Designing and Evaluating Multi-Representational Learning Environments for Primary Mathematics

Shaaron E. Ainsworth B.Sc. (Hons.) M.Sc.

Thesis submitted to the University of Nottingham, for the degree of Doctor of Philosophy, May 1997

Table of Contents

ABSTRACT

This thesis reports the design and evaluation of multi-representational learning environments that teach aspects of number sense. COPPERS is concerned with children's belief that mathematical problems can have only a single correct answer. CENTS addresses the skills and knowledge required for successful computational estimation.

Although, there is much multi-representational software and a significant body of research which suggests that learning with multiple external representations (MERs) is beneficial, little is known about the conditions under which MERs promote effective learning. To address this, a framework was proposed for considering MERs. It consists of a set of dimensions along which multi-representational software can be described and specifies learning demands of MERs. This framework was used to generate predictions about the effectiveness of different multi-representational systems.

Experiments investigated children's performance in multiple solutions and computational estimation before they received direct teaching and tested whether the learning environments could help children develop these skills. Each experiment examined how specific aspects of the learning environments contributed to learning outcomes.

Experiments with COPPERS showed that children's pre-test performance was generally poor. Improved post-test performance on multiple solutions tasks occurred when children gave substantially more answers on the computer than their pre-test base-line. They rarely chose this strategy for themselves. It was found that providing a tabular representation of solutions in addition to the familiar row and column representation improved learning.

Estimation is difficult for primary school children, but limited teaching led to substantial improvements in strategies and accuracy of estimates. Three experiments with CENTS addressed the effects of MERs on learning. When representations were too difficult to co-ordinate, then either children did not improve at understanding the accuracy of estimates, or focused their attention upon a single representation. Additionally, varying how information was distributed across representations influenced how representations were used.

These experiments show that when considering learning with MERs, it is not sufficient to consider the effects of each representation in isolation. Behaviour with representations changes depending on how they are combined. These findings are discussed in terms of their implications for the design of multi-representational learning environments.

ACKNOWLEDGEMENTS

Many people have helped bring this thesis to its conclusion.

First, my primary supervisor, Professor David Wood for the support and guidance he has given me during the course of the thesis and above all else for the enthusiasm with which he viewed this research and the confidence this gave me to continue.

Secondly, Dr. Peter Bibby who acted as a second supervisor for half of this thesis, but who provided advice, time and a sense of perspective throughout the whole time.

Many people commented upon drafts of this thesis. I would like to thank Dr Susan Cavendish, Dr. David Clarke, Dr. Peter Cheng, Dr. Sharon Ding, Dr. Jean Underwood, Dr. Heather Wood and Mrs Helen Woodiwiss for reading some of the chapters and especially Martin Oliver for commenting upon the whole draft - a thankless task for which I do thank him profusely.

I would like to acknowledge of the staff and students associated with the ESRC Centre for Research in Development, Instruction and Training for providing a stimulating environment in which to conduct research. Dr. Peter Cheng and Shirley Holst deserve particular mention.

The participation of many children and teachers from local schools made this research possible. I thank them for their enthusiasm and their tolerance.

My friends who have shared an office with me (Susannah Lamb, Lucy Copeland, and Carolyn Tan-Niam) and who have explained (on a weekly basis) the difference between a verb and noun.

Robert Gaizauskas has needed and displayed the patience of a saint. I do not have appropriate words to describe the difference his friendship and support have made to my life. Thank you, Rob.

Finally, I am indebted to Sharon Ding. She has provided friendship throughout by time in Nottingham, and in addition, a roof over my head, a child to torment and a cat to stroke.

CHAPTER ONE

Introduction

This thesis is concerned with the design of computer-based learning environments in relation to the issue of how different combinations of representations affect learning.

The domain explored is an aspect of mathematical understanding - *number sense.* Although difficult to define, number sense is commonly associated with certain mathematical skills and beliefs. Two pertinent areas have been addressed in this thesis. The first is how children's belief that mathematical problems have only a single correct answer can act as a barrier to the development of number sense. The second is computational estimation. Flexible estimation is seen as depending heavily on number sense. It is also thought that by helping children to become flexible estimators they further develop their number sense. Computer-based learning environments for both multiple solutions and computational estimation were designed, implemented and evaluated during the course of this research.

This thesis also addressed how the use of multiple external representations (MERs) influences learning. In recent years, there has been a large growth in the number of multi-representational learning environments. There is also an increasing body of research which suggests that providing MERs can confer significant benefits for learners. However, little is known about the conditions under which MERs promote effective learning. Consequently, designers and educators have few principles to guide their use of MERs.

A framework for considering MERs is proposed which consists of a set of dimensions (specific to MERs) along which multi-representational software can be described. In addition, the different learning demands of MERs are analysed. This framework was used to predict the conditions under which particular MERs facilitate learning. Empirical studies focused on two of these dimensions, the similarity of format and the degree of informational redundancy between representations.

Experiments performed with the learning environments had three basic aims: (1) to examine children's performance in multiple solutions and computational estimation before they had received direct teaching; (2) to test the claims that the learning environments could help children develop the skills they were designed to teach and; (3) to examine in detail how specific aspects of the learning environments contributed to learning outcomes. An important aspect of this research was the detailed quantitative measurement of children's actions upon the representations used within the learning environments. This was collected in order to explain how different combinations of representations affected the way that children met the various learning demands of the systems.

The remainder of this chapter gives an overview of the following eight chapters which describe the design of the learning environments and the experiments that were conducted with the systems.

Chapter Two begins by introducing examples of people solving mathematical problems in ways that demonstrate number sense and provides other examples that suggest the problems were solved with little regard to number sense. It then describes approaches to defining number sense. Mathematical skills and beliefs most closely associated with number sense are reviewed. The two areas of number sense that are addressed in this thesis are then discussed in detail. First, research on the generation of multiple solutions to mathematical problems is reviewed and then instruction aimed at developing this knowledge is discussed. The next section considers the conceptual and procedural aspects of computational estimation, along with related mathematical skills. Then, research that describes how children's knowledge of estimation develops is outlined. This section ends by considering approaches to the teaching of computational estimation. Finally, Chapter Two reviews the existing computer-based approaches to instruction in this area.

Chapter Three describes research on the role of MERs in learning. It begins by locating interest in learning with MERs within the context of research in how external representations influence learning and problem solving. In order to consider the effects of combining different representations, a review of some of the approaches to classifying representations is provided. The next section considers the advantages that MERs bring to learning, placing them in a framework which outlines the different purposes for which MERs can be used. The learning demands that must be met if MERS are to be used effectively are considered. The next sections describe in more detail a learning demand that is unique to MERS - translation between representations. It also describes previous approaches to measuring and assessing such translation. This chapter ends by proposing a set of dimensions which could serve as a framework for considering the design of multi-representational software.

Chapter Four describes the first learning environment, COPPERS. COPPERS sets children coin problems such as 'What is 3 × 20p + 4 × 10p ?' which must be answered by providing alternative decompositions of the total (*e.g.* '20p + 20p + 10p + 50p, or 10p + 2p + 2p + 1p + 5p + 10p + 10p + 5p + 5p + 50p' or '50p + 50p'). The design of this system is related to instructional methods in primary mathematics. Three aspects of the system design are considered in detail: problem representation and generation, the means by which problems are answered and representations used for feedback on answers.

Chapter Five reports two experiments with COPPERS. Experiment One examined three issues: (a) whether children need to be taught to give multiple solutions to mathematical problems; (b) whether COPPERS meets its educational objectives and; (c) how aspects of system design contributed to this goal. Two aspects of design were considered; the number of answers per question and the role of additional tabular feedback in supporting learning. This experiment found that children's pre-test performance was low, but that limited teaching led to substantial improvement. It also showed that children who saw an additional tabular representation of their answers performed significantly better at post-test than those who did not. Only lower performing children were found to benefit from giving multiple answers to problems during the computer intervention. Experiment Two addressed this issue further by asking children to give substantially more answers per question than they had at pre-test and by examining how many answers per question they would provide if they were given free choice during the intervention. The experiment suggested that to maximise learning outcomes children should be required to provide multiple decompositions on the computer and that they rarely invent or choose this learning strategy for themselves.

Chapter Six introduces CENTS - a computer-based learning environment that teaches aspects of computational estimation. It is also designed to provide experimenters with a large degree of flexibility about how information is presented and combined in multi-representational systems. The design of the system is discussed in relation to the research on computational estimation reviewed in Chapter Two and a detailed description of the representations provided by the system is given.

Chapter Seven presents the first experiment with CENTS. This study examined children's untaught estimation performance and was designed to establish whether CENTS could help children develop these skills. However, the primary focus of this experiment was on how multi-representational systems that differed in the way they displayed information influenced learning outcomes. This study confirmed that experience with CENTS did improve estimation performance. However, it showed that certain combinations of representations were better than others for teaching aspects of this knowledge. In particular, it was shown that representations of different formats were difficult to co-ordinate. The difficulties children had translating between representation on the computer were found to affect their subsequent post-test performance. An analysis of the different representations was performed to explain this result.

Chapter Eight reports experiments that examine the effects of presenting different combinations of representations over longer periods and also explores how varying the redundancy between representations influences learning outcomes. Experiment Four showed that children given representations that differed in format ultimately concentrated their attention upon only one of the representations presented. It was suggested that this strategy is highly adaptive when there is sufficient redundancy between representations so that one representation can carry all the necessary information. Therefore, Experiment Five examined what would happen when this was not the case. Representations were used which, in addition to manipulating similarity of format, also varied in redundancy (full redundancy between representations or no redundancy). There was no overall effect of redundancy on final learning outcomes, but some evidence that limiting redundancy aided initial task performance.

Chapter Nine begins by summarising and integrating the research reported in this thesis. It considers children's untaught performance at multiple solutions tasks and computational estimation and considers how the computer-based learning environments improved this performance. The next section reviews the aspects of system design examined by the experiments, concentrating primarily on the effects of different MERs. The limitations of the computer-based learning environments and the experiments are considered along with suggested improvements. The general issue of the design of multi-representational learning environments is discussed in terms of the dimensions first proposed in Chapter Three. In chapter nine, these are now reviewed in the light of the results from the experiments and the methods used to analyse them. Implications for future work are considered. The chapter ends with a short summary of the thesis.

CHAPTER TWO

Number Sense - Computational Estimation and Multiple Solutions

Focus in mathematics teaching has shifted from the learning of formal procedures and accepted facts to an emphasis on mathematics as flexible, insightful problem solving. Schoenfeld (1992) describes mathematics as a 'science of patterns' where the goal is to systematically study and explain the nature and principles of regularities in pure and applied systems. Consequently when solving mathematical problems, people should do so with a disposition to make sense of the problem - they should have *number sense*. Indeed, the development of number sense has been identified as the major objective of primary school mathematics in the U.S.A. (National Research Council, 1989). This chapter will review research on number sense and introduce two areas seen as important for this aspect of mathematical understanding - knowledge of multiple solutions and computational estimation. These aspects of number sense have been the focus of the two computer-based learning environments developed and evaluated during the course of the research conducted for this thesis.

The concept of number sense is introduced by providing examples of the types of solution which demonstrate both proficiency and lapses in number sense. Then, although a definition of number sense remains problematic, some approaches to describing it are presented. The remainder of the introductory section discusses mathematical skills and knowledge proposed as most pertinent to number sense. Two of these areas are then considered in detail. One barrier to the development of number sense is children's belief that mathematical problems have a single correct solution. Section 2.4 reports research that has examined knowledge and instruction in multiple solutions and strategies. Another key component of number sense is the ability to flexibly estimate answers to mathematical problems. The procedural and conceptual components are outlined along with related mathematical skills. Then research that has examined the development of computational estimation abilities and instruction in estimation is reviewed. The final section considers the potential of the computer for teaching these aspects of number sense.

