four item pairs were used to investigate further the impact of context on strategy use. One pair involved whole numbers (C4 & N10); the second, quarters and thirds (C14 & N15); the third, fifths (N21 & C24); and the fourth, decimals (C19 & N23). The results are outlined in Table 12.31. Note that the second column describes the sort of scale used in the item while the third column notes any differences between the items in the pair. The remaining columns summarise the results. Column 4 gives the number of students who correctly answered an item while column 5 notes the proportion of students who used the same strategy for a pair of items and column six notes any cases where it was unclear if a student had used the same strategy. The last two columns identify the proportion of students answering both items correctly. Column seven gives the proportion of the students who were correct and column eight the proportion of students who used the same strategy on an item pair (who are described as students who transferred their strategy from one situation to the other).

Table 12.31: Frequency of strategy transfer between number line and contextual problems

Items	Scale	Variations	Number correct	Transfer strategy	Unclear if strategy transferred	Both correct	Both correct given transfer strategy
C4	Reading a marked scale in twos.	Interval size. Different number.	19/32 (59%)	16/32 (50%)	2/32 (6%)	18/32 (56%)	12/16 (75%)
N10			27/32 (84%)				Unclear 2/2
C14	Locating a fraction on an unmarked scale	Different fraction.	14/32 (44%)	10/32 (31%)	0	8/32 (25%)	1/10 (10%)
N15			9/32 (28%)				
N21	Reading a marked scale in fifths.	Interval size. Different number.	13/32 (41%)	19/32 (59%)	0	12/32 (38%)	11/12 (92%)
C24			20/32 (63%)				
C19	Unmarked scale with a decimal.	Reading and locating.	25/32 (78%)	22/32 (66%)	1/32 (3%)	21/32 (66%)	15/21 (71%)
N23			22/32 (69%)				

Items C4 and N10

For this item pair the higher success rate with the number line can be explained by considering those students who changed their strategies for N10. For example, of the nine students who 'counted in ones' for C4, seven changed to a more sophisticated strategy for N10, six of them answering correctly. It seems the context did not trigger the use of a more sophisticated strategy in this case. Of the other students who changed strategies, the six who used 'half and a bit more' for C4 changed, possibly as this strategy was not appropriate for the number being found in N10. Three of these students correctly answered both items.

Items C14 and N15

While this item pair required students to locate similar fractions on scales (3/4 and 2/3), the different denominators led to strategy changes amongst students who had an understanding of fractions. For example, Student C appropriately changed strategies from using ½ and ½ again to locate ¾ in C14, to splitting the interval into three equal pieces for N15. This change of strategy is the opposite of the trend in the other three pairs, but is predictable according to the hierarchy of strategies, and is consistent with the findings for N13, N21, and N26 reported in Section 12.5.3.

Two students who had an appropriate strategy for C14 did not seem to realise it could also be used for thirds. For example, Student P partitioned the unit into four equal pieces for C14, but identified that for N15 'three is the whole, so two is 2/3 of it'. Both of these students only used '4 equal pieces' on C14, and were the only two students to have a strategy that they used solely on a contextual item but not on at least one (decontextualised) number line.

Items N21 and C24

This pair of items involved a number line with a large interval and a weighing scale with a small interval, so were visually very different. Even so, more than half of the students used the same strategy on both items. Furthermore, Table 12.31 shows that of those who answered both items correctly, almost all used the same strategy on both.

Three students incorrectly answering N21 swapped from counting to a successful additive strategy for C24. In each case, these students had used this 'new' strategy with other number line and contextual problems. Another three students consistently used a strategy that was potentially useful for both items, but only answered C24 correctly.

Items C19 and N23

This pair of items involved one item in which students had to read 1.6, and another in which they had to locate 1.6. It has already been shown that in some cases this change is enough to trigger a change in strategy use (see *reading and locating*, Section 12.5.3). Of the 21 students who used the same strategy for both items, five inaccurately located/read one of the numbers. Four of these students had one of the item pairs correct. Of the remaining six students who had both items correct but changed their strategy, two swapped to a more sophisticated strategy for the number line, while four swapped to a simpler strategy. All four of these students used their more sophisticated strategies on a different number line item. One student was able to answer the contextual item, but was unsure how to proceed with the number line. However, this student used their strategy for C19 on another item with a number line.

Discussion

The item pairs provide little evidence to support the hypothesis that contextual problems involving scale can help some students access informal knowledge that is not accessed by pure mathematical equivalents. In only two cases (both involving the same item) was a strategy identified for a contextual problem that was not used on a number line. Rather, some students did not always use the same strategy on the pair of items, but the 'new' strategy was commonly used elsewhere in the test on a less similar number line. This result provides evidence to support the suggestion that number lines should be used to teach scale. There was also no consistent trend for either the number line or context-based items used at School 4 having a higher success rate.

12.8.2 Transfer of strategies from whole number situations

Three item pairs were used to consider strategy transfer from whole number to decimal or fractional situations (Table 12.32). Note that the information provided by this table is similar to that provided by Table 12.31, even though there are slight differences in layout.

Table 12.32: Frequency of strategy transfer from whole number situations

First item in a pair	Scale	Second item in a pair	Scale	Both correct	Transfer strategy	Both correct given transfer strategy	Unclear if strategy transferred
C4	Marked scale in twos	C24	Marked scale in 0.2s	16/32 (50%)	15/32 (47%)	12/15 (80%)	1 (both correct)
C6	Marked scales in 25s	C13	Marked scale in quarters	12/32 (38%)	3/32 (9%)	3/3 (100%)	4 (all with both correct)
N12a	Partially marked scale, (0.5s)	N17	Unmarked scale in 100s	5/32 (16%)	7/32 (22%)	3/7 (42%)	0

In addition to the students transferring strategies identified above, another group was identified who appeared to swap strategies due to how they are named. On closer inspection they were found to be reciting a number of different counts. For example, seven students counted in ones for C4, five swapping to counting in points for C24. For N12a, seven made three skips of 0.1, five then swapped to three skips of ten for N17. Four of these students worked consistently by reciting counts of 0.1, one or ten on each of the four problems. Some also used a similar approach on at least one of the problem pairs. According to NUMPA results, while the students using this approach included the mathematically weakest, most had some part-whole strategies. This implies they were not strictly ‘counters’, but are not likely to have had the understanding of number to identify what each mark represents. This finding helps explain the results of the APU (that scales with subdivisions of 0.1, one and ten are associated with higher success rates) and the comment from Barcham (1996) that a common error by some students is to assume the scale involves a number like 0.1, one or ten. This approach would give these students some success, as many scales have these subdivisions, and may have the effect of masking the students’ lack of understanding from teachers.

Note that as some students have not progressed beyond counting in ones, while others work solely in points between whole numbers, the distinction between these counting strategies needs to be retained.

Discussion

These investigations show that when students attempt to solve scale-related problems, considerable strategy transfer takes place. The results provide more evidence that the students who were successful with such problems had a tendency to see the mathematical similarity between visually different situations, and used their understanding of the scale underpinning the task to help them answer. Likewise, they were more likely to see mathematical differences in situations that appeared visibly similar.

Minor changes to a problem, for example changing the problem from one involving reading to one requiring a number to be located (that an experienced student could identify did not affect the mathematics), caused some students to change their strategy. Some students (subconsciously perhaps) expected to use different skills for different problems. Successful users of scale seem to have a more unified concept. Figure 12.21 provides an excerpt from a teacher trial (a second can be found in Figure 9.12). In both cases, the teachers recognised the equivalence of the situations, using the same strategy to complete the problems.

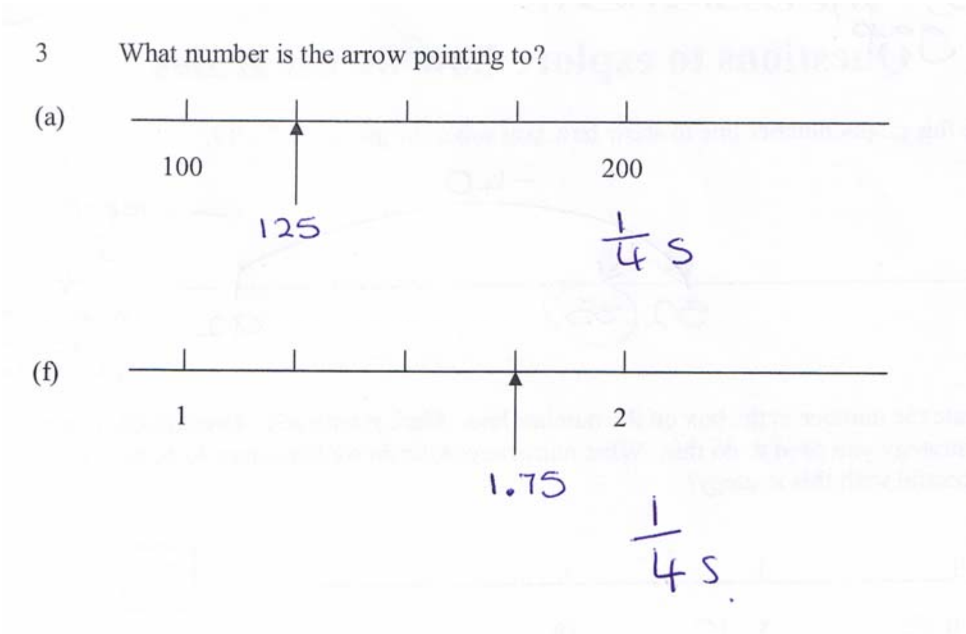


Figure 12.21. Teacher work showing consistent strategy use across different number sets

Such findings suggest that teachers looking to consciously develop students' understanding of scale need to do more than use scales in a variety of discrete contexts. The similarities between different situations need to be explored; a more unified approach to scale is needed.

12.9 The marked scales partitioned into five

Test 2 deliberately used a number of similar items to explore response trends. One set of items involved problems where each labelled interval was divided into five. Table 12.33 shows that there is a huge variation in success rates on this family of items, and that decimal items were generally harder than whole numbers ones.

Table 12.33: Success rates on items partitioned into five

Item	N10	C20	N25	C24	C4	N21a	N21b	C11
Scale	in 2s	in 2s	in 20s	in 0.2s	in 2s	in 0.2s	in 0.2s	in 0.2s
Success rate	27/32 (84%)	25/32 (78%)	23/32 (72%)	20/32 (63%)	19/32 (59%)	13/32 (41%)	13/32 (41%)	9/32 (28%)

Two items, C4 and C24, involved students reading from weighing scales where the spaces were roughly 1mm in width. The success rate is similar in spite of the use of a different number set. However, the success rate for C4 was lower than that for the other scales marked in twos, especially N10, where the interval was 9mm in width.

The hardest item was C11, the thermometer with 38 and 40 labelled, so the interval visually appeared to have been partitioned into ten. The cosmetic change of not labelling 39 had a definite impact on how some students perceived the problem, as nine of the 12 students who 'counted in points' did not do so on C24, where all whole numbers were labelled. That 28/32 students could identify half-way (using this strategy on C18 and other problems), including ten of those counting in points, again provides evidence that some students attend to the visual cues without considering the mathematics of the situation.

Eleven students used the same strategy on the four decimal items, eight of whom were the only students to get all of these correct. Of the other three, one had a good strategy but a naming problem (for example 38 and 4, 3 stone 4), the second had one answer incorrect and the third omitted all four. On these items there seems to be strategy stability amongst students who know what they are doing.

Finally, this set of results does not help clarify the impact of interval size on the strategies of students, even though there is a mix of items with 'small' and 'large' intervals like C4 and N10. Too many other variables have been shown to be in play to isolate this particular factor. Further research is required that involves large and small versions of identical scales if other variables are to be excluded.

12.10 Not all marks are equal

The responses to a range of items made it clear that some students had learned to recognise visual differences in the marks used on a scale. For example, for the broken ruler item C9, 9/32 recognised 'the bigger mark in the middle' was half-way, so started their millimetre count there at five (Figure 13.1). By identifying half-way (and the number it represented) these nine students were able to move beyond counting from one. In C11, others noted the darker line to show 39 on the thermometer in C11 (Figure 11.9). This may have helped these students identify that the unit was partitioned into five rather than ten.

12.11 Have curriculum expectations been met?

The analysis of the facilities undertaken in Section 12.3 shows that the students met the standard of evidence set as a measure of meeting curriculum expectations for only three contextual items (C18, reading half-way; C19, locating a decimal on an unmarked ruler; and C20 locating a number on a scale marked in twos). The above discussions also suggest that the students did not generally seem to benefit from the contexts, in that only one situation (involving a context-based fraction problem) brought to light a strategy not used elsewhere by the student on a number line. Furthermore, it has been identified that it is possible for students (especially those using counting strategies) to sometimes correctly answer scale-related problems involving both contexts and abstract number lines, but have little understanding of scale. This third finding provides evidence that correct answers and marks from a written test (of the form used at School 2) do not provide a reliable measure of scale understanding. When these factors are considered together it is reasonable to conclude that the students at School 4 have not met the curriculum expectations of being able to use a variety of scales in contexts by Year 7.

This section concludes the exploration of the student data. In the final chapter the findings of the study will be brought together, explanations sought for what has been observed, and wider implications considered.

Chapter 13
Concluding comments

13.1 Introduction

This study undertook a curriculum review in the content area of scale. This involved determining the scale understanding of students in Years 7 and 8. Because little research was found that directly studied scale and student understanding of it, the study was fundamentally explorative, mapping for the first time an area of mathematical importance in terms of the expectations that the curriculum and selected resources had of students. This was followed by an investigation, through interviews and testing, of the actual understanding of students and the various strategies they used to reach solutions to scale-related problems. In line with the explorative nature of the study, findings were identified in the early stages of the research that altered the course of the study.

This thesis reports findings from three case study schools and a series of teacher trials. The findings from Case 1 led to the introduction of the teacher trials. Their purpose was to challenge and refute those tentative findings. The data collected from the teacher trials supported the concept of a hierarchy of strategies to solve scale-related problems; the strategies used by the teachers were from the higher stages of the hierarchy and had previously been identified in the work of the students at School 1. Data collection from a second school was then undertaken to further challenge the hierarchy. Analysis of this larger data set supported the existence of the hierarchy, but demonstrated that it did not fully map student strategies. New questions about student understanding were raised. Further exploration was undertaken at a third school. Analysis of these data provided a much clearer picture of student strategies and student understanding, and led to a revision of the hierarchy.

Data from each of the cases and the teacher trials have been reported in detail. Both qualitative and quantitative approaches have been applied. Analyses to measure if the items were appropriately assessing the construct have been reported. Copies of written student work and verbatim verbal responses have been provided to exemplify student strategies, while simple counts present a sense of the variation in approaches to each item. Coherence between the findings from each data source and from the relevant literature has been reported.

This chapter summarises the findings of the study. Explanations are formed for what has been observed but not explained by the students themselves. Implications of the research for practice are considered, and recommendations made. Section 13.2 starts by addressing the research question.

13.2 What understanding of scale do the Year 7 and 8 students in the case study schools have?

In Chapter 2 it was suggested that to answer this research question it was necessary to answer five sub-questions. The answer to each sub-question is therefore outlined below.

13.2.1 What should Year 7 students know about and be able to do with scales?

The analysis of the mathematics curriculum document indicated that by the end of Year 6, students were expected to be able to use appropriate instruments to perform measuring tasks involving a range of metric units and scales in a variety of contexts. They were expected to use graphs to represent number and informal relationships, to read and discuss features in displays, and to represent and interpret statistical data in graph form, including pictograms, bar charts, dot plots, and strip graphs. Students were expected to learn about scale while working in contexts that utilised scales. The science, technology, and social studies curricula provided additional learning opportunities over and above those in mathematics. Measuring instruments and graphs were tools to be used to explore the scientific and social world. Students should use graphs to present information clearly and to identify trends, patterns, and relationships in data. In technology, teachers also were identified to have an active role in students' learning of such mathematics.

Three issues were identified as a result of the analysis of these documents:

1) The curricula did not identify what teachers needed to do for students to achieve these outcomes;
2) The number sets students should be working with were not explicitly outlined; and
3) There was some inconsistency between the different subject curricula about what students were expected to do.

13.2.2 How do some commonly available mathematics resources present the learning of scale?

In each of the three reviewed series of resources learning about scale was generally "within a range of meaningful contexts" (Ministry of Education, 1992, p. 66) such as measuring and graphing. This was in line with the expectations of MiNZC. In addition, some opportunities were provided for students to work with number lines. No explicit teaching about how to use scales was found, although some issues were raised where teachers' notes were supplied. Scale was at times used as a tool for other learning. In these cases an understanding of scale was assumed. If following a programme of learning based on any one of the resource series, Year 6 students should have worked on scales in a wide range of situations, using whole numbers, decimals, and at times fractions by the end of the year.

13.2.3 What did students in the case study schools actually know about scale?

One way used to measure what the students 'knew' about scale was to identify the test items for which the students had achieved the curriculum expectations. Facilities of 75% or over were chosen as the measurement criterion for this purpose. The contextual items in the tests generally assessed learning based on the AOs from CL2 and 3, and were intended to provide students with familiar contexts that matched the curriculum intent. In general the scales used in the test items were simpler than those found on many commercial measuring instruments or in graphs found in the media.

There was little evidence to suggest that the students from Schools 1, 2, and 4 had met the curriculum expectations in relation to the test items. At Schools 1 and 2 the required standard of evidence was met for only one of the 12 contextual tasks analysed. At School 4, the standard was reached for three of the 12 contextual tasks analysed. At all three schools students met the expectations more often for items involving number lines – that is, items without a context. Overall the students from the three schools could be said to know that on a horizontal number line, the numbers get bigger to the right and that when labelling a number line the numbers for labelling the scale should be aligned with any marks. They could also identify half-way between two whole number multiples (Schools 1 & 2 were assessed with a number line, School 4 with a context) and extend the labelling of a number line marked in sixes. Students at Schools 1 and 2 were also able to extend a scale in eights on a graph, but as this item was not used in Test 2 it is possible that the students at School 4 would have also met the 75% criterion.

By applying the 75% criterion to students' test scores as a whole, within Schools 2 (16%) and 4 (11%) of students could be said to have mastered the use of scale by Years 7 and 8, even though they were supposed to be working with standard measurement instruments and graphs from Year 3 (CL2).

The study has shown that there were few common understandings; rather what was known varied from individual to individual. In many situations, students performed better on a number line item than a similar contextual one, in spite of number lines not being mentioned in any AO within MiNZC (Ministry of Education, 1992). Some students struggled with scales involving multiples of two. Others could not successfully work between the numbers marked on the scale. Scales involving numbers other than the whole numbers caused problems; in particular, the majority of students throughout the study had problems relating fractions to scales. It was common to find that problems best solved with fractions were answered by using decimals, by converting to whole numbers, or with whole numbers. In some instances, for example on a bar graph, scales were not recognised for what they were. The role of zero on many scales, and in linear measurement, was not always recognised. Many students at School 4 did not do well on a version of the broken ruler problem. The item designed to test student understanding of the conventions of a number line showed that many students were not able to accurately iterate a unit. Others who could do this did not know that twice as much space should be left for an interval two units wide. Perhaps curricula phrased in terms of outcome statements do not provide sufficient guidance to teachers as to how those outcomes can be attained. Teachers may need more explicit guidance on the key understandings that students need, as well as ideas for how to develop these.

When working with scales, the interviews identified that regardless of a student's understanding, they used a range of mental strategies. For unmarked scales these involved partitioning. These strategies can be classified as being based on counting, adding, and multiplying, so seem to form a hierarchy (see Tables 12.2 & 12.3). This notion of a hierarchy is supported by evidence that the more effective users of scale, including the teachers from the teacher trials, tended to work with strategies classified in the higher stages of the hierarchy. The strategies used were not referred to in either the mathematics curriculum or the reviewed texts, rather they seem to have been picked up incidentally or invented by students.

13.2.4 What issues of understanding did the students have?

Many of the students in this study appeared not to have a consistent concept of scale, rather they saw scales in different situations as separate constructs, subject to different rules. The concept of contextual pollution was introduced to explain how the understanding some students had of a particular context restricted their ability to use that context to learn about scale.

The issues identified in the previous section are similar to those found amongst other groups of students as reported in the reviewed literature. They seem to be consistent across national borders and time, regardless of the curriculum being followed. The literature also showed that there are common issues related to scale that cross mathematical domains. It was not uncommon to find authors making comments about aspects of a particular context for scale being related to an understanding of measurement. However, the literature on linear measurement (the simplest form of measurement) identifies this concept is not understood well – a fact confirmed by student data from this study.

The students' ability to work with scale was found to involve a complex mix of factors. These were their:

- Understanding of the scale construct, and the conventions used on it.
- Understanding of number. This includes having a sound knowledge of the meaning of number symbols, and of where zero, integers, fractions, and decimals can be found in relation to the counting numbers. It also includes students being able to flexibly use multiplicative and fractional processes.
- Strategies to decide on the value of a partition in marked intervals;

- Strategies to partition unmarked intervals;
- Ability to iterate a unit. and
- Understanding of the role of marks and spaces.

The effective users of scale (both students and teachers) also tended to use the same solution strategy in relation to what they perceived as mathematically similar situations. Some other students used the same strategy when they identified a visually similar situation. Those who were less successful tended to use lower level strategies and changed strategies more; for example, minor changes to a item that did not affect the mathematics, so would be ignored by an effective user of scale, caused some students to change the way they approached a problem. An additional complicating factor was that some students did not have the language to describe the features of a scale so struggled when asked to articulate their thinking. It seems possible that this lack of language could also affect their ability to think and learn about scale.

Work at Schools 2 and 4 identified a link between number understanding and being able to successfully work with scales. Students at School 2 who demonstrated the ability to use multiplicative thinking in the NUMPA (Ministry of Education, 2005b) were generally successful with scales involving whole numbers and decimals. Success rates on these items tended to increase with NUMPA stage; however, success could not be associated clearly with a particular stage. This is because some students were able to answer certain items correctly by using particular counting-based or additive strategies (although evidence suggests that this was often more to do with luck than understanding). With items based on fractions, a different trend was found, even though understanding of fractions was not high. Students across a range of stages were able to answer the item if they used a fraction, but only those at the higher multiplicative stages were able to answer successfully using a decimal.

Which are easier – marked or unmarked intervals?

This investigation sought to identify if students found locating a fraction on a fully marked scale easier than locating the same fraction on a partially marked or unmarked scale, following a line of enquiry based on Behr et al. (1983). The investigation identified the following:

1) Students at Schools 1 and 4 generally used different strategies when dealing with marked and unmarked intervals. In relation to fractional items only:
    - A5a/A6 (reading ¾ from a marked number line and locating ¾ on an unmarked line): 0/13 at School 1 used the same strategy on both.
    - B3a/B3c (reading ¾ from a marked weighing scale and locating ¾ on an unmarked weighing scale): 1/13 copied the marking from the marked scale.
    - N15/N26 (locating 2/3 on an unmarked number line and reading 2/3 from a marked line): 1/32 at School 4 used the same strategy.
    - N13/C28 (reading a marked number line and unmarked weighing scale in quarters): 3/32 used the same strategy on both: halving and halving again.

- N23/N29 (unmarked interval reading 1.7, marked interval reading 1 14/20): 3/32 used the same strategy.

2) Slight variations to the numbers, the number line, and the question affected how some students approached a problem, so talking of one sort of cue being of different difficulty to others is not a helpful concept when working generally with scale. This finding is similar to that of Lesh, Behr, and Post (1987).

However, these results indicate that any assessment seeking to measure student understanding of scale needs to include items that utilise scales with both marked and unmarked intervals.

13.2.5 What does this research suggest as ways of addressing these issues?

To answer this sub-question literature, theory, and evidence from student responses are used to form explanations for observations made during the research. In doing this a way of addressing the four main issues is formed. For example, the first subsection uses schema theory (Skemp, 1971) to explain the observation that some students seem to treat scales as if they conform to a special set of rules that are particular to the situation.

a) Developing a concept of scale

A schema can be described as an individual's attempt to make sense of a concept. As such, it can be considered to be the sum of their internal representations and the external representations that are recognised as part of the concept. Given this, according to schema theory (see Section 3.5), the observations made during this study suggest that many of the students have not developed an appropriate schema for scale. Instead these students appear to have attached individual examples of scales to a number of inappropriate schemas. For example, instead of seeing the axes of a bar graph as scales, many of the students at School 2 saw them as something different, subject to their own set of rules. Schema theory thus helps explain the observation of substantial strategy transfer amongst 'successful students' and teachers, and the tendency for less successful students to approach each scale-related problem as a different one to which a different strategy applies. Schema theory also suggests that there is not likely to be a 'quick fix' for what has been observed. Skemp (1971) suggests that the "first part of the answer would seem to be to try to lay a well-structured foundation of basic mathematical ideas, on which the learner can build in whatever direction becomes necessary" (p. 53). As measurement and graphing are both applications of scale, and are important for school leavers and higher mathematics, this suggests that it is important to address the issue of scale understanding from a curriculum design perspective. It also suggests instead of considering a particular application for scale (e.g., measurement) to be related to an understanding of measurement, that understanding of the application should be related to an understanding of scale.

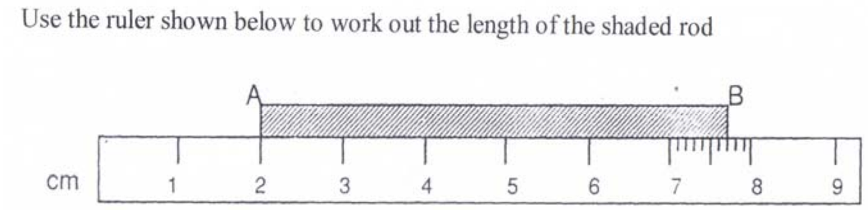
The analysis of the strategies used by students at School 4 showed that in general students used the same set of strategies on both the number line items and the items involving 'familiar' contexts. This suggests that using number lines could be an effective approach for developing students' understanding of scale, especially if these are used alongside, and are explicitly linked to, measurement scales and scales in graphs. In combination, this should help students develop an appropriate schema for scale.

Since the importance of highlighting the links between scale and measurement has been a recurrent theme in the literature reviewed for this study, a revised model for teaching measurement is introduced later in the chapter. Following Skemp's suggestion, this is designed to be used over time, starting with students at a young age. It has as a central purpose the use of measurement as a context that assists the development of a well-structured schema for scale, one that can be extended and adapted to accommodate new data as different situations are encountered. Prior to introducing the model, however, what this research tells us about measurement understanding is explored. In the next section student responses to the 'broken ruler' problem are considered, and how schema theory can suggest a way of improving students' understanding of measurement.

b) Counting and measuring

A common thread in the literature on linear measurement is that many children find the ruler difficult to use and understand (e.g., Hiebert, 1984; Irwin & Ell, 2002; Nunes & Bryant, 1998; Piaget et al., 1960). A suggestion commonly found in Chapter 3, not just in the measurement literature, was the need to focus students on counting the spaces not the marks to help address some of these difficulties. However, the wider perspective on scale in this study suggests different causes. These are outlined in the next paragraphs.

As noted in Chapter 12, when answering the 'broken ruler' problem, some students obtained answers between six and seven. These students started counting at the beginning of the rod, showing they had the interval concept of measurement, so may not have been counting marks erroneously.



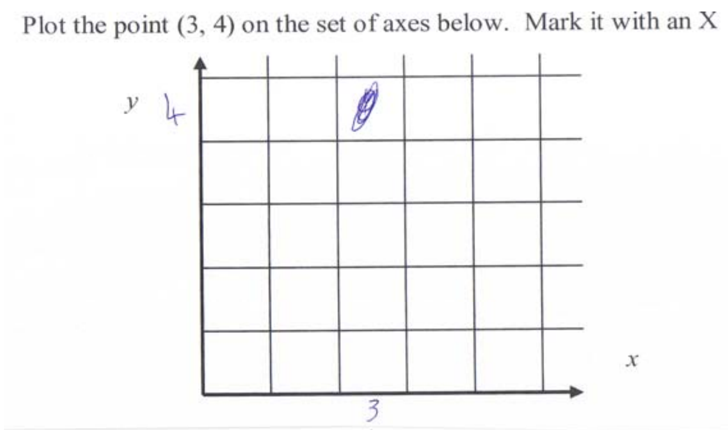
S(DD): Its 6cm and 7…mm.
I: … how did you work that out?
S(DD): I counted each [points to 2], 1, [points to 3] 2,… then I counted [points to millimetres], that will be 5, 6, 7.

S(G): Because I counted the ones until I got to 6 and then counted on until I got 7.

Figure 13.1. Responses to C9, a version of the broken ruler problem

Linear measurement is not simply a process by which we identify the length of objects by counting units. It also measures *distance* and in travel we have not made a journey of one kilometre until the process is ended. *The endpoint marks the kilometre*. Units can also be partitioned, so we talk about *the half-way point of the journey*. It seems likely that Year 7 and 8 students intuitively understand these ideas. That most of the students who were successful with the broken ruler problem used additive reasoning also provides an important insight. It suggests that the process of linear measurement requires more than an understanding of counting; the successful students appear to have developed the concept of *difference*. Therefore it is more likely that those students providing an answer between six and seven did not understand the role of zero in a measurement; that is, a measurement always has a start point to which we give the number zero. Instead, they have applied their knowledge of counting to the marks.