2.1 WHAT IS NUMBER SENSE

There is much disagreement about what constitutes an adequate definition of number sense. However, number sense can be recognised when it occurs in response to mathematical problems. To provide a framework for the following discussion, five answers to problems are presented which most people would agree show a demonstrated lack of understanding and sense of number, and five that demonstrate number sense are given.

(1) 'Estimate 789 $\times$ 0.52' (quoted in Threadgill-Sowder, 1984)

Common responses were either 800 '0.52 is nearly 1, 1 $\times$ 789 is roughly 800' or 0 '0.52 is very small call it zero, 0 $\times$ 789 = 0'

(2) 'Which of the following is an estimate of $\frac{9}{10}+\frac{11}{12}$ (1, 2, 19, 21, I don't know)' (quoted in Sowder, 1995)

This question formed part of the large scale National Assessment of Education Progress (NAEP) in the USA - only 24% of 13-year-olds and 37% of 17-year-olds correctly selected 2.

(3) 'How many buses will be required to take 1228 soldiers if each bus hold 36 soldiers?' (another NAEP problem quoted in Schoenfeld, 1988).

This is commonly answered with 31 remainder 12.

(4) '15.24 $\times$ 4.5 is 6858, but the decimal point is missing from the solution. Place the decimal point where it should be.'

Markovits (1989) reports that out of a sample of 49 trainee elementary school teachers, 79% responded with 6.858.

(5) The literature on buggy procedures (*e.g.* Young & O'Shea, 1981; Brown & Burton, 1978) in subtraction abounds with responses which demonstrate little sense of the numbers being operated on. For example, the infamous 0-N= N bug.

$$\begin{array}{r} 230 \\ -\underline{236} \\ \underline{6} \end{array}$$

There are also example from a wide range of areas which demonstrate people acting on mathematical tasks with number sense.

(1) 'Estimate 482 × 51.2'

Dowker's studies of expert mathematicians (*e.g.* Dowker, 1992) yielded responses such as '$482 \times \frac{1}{2} \times 100$'

(2) 'Add 159 and 142' (presented orally)

Resnick's (reviewed in Resnick, 1992) longitudinal analysis of one seven-year-olds performance illustrates the flexibility of some children's informal arithmetic. The child's answer to this problem involved the following steps: 2 × 100 = 200; 50 + 40 = 90; 9 + 2 = 11; 11 + 90 = 101; 200 + 101 = 301.

(3) '200 - 35 = ?'

Using problems such as these research in the ethnomathematic tradition also provides plenty of examples of unschooled children and adults operating on numbers in ways that demonstrate their number sense. One example, described by Nunes (1992) is in response to the above problem "if it was 30, then the result would be 70, but it is 35. So, its 65, 165".

(4) '5 × 29 = ?'

Problem such as these can be solved by using well known numbers to figure out facts (Resnick, 1989). This problem might be solved remembering that "I can buy 5 comics at 30 pence each with my pocket money of £1.50. So five less than is £1.45, therefore 5 × 29 = 145"

(5) '$\frac{3}{4}+\frac{2}{8}$'

If children have rational number sense, they should see that $\frac{3}{4}$ can be expressed as $\frac{6}{8}$ or $\frac{2}{8}$ as $\frac{1}{4}$, thus, making the problem of adding these fractions much simpler. On others occasions, fractions could be converted to decimals or percentages in order to allow numbers to be operated on flexibly.

2.2 DEFINING NUMBER SENSE

Number sense has been defined as involving: *a sound understanding of the meaning of a number and of relationships between numbers, a good understanding of the relative magnitudes of numbers and awareness of numbers used in everyday life* (National Council of Teachers of Mathematics, 1989). Sowder (1989) characterises number sense as *a well organised conceptual network that enables a person to relate numbers and operation procedures. It can be recognised in the ability to use number magnitude, both relative and absolute, when making qualitative and quantitative judgements necessary for, but not restricted to, number comparison, recognition of unreasonable results for calculation and the use of non-standard algorithms for mental computation and estimation. It is demonstrated by flexible and creative ways of solving numerical problems.* She warns that it is neither easily taught or measured. One further definition is given by Reys *et al.* (1991) who describe number sense *as an intuitive feeling for numbers and their various uses and interpretations; an appreciation for various levels of accuracy when figuring; the ability to detect arithmetical errors; and a common sense approach to using numbers* (quoted in Sowder, 1995).

These definitions have common aspects showing some convergence of thought about number sense. At the same time, there is still doubt that such definitions have as yet captured the necessary and sufficient features that define number sense (Sowder, 1989). Resnick (1989) doubts that any traditional approach to defining number sense is possible. As number sense is inherently contextualised, there will be no possibility of producing a definition that is decontextualised. Therefore by comparison to a characterisation of higher order thinking, she proposes the following dimensions of number sense:

- Number sense is non-algorithmic
- Number sense tends to be complex
- Number sense often yields multiple solutions, each with costs and benefits, rather than unique solutions
- Number sense involves nuanced judgement and interpretation
- Number sense involves the application of multiple criteria
- Number sense often involves uncertainty
- Number sense involves self regulation of the thinking process
- Number sense involves imposing meaning
- Number sense is effortful

As many researchers have pointed out number sense is not a discrete quality (*e.g.* Reys, 1989). It will not be possible to state that a learner does or does not have number sense. Instead number sense is a continuous quality that could be apparent at a number of different levels. It is possible to exhibit number sense for some aspect of mathematics and not for others (*e.g.* when dealing with fractions children's behaviour often exhibits less number sense than when dealing with whole numbers). This has implications for the teaching of number sense. Instruction in number sense, as all the participants of the 1989 conference on Number Sense agreed, must not be restricted to discrete lessons. It should be applied to the whole of mathematics learning.

2.3 KNOWLEDGE AND BELIEFS ASSOCIATED WITH NUMBER SENSE

Certain aspects of mathematical skills and knowledge seem more intimately tied to the development of number sense than others. In particular, the following areas are often identified as both leading to improved number sense and also to depend most heavily on number sense: numeration, number magnitude, mental computation and computational estimation (Sowder, 1992b). A few examples of the relation between these areas and number sense are discussed.

Numeration is fundamental to number sense. Sowder (1992b) identifies cardinal, ordinal and place value understanding as the most important components. Studies performed by a number of researchers have demonstrated that children often have sophisticated competencies in counting and additive composition when they enter school. For example, Nunes & Bryant (1996) describe a study that showed that nearly 40% of Brazilian pre-school children could use the principle of additive composition when using coins. Given four 10 cents coins and four 1 cents coins (a total of only eight coins), these children understood how to buy an item costing 13 cents.

Place value is more complex and children often fail to develop a competent understanding of it. Resnick (1983) identifies three levels of understanding of place value. At stage one, children can identify tens, units, *etc.*, in relation to concrete objects and then can progress to mental computation. At stage two children are capable of producing non-canonical decompositions (*e.g.* 22 = 1 ten and 12 ones). Again this is dependent on a physical representation, such as Dienes blocks. The final stage of place value development is the semantic linking of this partitioning ability with written algorithms. Resnick & Omanson (1987) show that even when children are given mapping instruction which explicitly addresses this issue, they still can have problems with written arithmetic (described in more detail in section 3.5).

A number of studies have shown that children who appear to have an intuitive grasp of aspects of the number system, can fail to apply them to school mathematics (*e.g.* Nunes, Schielerman & Carraher, 1993). One of the problems facing teaching aimed at developing numeration and number sense is how to get children apply their out of school mathematical competencies and build upon these to develop more sophisticated mathematical understanding.

Number magnitude can be either relative or absolute. Relative magnitude involves the ability to compare and order number. It is ultimately dependent on a good understanding of place value. For example, a common problem is demonstrated when children identify .1814 as larger than 0.9 (the first number has more digits and so must be bigger). Analysis of school textbooks suggests that although a sense of relative number magnitude is very closely linked to successful performance in a wide range of mathematical areas (*e.g.* fractions, decimals, place value) only limited attention is paid to this topic (Sowder, 1992b). Absolute number magnitude involves understanding what a number might 'mean'. Sowder (1992b) reports a task asking children (4th, 6th 8th and 10th grades) which of the following was a good estimate of the number of people at a concert; '65, 380, 40,000 5,000,000'. Each was selected as reasonable by 35% of 8-9 year-old children (*i.e.* some selected more than one as reasonable). This type of knowledge is often poor even in adults and is more obvious with larger numbers. Hofstader (1985) refers to this failing as 'number numbness'. His suggestion for encouraging people's sense of large numbers involves developing prototypes in a number of domains (populations, budgets, ants, coins, *etc.*).

Mental computation involves number sense when students use and invent strategies that take advantage of numerical and operational properties rather than relying on rote learned procedures or mental versions of written symbolic manipulation. Markovitz & Sowder (1988) showed that when instruction in mental calculation did not involve rote learning or rule memorisation, children abandoned analogues of pen and paper techniques (impossibly unwieldy for mental calculation) and used more non-standard procedures (*e.g.* a left to right strategy, decomposition, counting up rather than subtracting). This approach to mental calculation has been said to have developed children's number sense.

Computational estimation can be defined as the process of simplifying an arithmetic problem using some set of rules or procedures to produce an approximate but satisfactory answer through mental calculation (Dowker, 1992). The crucial aspects of computational estimation in relation to number sense involve flexible approaches to approximation rather than reliance of simple inappropriate algorithms. It is also necessary for recognising when an estimate is satisfactory (Sowder 1992a). A further study by Sowder and Markovits (Markovits & Sowder, 1994) using a similar approach to that of the mental calculation study showed that children could develop flexible estimation skills.

In addition to these mathematical skills, learners' beliefs about the nature of mathematics have been identified as influencing their ability to develop number sense. Teachers have identified learners' mathematical beliefs as the biggest barrier to encouraging number sense in the mathematics curriculum. The teachers interviewed by Phillip, Flores, Sowder & Schappelle (1994) worry that the children in their classes believe in, and practise, a mathematics that is devoid of questioning, creativity and of sense making. They suggest that children believe that there is a single right answer to a mathematical problem, there exists only one right way to obtain that answer and that teachers should provide them with the appropriate rule.

Schoenfeld's review of children's mathematical beliefs (Schoenfeld, 1992) suggests that typically, pupils believe that:

- Mathematics problems have one and only one right answer
- There is only one correct way to solve any mathematics problem - usually the rule that the teacher has most recently demonstrated
- Ordinary students cannot expect to understand mathematics; they expect to simply memorise and apply what they have learned mechanically and without understanding
- Mathematics is a solitary activity done by individuals in isolation
- Students who have understood the mathematics they have studied will be able to solve any assigned problem in five minutes or less
- The mathematics learned in school has little or nothing to do with the real world
- Formal proof is irrelevant to processes of discovery or invention

Baroody (1987) identifies four ways that children's beliefs about mathematics impinges upon their views of themselves:

- An inability to learn facts or procedures quickly is a sign of inferior intelligence and character
- An inability to answer quickly or use a procedure efficiently indicates 'slowness'
- An inability to answer correctly denotes a mental deficiency
- An inability to answer at all signals real stupidity

These mathematical beliefs also have a behavioural corollary. For example, Schoenfeld (1992) reports that students would give up on problems quickly even though perseverance may well have led to a successful answer.

Yet these beliefs and mathematical behaviour seem at odds with children's pre-school behaviours. Fuson (1992) reports that pre-school children appear to naturally use multiple approaches to answer simple arithmetic problems. Nunes & Bryant (1996) review an impressive range of pre-school competencies. These researchers conclude that it is the school and cultural environment that shapes children's overall beliefs about the nature of mathematics. Baroody (1987) proposes that it is classroom emphasis of getting *the* right answer using *the* correct procedure that creates what could be termed the 'right answer hypothesis': children commonly believe that all problems must have a correct answer, that there is only one correct way to solve a problem, and that inexact answers (such as estimates) or procedures (trial & error problem solving) are undesirable.