Student work has also shown how zero can be missed from number lines, graph scales, and measurement devices. Issues around the role of zero have also been identified in the literature. Figure 13.2 shows how a counting approach based on spaces led Student FF to not only incorrectly mark a square, but also to omit zero.



S(FF): There's the 3 okay [starts in bottom-left square and points to adjacent squares in turn] and there's 4 – 1, 2, 3, 4 [moves vertically four squares].

Figure 13.2. Student FF counts squares to plot a point

Finally, it should be noted that counting spaces can create its own misconceptions. Stephan et al. (2001) and Irwin and Ell (2002) identified that when measuring, some students 'forgot to count the first tile laid'. To explain this error, Stephan et al. (2001) conjectured that students were not viewing measurement as covering the defined space, rather as a count of how many it took to reach the end of the object. However, such an error can also be explained by reflecting on how students play simple board games like snakes and ladders. They learn that when throwing a six you are expected to travel/move a distance of six squares, and that *the next square, not the one you are on*, is counted as one. In doing this they seem to be considering distance (from here to there) – an alternative use for measures of length, and one that can be perceived by the learner as an active process. If such understanding is applied to the measurement of a rug with one's feet (or a ruler, or number line), the student will count spaces, treat the first square (foot placement) as 'here' or 'the square you are on' and omit it from the count. This does not seem to be an incorrect process for measuring length or distance, for if everyone followed this convention there would be no issue. It seems more a consequence of encouraging the use of using space counting for measuring length.

The above evidence suggests that it is important to use the conventions of measurement with a scale, part of which involves starting at zero and counting the marks. Previous recommendations about counting spaces may help students answer correctly, but is it enough to get students to answer correctly when there is still evidence of the need for greater conceptual understanding? Perhaps we should try to emphasise that we use different conventions in measuring and counting; provide a language that students can use to discuss scale and align the introduction of scales with the development of additive reasoning. This, at least, would avoid the current situation in which students who are limited to an understanding of counting strategies when working with numbers are sometimes asked to start counting at one, and at other times at zero, with no clear indication when each should happen.

c) Not all marks are equal

It was noted in Chapter 12 that some students had learned to identify visual differences in the marks on a scale, and that these provided scaffolding information that helped them work with the scale. This provides another reason for focusing students' attention on counting marks, and suggests that once students have learned to count them correctly, the different information gained by attending to a particular subset of marks may be a useful way to help students apply more sophisticated number strategies to scale. However, it is important to remember that some of the Year 7 and 8 students predominantly used counting-based strategies, and it was suggested that some of these strategies may have developed when learning to use the ruler.

d) The problem with the ruler

Chapter 4 identified that when following a learning progression based on MiNZC (Ministry of Education, 1992) students move from counting non-standard units at CL1 to using a ruler as a tool for practical measuring tasks at CL2. Observations of a standard classroom ruler suggests that when following this progression students may become accustomed to 'compound scales' before they learn the conventions of simple scales. The ruler in Figure 13.3 shows a simple scale, where the conventions are clearly visible, including zero, something not shown on a New Zealand school ruler (Figure 13.4). (Note that this common New Zealand school ruler partially follows the conventions suggested by Hiebert (1984), starting the scale a short space from the end to encourage children to think about both endpoints of the measure when using a ruler.)

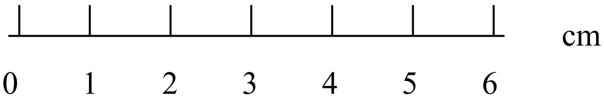


Figure 13.3. A simple ruler

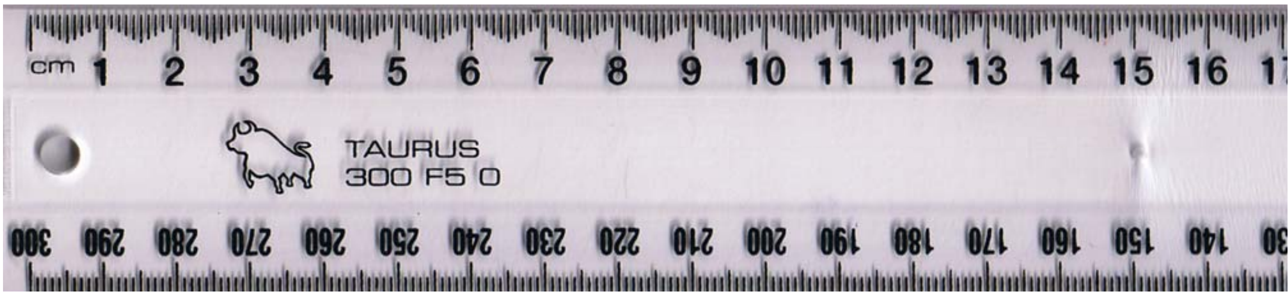


Figure 13.4. Part of a standard school ruler

Figure 13.5 represents a section of such a ruler with its compound scale. Here, two different scales operate simultaneously within the same space, a centimetre scale and a millimetre scale. A young seven year old student, possibly meeting a scale for the first time and unfamiliar with the relationship between millimetres and centimetres, may come to accept that the numbers on such a ruler are not always placed in counting order. Another student who predominantly solves numerical problems by counting may learn to deal with the ‘small pieces’ by ‘counting in points’.

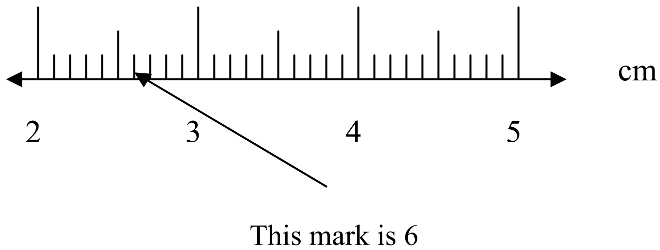


Figure 13.5. Ruler with a compound scale

Such possibilities help explain why some students were comfortable with the idiosyncratic placement of numbers illustrated in Figure 13.6. In responding thus students are not faced with contradictions of their schema as this is how they understand that such instruments work – two independent systems that involve different counts rather than two interrelated systems. In this schema, fractions have no place on the scale, which helps explain why so many students were not able to locate a number like ¾ between zero and one – including some of those who could read the number as three quarters.

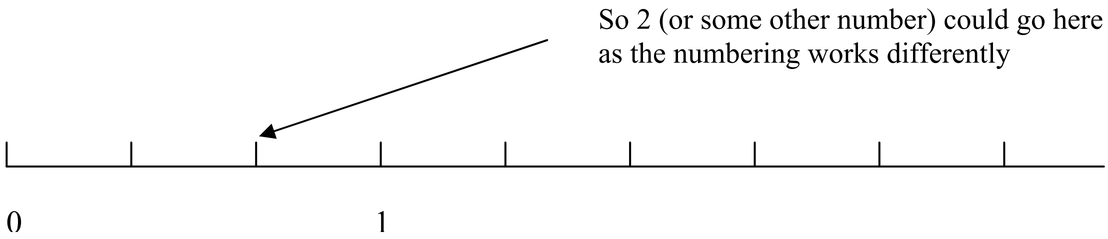


Figure 13.6. Misinterpretation of a fractional scale

This hypothesis suggests that the learning sequence is at the core of the problems some students have with understanding not only the ruler but also scales more generally; using a standard school ruler too early can lead to poor conceptions forming. Perhaps a way forward is to change the learning sequence for developing an understanding of measurement by initially focusing on creating and learning to use a number of simple scales that include zero. Each developed scale would have a different unit. For example, big items and distances are measured with a scale based wholly on metres or some class-constructed unit. Once such scales are understood then the focus could shift to measuring more accurately and working between the marks on the scale. At this stage an emphasis on the interval nature of the measurement could link to the development of student understanding of addition and subtraction. This would allow meaning to develop for difference or subtraction based strategies for working with 'broken' rulers. In the next section, these ideas are explored in more detail.

13.3 Developing measurement: a revised model

When evaluating a curriculum it is useful to review the learning progressions that are embedded in that curriculum. It follows that it will also be useful to identify how that learning progression may be changed to better reflect student understanding and development of that understanding. In Chapter 4 a common conceptual model for developing measurement was introduced (Zevenbergen et al., 2004) and was identified as being similar to the progression implied by the MiNZC curriculum statement (Ministry of Education, 1992). At that time it was suggested that the MiNZC curriculum statement did not adequately address the development of an understanding of scale, upon which an understanding of measurement relies. The findings from this research provide evidence to support the need for a better defined model, one that addresses the need to develop the forms of (non-unit) scale commonly used in measuring attributes other than length; for example, in weighing and when measuring temperature or capacity. The curriculum evaluation, the reviewed areas of scholarship, the table of differences between measurement and counting (see Table 3.2), and the concept map developed for this research (see Figure 11.21) provided sufficient guidance on issues and potential remedies for such a revised model to be developed. This is shown in Table 13.1.

Table 13.1: Revised model for developing measurement concepts

Stage	Important aspects to highlight	Use of number
Identify the attribute.	• Work out what is to be measured.	None.
Compare and order. Initially directly, later indirectly.	• Identify what can happen. • Develop/agree some language to explain what you see. • Where appropriate, develop the need for a common baseline.	None.
Use non-standard units to compare.	• Put numbers to the measurement. • Develop the concept of the unit. • Where appropriate, develop the concept of unit iteration – no gaps and no overlaps. • Counting to the nearest whole number to identify the total number of units (for example nearly four).	Counting numbers.
Develop standard units.	• Adopt standard units to allow communication about measurements. • Size of the unit is related to the context. • Develop a 'sense' of common units. • Measurement through counting, but to the nearest half.	Counting numbers.
Create a scale.	• The unit as an interval. • Development of the unit scale. • Explicitly highlight the change from counting to measurement conventions, and the reasons for this change. • The role of zero on the scale. • Units can be changed (reunitising).	Numbers as measurements.
Use standardised scales.	• Develop an aggregated unit scale. • Work with partitioned intervals in an aggregated unit scale (identifying the value of the marks). • Continuity of the number system – fractions and decimals live between the whole numbers. • Subdivision or partitioning of the interval/unit to create whole number, fractional, or decimal answers. • Conversion between the standard metric units.	• The continuous number system. • Multiplication to convert between units (so 62mm = 6.2cm).
Multi-conceptual thinking.	• Recognise the existence of different measuring systems. Work within and translate between systems as appropriate. • Use of real measurement instruments with alternate scales.	Multiple measurement concepts – including conversion factors between units in different systems.

The early stages of this conceptual model are little changed, though there is more detail. For example, when applying this model to linear measurement, the processes based on iterating non-standard units are now firmly tied to counting, so students are expected to undertake tasks like measuring the width of a table using their hands. (Because it is an object being counted (the hand), it cannot be assumed the concept of an interval has developed.) To establish the need for a consistent counting method and standard units, discussion of the different answers reached by children is required (e.g., see Stephan et al., 2001). At this stage, several standard units can be developed: a big unit to measure big things, smaller units to measure small things. Later, students can learn to partition the unit, initially with halves, then quarters.

The stage 'create a scale' is a subdivision of 'creating standard units'. It is separated as scales are used to measure many concepts – time, temperature, mass, speed, voltage, amount of petrol in a tank, engine revs, capacity – as well as length. For each, once a standard unit is established a scale is developed to measure the concept. For length it is at this stage that, as suggested by Rangecroft (1994) in relation to statistical graphs, counting conventions should be firmly left behind and measurement conventions adopted and explained. This emphasises the concept of the unit as an interval, placing labels on the marks, and having a starting point for the measurement called zero. Students also should make (and use) their own scales if they are to understand how scales operate, and once they have practice doing this they can also learn to use the different units to measure more accurately, creating answers like 4cm and 6mm. Such reunitising is an early problem-solving process with fractions (Mack, 1993, 1995), and is found with money, where $2.50 can be seen as 2 whole dollars and 50 whole cents with the decimal point acting as a separator between whole number systems (Swan, 1983). Reunitising is common practice in measurement, so the revised model acknowledges its use, and its importance as a strategy to avoid working with fractions or decimals in the early stages of learning measurement.

The second departure is 'use standardised scales'. Many commonly used scales, like kitchen scales, do not use unit scales, which the literature identifies causes problems for students (Dunham & Osborne, 1991, for algebraic graphs; Hart (1981) with fractions and decimals; Rangecroft (1994) for statistical graphs). This study also identifies that non-unit scales link to multiplicative thinking, and that a number of factors impact on students' ability to use these scales.

The stage 'use standardised scales' suggests that the multiplicative process of converting between units and creating decimal answers is left until students have some understanding of decimals and multiplication. This resolves the problem identified in the present research and other studies that many students (even at Years 7 and 8) show little understanding of a partitioned unit so 'count in points' (e.g., Swan, 1983). These issues suggest explicit instruction is needed to support the move from unit scales.

The final stage involves *multi-conceptual thinking*. *Multi-concept alignment* is a process where one set of concepts is aligned to another set (Wang, Isaac, van der Meij, & Schlobach, 2007). Multi-conceptual thinking is a related idea but requires an understanding of the alignment between the concepts. In terms of mathematics it involves identifying and accepting the existence of multiple mathematical realities, and being able to work in, and convert between these realities as appropriate. For example, in mathematics it is important to recognise that the conceptual systems of decimals, fractions, and percentages (while coherent on their own) are all able to provide different representations of the same number. It is also usually beneficial to use the conceptual system that makes communicating or working with the numbers in a particular problem easiest. For this to occur, a person first needs to understand each of the individual conceptual systems (have a well-developed schema for each), recognise that these align, and be able to convert from one to another. That this is a high level skill is supported by the results of this study; only the most able were able to recognise that both fractions *and* decimals are found between the whole numbers. A number of students were only comfortable with working within one system (often decimals), even though this appeared to create logical inconsistencies.

Multi-conceptual thinking is an important skill for working with many commercially produced measurement instruments, and it is at this stage that students become confident and competent users of scales to solve measurement problems. Multi-conceptual thinkers can do more than simply read a metric weighing scale. They can look at a commercial instrument and recognise that there are two alternatives provided (e.g., kilograms and pounds), and understand that these are appropriate in different cultures. They can use either unit/scale as appropriate, can work out what the divisions within each scale represent, and convert not only between units on a particular scale but also *between* one scale and the other. They recognise the equivalence between the two scales and are comfortable working with both, which is why they are capable of using commercial measuring instruments effectively.

13.3.1 Uses of the revised model

This model aims to help establish a mental structure (schema) that can be used when developing a much wider range of measurement processes than traditionally considered. As such it conforms to the requirements of Skemp (1971) who states that when we develop a mental structure to work within we create a mental tool for applying the same approach to future tasks.

This model can also be applied at two different levels. On one level it helps establish learning progressions for the development of measurement concepts. On another, it helps teachers structure the tasks that they develop for students so the potential for learning about measurement is maximised. At this level, the model is not supposed to be a rigid structure, rather a lens through which tasks can be interpreted. Stages should be merged or omitted so a framework that suits a particular task is developed.

An example of how this conceptual model can be applied is outlined below. It is provided to demonstrate how creating and using this model supports, among other things, the development of a unified concept of scale. The teaching of statistics is an area of the curriculum where the measurement aspect of the tasks has often been overlooked, so the example is from that strand.

The revised model in practice: investigating foot size

The model suggests that the task of investigating foot size is fundamentally a measurement one as it involves the development (or use) of a scale. Here, however, it needs to be noted that to maximise the learning from its measurement aspects, it is important for students to collect sufficient data to work with non-unit scales on any graph they produce.

Consider the attribute
When considering foot size it is important to ask why the information is being gathered as this helps frame the question being asked. It can also help decide how foot size will be measured. For example, the length, the width, the length and width, the foot area, or shoe size can all be used to measure feet. Considering the purpose of the investigation will also help establish whose feet should be measured.

Compare and order
Here the measurement strategy is devised. For example, when measuring length a standard start point (zero) is created and measurement is from this point, so for foot length this may involve standing (with shoes off) with back and heels to a wall.
At this stage it is also important to develop the language to discuss foot size and the measuring of feet.

Decide what units to use
If measuring length, students could choose to make their own unit, or choose one of the existing international units (European, UK, or US).

Create a scale
To measure foot length, a ruler that uses the chosen scale is needed.

Use standardised scales
Groups compare what they have done, and discuss results. At this point, good statistical discussions would be that different samples would produce different sets of results. Good measurement discussions would involve the need for each group to use a similar measurement system so the results can be compared and generalisations made.

Multi-conceptual thinking
Learn to translate between systems (e.g., the systems students have used in class).
Create a scale converter (e.g., a double number line or a conversion table).
Students could be challenged to explore the different international systems for measuring foot size, their history, and be required to create scale converters.

Note that the model can also be applied to the graphing aspect of result reporting, which in this case is the second level of application for the model. The following discussions or actions can be applied to the foot length example.

Compare and order
Decide what will be shown on each axis of the graph and consider the spread of data that needs to be shown.
Group the data appropriately.

Units and scale
Decide a scale for each axis (and the size of the interval for each unit). 'What is an appropriate scale to show all this information?'
Use appropriate measurement conventions to create the chosen scale.

> **Use standardised scales**
> Discuss the need for standard units and identical scales if graphs are to be compared. An important part of this stage is to discuss how the picture a graph shows can change if the data are regrouped, or the interval size or scales are changed.

13.3.2 The revised model and the revised curriculum

In Chapter 1 it was noted that a key aim of this research was to evaluate the effectiveness of MiNZC for teaching scale and use this knowledge to inform the development of the revised mathematics curriculum. To this end a short submission based on the early findings was developed and submitted as feedback on the curriculum draft (Ministry of Education, 2006). This was published on the world wide web with other 'long submissions' and is retrievable from http://www.tki.org.nz/r/nzcurriculum/docs/feedback/sec14.pdf. This contained a commentary on the proposed measurement progression, along with suggested changes, based on the model in Table 13.1.

13.4 Implications

While this study has focused on issues related to the understanding of scale in the context of mathematics, the findings may be of wider interest. For example, Chapter 4 showed that scale is also used in technology education as well as in the physical and social sciences.

This study has identified that not all students develop an understanding of scale through working with scale in situations involving measurement and graphing. It has also identified a series of understandings about scale not previously reported in the research literature. As noted in Chapter 4, MiNZC (Ministry of Education, 1992) has not required mathematics teachers to focus explicitly on developing an understanding of scale. Neither did the reviewed resources highlight aspects of scale to which teachers should attend.

> A teacher with the intention of developing student understanding will not necessarily produce the desired effect. Unless teachers make good sense of the mathematical ideas, they will not develop the flexibility they need for spotting the golden opportunities and wise points of entry that they can use for moving students towards more sophisticated and mathematically grounded understandings. (Anthony & Walshaw, p. 206)

It seems likely that a lack of teacher knowledge about how students develop an understanding of scale, and what students know about scale, has contributed to the state of knowledge identified in this study. This lack of knowledge may have had an impact on the learning of students in a number of curriculum areas, especially where it is assumed that students come to the subject with an existing understanding of scale which is then expected to be applied to subject-related learning. The findings of this study therefore have implications for the future teaching, learning, and assessment of scale. These implications are addressed in the next sections.

13.4.1 Towards an assessment for scale

This study has shown that when responding to a question involving a scale a correct answer on its own does not provide information about a student's solution strategy. The analysis of the interviews also identified that students may not have adequate language to discuss scale. This suggests that a teacher wishing to ascertain a student's understanding of scale should initially use a diagnostic interview (although several items have been identified that are appropriate to give in written form). However, when using this form of assessment, it is important to recognise that a student's physical strategy (what they do with their hands, lips, etc.) sometimes does not match their explained strategy. Often the explained strategy is their final strategy, rather than an accurate report of what they did. Further probing may be needed to properly identify the thinking used and the range of strategies attempted, and the responses to some items may have to be cross-referenced with others. The teacher trials showed that it is possible to collect strategy-based information from a written test that requires an explanation of strategy to be given. It is therefore possible that once a language of scale and an understanding of the construct have been developed, a form of written assessment might be constructed to elicit similar information to a diagnostic interview.

In the process of identifying student understanding at the case study schools, a number of items were developed and findings were made relating to effective practice for assessing scale understanding. These can be used to inform the future development of a diagnostic interview for scale so are summarised below.

- The majority of the items developed for the assessments used in this study provided good discrimination between the correct answer for an item and the total in the test. The exceptions were some of the items with a high facility, and items requiring more spatial ability. However, the spatial ability to iterate a unit is an important aspect of working with unmarked intervals, so needs to be assessed.
- Both number lines and scales presented in 'familiar' contexts are effective when investigating student understanding. The items involving contexts rarely evoked strategies not used on number lines. What seemed to happen is that some students used different strategies on different problems, and that the use of a context or a number line could cause a strategy change.
- Lines with marked and unmarked intervals generally evoked different solution strategies, so an effective assessment requires items that utilise both of these types of scale.
- Items in which students were asked to read a scale were generally more effective at identifying understanding than items asking students to locate a number. This second form of item has the potential to eliminate strategies a student may naturally wish to use.
- Intervals partitioned into both four and five are needed to identify whether or not some students have an understanding of the strategy they use, or are simply reciting a count.
- Items based on scales can be a good test of number understanding.

13.4.2 Towards more effective teaching and learning for scale

The identification of the mental strategies used when answering scale-related problems can be seen as the first step in an iterative process designed to improve student understanding of scale. A number of suggestions that may improve teaching and learning have also been made. These were based on both the reviewed literature and the findings of this study. These suggestions could form the basis of resources and a teaching experiment that focuses on improving student understanding of scale amongst Year 7 and 8 students so are summarised below.

- Students need to be explicitly taught about scale.
- Scales are a common measurement representation. To use them requires measurement conventions, and an understanding that measurements have a start (usually called zero) and a finish.
- Students need a language they can use to talk and think about scale.
- Number lines remove the complexity of context. Simple number lines could make the scale construct more understandable and accessible to students as they are a mathematical situation that can be seen and meaningfully discussed by students.
- The revised model for developing an understanding of measurement concepts suggests that the development of the concept of a scale in the context of measurement needs to take place over a number of years.
- While short interventions may be effective in addressing particular issues, some issues relating to learning about scale involve the development of a different understanding of number so require a transformation in thinking. For example, the responses to fraction items suggest that some students need to understand that many fractions consist of *three* numerals, with a whole number and a fractional part, as well as to develop the understanding that *both* fractions and decimals are found between the whole numbers.
- When working with a marked interval, students need to learn to count the marks, and when working with unmarked intervals to use plentiful ‘working’ marks. Students need to be exposed to both situations.
- Students could benefit from learning about coordinates and the process of plotting points in an algebraic context before meeting bivariate data in graphing situations. This would allow them to learn that both of the axes on a graph can act as a scale, that conventions for plotting points should be followed, and that the numbers in a pair cannot be separated. This environment could also be used to explore the impact that changing the numbering and unit intervals has on the visual image (i.e., that a graph is a scaled representation).

The findings of this study can also be used by current teachers of Years 7 and 8 seeking to identify and address the problems they see students facing. Firstly, observing students trying to answer individual items can help a teacher identify the issues that a student is having with scale (e.g., when trying to work out how much water a measuring jug has in it). Secondly, understanding and following the progressions in the strategy hierarchy and the model for developing measurement processes may enable teachers to make measurement more accessible to students. For example, when working with marked intervals teachers could use the following approaches for students at the different stages.

- *For students whose conceptual understanding of number is restricted to counting.*
  Emphasise that all measuring processes have a start and a finish, that we use the number zero to show the start place, and that between the start line and the finish we include all 'units' in the count. If the students already have this understanding, teachers could encourage students to construct and use simple unit scales like a paper ruler and emphasise counting marks from the start point (0) to the finish point. They could also provide situations which naturally lead students to wanting to develop 'different sized units for different sized objects' and they could develop greater accuracy from students as they read scales by establishing the concept of half-way as the midpoint of a unit.
- *For students whose conceptual understanding of number involves adding and subtracting.*
  Use non-unit scales and highlight the need to count how many pieces each interval has been cut into. They could emphasise the use of skip counting to fill the gap, and could get students to check that their count finishes with the right number at the end of the interval. Teachers could also encourage students to notice any differences in the marks, and to reflect on how this might provide scaffolding information that might help interpret the scale.
- *For students whose conceptual understanding of number includes multiplication and division*.
  Emphasise the need for students to count both the number of marks, and calculate the width of the interval, then use either multiplication or division to establish the number associated with each mark.

13.4.3 Unintended consequences? Towards more effective curriculum change

In pre-decimal and pre-calculator days, more time was spent on developing the ability to work with scale, as it was a key part of the slide rule, an important mathematical tool. Rulers were also marked in inches, which were subdivided into halves, quarters, and eighths, with these fractions being located between the whole numbers. Fifths and tenths were on the other side of the ruler. Developing an understanding of these partitions, starting with a half, was a central part of measurement. Work was also done with fractions and decimals on number lines.

Under 'new maths' Ernest (1985) identifies that number lines were used as a teaching aid for three major purposes:

- as a model for ordering numbers;
- as a model for the operations; and
- as an item of content in itself.

A review of some of the standard primary text books for Years 3 and 4 from this era shows that common use was made of measurement-based representations such as cuisenaire, number tracks with rods, and number lines. However, the number lines and other linear models tended to mix how operations were represented; sometimes starting from zero, and sometimes ‘counting on’ from the location of the first number. They were also used alongside (and were treated as being interchangeable with) set models based on counting from one (Department of Education, 1974a, 1974b, 1983). This suggests that some students may have been confused by the range of representations, and the lack of clear conventions for their use. This may account for Ernest’s comment that research suggested the number line was not very useful for modelling the operations, but was a direct contributor to linear measurement and reading scales, and to graphical work as axes in statistics and algebra.

MiNZC (Ministry of Education, 1992) reflected a new era, based on the decimal system and modern technology, and used a problem-solving approach with meaningful contexts for mathematical learning. The concept of a number line did not fit with this approach. However, students were still expected to develop the same skills for using scales but from experiences directly involving measurement, graphing, and scale drawing (see Chapter 4). Perhaps one of the unintended side-effects of this curriculum change was to reduce understanding of scale and aspects of mathematics based upon it. If so, this provides a strong argument for basing curriculum change on evidence of what is known to be effective in the teaching and learning of mathematics, rather than what might conceivably be effective.

This research has focused upon the understanding that Year 7 and 8 students have of scale. In the next section recommendations for follow-up work are made.

13.5 Recommendations

1) Further research should be undertaken into how students develop an understanding of scale.

This study focused on the understanding of students in Years 7 and 8. A study of the understanding of students in Years 1 to 6 could challenge the findings made here, or help develop a detailed learning progression for students.

2) A diagnostic interview to assess scale understanding should be developed, and psychometrically tested.

A selection of the items used in this study could form the starting point for this development.

3) Because a better understanding of scale could improve students’ understanding of number, measurement, and graphing, a framework should be developed to support the teaching of scale.

A draft of this framework could be linked to the number framework (Ministry of Education, 2005a) and be based on the ‘typical student strategies’ used by students in this research (Tables 12.2 and 12.3) although it should be noted that not all of the strategies are ones that should be encouraged. The model for developing measurement concepts (Table 13.1), and the results from individual items in this study could also inform such a framework which could then be empirically tested through the development of teaching interventions.

4) Teaching resources that focus on developing scale understanding should be developed and trialled.

This research indicates a number of directions that, if followed, may improve student understanding of scale. For example, a measurement-based ‘count the marks’ approach with an emphasis on students learning ‘how many cuts make how many pieces’ could be adopted to strengthen understanding of scales involving multiples or fractions. The effectiveness of the resources based on such approaches could then be evaluated experimentally.

5) The findings of this study should be made available to teachers in an accessible form.

As scale underpins a significant body of mathematics that seems to be poorly understood and not well taught, professional development should be given to teachers in schools. Putting this in place could be one of the next foci for educational professionals working within the NDP. The support material for the mathematics learning area of the revised curriculum should also explicitly state what students need to learn and understand about scale, and by what curriculum level they need to know this. Suggestions about how they may come to know this also need to be made.

13.6 Final words

In conclusion, this study has shown that scale and the learning of scale has been taken for granted, and that it is not sufficient to assume that students develop an understanding of scale through exposure to situations involving measurement and graphing. Rather, it is important to remember that a scale is not a simple object, nor a single object, but a common mathematical tool whose great mathematical power stems from its flexibility and variability. The results and findings from this study clearly demonstrate that scale is a complex tool requiring the application of multiple skill sets in a coordinated manner. This means that while it is appropriate to identify that all scales follow a particular set of rules and to talk about ‘the scale on a thermometer’, students also need to understand that different thermometers can have different scales and that when working with a particular scale they must first work out how that scale has been constructed. Such complexity should be argument enough for the need for curriculum developers, professional development designers, and teachers to take time to nurture and encourage the level of understanding required for using scale successfully.

References

Adjiage, R., & Pluvinage, F. (2007). An experiment in teaching ratio and proportion. *Educational Studies in Mathematics, 65,* 149-175.

American Education Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington DC: American Education Research Association.