In this section, number sense has been associated with a variety of mathematical skills and beliefs. Aspects of mathematics such as mental calculation, numeration, relative and absolute magnitude and computational estimation have been related to number sense. One of the barriers to the development of number sense *i.e.* children's beliefs about mathematics, has been briefly reviewed. The learning environments developed and evaluated during the course of this research have examined two aspects of a number sense approach to mathematics. They have focused on challenging the right answer hypothesis and have been concerned with developing understanding that mathematical problems can have many correct solutions and that inexact answers (estimates) are an important part of mathematics. In the following sections, research in these areas is reviewed in more detail.

2.4 MULTIPLE SOLUTIONS AND STRATEGIES

Little research has specifically addressed the role of multiple solutions in mathematics. However, encouraging children to believe that there can be multiple answers and ways to solve a mathematical problem has been identified as crucial for performance in a number of areas and for developing number sense.

The term multiple solutions can have a number of different meanings. Firstly, it can mean that there are multiple correct outcomes to a mathematical problem. The computational estimation domain is an obvious case where there can be lots of 'right' answers. Indeed, Sowder & Wheeler suggested that understanding that problems can have multiple answers was crucial for developing estimation abilities. A second meaning is where there may be a single correct answer, but multiple strategies for obtaining that answer. For example, an integral calculus problem such as $\int \frac{x}{x^2 - 9} dx$ could be solved in a variety of ways with varying degrees of complexity. The third type of multiple solutions problem is where although there might be a single correct answer, there could be lots of ways of expressing that answer. Thus, 'what's 2 +7?' always equals nine, but could be expressed as '1 + 8', '3 + 6', '4 +5', *etc.* For clarity, the first case will be referred to multiple answers, the second, multiple strategies and the third, multiple solutions. A further example of multiple answers/solutions, where different representations are used, is reviewed in depth in Chapter Three.

2.4.1 Multiple Answers

There are multiple right answers to estimation problems. For example, given '442 × 362', reasonable answers include but are not limited to 12,000, 20,000 16,000, 15,000, *etc.* Values in equations often have multiple correct answers. For example, given a problem such as: 'find a value for x that makes the following equation true, $x^2+14=18$'; this can obviously be true for x = 2 and x = -2. The CSMS Mathematics Team (1981) suggest that children may interpret letters in algebraic expressions in six different ways, but in only two of these do they understand that a letter can represent more than one value.

2.4.2 Multiple Strategies

Research that has examined the role of multiple strategies in mathematics includes Tabachneck, Koedinger & Nathan (1994). They examined the role of multiple strategies in solving word algebra problems and identified four broad classes of category: *algebra* where the problem statement is translated into algebraic assignments and equations; *guess and test* involving translating into calculation statements; *verbal-math* where the original statement is recoded into another verbal statement and; *diagrammatic* where the problem statement in translated into a diagrammatic representation. They found no difference between the success rates of different individual strategies, but that multiple strategy use was about twice as effective as any single strategy. They propose that this effect occurs because these different strategies were associated with different types of errors. This allowed impasses to be bypassed or overcome.

Santos (1994) gave 14 to 15-year old children problem solving tasks which had various methods of solution. For example, one problem was to decide if two whole numbers whose product is one million could have factors that did not include a number containing a zero. He found that (different) children used either trial divisors, prime factors or simple problems to obtain a solution. Individual children, however, did not consider multiple strategies easily. They stuck primarily to an algebraic re-statement. Thus, the multiple strategy effect observed by Tabachneck *et al.* could not operate and these children did seem to arrive at impasses. Similarly, Stacey & MacGregor (1995) found evidence that children may have difficulty in adjusting their solution strategy over the course of a set of problems. They may get fixed on solution methods which worked with earlier problems, rather than adjusting their solutions to the problem in hand. This result accords with the psychological literature on functional fixedness (Duncker, 1945; Luchins & Luchins, 1950). Stacey & MacGregor conclude that students need to know that there are alternative models of a situation and that their initial perception of underlying structure may not be the most useful

The problem is not only limited to primary mathematics. It has been demonstrated by much older students working in more complex domains. For example, Schoenfeld (1987) reports that his students would commonly pick difficult techniques to solve calculus problems when much simpler ones existed that would have allowed faster solutions. It would seem his students failed to consider whether alternative (simpler) solutions existed and that this failure was not due to poor mathematical knowledge since these more difficult techniques showed that students had proficient mastery of the domain.

2.4.3 Multiple Solutions

'How many different ways can you?' problems seem relatively common in the primary mathematics classroom, yet little research seems to have been done of them. As part of the PrIme project, Price and Foreman (1989) developed a number of different problems that asked this question. 'Ice cream cone problems' required children to make cones with different combinations of flavours. They were told to explore how many solutions could be found and to find out what the effect on price would be of various combinations. A similar problem was the 'Witch's spell problem': 'animals in the cupboard had different number of legs and the witch need's 24 legs in her pot, how many spells could she create from spiders, lizards, bats?.' This study was not formally evaluated but there was informal evidence that children responded well to these problems. In particular, some children poor at 'normal' mathematics responded well to these problems. Children's attention span on these problems was also good. One extension to this type of work would be to set problems where children could not be expected to find all the solutions. This would mean that 'the right answer' in these cases is not all of the possible correct solutions.

Lampert (1986a,b) set her students 'how many different ways' problems using money. Lampert's aim was to teach principled knowledge about multiplication (*e.g.* additive composition, commutativity of addition and multiplication) by linking it to operations on familiar objects. She favours collaborative teaching where students and teacher work together to make sense of mathematical problems. A typical question set to her pupils was to find out how many different combinations of nickels and pennies can make 82 cents. Children were encouraged to record all their different tries (not just correct ones) and were rewarded for this. She proposes that this approach allows children to develop strategies which can then be discussed with a class. She reports that practice with these types of problems resulted in children becoming more inventive in seeking out different decomposition and recomposition strategies.

2.4.4 Instruction in Multiple Solutions and Methods

Baroody (1987) has suggested that current classroom practice does not encourage children to consider multiple solutions to mathematical problems as legitimate. A number of researchers have considered how the teaching of mathematics might be adapted to consider multiple solutions. Two proposals are (a) to consider a wider variety of problems which require or invite different methods of solution and (b) to consider a variety of solutions to a single problem.

Fuson (1992), in the context of addition and subtraction, and Greer (1992), for multiplication and division, review the wide range of situations that problems involving these operations can model. They show how superficially similar problems invite different conceptualisations. Differences in semantic structure, mathematical structure, numbers and children's intuitive models have all been related to different solution methods (*e.g.* Bell *et al.*, 1989; De Corte & Verschaffel, 1987; Fischbein *et al.*, 1985; Mulligan, 1992). An example given by Greer is (a) A painter mixes a colour by using 3.2 times as much red as yellow. How much red does he need with 4.6 pints of yellow? or (b) A painter mixes a colour by using 3.2 pints of red for each pint of yellow. How much red should he use with 4.6 pints of yellow? The first is biased towards a multiplicative comparison and the second a rate conceptualisation. Equally, the numbers used in the situations can bias interpretation of the problem. One common finding (*e.g.* Bell *et al.*, 1989) is that when the multiplier is a decimal less than 1, children and adults are much less likely to accept that the situation calls for multiplication. However, as a number of researchers point out, children in Western classrooms are rarely exposed to the full variety of these different situations and are not given the opportunity to work with problems that provide counter-examples to their misconceptions (*e.g.* Nesher, 1987). This contrasts with approaches to word problems within the Soviet Union which cover a much broader range of types and mix the type of problems within a page (Fuson, 1992).

Other research has examined different methods of solution to the same problem. Cross-cultural comparisons suggest that in contrast to British and American classrooms Japanese and Taiwanese classrooms emphasise multiple solutions and strategies to a single problem (reviewed in Fuson, 1992). Schoenfeld (1992) reports that students in an American high school class were expected to work through 25 problems in a 54 minute class (just over two minutes per problem). Western approaches to teaching often involve a large number of different problems each to be solved by a single method. Japanese teachers spent much longer on a single problem and emphasised a variety of solutions. Individuals or groups of students were also often invited to present their solutions to problems to the class. This approach is similar to the one favoured by Lampert (described above).

This section has considered research on the generation of multiple solutions/answers to mathematical problems and has reviewed how this might best be taught. Section 2.6 considers how computer-based learning environments can teach children to consider multiple solutions. In the next section, the second aspect of number sense explored by this thesis is discussed

2.5 COMPUTATIONAL ESTIMATION

To successfully perform computational estimation a wide range of mathematical knowledge is required. LeFevre, Greenham & Waheed (1993) propose that three types of knowledge are necessary for computational estimation - conceptual, procedural and factual. The factual knowledge required for mental calculation includes knowledge of place-value and memorised number facts. Conceptual knowledge is needed to choose an estimation strategy that will produce approximate numbers to facilitate computation. Procedural knowledge is required to perform the approximation.

The following two sections describe procedural and conceptual knowledge necessary in estimation. Further aspects of mathematics that have been related to computational estimation are reviewed in section 2.5.3. The final parts of this section discusses the development of computational estimation and previous approaches to instruction in estimation.

2.5.1 Conceptual Components

Research into the conceptual knowledge necessary for successful estimation has been limited. An exception is a study by Sowder & Wheeler (1989) who examined how conceptual understanding develops through the school years without explicit instruction. They gave children descriptions of situations involving computational estimation. Children were shown responses by hypothetical students and then interviewed about these tasks. For example, to examine whether children recognised the need for approximate numbers, the subjects were presented with the problem of '9 × 52', together with alternative solutions of either '10 × 50 = 500', or '9 × 52 = 468 round to 500' (these were actually given in cover stories *e.g.* 9 boxes of candy which contain 52 pieces).

Their study outlined three main concepts which are directly implicated in successful performance of computational estimation: the role of approximate numbers, multiple processes and outcomes and appropriateness of estimates. All of these concepts have two aspects: the process of estimation and the outcome of an estimation procedure.

The first concept is the *role of approximate numbers*. Students must accept that approximate numbers are used to compute and recognise that the outcome of the computation can be approximate. For example, children were judged to lack this conceptual knowledge if they accepted that a viable estimation strategy was to use the exact numbers and then round (*e.g.* '19 × 31 = 589' so the answer is roughly 600). The majority of 7 to 8-year old children accepted approximate answers as valid. However, even 13 to 14-year-old children still preferred to compute an answer and then round it.

The second concept identified by Sowder & Wheeler is that of *multiple processes and outcomes*. To understand estimation thoroughly, students must accept that there can be more than one process for obtaining an estimate and more than one value for an estimate. To test this, Sowder & Wheeler gave children two different estimates to problems and asked them about the acceptability of these estimates. Their results suggest a dissociation between outcome and process. The majority of children accepted that there could be alternative right methods to solve a problem, but only a minority accepted that there could be different correct answers (one from twelve at ages 7 to 8 rising to six from twelve by age 14).

The final concept is *appropriateness*. Students should recognise that the appropriateness of the process depends on the context. Furthermore, the appropriateness of the estimate depends on the desired accuracy. Sowder & Wheeler identified this concept but did not examine it empirically.

Conceptual Principles

LeFevre *et al.* (1993) propose that there are two principles that summarise the conceptual knowledge required for estimation: proximity and simplicity. Proximity reflects the knowledge that the estimation should be reasonably close to the answer. Simplicity refers to knowledge of the best way to modify a given problem to produce a solvable intermediate solution. These two principles are interesting in that they will often be antagonistic - the simplest solution will not always be the closest.

LeFevre *et al.* gave estimation problems to children of various ages (8 to 9 year-olds, 10 to 11 year-olds and 12 to 13-year-olds) and also to adults. Amongst other aspects of this research, they looked for evidence of the conceptual principles in subjects' answers to problems and in their descriptions of estimates. Even the youngest children seemed aware of the simplicity principle of estimation. However, proximity was much less apparent at all ages. This could be due to a less developed awareness of this concept. Alternatively, limitations of processing abilities or other mathematical skills may have prevented application of this conceptual principle. Evidence that this is at least partly due to under-developed conceptual knowledge is provided by children's descriptions of estimation which rarely mentioned proximity. In contrast, adults showed awareness of both principles. Proximity seemed more important than simplicity. It guided their choice of strategy and was mentioned in their definitions of estimation.