Anthony, G., & Walshaw, M. (2007). *Effective pedagogy in mathematics/pāngarau*. Wellington: Ministry of Education.

Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Melbourne: Australian Education Council and Curriculum Corporation.

Barcham, P. (1996). Student interpretation of graphs. In H. Forgasz, T. Jones, G. Leder, J. Lynch, K. Maguire & C. Pearn (Eds.), *Mathematics, making connections: MAV 1996*. Brunswick: Mathematical Association of Victoria.

Baturo, A., & Cooper T. (1999). Fractions, reunitisation and the number line representation. In O. Zaslavsky (Ed.), *Proceedings of the 23rd conference of the International Group for the Psychology of Mathematics Education.* Rehovot: Department of Science Teaching, The Weizmann Institute of Science.

Beatty, P. (2004). The dynamics of cognitive interviewing. In S. Presser, J. Rothgeb, M. Couper, J. Lessler, E. Martin, J. Martin et al. (Eds.), *Methods for testing and evaluating survey questionnaires.* Hoboken, NJ: John Wiley.

Behr, M., Lesh, R., Post, T., & Silver, E. (1983). Rational number concepts. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp.92-127). New York: Academic Press.

Berg, B. (2007). *Qualitative research methods for the social sciences* (6th ed.). Boston: Pearson Educational.

Bishton, S. (1998). *Interpretation and clarification of "what is a problem".* Wellington: NZQA.

Boaler, J. (1993). The role of contexts in the mathematics classroom: Do they make mathematics more "real"? *For the Learning of Mathematics, 13*(2), 12-17.

Bobis, J., Mulligan, J. & Lowrie, T. (2004). *Mathematics for children* (2nd ed.). Frenchs Forest, NSW: Pearson Education Australia.

Bragg, P., & Outhred, L. (2000a). What is taught versus what is learned: The case for linear measurement. *Mathematics education beyond 2000: Proceedings of the 23rd annual conference of the Mathematics Education Research Group of Australasia, Freemantle, Western Australia*. Perth: MERGA.

Bragg, P., & Outhred, L. (2000b). Students' knowledge of length units: Do they know more than rules about rulers? In T. Nakahara & M. Koyama (Eds.), *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education* (pp.97-104). Hiroshima: Hiroshima University.

Bright, G., Behr, M., Post, T., & Wachsmuth, I. (1988). Identifying fractions on number lines. *Journal for Research in Mathematics Education*, *19*(3), 215-232.

Britt, M., & Irwin, K. (2008). Algebraic thinking with and without algebraic representation: A three year longitudinal study. *ZDM Mathematics Education, 40*(1) 39-53.

Brown, M. (1981). Place value and decimals. In K. Hart (Ed.), *Children's understanding of mathematics* (pp. 11-16). Newcastle upon Tyne: Athenaeum Press.

Carr, K., & Katterns, B. (1984). Does the number line help? *Mathematics in School, 13*(4), 30-34.

Chamberlain, M., & Walker, M. (2001). *Trends in year 9 students' mathematics and science achievement. Results from New Zealand's participation in the repeat of the Third International Mathematics and Science Study.* Wellington: Ministry of Education.

Cobb, P., Yackel, E., & McClain, K. (2000), *Symbolizing and communicating in mathematics classrooms*. Mahwah, NJ: Lawrence Erlbaum.

Cockcroft, W. (1982). *Mathematics counts*. London: DES/HMSO.

Cohen, L., & Manion, L. (1994). *Research methods in education* (4th ed.). London: Routledge.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London: RoutledgeFalmer.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Oxon: Routledge.

Cooper, B., & Dunne, M. (2000). *Assessing children's mathematical knowledge: Social class, sex and problem solving.* Buckingham, UK: Open University Press.

Costello, J. (1991). *Teaching and learning mathematics: 11-16*. London: Routledge.

Crooks, T., & Flockton, L. (1996). *Graphs, tables and maps assessment results 1995.* Dunedin: University of Otago, Educational Assessment Research Unit.

Crooks, T., & Flockton, L. (2000). *Graphs, tables and maps assessment results 1999.* Dunedin: University of Otago, Educational Assessment Research Unit.

Crooks, T., & Flockton, L. (2002). *Mathematics assessment results 2001.* Dunedin: University of Otago, Educational Assessment Research Unit.

Davidson, E. (2005). *Evaluation methodology basics*. Thousand Oaks, CA: Sage.

Denscombe, M. (1998). *The good research guide for small scale social research projects*. Buckingham, UK: Open University Press.

Department of Education. (1974a). *School mathematics: Book 1A*. Wellington: Government Printer.

Department of Education. (1974b). *School mathematics: Book 1B*. Wellington: Government Printer.

Department of Education. (1983). *School mathematics 2*. Wellington: Government Printer.

Drake, M. (2007). Informal knowledge and prior learning: Student strategies for identifying and locating numbers on scales. In J. Watson & K. Beswick (Eds.), *Mathematics: Essential research, essential practice: Volume 1. Proceedings of the 30th annual conference of the Mathematics Education Research Group of Australasia.* Adelaide: MERGA.

Dufour-Janvier, B., Bednarz, N., & Belanger, M. (1987). Pedagogical considerations concerning the problem of representation. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics.* Hillsdale, NJ: Lawrence Erlbaum.

Dunham, P., & Osborne, A. (1991). Learning how to see: Students' graphing difficulties. *Focus on Learning Problems in Mathematics*, *13*(4), 35-49.

Eather, J. (2010). *A mathematics dictionary for kids.* Retrieved February 8, 2010, from http://www.amathsdictionaryforkids.com

Education Review Office. (1996). *Improving schooling in Mangere and Otara.* Education Evaluation Reports, Crown.

Education Review Office. (1997). *Improving schooling on the east coast.* Education Evaluation Reports, Crown.

Education Review Office. (1998a). *Schooling in the far north.* Education Evaluation Reports, Crown.

Education Review Office. (1998b). *Schooling in Mangere and Otara: Progress since 1996: Follow-up report.* Education Evaluation Reports, Crown.

Education Review Office. (2000). *In time for the future: A comparative study of mathematics and science education.* Education Evaluation Reports, Crown.

Education Review Office. (2002). *Maths now: Mathematics education for students in years 1 to 8.* Education Evaluation Reports, Crown.

Ernest, P. (1985). The number line as a teaching aid. *Educational Studies in Mathematics, 16*(4), 411-424.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.

Flockton, L., & Crooks, T. (1998). *Mathematics assessment results 1997.* Dunedin: University of Otago, Educational Assessment Research Unit.

Flockton, L., Crooks, T., & Gilmore, A. (2004). *Graphs, tables and maps assessment results 2003.* Dunedin: University of Otago, Educational Assessment Research Unit.

Fowler, H. & Fowler, F. (1964). *The concise Oxford dictionary* (5th ed.). Oxford: Clarendon.

Freudenthal, H. (1983). *Didactical phenomenonology of mathematical structures*. Dordrecht, Holland: D. Reidel.

Friel, S., Curcio, F., & Bright, G. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 23*(2), 124-158.

Garden, R. (Ed.). (1996). *Mathematics performance of New Zealand form 2 and form 3 students*. Wellington: Ministry of Education.

Garden, R. (Ed.). (1997). *Mathematics and science performance in middle primary school.* Wellington: Ministry of Education.

Goldenberg, E. (1988). Mathematics, metaphors, and human factors: Mathematical, technical and pedagogical challenges in the educational use of graphical representation of functions. *Journal of Mathematical Behaviour, 7,* 135-173.

Goldin, G., & Shteingold, N. (2001). Systems of representation and the development of mathematical concepts. In A. Cuoco & F. Curcio (Eds.), *The roles of representation in school mathematics: 2001 yearbook.* Reston, VA: NCTM.

Gould, P., Outhred, L., & Mitchelmore, M. (2006). One-third is three-quarters of one-half. In P. Grootenboer, R. Zevenbergen & M. Chinnappan (Eds.), *Identities, cultures and learning spaces: Proceeding of the 29th annual conference of the Mathematics Education Research Group of Australasia.* Adelaide: MERGA.

Government Social Research Unit. (2007). *The magenta book: Guidance notes for policy evaluation and analysis. Background paper 1: What is policy analysis?* London: HM Treasury.

Hall, C. (2000). National Certificate of Educational Achievement: Issues of reliability, validity and manageability. *New Zealand Annual Review of Education, 9,* 173-196.

Hall, C. (2002). *Evaluation as a method of educational enquiry*. Wellington: Victoria University of Wellington, School of Education (unpublished class handout).

Hall, C. (2007a). *Normative and interpretive research paradigms*. Wellington: Victoria University of Wellington, School of Education Studies (unpublished class handout).

Hall, C. (2007b). *Quantitative research designs*. Wellington: Victoria University of Wellington, School of Education Studies (unpublished lecture notes).

Hart, K. (1978). Mistakes in mathematics. *Mathematics Teaching, 85, 38*-40.

Hart, K., & Johnson, D. (1980). *Secondary school children's understanding of mathematics: A report of the mathematics component of the concepts in secondary mathematics and science programme*. London: Mathematics Education Centre for Science Education, Chelsea College, University of London.

Hart, K. (1981). *Children's understanding of mathematics: 11-16*. Newcastle upon Tyne: Athenaeum Press.

Hart, K. (1989). Fractions: Equivalence and addition. In D. Johnson, (Ed.), *Children's mathematical frameworks 8-13: A study of classroom teaching.* Windsor, England: Nfer-Nelson.

Hiebert, J. (1984). Why do some children have trouble learning measurement concepts. *Arithmetic Teacher,* March, 19-24.

Higgins, J. (1999). Pedagogical content knowledge in mathematics. *Exploring issues in mathematics education*. Wellington: Ministry of Education.

Higgins, J. (2001). *An evaluation of the year 4-6 numeracy exploratory study*. Wellington: Ministry of Education.

Higgins, J. (2002). *An evaluation of the Advanced Numeracy Project 2001*. Wellington: Ministry of Education.

Higgins, J. (2003). *An evaluation of the Advanced Numeracy Project 2002*. Wellington: Ministry of Education.

Higgins, J., Bonne, L., & Fraser, K. (2004). *An evaluation of the Advanced Numeracy Project 2003*. Wellington: Ministry of Education.

Higgins, J., Irwin, K., Thomas, G., Trinick, T., & Young-Loveridge, J. (2005). *Findings from the New Zealand Numeracy Development Project 2004*. Wellington: Ministry of Education.

Hunting, R. (1989). Beginning to learn fractions. In I. Livingstone & J. Izard (Eds.), *Best of set: Mathematics.* Wellington: NZCER.

Irwin, K (2003). *An evaluation of the Numeracy Project for years 7-10, 2002*. Wellington: Ministry of Education.

Irwin, K. (2004). *An evaluation of the Numeracy Project for years 7-9, 2003*. Wellington: Ministry of Education.

Irwin, K., & Britt, M. (2005). Algebraic thinking in the numeracy project: Year one of a three-year study. In J. Higgins, K. Irwin, G. Thomas, T. Trinick & J. Young-Loveridge, (Eds.), *Findings from the New Zealand Numeracy Development Project 2004*. Wellington: Ministry of Education.

Irwin, K., & Ell, F. (2002). Visualising and the move from informal to formal linear measurement. In B. Barton, K. Irwin, M. Pfannkuch & M. Thomas (Eds.), *Mathematics Education in the South Pacific*: *Proceedings of the 25th annual conference of the Mathematics Education Research Group of Australasia, Auckland* (pp.358-365). Sydney: MERGA.

Irwin, K, & Niederer, K. (2002). *An evaluation of the Numeracy Exploratory Study (NEST) and the associated Numeracy Exploratory Study Assessment (NESTA) years 7-10, 2001*. Wellington: Ministry of Education.

Izsák, A., Tillema, E., & Tunç-Pekkan Z. (2008). Teaching and learning fraction addition on number lines. *Journal for Research in Mathematics Education, 39*(1) 33-62.

Janvier, C. (Ed.). (1987). *Problems of representation in the teaching and learning of mathematics.* Hillsdale, NJ: Lawrence Erlbaum.

Kaput, J. (1987). Representation systems and mathematics. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics.* Hillsdale, NJ: Lawrence Erlbaum.

Kaput, J. (1992). Technology and mathematics education. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*. New York: Macmillan.

Kaput, J. (1998). Representations, inscriptions, descriptions and learning: A kaleidoscope of windows . *Journal of Mathematical Behavior, 17*(2), 265-281.

Kerslake, D. (1981). Graphs. In K. Hart (Ed.), *Children's understanding of mathematics: 11-16*. Newcastle upon Tyne: Athenaeum Press.

Kieran, T. (1976). On the mathematical, cognitive, and instructional foundations of rational number. In R. Lesh, & D. Bradbard, (Eds.), *Number and measurement: Papers from a research workshop* (pp.101-144). Columbus, Ohio: ERIC/SMEAC.

Knight, G., & Meyer, D. (1996). *Critical factors in the implementation of the new mathematics curriculum*. Palmerston North: Massey University, Faculty of Information and Mathematical Sciences.

Kuchemann, D. (1981). Positive and negative numbers. In K. Hart (Ed.), *Children's understanding of mathematics: 11-16*. Newcastle upon Tyne: Athenaeum.

Lehrer, R., Jenkins, M., & Osana, H. (1998). Longitudinal study of children's reasoning about space and geometry. In R. Lehrer& D. Chazan (Eds.), *Designing learning environments for developing understanding of geometry and space.* Mahwah, NJ: Lawrence Erlbaum.

Lehrer, R. & Schauble, L. (2000). Modeling in mathematics and science. In R. Glaser (Ed.), *Advances in instructional psychology: Volume 5* e*ducational design and cognitive science*. Mahwah, NJ: Lawrence Erlbaum.

Lehrer, R., Schauble, L., Carpenter, S., & Penner. D. (2000). The interrelated development of inscriptions and conceptual understanding. In P. Cobb, E. Yackel & K. McClain, (Eds.), *Symbolizing and communicating in mathematics classrooms*. Mahwah, NJ: Lawrence Erlbaum.

Leinhardt, G., Zaslavsky, O., & Stein, M. (1990). Functions, graphs and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60*(1), 1-64.

Lesh, R., Behr, M., & Post, T. (1987). Rational number relations and proportions. In C. Janvier, (Ed.), *Problems of representation in the teaching and learning of mathematics.* Hillsdale, NJ: Lawrence Erlbaum.

Lesh, R., & Doerr, H. (2000). Symbolizing, communicating and mathematizing: Key components of models and modelling. In P. Cobb, E. Yackel & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms*. Mahwah, NJ: Lawrence Erlbaum.

Lesh, R., Post, T., & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier, (Ed.), *Problems of representation in the teaching and learning of mathematics.* Hillsdale, NJ: Lawrence Erlbaum.

McGatha, M., Cobb, P., & McClain, K. (1998, April). *An analysis of students' statistical understandings.* Paper presented to the annual meeting of the American Educational Research Association, San Diego, California.

Mack, N. (1993). Learning rational numbers with understanding: The case of informal knowledge. In T. Carpenter, E. Fennema & T. Romberg (Eds.), *Rational numbers: An integration of research*. Hillsdale, NJ: Lawrence Erlbaum.

Mack, N. (1995). Confounding whole-number and fraction concepts when building on informal knowledge. *Journal for Research in Mathematics Education, 26*(5), 422-441.

Mathematics and Science Taskforce. (1997). *Report of the mathematics and science taskforce*. Wellington: Ministry of Education.

Ministry of Education. (1992). *Mathematics in the New Zealand curriculum*. Wellington: Learning Media.

Ministry of Education. (1993a). *Science in the New Zealand curriculum.* Wellington: Learning Media.

Ministry of Education. (1993b). *The New Zealand curriculum framework.* Wellington: Learning Media.

Ministry of Education. (1995). *Technology in the New Zealand curriculum.* Wellington: Learning Media.

Ministry of Education. (1997). *Social studies in the New Zealand curriculum*. Wellington: Learning Media.

Ministry of Education. (1999a). *Figure it out: Statistics levels 2-3*. Wellington: Learning Media.

Ministry of Education. (1999b). *Figure it out: Statistics levels 2-3: Answers and teachers' notes*. Wellington: Learning Media.

Ministry of Education. (1999c). *Figure it out: Measurement levels 2-3*. Wellington: Learning Media.

Ministry of Education. (1999d). *Figure it out: Measurement levels 2-3: Answers and teachers' notes*. Wellington: Learning Media.

Ministry of Education. (1999e). *Figure it out: Algebra levels 2-3*. Wellington: Learning Media.

Ministry of Education. (1999f). *Figure it out: Algebra levels 2-3: Answers and teachers' notes*. Wellington: Learning Media.

Ministry of Education. (2000a). *Figure it out: Statistics level 3*. Wellington: Learning Media.

Ministry of Education. (2000b). *Figure it out: Statistics level 3: Answers and teachers' notes*. Wellington: Learning Media.

Ministry of Education. (2000c). *Figure it out: Measurement level 3*. Wellington: Learning Media.

Ministry of Education. (2000d). *Figure it out: Measurement level 3: Answers and teachers' notes*. Wellington: Learning Media.

Ministry of Education. (2000e). *Figure it out: Algebra level 3*. Wellington: Learning Media.

Ministry of Education. (2000f). *Figure it out: Algebra level 3: Answers and teachers' notes*. Wellington: Learning Media.

Ministry of Education. (2001a). *Figure it out: Statistics levels 3-4*. Wellington: Learning Media.

Ministry of Education. (2001b). *Figure it out: Statistics levels 3-4: Answers and teachers' notes*. Wellington: Learning Media.

Ministry of Education. (2001c). *Figure it out: Measurement levels 3-4*. Wellington: Learning Media.

Ministry of Education. (2001d). *Figure it out: Measurement levels 3-4: Answers and teachers' notes*. Wellington: Learning Media.

Ministry of Education. (2001e). *Figure it out: Algebra levels 3-4*. Wellington: Learning Media.

Ministry of Education. (2001f). *Figure it out: Algebra levels 3-4: Answers and teachers' notes*. Wellington: Learning Media.

Ministry of Education. (2001g). *The New Zealand curriculum exemplars: Mathematics*. Wellington: Learning Media.

Ministry of Education. (2004). *Book 9: Teaching number through measurement, geometry, algebra and statistics.* Wellington: Ministry of Education.

Ministry of Education. (2005a). *Book 1: The number framework*. Wellington: Ministry of Education.

Ministry of Education. (2005b). *Book 2: The diagnostic interview*. Wellington: Ministry of Education.

Ministry of Education. (2006). *The New Zealand curriculum: Draft for consultation 2006.* Wellington: Learning Media.

Ministry of Education. (2007). *The New Zealand curriculum.* Wellington: Learning Media.

Ministry of Education. (2008). *Intent: Statement of intent 2008-2013.* Wellington: Ministry of Education.

Mitchell, A. (2005). Measuring fractions. In P. Clarkson, A Downton, D. Gronn, M. Horne, A. McDonough, R. Pierce & A. Roche (Eds.), *Building connections: Theory, research and practice*: *Proceedings of the annual conference of the Mathematics Education Research Group of Australasia, held at RMIT, Melbourne 7th-9th July, 2005* (pp.545-552). Sydney: MERGA.

Moritz, J. (2003). Constructing coordinate graphs: Representing corresponding ordered values with variation in two-dimensional space. *Mathematics Education Research Journal, 15*(3), 226-251.

Ni, Y. (2000). How valid is it to use number lines to measure children's conceptual knowledge about rational number? *Educational Psychology*, *20*(2), 139-152.

Nisbet, S. (2002). Year seven students' representation of numerical data: The influence of sample size. In B. Barton, K. Irwin, M. Pfannkuch & M. Thomas (Eds.), *Mathematics education in the South Pacific: Proceedings of the 25th annual conference of the Mathematics Education Research Group of Australasia, Auckland* (pp.520-527). Sydney, MERGA.

Nisbet, S. (2003). Organising and representing grouped data. In L. Bragg, C. Campbell, G. Herbert & J. Mousley (Eds.), *Mathematics education research: Innovation, networking, opportunity: Proceedings of the 26th annual conference of the Mathematics Education Research Group of Australasia, Geelong.* Melbourne: Deakin University.

Nelson, D. (Ed.), (2003). *The Penguin dictionary of mathematics*. London: Penguin.

Nunes, T., & Bryant, P. (1998). *Children doing mathematics*. Massachusetts: Blackwell.

NZARE. (1998). *New Zealand Association for Research in Education, ethical guidelines*. Retrieved May 22, 2007 from http://www.nzare.org.nz/pdfs/NZARE_ethical_guidelines.pdf .

Olive, J., & Steffe, L. (2002). Schemes, schemas and director systems (an intergration of Piagetian scheme theory with Skemp's model of intelligent learning). In D. Tall & M. Thomas (Eds.), *Intelligence, learning and understanding in mathematics.* Flaxton, Queensland: Post.

Outhred, L., & McPhail, D. (2000). A framework for teaching early measurement. In *Mathematics education beyond 2000: Proceedings of the 23rd annual conference of the Mathematics Education Research Group of Australasia, Freemantle, Western Australia*. Perth: MERGA.

Owen, J. (2007). *Program evaluation: forms and approaches* (3rd ed.). New York: Guilford.

Parsons, R. (2005, March). *Numeracy development project: Scope and scale*. Paper presented at the Numeracy Development Project Reference Group, Wellington.

Pearn, C., & Stephens, M. (2004). Why you have to probe to discover what year 8 students really think about fractions. In I. Putt, R. Faragher & M. McLean (Eds.), *Mathematics education for the third millennium: Towards 2010: Proceedings of the 27th annual conference of the Mathematics Education Research Group of Australasia, Townsville*. Sydney: MERGA.

Pearn, C., & Stephens, M. (2007). Whole number knowledge and number lines help to develop fraction concepts. In J. Watson & K. Beswick (Eds.), *Mathematics: Essential research, essential practice: Volume 2. Proceedings of the 30th annual conference of the Mathematics Education Research Group of Australasia.* Adelaide: MERGA.

Pereira-Mendoza, L., & Mellor, J. (1991). Students' concepts of bar graphs: Some preliminary findings. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics: Vol. 1. School and general issues* (pp. 150-157). Voorburg, The Netherlands: International Statistical Institute.

Piaget, J., Inhelder B., & Szeminska A. (1960). *The child's concept of geometry*. London: Routledge & Kegan Paul.

Pirie, S. & Martin, L. (1997). The equation, the whole equation and nothing but the equation! One approach to the teaching of linear equations. *Educational studies in mathematics, 34*(2), 159-181.

Presser, S., Rothgeb, J., Couper, M., Lessler, J., Martin, E., Martin, J., & Singer E. (Eds.). (2004). *Methods for testing and evaluating survey questionnaires.* Hoboken, NJ: John Wiley.

Rangecroft, M. (1994). Graphwork: Developing a progression. In D. Green & Teaching Statistics Trust (Eds.), *Teaching statistics at its best: A series of articles for teachers.* Sheffield: Teaching Statistics Trust.

Robson, C. (2002). *Real world research* (2nd ed.). Oxford: Blackwell.

Sharma, S. (2003). High school students' interpretation of tables and graphs: Some findings from Fiji. In L. Bragg, C. Campbell, G. Herbert & J. Mousley (Eds.), *Mathematics education research: Innovation, networking, opportunity: Proceedings of the 26th Annual Conference of the Mathematics Education Research Group of Australasia, Geelong.* Melbourne: Deakin University.

Siegal, S., & Smith, J. (1997). Towards making representation count in children's conceptions of fractions. *Contemporary Educational Psychology, 1*(22), 1-22.

Skemp, R. (1971). *The psychology of learning mathematics.* Harmondsley, Middlesex: Penguin Books.

Stake, R. (1998). Case studies. In K. Denzin & Y. Lincoln (Eds.), *Strategies for qualitative inquiry.* Thousand Oaks, CA: Sage.

Stephan, M., Cobb, P., Gravemeijer, K., & Estes, B. (2001). The role of tools in supporting students' development of measurement conceptions. In A Cuoco & F. Curcio (Eds.), *The roles of representation in school mathematics: 2001 yearbook.* Reston, VA: NCTM.

Stevens, S. (1946). On the theory of scales of measurement. *Science 103*(2684), 677-680.

Swan, M. (1983). *Teaching decimal place value.* Nottingham: Shell Centre for Mathematical Education, Nottingham University.

Tall, D., & Thomas, M. (2002). Richard Skemp March 10, 1919 – June 22, 1995. In D. Tall & M. Thomas (Eds.). *Intelligence, learning and understanding in mathematics.* Flaxton, Queensland: Post.

Thomas, G., & Ward, J. (2001). *An evaluation of the Count Me In Too Pilot Project*. Wellington: Ministry of Education.

Thomas, G., & Ward, J. (2002). *An evaluation of the Early Numeracy Project*. Wellington: Ministry of Education.

Tipler, M., & Catley, K. (1998). *National curriculum mathematics level 3 book 1*. Christchurch: Caxton Educational.

Tipler, M., & Catley, K. (1999). *National curriculum mathematics level 3 book 2*. Christchurch: Caxton Educational.

Tipler, M., & Douglas, J. (2000a). *National curriculum mathematics level 2 book 1*. Christchurch: Caxton Educational.

Tipler, M., & Douglas, J. (2000b). *National curriculum mathematics level 2 book 2*. Christchurch: Caxton Educational.

Van den Heuvel-Panhuizen, M. (2005). The role of contexts in assessment problems in mathematics. *For the Learning of Mathematics, 25*(2), 2-9.

Wang, S., Isaac, A., van der Meij, L., & Schlobach (2007). *Multi-concept alignment and evaluation*. In *Second International Workshop on Ontology Matching, ISWC 2007, Busan, Korea.* Heidelberg: Springer.

Ward, M. (1979). *Mathematics and the 10-year-old*. Suffolk: Evans/Methuen Educational.

Watson, A. (2004). Red herrings: Post-14 'best' mathematical teaching and curricula. *British Journal of Educational Studies*, *52*(4), 359-376.

White, A. (1994). Hurdles: Who put them on the directed number line? *The Australian Mathematics Teacher, 50*(1), 14-18.

Wikipedia, (2008a). *The Beaufort scale*. Downloaded August 16, 2008, from http://en.wikipedia.org/wiki/Beaufort_scale

Wikipedia, (2008b). *Level of measurement*. Downloaded August 19, 2008, from http://en.wikipedia.org/wiki/Level_of_measurement

Wilkinson, C. (2000a). *Pearson primary maths 2a*. Auckland: Pearson.

Wilkinson, C. (2000b). *Pearson primary maths 2c*. Auckland: Pearson.

Wilkinson, C. (2000c). *Pearson primary maths 2: Teacher's guide*. Auckland: Pearson.

Wilkinson, C. (2001a). *Pearson primary maths 2b*. Auckland: Pearson.

Wilkinson, C. (2001b). *Pearson primary maths 3a*. Auckland: Pearson.

Wilkinson, C. (2001c). *Pearson primary maths book 2b: Teacher's guide*. Auckland: Pearson Education New Zealand.

Wilkinson, C. (2001d). *Pearson primary maths book 3a: Teachers guide*. Auckland: Pearson.

Wilkinson, C. (2001e). *Pearson primary maths 3c*. Auckland: Pearson.

Wilkinson, C. (2001f). *Pearson primary maths book 3c: Teacher's guide*. Auckland: Pearson.

Wilkinson, C. (2002a). *Pearson primary maths 3b*. Auckland: Pearson.

Wilkinson, C. (2002b). *Pearson primary maths book 3b: Teacher's guide*. Auckland: Pearson.

Willis, G. (2004). Cognitive interviewing revisited: A useful technique in theory. In S. Presser, J. Rothgeb, M. Couper, J. Lessler, E. Martin, J. Martin & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires.* Hoboken, NJ: John Wiley.

Wright, R. J. (1996). First-graders' progress in an experimental Mathematics Recovery program. In J. Mulligan & M. Mitchelmore (Eds.), *Children's number learning* (pp. 55-72). Adelaide: The Australian Association of Mathematics Teachers.

Wright, R. J. (1998). An overview of a research-based framework for assessment and teaching early number. In C. Kanes, M. Goos, & E. Warren (Eds.), *Teaching mathematics in new times: Proceedings of the 21st annual conference of the Mathematics Education Research Group of Australasia: Volume 2 (*pp. 701–708). Brisbane: MERGA.

Wright, R. J. & Gould, P. (2000). Reviewing literature relevant to a systemic early numeracy initiative: Bases of CMIT. In J. Bana & A. Chapman (Eds.), *Mathematics education beyond 2000: Proceedings of the 23rd annual conference of the Mathematics Education Research Group of Australasia: Volume 1* (pp. 56–63). Freemantle: MERGA.

Yin, R. (2003). *Case study research: design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Young-Loveridge, J. (2004). *Patterns of performance and progress on the numeracy projects 2001-2003.* Wellington: Ministry of Education.

Zevenbergen, R., Dole, S., & Wright, R. (2004). *Teaching mathematics in primary schools*. Crows Nest, NSW: Allen & Unwin.