2.5.2 Procedural Components

A number of studies have examined strategies that are used by successful estimators (*e.g.* Dowker, 1992; Reys *et al.*, 1991). The principal study is that of Reys, Rybolt, Bestgen & Wyatt (1982) who gave computational estimation tests to 1200, 11-17 year-old children. They selected the children who scored in the top 10% of each year group for further interviews. They presented these subjects with further problems and attempted to classify their strategies. These were categorised into three broad classes - reformulation, translation and compensation.

Reformulation involves altering numerical data to produce a more mentally manageable form without altering the structure of the problem. A number of reformulation strategies have been observed, the most common of which is rounding, but truncation, averaging and changing the numerical form are all common. An example of each is given below:

Rounding: where the number is transformed to the nearest multiple of 10, 100, *etc.*

e.g. estimate '283 × 178'

283 is closest to 300, 178 is closest to 200

so 300 × 200 = 60000

Truncation: where the right-most digits are ignored

e.g. estimate '283 × 178'

283 is changed 200, 178 is changed to 100

so 200 × 100 = 20000

Averaging: noticing that a set of multiplicands are all close to one number and then using that number.

e.g. estimate '253 × 168'

200 is roughly halfway between 253 and 168

so 200 × 200 = 40000

Compatible numbers: transforming a number to one more compatible with others in the problem.

e.g. estimate '$\frac{347 \times 6}{43}$'

347 is roughly 350, 43 to 42 so that you can cancel leaving $\frac{350}{7}$

Changing the form: using an approximately equivalent form of a number *e.g.* conversion between decimal and fraction

e.g. '0.3 × 100' could be changed to a percentage - 30% of 100

Translation strategies are the second kind of processes noted by Reys *et al.* (1982) This refers to the action of changing the mathematical structure of the problem to a more mentally manageable form. This form is then used computationally.

Order Changing: changing the order in which the numerical values are processed *e.g.* estimate '$\frac{347 \times 6}{43}$'

divide the 6 and 43 first and get 7 , so 347 divided by 7 is about 50

Operation Changing: changing the operations in a problem

e.g. estimate '8700 + 9200 + 9500'

9000 × 3 = 27000

The final strategy identified is *compensation*, where adjustments are made during or after computation.

Intermediate compensation: adjustments are made during compensation

e.g. estimate '2500 + 2100+ 2600 + 2500'

3000 + 3000 + 3000 = 9000 "round them all up except 2100 which is dropped to make up for the rounding"

Final compensation: adjustments are made after computation

e.g. estimate '3.2 + 2.7 + 1.3'

= 3 + 2 + 1 = 6 but then add a final 1 to get 7

Levine (1982) gave college students estimation problems involving multiplication and division. Her classification of strategies included:

- use of fractional relationships, *e.g.* '482 × 51.2' is transformed to '$482 \times \frac{1}{2} \times 100$'.
- exponents, *e.g.* '0.47 × 0.26' becomes '$(5 \times 10^{-1}) \times (3 \times 10^{-1})$'
- rounding both numbers
- rounding one number
- powers of ten (*e.g.* '76 × 89' is '100 × 100')
- known numbers

- incomplete partial products/quotients (where '$25410 \div 65$' would be changed to '$25400 \div 60$' and '$10 \div 5$')
- proceeding algorithmically, *i.e.* using known algorithms to calculate roughly, estimate and then combine all partial products or quotients

She found that rounding both numbers and using algorithms were the most common strategies. This probably reflects the lack of instruction in estimation reported by her subjects - commonly they reported either no instruction or just instruction in rounding.

Dowker (1992) examined a class of people who would be expected to be good estimators. She gave Levine's battery of multiplication and division problems to 44 academic mathematicians. In addition to the examples already given, she also identified a number of more unusual strategies. For example, she noted the use of the powers of 2 where each number is converted to 2 raised to a given power. So the problem '64.6×0.16' becomes '$2^6 \times 2^4 = 1024$' and then adjusted to give 10.24. Another is to use the rule $(a + b)(a-b) = a^2 - b^2$. The problem '$12.6 \times 11.4 = ?$' is converted to $(12 + 0.6)(12 - 0.6) = 12^2 - 0.6^2 = 144 - 0.36$.

Her results suggest that indeed academic mathematicians are excellent estimators, both flexible and accurate (1030 from 1270 solutions scored within 10% of the correct answer). The most common strategies were to exploit fractional relations and the use of known and nice numbers. However, the mathematicians also used a wide range of strategies. Upon re-testing a portion of the subjects, this flexibility was noticeable, as many used different strategies to those they had used originally. They also had much less reliance on school taught strategies than Levine's sample.

2.5.3 Related Components

A number of concepts and skills have been related to ability in computational estimation. These can be characterised as those concerned with mathematical skills and those concerned with affect and beliefs.

Rubenstein (1985) was interested in the relation between estimation and other mathematical skills. She developed a number of different tasks - an open-ended estimation scale, a reasonable vs. unreasonable estimation scale, a reference number estimation scale and an order of magnitude estimation scale. These tasks included whole numbers and decimals, numerical or verbal descriptions and all four operations (addition, subtraction, multiplication and division). A test designed to examine factors related to computational estimation looked at selection of the right operator, relative number magnitude, known number facts, operating with tens and multiple of tens, place value and rounding. She also used the Iowa Problem solving test which involves three separate subtasks: getting to know the problem, solving the problem and looking back.

Rubenstein gave these tests to 309 12 to 13-year-old children. She found no difference in performance depending on how the estimation tasks were described (*i.e.* verbal, numerical). This result differs from that of Morgan (1990) who found that estimation problems were generally answered more successfully if presented in a context as this encouraged children to abandon algorithmic strategies. Rubenstein did find that decimal numbers and multiplication and division increased the difficulty of problems. This replicates Bestgen *et al.* (1980) study of pre-service primary teachers. When examining the relation between estimation performance and other mathematical skills, she found that the most important dimensions were operating with tens, number magnitude and getting to know the problem scales. Surprisingly, place value, operating with multiples of ten, number facts and rounding were found to have almost no relation to estimation performance.

Other research has examined affective components that influence performance at computational estimation. Reys *et al.* (1982) identified confidence in one's own ability to do mathematics and estimation as important for successful performance in estimation. Another key dimension is the belief that estimation is useful. Morgan (1990) found that many children thought estimation pointless, almost invariably preferring an exact solution. This attitude was also prevalent amongst the younger children in Sowder & Wheeler's study. A final component identified is tolerance for error. Good estimators in Reys *et al.*'s study were more comfortable with some error and did not see inexact solutions as wrong. Dowker's study of academic mathematics also indicated that they were tolerant of error. She believes that mathematicians were comfortable with deviation because they were confident that they would be able to solve any problems caused by such error. Reys *et al.* (1992) in a study of Japanese children's estimating abilities found that tolerance for error was low and as a result the children in their study tended to use algorithmic approaches to exact computation rather than estimation strategies. One encouraging finding is that instruction in estimation strategies has been shown to increase favourable views of estimation (Bestgen *et al.* 1980). Again, the prevalence of these beliefs has been linked to common classroom practice stressing the importance of exact computation in order to calculate *the* right answer.

2.5.4 The Development of Computational Estimation

A number of researchers have been interested in the development of computational estimation abilities. Some of the research already discussed has a developmental component (Sowder & Wheeler, 1989; LeFevre *et al.*, 1993). Dowker has looked at very young children's estimation strategies (Dowker, 1989; Dowker, 1996). She showed that children as young as five can give reasonable estimates to simple addition problems that are just beyond what they could calculate (*Zone of Partial Knowledge* in her terms). However, these young children often gave unreasonable estimates; producing answers that are less than one of the addends or more than twice their sum. This tendency became much more pronounced as they were given sums that were further from their calculation competencies (as they moved through the zone of partial knowledge). The educational implications of such a result seem to agree with a Vygotskian approach to instruction. It would seem profitable to set children problems of a level of difficulty just beyond where they perform without help.

Case & Sowder (1990) used Case's theory of cognitive development (Case, 1985) to make predictions about the types of estimation tasks that could successfully be undertaken by children of different ages who had *not* been taught estimation. Case's theory extends from age 0 to age 19, but the two periods relevant to school age children are the dimensional stage (age 5-10) where children can only focus on one dimension at a time and the vectorial stage (ages 11 to 18) where children can co-ordinate two or more dimensions of a task simultaneously. Each of these periods involves three stages. Case & Sowder examined computational estimation and identified two subcomponents: (a) the ability to convert an exact multidigit number into an approximate number and; (b) the ability to add a reasonably large column of numbers.

They made a series of predictions concerning how children at different substages should perform at various mathematical tasks. At stage one of the dimensional period children should be able to compute single-digit sums and make single column nearness judgements. At stage two, they should be able to extend this to two digit problems. At stage three, children should be capable of computing two digit problems which require carrying or regrouping. At stage one of the vectorial stage, children should be capable of multi-digit estimation. This predictions rests on the premise that estimation requires co-ordinating double-digit rounding with double-digit calculation and as such requires vectorial competence. At stage two, some compensation should be possible. The final stage suggests that adolescents should have generalised competence at estimation.

Case & Sowder then gave theoretically appropriate tasks based on these predictions to children at ages, 6, 8, 10, 12, 14.5 and 16.5 who were judged by their teachers to have average mathematical ability. These children had not received specialised instruction in estimation. They found that performance on these tasks was very close to that predicted both in terms of the tasks that the majority of subjects passed at each level and in that the same tasks could not be passed by children at the previous level. For example, a task at the second dimensional stage asked children 'which is $25.85

closer to \$20 or \$30 ?' which 83% of 10-year-olds but only 33% of 8-year-olds passed. Problems at the first vectorial stage asked children to estimate '\$2.25 + \$3.42 + \$1.25'. No ten-year-olds passed this test; they attempted to calculate the answer exactly but 83% of 12-year-olds did. The authors conclude that these results should signal caution for the teaching of estimation. They propose that if estimation is taught too early, it may become divorced from meaning.

2.5.5 Instruction in Computational Estimation

Curriculum developers in many countries are calling for instruction in computational estimation (*e.g.* USA National Council of Mathematics 1989; Japan Ministry of Education 1989). In the UK, the National Curriculum (1994) also recognises its importance. The second (of five) attainment targets is *Number*. The National Curriculum states that pupils should come to understand and use number, including estimation and approximation, interpreting results and checking for reasonableness. The first mention of estimation is at Level 3 (within Keystages 1, 2 and 3). It has two components: the first is measurement estimation and the second computational estimation. For computational estimation, the National Curriculum states that children should recognise that the first digit is the most important in indicating the size of a number and children should approximate to the nearest 10 or 100. At level 4, children are expected to be able to check the validity of addition and subtraction calculations and at level 6, multiplication and division, by estimation. At level 7, children must accept that measurement is approximate and choose the appropriate degree of accuracy. At level 8, children should check that the magnitude of answers to problems are in order.

Little research has been published which describe how computational estimation might be taught. Schoen, Freisen, Jarret & Ursbatch (1981) describes two studies with 8 to 9-year-old children that taught front-end estimation and rounding. These researchers developed worksheets which teachers administered and the analogy of shooting at a target was used. Schoen *et al.* found that children could use the strategies that they were taught (as measured by pre-test to post-test performance) and that they retained these skills. However, only the skills of estimation were taught, concepts and beliefs were not addressed.

Trafton (1986) is concerned with how estimation might be taught to encourage an estimation mind-set. Essentially, these are the concepts and affective components related to estimation (described above). For example, he suggests that instruction should include accepting that estimation is useful, recognising when an estimate is appropriate and tolerance of error. He states that developing children's thinking and reasoning ability in estimation is as important as teaching estimation strategies. He proposes six routes by which this may be achieved: introduce estimation with examples where estimated amounts are used; emphasise situations where only an estimate is required; use real world application extensively; use easy examples in the early stages avoiding precision; emphasise the language of estimation; accept a variety of estimates; use oral work and group discussion and; emphasise estimation regularly. Other important aspects are to encourage children to have a sense of the relationship between the estimate and the exact answer; develop flexible thinking and decision making abilities by presenting situations where students can analyse what type of estimate should be used; and show them different approaches to the same problem.

Markovits & Sowder (1994) taught estimation skills to children in grade 7 (11 to 12 year-olds) as part of an on-going program to develop number sense (the other units were mental computation, fractions and number size). The estimation component consisted of seven lessons and work extended over nine class periods. They addressed a number of issues including the appropriateness of an estimate, degree of accuracy required in particular situations, absolute and percentage errors, compensation and reasonableness of an answer. Units included multiplication and division by numbers greater than and less than one. Lessons were led by a class teacher and were designed to allow children to actively question and explore their knowledge.

They found a number of positive effects of instruction. For example, before the study children had little understanding of relative error. If two estimate produced the same absolute error, they were considered identical even if the relative errors differed. After the intervention, 60% of children were able to judge relative error. They used compatible numbers strategies when estimating and many had overcome the pervasive 'division means make smaller' belief. This understanding was still present at a retention interview six months later.

This research suggests that while computational estimation may be difficult for children, that appropriate instruction can help them develop effective and flexible estimation skills. The next section considers how computer-based learning environments might be used to teach computational estimation.

2.6 COMPUTER-BASED APPROACHES

Little research has directly addressed the role that computer-based learning environments could play in supporting these aspects of number sense. This section will briefly review some relevant systems and consider the advantages that computer-based learning environments can bring to learning in these areas. In doing so, it is important to emphasise what computers do badly - and that is provide the discussion of concepts, strategies and solutions that teachers and peers do. However, they can be used as a tool to stimulate such discussion. Thus, the systems developed for this thesis are not intended to be used as stand-alones. They are designed to supplement classroom teaching, rather than replace it.

Some computer-based learning environments have implicitly supported the development of understanding that there can be multiple correct solutions to single problems. The main purpose of Shopping on Mars (Hennessy, O'Shea, Evertsz & Floyd, 1989) is to help children realise that different types of problems are sometimes best solved using different methods. It takes the form of a (non-violent) adventure game, in which two players land on Mars with no fuel and must negotiate a series of obstacles before reaching a fuel shop. Obstacles can be overcome by means of items purchased from nearby shops. The computer acts as shopkeeper and controls the level of difficulty. It can also intervene to encourage the use of efficient calculation tools and informal methods of calculation. Broken Calculator (Xploratorium 1991) asks children to solve calculations when various keys on the calculator are 'broken'. It keeps logs of the various problem solutions that learners attempt.

Chapter four describes COPPERS - a computer-based learning environments which looks at the simplest of the multiple solutions/answers/strategies problems. It is based on aspects of Lampert's teaching (reviewed above) and asks children to find different solutions to simple coin problems. Chapter Five reports two evaluation studies that examined whether COPPERS could effectively teach six to nine-year-old children to give multiple solutions to coin problems.

In the realm of computational estimation, it has been shown that children can respond well to developmentally appropriate instruction. However, with the exception of a few studies (*e.g.* Schoen *et al.*, 1981), little has been done to exploit the computer as a tool for developing children estimation skills. Yet, computer-based learning environments have a number of properties which make them highly appropriate for developing understanding of computational estimation. Computers can differentiate problems given to children adjusted through their zone of partial knowledge. They can log estimates of current and previous users allowing comparisons between different methods. Different solutions to the same problems can be presented allowing comparisons between estimation strategies. They can show the relation between an estimate and an exact answer and highlight differences between absolute and relative error. They can support other mathematical demands such as number facts and order of magnitude correction so that children can be prevented from failing due to slips which might otherwise be internalised as "I'm bad at estimating". Learning environments that are based on a guided discovery approach (Elsom-Cook, 1990) can also fade this support as children's competencies develop.

Chapter six reports on how a computer based learning environments - CENTS was designed and implemented by drawing on the research on computational estimation. Chapters seven and eight report on evaluation studies with CENTS.

2.7 CONCLUSION

This chapter has reviewed research which addresses the role of number sense in mathematical understanding. Two areas considered central to developing number sense are computational estimation and understanding that mathematical problems can have multiple correct solutions. Research on these aspects of mathematics has been described in detail. This research will be related to the design of two learning environments that have been developed and evaluated during the research conducted for this thesis (Chapters Four and Six). The second aspect of the design of these environments is the use of multiple external representations in supporting learning. Chapter Three describes research in this area.

CHAPTER THREE

Learning with Multiple External Representations

The use of multiple external representations (MERs) to support learning does not begin with the advent of computer-based learning environments. Teachers use MERs explicitly in order to make abstract situations more concrete. For example, children are often given a percentage such as 33% alongside a drawing of a pie chart with one third shaded. MERs may be used implicitly such as when a book contains pictures or explicitly, as with algebra word problems when learners must translate from one representation to another.

In addition, software that employs MERs has become increasingly available at all levels of education. For example, geometry packages such as Geometry Inventor (LOGAL / Tangible Math) allow tables and graphs to be dynamically linked to geometrical figures. Function Probe (Confrey, 1992) provides graphs, tables, algebra and calculator keystroke actions and allows students to act upon any of these representations. One of the biggest areas of expansion in educational software is with multi-media technologies. By definition, these systems involve MERs, often including video and spoken text. Even traditional classroom uses of MERs, such as using an equation to produce a table of values which can then be plotted as a graph, have been significantly altered by the introduction of graphical calculators.

Given this growth of multi-representational software, it is appropriate to ask what evidence is there that providing learners with MERs facilitates understanding. There is an increasing body of research which suggests that MERs can confer significant benefits when learning. However, much less is known about the conditions under which MERs are beneficial. Consequently, designers and educators have few principles to guide their use of MERs. One aim of this thesis is to identify under what conditions particular MERs facilitate learning.

This chapter will review the research on learning with MERs. This research is first discussed in relation to the basic premise of learning with external representations -

that the way information is represented has significant consequences for learning. In order to predict the effects of a particular representation or combinations of representations upon learning, we need a means of analysing particular representations. Hence, different approaches to classifying representations will be discussed in section 3.2. The role of MERs in supporting learning is reviewed and synthesised and considers both their advantages and disadvantages. Particular attention is paid to translation between representations as this issue is unique to learning with MERs. Finally, this chapter will end by proposing a set of core issues which could form the basis of a framework for designing multi-representational learning environments.

3.1 LEARNING WITH AN EXTERNAL REPRESENTATION

Interest in MERs has followed from the abundant evidence that the way in which information is presented in a single representation affects learning and problem solving. It is known that different representations of the same information can result in different inferential processes. Larkin & Simon (1987) contrasted interpretation of diagrammatic and sentential representations in terms of search, recognition and inference. Analysing a physics pulley example, they show how search processes are considerably more efficient in diagrammatic rather than sentential representations. With a more complex geometry example, the representations differ in terms of cost of recognition. They propose that sentential representations have a high cost of perceptual enhancement when compared to the diagrammatic representations.

Many studies have found that how information is presented affects what people learn. For example, Bibby & Payne (1993) studied the effect of different instructional representations (table, procedure, diagram) upon learning to use a simple control panel device. They found that users could learn to perform certain tasks more easily with one representation than another. Furthermore the effects of the different representations continued even after substantial practice (Bibby & Payne, 1996).

External representations have been shown to influence the efficiency of problem solving. Zhang & Norman (1994) manipulated different representations of the Tower of Hanoi problem. In some cases subjects needed to internalise the rules for manipulating the problem, in others they were embedded in the external representation. The fewer rules that needed to be internalised the better the subjects performed.

In general, there has been much research showing that the nature of an external representation can facilitate learning and problem solving - although there is still much debate about exactly how and why this effect occurs (for a recent critique of the arguments, see Scaife & Rogers, 1996).

Research on learning with an external representation has examined many different issues. Some of the dimensions that have been considered include: (a) the relation between the modality of the representation and learning (*e.g.* Larkin & Simon, 1987; Stenning & Oberlander, 1995); (b) the learner's expertise and experience with particular representations (*e.g.* Anzai, 1991); (c) whether representations with certain fundamental properties better support learning (*e.g.* Law Encoding Diagrams - Cheng, 1996c; representations at intermediate levels of abstraction - White, 1993); (d) whether representations are self-constructed or presented to learners (see Cox, 1996 for a review) (e) the media in which a representation was instantiated - such as computer-based or physical manipulables (*e.g.* Thompson, 1992), *etc*. The research reviewed in this chapter builds upon this to address a further issue - learning with more than one representation.

3.2 CLASSIFYING REPRESENTATIONS

Descriptions of representations are commonly based upon Palmer's analysis (Palmer, 1978). He proposes that any particular representation should be described in terms of (1) the represented world, (2) the representing world, (3) what aspects of the represented world are being represented, (4) what aspects of the representing world are doing the modelling and (5) the correspondence between the two worlds. Using this definition of a representation, a number of different approaches have been developed to classify representations. Although modality (*i.e.* graphical or propositional representations) is most commonly considered, there are many other ways of distinguishing representations. These different approaches can be described in terms of three alternative methods: (a) to identify equivalences between representations; (b) to propose taxonomies of representations and; (c) to distinguish fundamental properties of representations. These approaches are, to a large extent, complementary.

3.2.1 Equivalence

One fundamental property of representations already touched upon is the notion of equivalence. Two representations are said to be *informationally* equivalent if all the information provided by one representation is available from the other. Each could be constructed from the other. Two representations are said to *computationally* equivalent if the information directly inferable from one is as easily and quickly drawn from the other (Larkin & Simon, 1987). Representations such as graphs, tables and equations can therefore be said to be informationally equivalent (given a graph, one could construct the equation, given a table of value, one could construct a graph, *etc.*). However, there are obvious differences between the inferences they support. For example, variation is more implicit in the equation '$y=x^2+6$' than in the resulting graph where all values for x and y can be seen. A graph automatically orders values, a table need not order the values it contains.

Another form of equivalence identified by Kaput (1987) is that of semantic and syntactic equivalence. Two entities can be said to be *semantically* equivalent if they correspond to the same element of a reference field. If their equivalence can be defined solely in terms of the symbol scheme and its syntactic rules, then they are said to be *syntactically* equivalent. This allows us to distinguish between '$y=2x+3$' and '$u=2v+3$' (syntactically equivalence) and '$y=2x+3$' and '$2y=4x+6$' (semantically equivalent).

3.2.2 Taxonomies

An alternative way of analysing representations is to produce a taxonomic description of representation types. For mathematics, Lesh, Post & Behr (1987) have identified five distinct types of representation system

- experience-based 'scripts' - real world events serve to structure knowledge so that it serves as general contexts for interpreting other kinds of problem situations;
- manipulable models (*e.g.* Dienes blocks, number lines) which have 'built in' relations and operators which fit everyday situations;
- pictures and diagrams - static models that can be internalised as images;
- spoken languages, including specialised sub-languages like logic and;
- written symbols such as 'x+3 =7'.

At a higher level of granularity, Kaput (1987) identifies four classes of representation: (a) cognitive and perceptual representations, (b) explanatory representation involving models, (c) mathematical representations and (d) external symbolic representations. Taxonomies such as these two are based upon analyses of domains and, presumably, the authors' intuitions.

An alternative approach to producing a taxonomy of representations was taken by Lohse, Biolsi, Walker & Rueler (1994). They selected 60 different graphical representations and asked subjects to rate them on ten scales of properties (previously identified by subjects as relevant dimensions) such as attractiveness, difficulty, numerical, *etc.* These items were classified into eleven major clusters: graphs, numerical and graphical tables, time charts, cartograms, icons, pictures, networks, structure diagrams, process diagrams and maps clusters. Cox (1996) asked subjects to sort 87 different representation and identified 18 principle clusters. Similar categories to those of Lohse *et al* emerged allowing for the different corpora used (*e.g.* the addition of musical notation in Cox's study).