Appendix 1
(Test 1)

Scale diagnostic
Part A

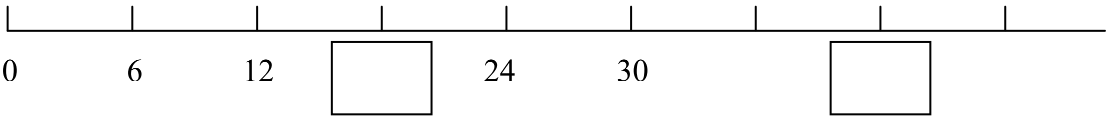
Name: ___________________________ Year: ___________

Room: ___________

1) Change the line below into a number line. Show on it where the numbers 2, 3, 5, 8 and 10 go

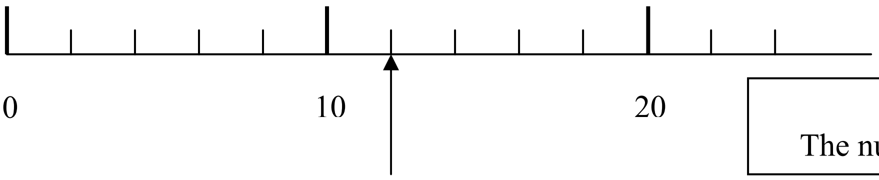
____

2) The drawing below shows numbers on a number line. What missing numbers should go in the boxes?



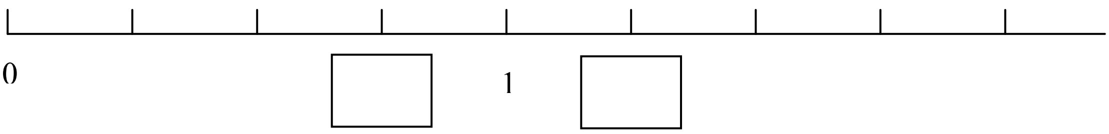
3) On the number line above put a cross where 11 would be

4) What number is the arrow pointing to?

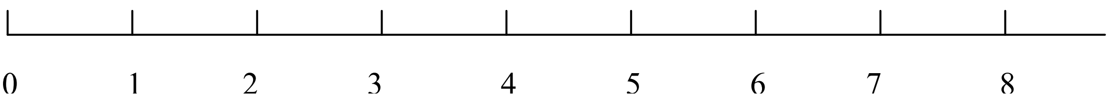


The number is _________

5) What missing numbers go in the boxes?



6) Put a cross where the number ¾ goes on this number line



7) Using a cross, show where the number 0.4 goes on the number line below

8) Use the line below to draw another number line. On it show where –3 should go

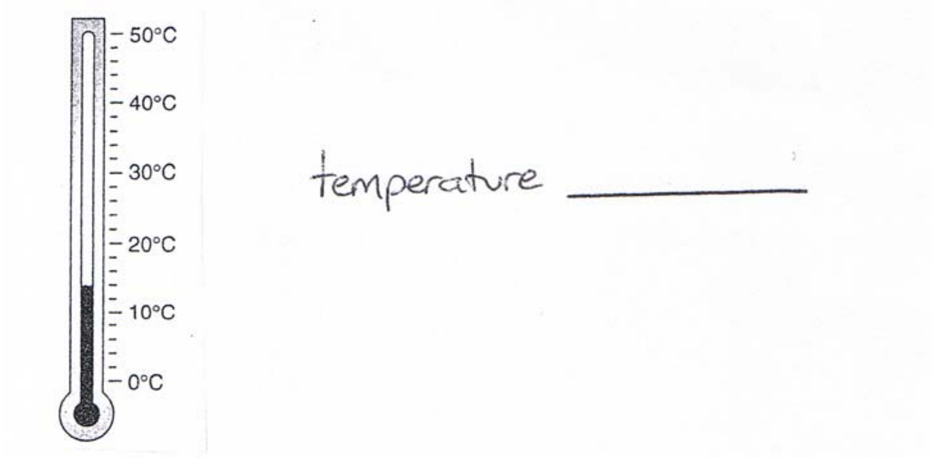
____

Scale diagnostic
Part B

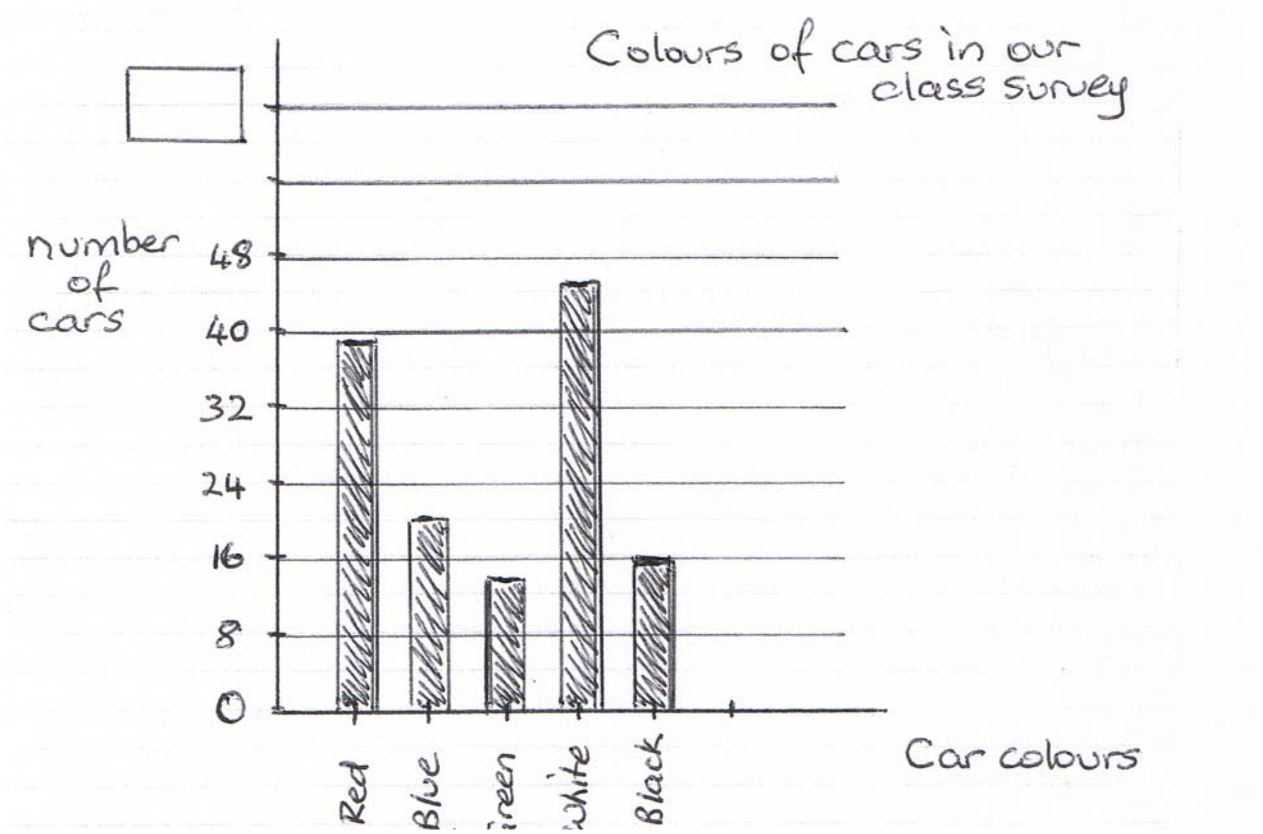
Name: __________________________ Year: ___________

Room: ___________

1) What temperature is the thermometer showing?



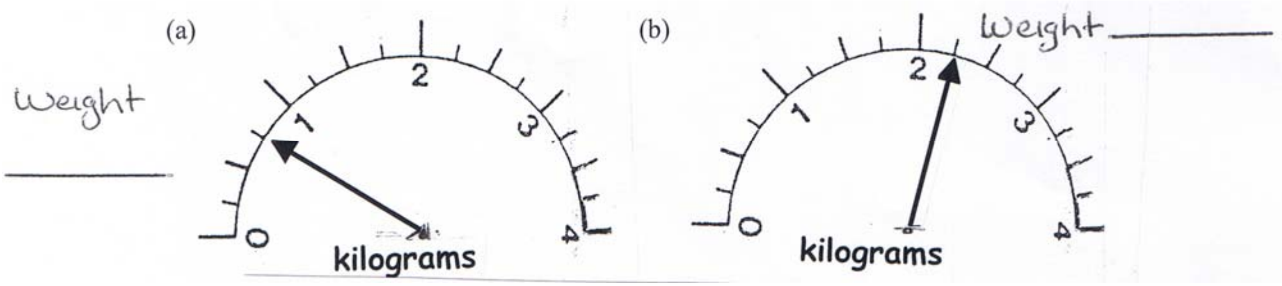
2) This graph shows the results of a survey a class did into favourite car colours.



(a) What missing number should go in the box?

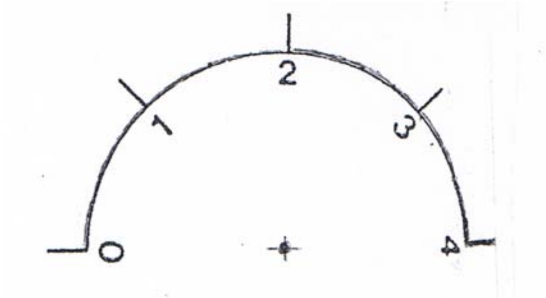
(b) How many red cars did the class count?

3) Gemma's mum has a set of scales in her kitchen. The scale looks like the ones below. What weights are the arrows pointing to?

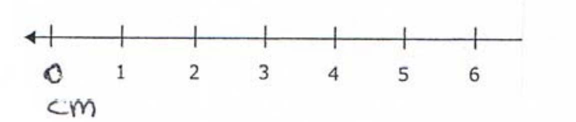


(a) (b)

(c) Ross's mum needs to weigh ¾ of a kilogram of flour for a bread recipe. Draw an arrow to show where ¾ of a kilogram is on her set of scales



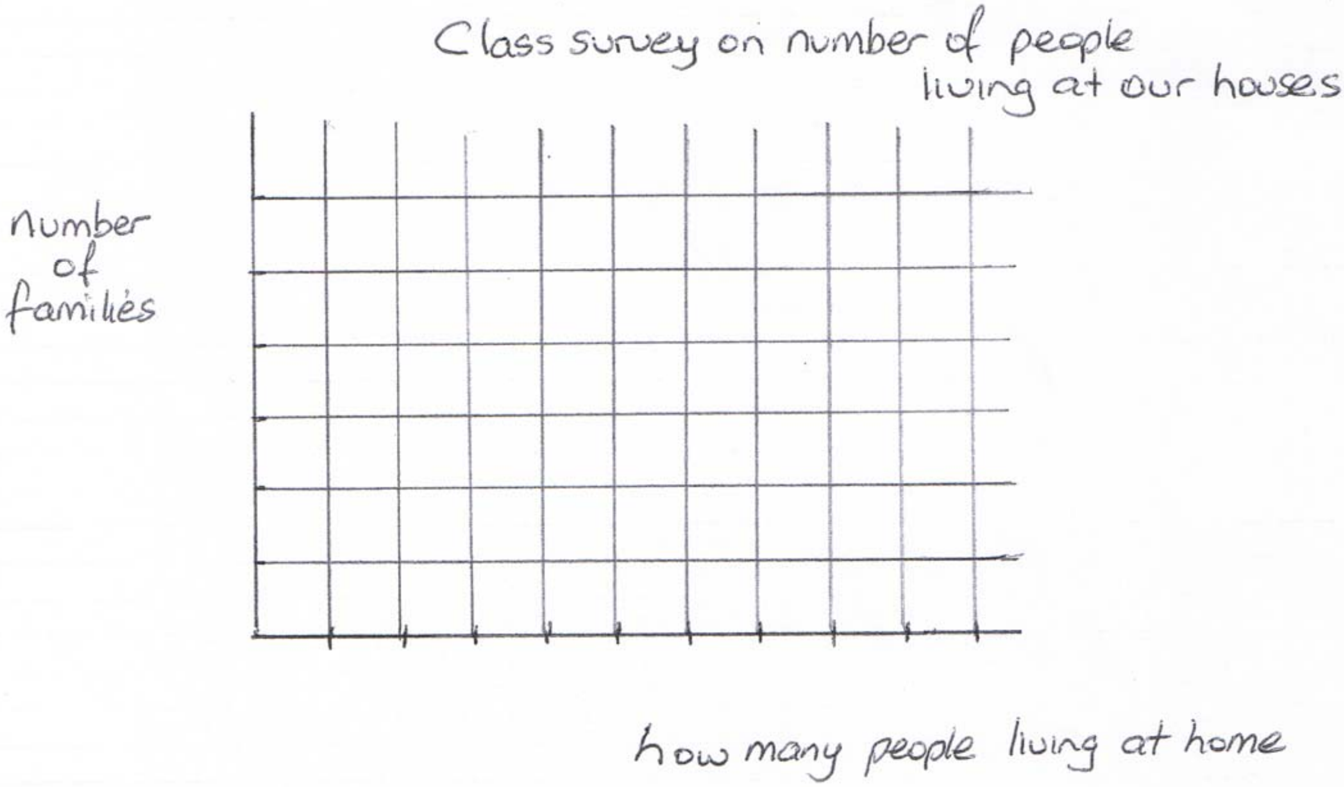
4) Below is part of a ruler marked in centimetres. Put a cross where you think 0.4 cm is



5) The teacher of a year 7 class wants to find out about the families of her students. One question she asks is how many people live at home. Here are the results.

How many people live at home	2	3	4	5	7	9
Number of families	3	8	11	4	2	1

Finish the bar graph started below to show what she found out



Appendix 2
(Test 2)

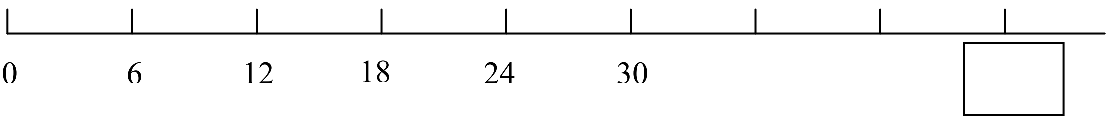
Understanding number lines and scales
Diagnostic test

Name: ___________________________ Year: ___________

Room: ___________

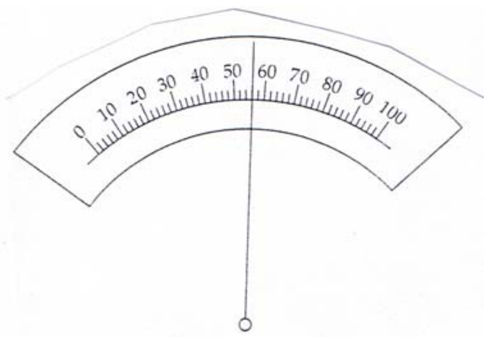
1) On the line provided below, complete a number line to show where the numbers 2, 3, 5, 8 and 10 would go

2) The drawing below shows numbers on a number line. What missing number should go in the box?