These sorts of taxonomies may be useful within a particular community to provide a commonly agreed set of terms. However, descriptions based upon analysis of visual similarity alone ignores function. Classifying steering wheels with Venn diagrams becomes much more sensible if the task given is to represent circularity. Completely different classifications might emerge if the task emphasised the function of the representations in relation to particular tasks. Cox's task emphasised problem solving and hence may be more useful for instructional situations, whereas Lohse *et al*'s concerned visual display. In addition, there will almost certainly be novice-expert differences in how these representations are perceived and used (*e.g.* Anzai, 1991; Koedinger & Anderson, 1990). This is particularly important when representations are used to support learning.

3.2.3 Fundamental Properties

Another approach to classifying representation is to isolate fundamental properties which define representations and then describe each representation in these terms. For example, Lohse *et al* (1994) analysed words and phrases people used when sorting representations and produced 10 different clusters of properties. These were spatial, temporal, difficulty, concreteness, continuous, attractive, part-whole emphasis, numeric, static, informationally rich.

All of Lohse *et al*'s representations were used to display information. However, Kaput considers whether a representation is used for action or display to be a further fundamental property. This difference is due not to absolute properties, but to features that evince different patterns of use. Display representations are not intended to be acted upon by users, except to build them initially. Action notations support a variety of transformations and actions. For example, transforming equations, substituting values for variables and extending tables are all examples of actions upon representations. The medium in which a representation is instantiated also affects the degree to which it can be used for display or action. The computer offers the potential to use traditional display notations with new forms of actions.

One of the most widely applied approaches to describing fundamental properties of representations is Green's Cognitive Dimensions (*e.g.* Green, 1989; 1990; Green & Petre, 1996). A cognitive dimension of a notation is a characteristic of the way that information is structured and represented. It interacts with the human cognitive architecture to influence the way people use the notation for a given activity. Examples include:

- *Hidden/Explicit dependencies* which describes the extent to which a representation hides important relationships between states
- *Viscosity* captures how much a notation resists change
- *Abstraction gradients* describes the maximum and minimum levels of abstraction
- *Role Expressiveness* describes how much the structures of the notations display their functional role
- *Secondary notation* concerns cues such as colour and layout that are not described by the formal semantics of a representation
- *Diffuseness* is a measure of the number of symbols per idea

Green proposed cognitive dimensions as discussion tools, available to the non-specialist as well as the HCI expert. The aim is to make clear the cognitive consequences of design choices. Green shows that representations designed to be ideal on one dimension (*e.g.* reduced viscosity) will have consequent effects on another (*e.g.* increased abstraction). This point in interesting in the context of MERs as one potential (partial) solution to a problem (*e.g.* hidden dependencies) may be to provide an alternative representation which makes this information salient.

A very different approach to analysing properties of representation is that of Stenning & Oberlander (1995). They identify *specificity* as a fundamental property of a representation that has direct ramifications for processing efficiency. Specificity is the demand by a system of representation that information in some class be specified in any interpretable representation. The specificity of a representation determines the extent to which the representation permits expression of abstraction. Based upon this property, Stenning and Oberlander propose that there are three main classes of representation: Minimal Abstraction Representational Systems (MARS), Limited Abstraction Representational Systems (LARS) and Unlimited Abstraction Representational Systems (UARS) (in increasing order of expressiveness).

MARS are representations in which there is exactly one model for each representation in the system, under the intended interpretation. An example of a MARS is a tabular representation where each object must be described as possessing or not possessing each property listed. To be minimally abstract the representation must always represent each of the objects and dimensions and must assign each object exactly one value on each dimension. Such representations are extremely restricted in their expressive power. LARS allow for more than one model for each representation. They remain limited in that abstraction is only permitted over models which differ with regard to an object's value on one dimension. So, for the table example, an object need not be described as possessing / not possessing a value (say by leaving a cell blank). Stenning & Oberlander propose that most graphical representations could be considered as examples of LARS. A representation can be classified as an UARS if its expressiveness depends upon equations or arbitrary dependencies. So a value in a table could now be expressed by an equation.

It is proposed that the class that each representation belongs to will allow us to predict their cognitive computational properties, with a LARS being more computationally effective than a UARS as these systems are syntactically constrained and limit the number of cases that must be computed over. This analyses have been used to describe the effectiveness of Euler Circles for solving syllogisms (Stenning & Oberlander, 1995) and to explain the failure of certain VCR interfaces (Williams, Duncomb & Alty, 1996).

This approach represents one of the most principled approaches to specifying how a fundamental property of a representation affects cognitive processes. It provides a idealised model from which to make predictions about how particular representations will support understanding. A remaining concern for learning and instruction, however, is how useful an analysis of the intended interpretation will be when learners will almost certainly have completely different interpretations of the given representation.

Palmer (1978) identifies three fundamental properties of non informationally equivalent representations. The *type of information* is simply the dimension(s) of the represented world that a representation encode. *Resolution* refers to the grain size of a representation. If a dimension describes n relations, the higher the value of n, the higher the resolution and the smaller the grain size. For example, some could be described as either short or tall (2 relations) or 5 feet 2 inches, 6 feet inch, 5 feet 11 inches, *etc*. *Uniqueness* refers to whether the representation directly supports the required inference or whether additional information is needed to construct the right interpretation. Palmer uses the example of a map where city size is either given by the size of black dots (the former case) or whether it is given using colour codes (the latter).

The final approach considered is that of Cheng's functional roles analysis (Cheng 1996a). Functional roles are capacities or features that a diagram may possess which can support particular forms of reasoning or specific problem solving. They do this by making the relevant information salient such that little computation has to be done. Examples of such roles are: showing spatial structure and organisation, capturing physical relations, displaying states or values and encoding temporal sequences. This level of analysis (between consideration of a diagram as a whole entity and analysis of elementary diagrammatic components) seems ideally suited to the task of designing external representations to support learning.

Given the breadth of the previous discussion, it is obvious that there is no single 'correct' way of classifying representations. Yet, in order to make predictions about combining different representations, it is necessary to describe the dimensions that makes each representation different. This thesis will utilise these three different approaches describing representations in terms of equivalence, taxonomy and property. In particular, representations will be described in terms of informational and computational equivalence. This has been manipulated over experiments to explore cases when representations that are fully, partially or non informationally redundant. Taxonomies of representations are used to identify each representation as belonging to a particular class of representations, in particular Kaput's and Lesh's models are used to distinguish between mathematical and pictorial representations. At more fine-grained level, reference is made to Lohse *et al*'s taxonomy. Four experiments have been concerned with the effects of combining representations of different types. Finally, properties of representations are considered along the relevant dimensions of Palmer's and Cheng's analysis. These approaches were considered to be the most appropriate to the fairly simple representations used within the learning environments and have been used to ensure constancy across representations that vary along other dimensions. Further detailed description of the representations used in COPPERS can be found in section 4.4.3 and those for CENTS in section 6.5.

The effects of providing MERs on learning will be discussed in depth in the next two sections. Firstly, by considering the advantages that may be provided by MERs and secondly by analysing the difficulties learning with MERs presents.

3.3 ADVANTAGES OF LEARNING WITH MERS

It is proposed that one significant factor hampering the development of generalised principles for learning with MERs has been the failure to recognise that MERs are used for quite distinct purposes. Consequently, while there has been much research on individual examples and some theoretical explanation, an integrative framework has been slow to develop. This section will review the evidence which suggests that the use of MERs can provide a number of benefits for learning. In order to begin to specify such a framework, research reviewed will be characterised in terms of three fundamental uses of MERs. These are proposed and discussed in turn:

- MERs support different ideas and processes
- MERs constrain interpretations
- MERs promote a deeper understanding of the domain

3.3.1 Different Ideas and Processes

Essentially, the basis of this assertion is that by combining representations that differ in either informational or computational properties, we can exploit the advantages of each representation in the representational system.

A common use of MERs is when the information varies between the representations in the multi-representational system. Thus, quite simply, each representation serves a distinct purpose. Sometimes, information is partially redundant between representations, in other cases, there is no redundancy between the information expressed by each representation. Tarski's World (Barwise & Etchemendy, 1992) provides a graphical display of elements of world (tetrahedrons, spheres, cubes) and a sentential representation of the logical description of the world. 'MoLE', Oliver & O'Shea (1996) expands on this by providing one representation to express the relation between different modal worlds, and another to illustrate each world's content (illustrated in Figure 3.1). In this case, there is no redundancy between the two representations.

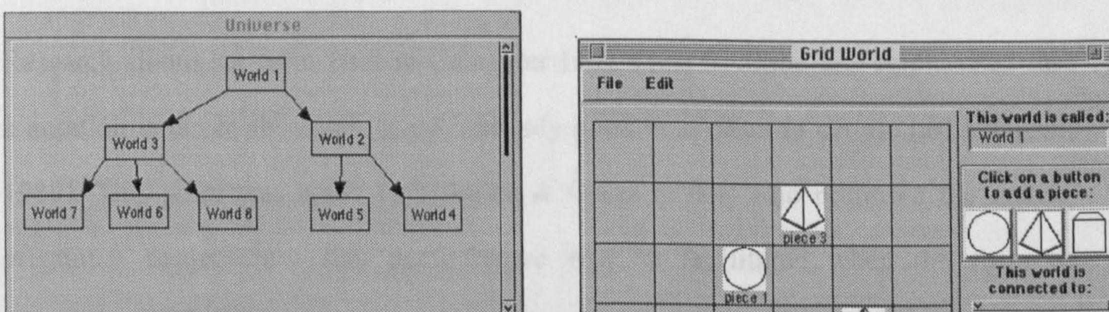


Figure 3.1 The relation and world descriptions representations in MoLE

Within the Internet Software Visualisation Laboratory (Mulholland & Domaigne, in press), one representation shows the search path that a PROLOG interpreter takes when satisfying a subgoal. A second textual representation describes the detail of each predicate in turn.

These approaches to using MERs are ideal if a single representation would be insufficient to carry all the information about the domain or would be too complicated for people to interpret if it did so.

A second use of MERs follows from Larkin & Simon's analysis that representations that are informationally equivalent still differ in their computational properties. For example, they proposed that diagrams exploit perceptual processes, by grouping together relevant information, and hence make processes such as search and recognition easier. Further research has shown that other common representations differ in their inferential power (*e.g.* Cox & Brna, 1995; Kaput, 1989). For example, tables tend to make information explicit, emphasise empty cells (thus directing attention to unexplored alternatives) and highlight patterns and regularities. The quantitative relationship that is compactly expressed by the equation '$y=x^2+5$' fails to make explicit the variation which is evident in an (informationally) equivalent graph. Therefore, MERs can be used to obtain the different computational properties of the individual representations.

Research stemming from Human-Computer Interaction tradition has emphasised that a notation is never absolutely good, but only good in relation to certain tasks (Green, 1989). This point was made by Gilmore & Green (1984) who proposed the match-mismatch conjecture - that performance will be facilitated when the form of information required by the problem matches the form provided by the notation. This analysis has subsequently been applied to a number of domains (*e.g.* comparing visual and textual programming languages; Green, Bellamy & Petre, 1991). Bibby and Payne (1993) examined how different, informationally equivalent, representations (table, procedure, diagram) supported acquisition of various aspects of device knowledge. Looking at performance on a simple control panel device, they found cross over effects. Subjects given tables and diagrams identified faulty components faster. However, those given procedures were faster at deciding which switches were mispositioned.

Therefore, a further reason for using MERs is when a learning goal requires different tasks to be undertaken. MERs can be designed so that an appropriate representation could be supplied for each of the tasks.

Research that examined the relation between different representations and strategies has also provided support for the use of multiple representations. Tabachneck, Koedinger & Nathan (1994) showed that the different representations used to solve algebra word problems were associated with different strategies. No single strategy was more effective than any other, but the use of multiple strategies was about twice as effective as any strategy used alone. As each strategy had inherent weaknesses, switching between strategies made problem solving more successful by compensating for this. Cox (1996) observed a similar effect when students solved analytical reasoning problems. He found that subjects tended to switch between representations at impasses and on difficult problems.

A further rationale often provided for the use of MERs is that there are individual differences in representational and strategic preference. Thus, if two alternative representations are provided, users could act upon the representation of their choice. Research examining the impact of various personality or cognitive factors in relation to learning with external representations has proposed differential effects of IQ, spatial reasoning, locus of control, field dependence, verbal ability, vocabulary, gender and age (see Winn, 1987). A common (although by no means invariant finding) is that learners defined as showing less aptitude in the domain benefit from graphical representations of the task (see Cronbach & Snow, 1977; Snow & Yalow, 1982).

This explanation for using MERs is often given by those who believe in distinct cognitive styles. However, cognitive style remains somewhat of a contentious issue as there are noted intra-individual differences as well as inter-individual differences. Furthermore, preferred style may not directly relate to task performance. Roberts, Wood, & Gilmore (1994) showed that when solving problems which involved spatial reasoning (*e.g.* mentally working out compass points in a route planning task), people high on spatial reasoning were more successful than people who scored low on spatial reasoning. Contrary to the simple prediction, this effect occurred as high spatial subjects typically used non-spatial strategies whereas people with low spatial ability struggled unsuccessfully to solve problems by mentally constructing spatial images. However, cognitive style is not the only reason why individuals may prefer certain representations. An account based simply on the proposition that learners will often having varying experience and expertise with different representations would also suggest that MERs would be beneficial.

It can be seen that there may be considerable advantages for learning with MERs. By combining representations with different informational and/or computational properties, learners are no longer limited by the strengths and weaknesses of one particular representation.

3.3.2 Constraints on Interpretation

A second use of MERS is to help learners develop a better understanding of a domain by constraining interpretation. This can be achieved in three ways.

Firstly, an additional representation may be employed to support the interpretation of a more complicated, abstract or unfamiliar representation. Thus, the second representation can provide support for a learner's missing or erroneous knowledge. For example, microworlds such as DM^3 (Henessey *et al.,* 1995) provide a simulation of a skater alongside a velocity-time graph (amongst other representations). Two misconceptions common to children learning Newtonian mechanics are that a horizontal line on a velocity-time graph must represent a stationary object and that negative gradient must entail negative direction. These misinterpretations of the line-graph are not possible, however, when the simulation shows the skater still moving forward. ReMIS-CL (Cheng, 1996b) teaches about the physics of elastic collisions. Law Encoding Diagrams (LEDs) are presented for learners to reason with and act upon. User's reasoning about information presented in the LEDs (*e.g.* initial and final velocities) can be debugged by comparison to an animated simulation of the collision.

A further example is that of Yerushalmy (1989) who describes a multi-representational learning environment for teaching algebraic transformations. It presents users with an algebraic window where they transform algebraic expressions. It also provides three graphs: the first displays a graph of the original expression; the second displays the current transformed expressions and; the third describes any difference between the two expressions. Consequently, learners are encouraged to check that their transformations are correct as graphs should not change if a transformation was legal.

Multimedia systems often exploit this aspect of MERs (*e.g.* Millwood, 1996), for example, by providing written and spoken text simultaneously. If children are developing reading skills and find the written text difficult, or if the spoken text is hard to understand (*e.g.* Shakespearean language, speech with a broad regional accent), then presence of the second representation may help support understanding of the first.

A second use of MERs to constrain interpretation is when one of the representations permits less expression of abstraction. To use an example based on Johnson-Laird's research (*e.g.* Ehrlich & Johnson-Laird, 1982), the ambiguity in the propositional representation 'the knife is beside the fork' is completely permissible. However, an equivalent image would have to picture the fork as either to the left or to the right of the knife. Thus, when these two representations are presented as a multi-representational system, interpretation of the first representation must be constrained by the second when the representational system is considered as a whole.

Finally, information expressed in each representation in a multi-representation system could describe different aspects of the same situation. Together this information may constrain interpretation about a domain. For example, the representation of an abstract sentence 'L(a) or L(b)' permits three valid interpretations (*e.g.* Adam is by the lake, or Bill is by the lake or both are by the lake). The second representation denotes 'L(a) or not L(b)' (*e.g.* Adam is by the lake, or Bill is not by the lake, or Adam is by the lake and Bill is not by the lake). Together, the representations constrain the interpretations about the situation. By reasoning about the conjunction of the representations, we know that the only situation that makes both these sentences true is L(a), (*i.e.* Adam is by the lake).

Thus, there are a variety of ways that MERs may constrain interpretation either by supporting missing knowledge, through providing representations which permit different interpretations, or through providing representations which provide mutually constraining information. The first use of constraint is likely to be the most common in learning environments.

3.3.3 Deeper Understanding of the Domain

Kaput (1989) proposed that multiple linked representations may allow learners to perceive complex ideas in a new way and to apply them more effectively. By providing a rich source of representations of a domain, learners can be provided with opportunities to build references across these representations. Such knowledge can be used to expose the underlying structure of the domain represented. On this view, mathematics knowledge can be characterised as the ability to construct and map across different representations. Similarly, Resnick & Omanson (1987) suggested that mapping between representations plays an important role in developing a more abstract representation that encompasses both. When they describe the process of abstracting over Dienes blocks and written numerals, it is the quantities that both representations express that permit mapping. Schwartz (1995) provides interesting converging evidence that multiple representations can generate more abstract understanding. In this case, the multiple representations are provided by different members of a collaborating pair. With a number of tasks (the rotary motion of imaginary gears, text from biology tasks where inferences must be made), he showed that the representations that emerge with collaborating peers are more abstract than those created by individuals. One explanation of these results is that the abstracted representation emerged as a consequence of requiring a single representation that could bridge both individual's representation.

Therefore, although research with this aspect of MERs seems more speculative than research on the first two purposes of MERs, evidence from both individuals and pairs suggest that an abstracted understanding can result from working with MERs.

3.3.4 Summary

There are many different reasons why MERs should be beneficial for learning. Research was reviewed and it was suggested that MERs are commonly used for one of three main purposes (*i.e.* that MERs support different ideas and processes, can constrain interpretations and promote a deeper understanding of the domain). For each of these uses, multiple sub-components were identified. For example, three different mechanisms by which MERs could support constraint were outlined. Furthermore, MERs used in a single system may fulfil two or more of these purposes simultaneously. For example, representations used to describe different aspects of a domain may also encourage abstraction if learners can map over them.

However, for these objectives to be met, learners must meet a number of significant learning demands. These are discussed in the next section.

3.4 DISADVANTAGES OF LEARNING WITH MERS

These potential advantages of MERs do not come without associated costs. Learners are faced with three learning tasks when they are presented with MERs. Firstly, they must learn the format and operators of each representation. Secondly, learners must come to understand the relation between the representation and the domain it represents. Finally, and uniquely to MERs, learners must come to understand how the representations relate to each other.

The following section will give examples of each of these learning demands and the problems associated with them.

3.4.1 Learning to Understand a Representation

The first learning task facing any user of a representation is to ensure that they understand each representation. They must understand how a representation encodes and presents information (the 'format'). In the case of a graph, the format would be attributes such as lines, labels, and axes. They must also learn what the 'operators' are for a given representation. For a graph, operators to be learnt include how to find the gradients of lines, maxima and minima, intercepts, *etc*. At least initially, such learning demands will be great, and will obviously increase with the number of representations employed.

Petre (1993) provides evidence for the effects of learning to understand a representation in regard to visual interfaces - countering the notion that graphical representations are inherently better than textual ones as they require no learning in order to use them. In observing differences between novices and experts, she showed that novices lack proficiency in secondary notation (*i.e.* perceptual cues that are not described by the formal semantics of a representation). Novices may find navigation of graphical representations difficult as they don't have the required reading and search strategies. In contrast to expert performance, they tend not to match strategies to the available representations.

3.4.2 Learning the Relation Between the Representation and the Domain

Learners must also come to understand the relation between the representation and the domain it is representing. This task will be particularly difficult for learning with MERs as opposed to problem solving or reasoning, as learners will also have incomplete domain knowledge.

Brna (1996) provides details from a numbers of domains about the difficulties learners face when attempting to relate a representation to a domain. For example, even fairly competent programmers who had received information about the elements of a new (visual programming) representation failed to clearly map the format of the new representation onto their existing domain knowledge.

Learners will not just have problems relating the format of a representation to the domain; they must also learn which operators to apply to the representation to retrieve the relevant domain information. To return to the graph example, children must learn when it is appropriate to examine the slope of a line, the height of a line, or the area under a line. For example, when attempting to read the velocity of an object from a distance-time graph, children often examine the height of line, rather than the gradient.

Laborde (1996) discusses the difficulties that students had in connecting geometrical properties to spatial properties when learning with Cabri-géometrè. Encouragingly, though, she believes that the computer environment acted to help children learn these relations by enlarging the range of visual phenomena possible (for example by dragging circles, tangents, *etc.)* whilst at the same time constructing these visualisations in a theoretically meaningful way.

Additionally, the operators of one representation are often used inappropriately to interpret a different representation. A representation of graph may be interpreted using the operators for pictures. This behaviour is seen when children are given a velocity-time graph of a cyclist travelling over a hill. Children should select a U shaped graph, yet they show a preference for graphs with a hill shaped curve (*e.g.* Kaput, 1989).

These problems do not only arise with abstract representation such as graphs, visual programming languages or geometric objects. Boulton-Lewis & Halford (1990) point out that even concrete representation such as Dienes blocks and fingers still need to be mapped to domain knowledge. Processing loads may still be too high for children to obtain the anticipated benefits of such apparently simple representations.

3.4.3 Learning to Understand the Relation between Representations

The final learning demand, unique to multi-representational situations, is that when MERs are presented together, learners must come to understand how representations relate to each other. Without abstraction across representations, any invariances of the domain may remain hidden.

Some multi-representational software has been designed expressly to teach such relations. For example, Green Globs (Dugdale, 1982) provides opportunities for learners to relate graphs to equations. A computer displays co-ordinate axes and 13 'green globs'. Students must generate equations that hit as many of these points as possible. This type of learning environment is common when the relation between representations is difficult. Grapher (Schoenfeld, Smith & Arcavi, 1993) consists of three micro-worlds: (a) Black Globs (similar to Green Globs described above), (b) Point Grapher which allows students to define function (*e.g.* 'y = 2x + 3') and produces tables and graphs and, (c) Dynamic Grapher, where families of function (*e.g.* 'y = mx + b') can be explored graphically. The instructional goal of the micro-world is essentially to develop the complex set of mappings that describe the relation between graphs and algebraic expressions ('the Cartesian Connection' in Schoenfeld's terms). A further example from the function domain is Confrey's Function Probe (*e.g.* Confrey & Smith, 1992). This provides students with graphs, tables and equations, plus a calculator keystroke representation which allows buttons to be built to generalise procedures. Again, the stated instructional aim of this program is to teach students to co-ordinate their actions on these different representations.

Other environments have been designed to exploit translation to some other instructional end. One example for the primary classroom is that of the Blocks World (Thompson, 1992) which combines Dienes blocks with numerical information. Users act in one notation (such as the blocks) and see the results of their actions in another (numbers). Thompson found that average and above average students developed better understanding of the number system structure and algorithms than students who had used a non-computerised version. Kaput (1992) proposed that the automatic translation provided by the learning environment supported the development of this knowledge by reducing cognitive load. MathsCar (Kaput, 1994) presents learners with a vehicle whose motion is mapped onto a variety of representations (*e.g.* distance-time and velocity-time graphs, odometers, clocks, auditory feedback). The aim of the system is to support understanding of calculus by making accessible core ideas such as the relation between change and accumulation. Again, the key idea is that translation over these representation will develop understanding when supported (for example, by making translation activities time-independent).