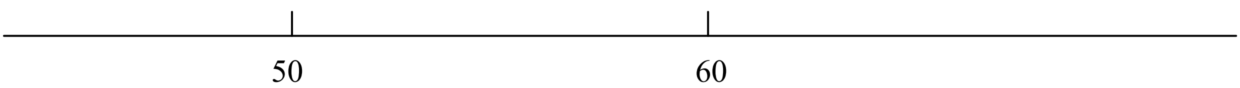


3) (a) **On the number line above** put a cross where 23 would be
(b) **Also on the number line above**, use an arrow to show where 14 is

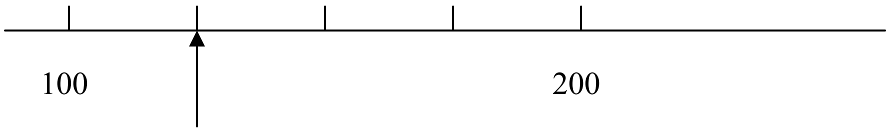
4) Tom is weighing himself on a set of bathroom scales. How heavy is Tom? (The scales measure people's weight in kilograms.)



5) Put a cross where 52 is on this number line

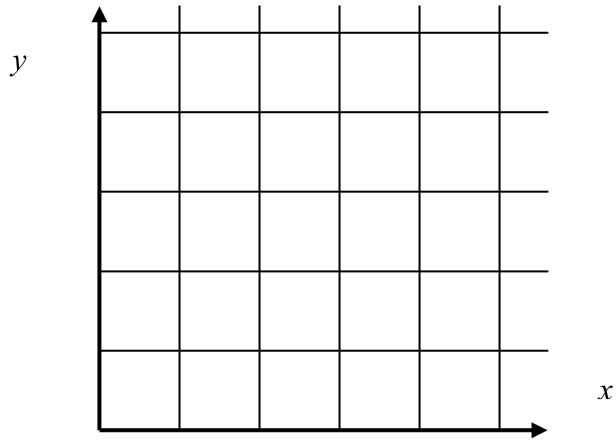


6) What number is the arrow pointing to?

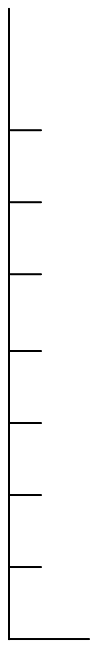


The number is _______________

7) Plot the point (3, 4) on the set of axes below. Mark it with an X

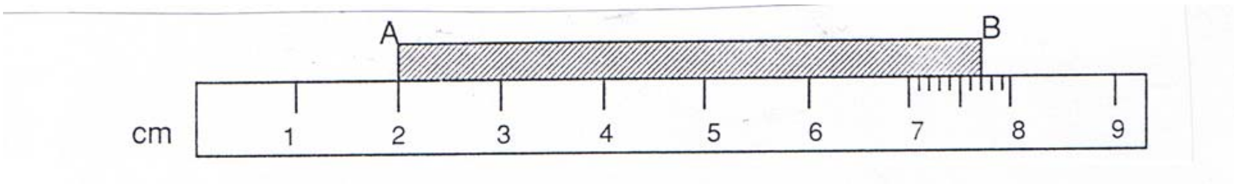


8) The number line below is the vertical axis from a graph.



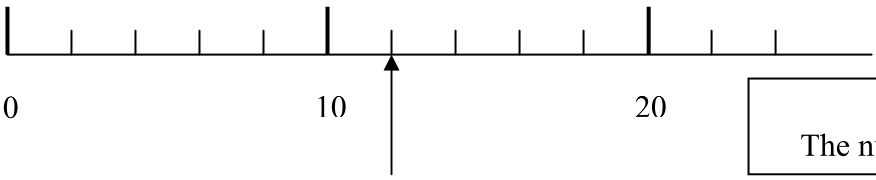
Show where 4, 8, and 37 are on the axis, and mark what your scale is

9) Use the ruler shown below to work out the length of the shaded rod



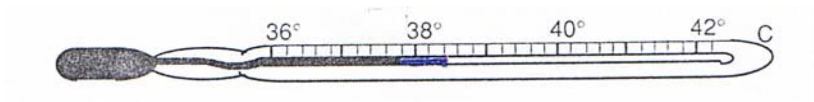
The length of the rod is _______________

10) What number is the arrow pointing to?



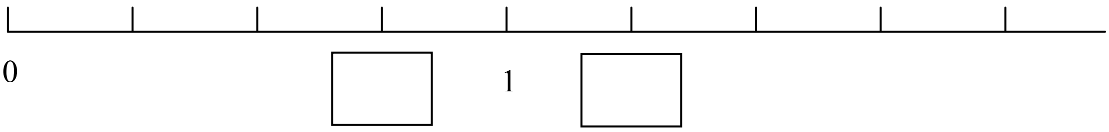
The number is ________

11) Here is a thermometer. What temperature does it show?

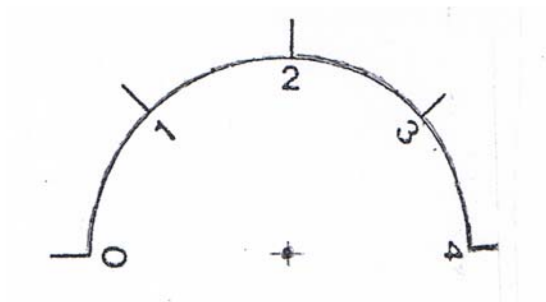


12) (a) On the number line below use a cross to show where the number 0.3 is.
(b) Also on the number line below draw an arrow to show where the number 0.21 is

13) What missing numbers go in the boxes?

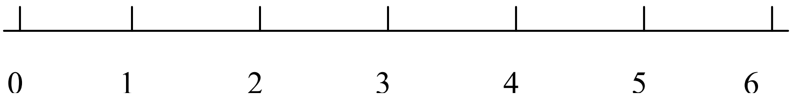


14) Ross's mum needs to weigh ¾ of a kilogram of flour for a bread recipe. Draw an arrow to show where ¾ of a kilogram is on her set of scales



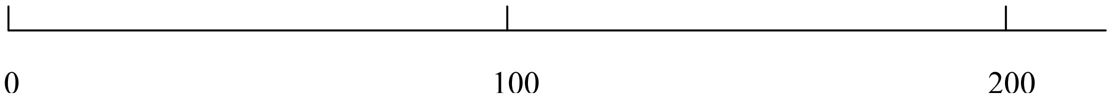
kilograms

15) Put a cross where the number $\frac{2}{3}$ goes on this number line

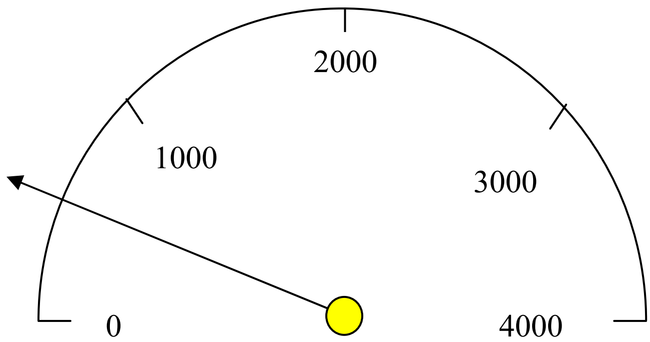


16) Use the line below to make a number line. On it show where –1 and –4 would go

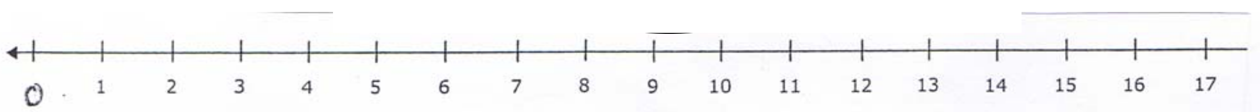
17) Put a cross to show where 130 is on this number line



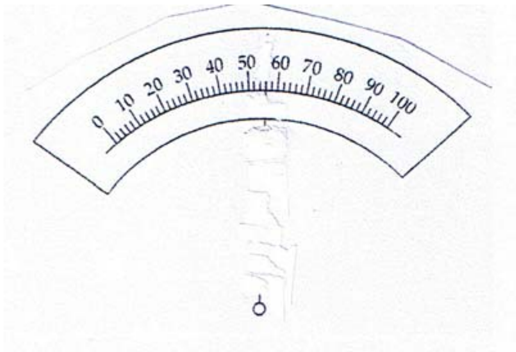
18) On the dashboard of a car, there are two large dials, a speedo and a rev counter. Here is a rev counter. How many revs is it showing?



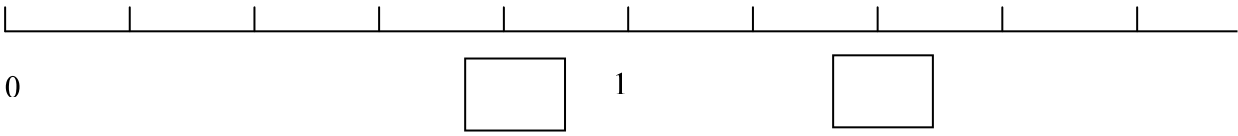
19) Below is part of a ruler marked in centimetres. Put a cross where you think 1.6cm is



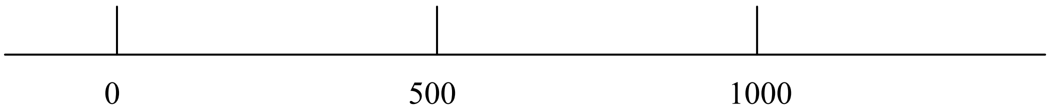
20) Noema wants to weigh 56 grams of butter for a recipe. Her scales weigh things in grams. Where is 56 grams on this scale?



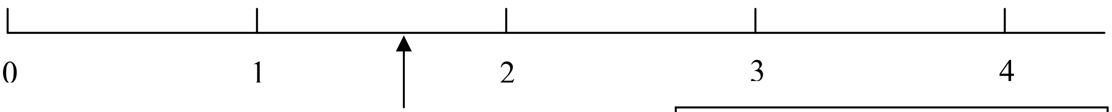
21) What missing numbers go in the boxes?



22) Draw an arrow to show where 625 can be found on this number line

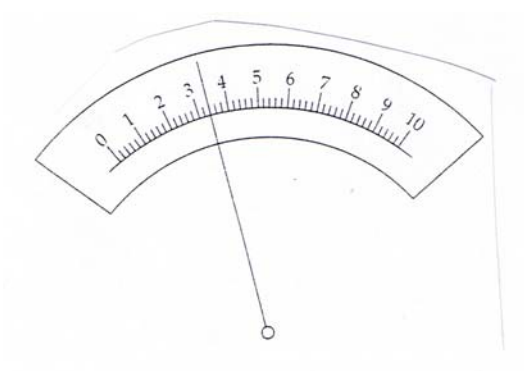


23) What number is the arrow pointing to?

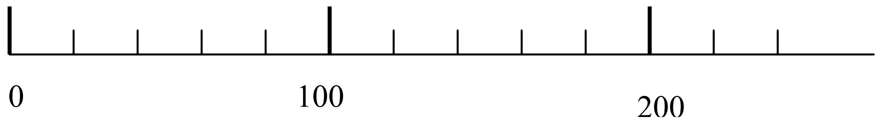


The number is ___________

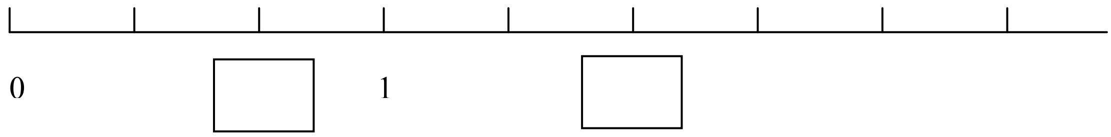
24) Jane is weighing a stone for a science project. How heavy is the stone? (The scales are in kilograms.)



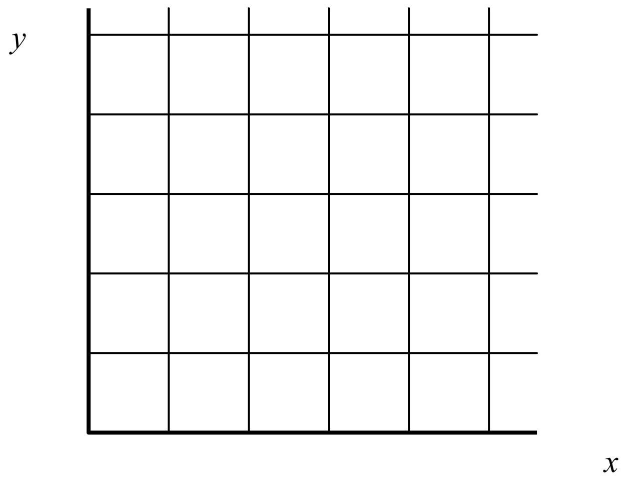
25) Draw an arrow to show where 120 is



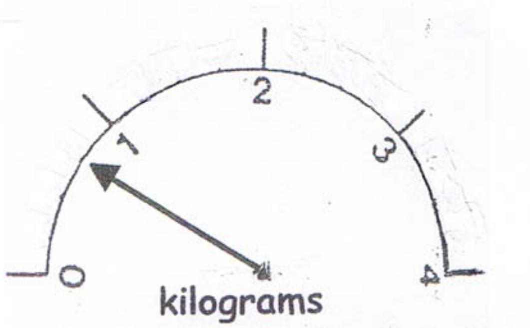
26) What missing numbers go in the boxes?



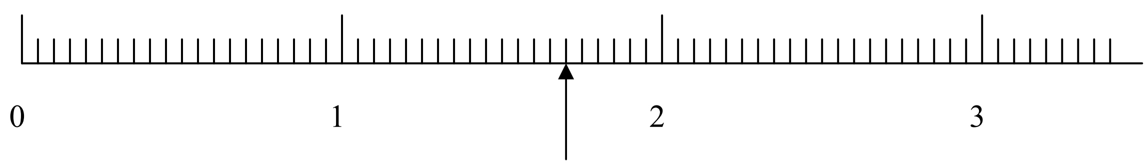
27) The point (8, 15) needs to be shown on this graph. Plot the point (8, 15) and mark it with a cross. Make sure you show your scales.



28) Pua is weighing a book as part of a mathematics test. How heavy is the book?



29) What number is the arrow pointing to?



30) Put a cross where the number ¾ goes on this number line