Not all researchers are optimistic about the potential for using multi-representational software to teach about translations between representations. Pimm (1995) warns that linking representations may not be neutral. He suggests that one representation will come to predominate and that by doing so it will no longer be viewed as a representation. Thus, meaning will not be associated with the relation between representations, but with the one dominant representation.

In addition, a number of researchers have noted the problems that novices have in learning the relation between representations. Tabachneck, Leonardo & Simon (1994) report that novices learning with MERs in economics did not attempt to integrate information between line graphs and written information. Students' performance on quantitative problems, where answers could be read off from graphs, was good, but it was poor on problems requiring explanation and justification. A similar pattern of results was found for graph generation as well as interpretation. This contrasted with expert performance where graphical and verbal explanations were tied closely together. Similarly, Yerushamly (1991) examined 35 fourteen year olds understanding of functions after an intensive three month course with multi-representational software. In total, he found that only 12% of students gave answers which involved both visual and numerical considerations. Lesh, Post & Behr (1987) provide more examples of the difficulties that children have in translating between representations. In an apparently simple problem of choosing which of three pictures showed 1/3rd shaded, grade school pupils' and even college students' performance was surprisingly poor. For example, only 25% of 12 to 13-year-old children could select the right answer.

Borba (1994) emphasises the importance for competent performance with MERs of noticing both regularities and discrepancies between representations. Confrey (1994) also highlights the importance of contrasts in addition to convergence between representations. Yet, Yerushamly (1991) found that the few children in his study who used two representations were just as error prone as those who employed a single representation. He found that students seemed unaware of contradictions between answers in the different representations. DuFour-Janvier, Bednarz & Belanger (1987) report a similar phenomenon. When children were asked to subtract using both an abacus and conventional written symbols, they commonly did not recognise the correspondence between the two representations and were unconcerned if they obtained different answers from each representation.

Research on the components of expertise is physics, chess, programming, *etc.* is also relevant to this debate. Generally, it has been shown that learners tend to characterise problem representations by their surface features, not their deep structure (*e.g.* Chi, Feltovich & Glaser, 1981; Adelson, 1981). Consequently, learners may find it difficult to translate between two representations of the same or similar information, if the surface features differ.

Thus, there is considerable evidence that learners find translating between representations difficult. They frequently do not use more than one representation, even after extensive training with multi-representational software. Even when they are required to do so, they seem to treat each representation in isolation, not noticing the regularities and discrepancies between the representations which would have aided their understanding.

Three different learning demands of presented MERs have been described. It is obvious from this discussion that learners will not be able to benefit from the proposed advantages of MERs if they can not meet these demands. Each time a new representation is introduced to a multi-representational system, these demands increase. In all cases, the format and operators of a representation must be understood as must the relationship between the representation and the domain. In addition, as translation between the different representations is required for many of the uses of MERs, increases in learning demands will not be simply additive.

These first two learning demands are present when any external representation is used to support problem solving or learning. However, translation between representations is unique to MERs. In the next section, this learning objective will be discussed in greater depth.

3.5 TRANSLATION IN MORE DETAIL

A number of researchers in the field of mathematics have stressed the importance of translation between different representations for understanding. For example, Kaput (1987) argues that meaning in mathematics is constructed in four ways:

- via translation between mathematical representation systems;
- via translation between mathematical and non-mathematical systems;
- via a pattern of syntax learning through transformations within and operations upon the notations of a particular representational system;
- via mental entity building through the re-ification of actions, procedures, and concepts into phenomenological objects. These can serve as the basis for new actions, procedures, and concepts at a higher level of organisation.

Meaning is said to be developed in different ways by the four activities: the first two are said to promote 'horizontal growth' by extending referential meaning; whilst the second two promote 'vertical growth' by transforming actions at one level into objects and relations that serve as inputs for a higher level. Kaput claims that while the third form of learning is the most shallow, it receives the most attention within the school curriculum.

Behr, Harel, Post & Lesh (1992) proposed a model based in part on Bruner's enactive, iconic and symbolic modes of representation (Bruner, 1966). Figure 3.2 describes these relationships. The representations were originally described in section 3.2) Arrows denote translations between modes (translation within modes are obviously also possible but have not been represented in this diagram for the sake of clarity). At the moment, as the researchers note, this model does not specify which paths are necessary or crucial to developing meaning. It also does not attempt to describe the nature of the translation between different representations.

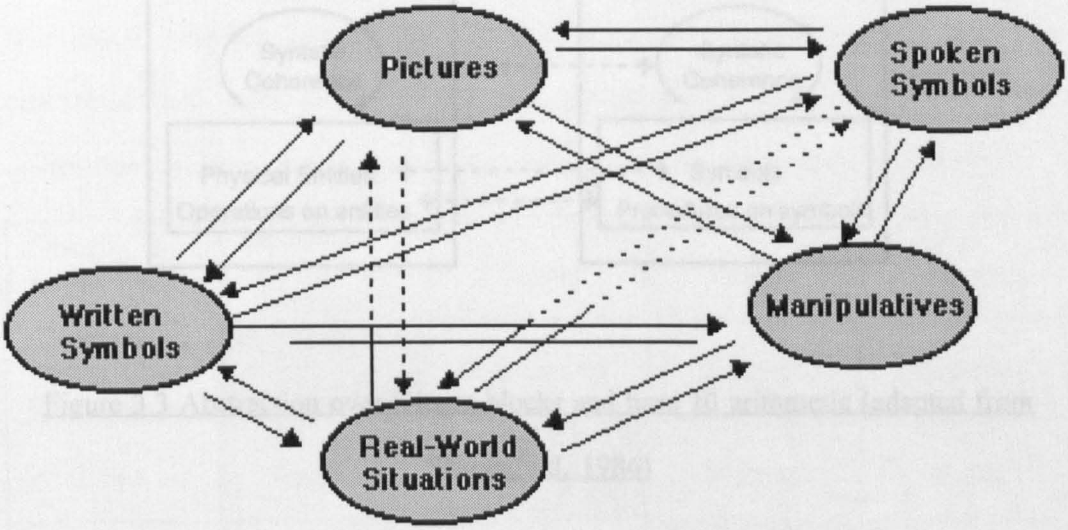


Figure 3.2 Lesh's model for translations amongst representations (adapted from Behr, Harel, Post & Lesh, 1992)

Resnick & Omanson (1987) examined the mapping between two of these representations in detail when considering instruction in arithmetic. They looked at the written symbols of subtraction (in the canonical row and column notation) and the manipulatives, Dienes blocks. Mapping instruction was given to nine to twelve year-old children which aimed to link principled knowledge of mathematics (*e.g.* composition, partitioning, compensation) to symbolic manipulation through manipulation of Dienes blocks. Children were required to keep a step by step correspondence between their manipulation of the blocks and written symbols as they solved problems. The aims of this instruction were summarised by Schoenfeld (1986).

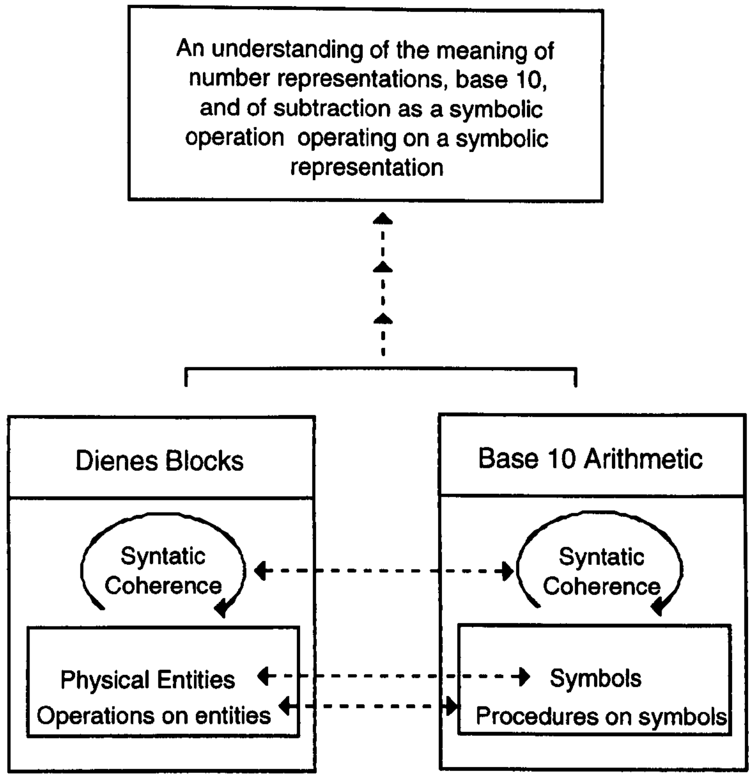


Figure 3.3 Abstraction over Dienes blocks and base 10 arithmetic (adapted from Schoenfeld, 1986)

Resnick & Omanson found that contrary to prediction this instruction was not successful at eradicating children's buggy procedures, but that it did lead to improved understanding of their knowledge of principles. They propose this finding rests on the fact that understanding how these principles may apply to borrowing does not mean that children will then apply them when performing borrowing operations. The children who were successful were those who had explicitly made verbalisations about the quantities involved in borrowing during the intervention phase - a result consistent with research into the self-explanation effect (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). In conclusion, they note how the automated performance of symbolic manipulation does not easily allow for application of principled knowledge. This has strong implications for the timing of instruction in translation between representations. There may be little point helping children learn some aspect of mathematics by teaching them to translate between the representations if they have already achieved automated performance with one of the representations.

Janvier (1987) provides a description of the nature of the translations between different representations (figure 3.4). He also indicates that translations between two representations are commonly achieved via a third (for example, formulae through tables to graphs). Interestingly, this may be changing with the advent of computer tools for manipulating representations; whether this change is beneficial or not is yet to be resolved. Using this table as an analytic tool, he also argues that when teaching translation between representations, that processes should be considered as complementary pairs (*e.g.* the interpretation of graphs as situations and verbal descriptions and the complement of sketching graphs from verbal descriptions).

From \ To	Situations Verbal Descriptions	Tables	Graphs	Formulae
Situations Verbal Descriptions		Measuring	Sketching	Modelling
Tables	Reading		Plotting	Fitting
Graphs	Interpretation	Reading Off		
Formulae	Parameter Recognition	Computing	Sketching	

Figure 3.4 Janvier's model of translation process between different representations (adapted from Janvier, 1987)

Research that has found that learners may not be able to integrate information from different representations presented in multi-representation software is not surprising given the complexity of translation between representations.

3.6 MEASURING TRANSLATION BETWEEN REPRESENTATIONS

Before leaving the issue of translation between different representations, two further studies will be discussed which illustrate different approaches to assessing translation between representations. The first illustrates a qualitative approach and the second, a quantitative approach.

Schoenfeld, Smith & Arcavi (1993) examined micro-genetically one student's understanding of function using the Grapher environment (described above). They described in detail the mappings between the algebraic and graphical representation in this domain. For example, figure 3. 5 illustrates just one of these connections (not all branches are shown).

2-Dimensional Graphs	The Cartesian Connection	Algebra
Slope is inclination, relative steepness.	The slope of the line segment that passes through the point (x_1-y_1) & (x_2-y_2) is given by the ratio of the line segments's vertical distance VD/HD = $(y_2-y_1)/(x_2-x_1)$	If (x_1-y_1) & (x_2-y_2) are co-ordinates on the line L, then $m=(y_2-y_1)/(x_2-x_1)$

Figure 3.5 An aspect of the Cartesian Connection (adapted from Schoenfeld, Smith & Arcavi, 1993)

Working with one student over a number of sessions, they showed from a detailed analysis of her transcript how a student could appear to have mastered fundamental components of a domain both in terms of algebra or graphs. However, as some of the connections between these modes of representation were missing, her behaviour with the representations was often misguided . For example, she could generate the slope-intercept equation for a line, yet not realise that the x value in '$y = x + 8$' would give the y value. Schoenfeld *et al.*'s analysis reveals the complexity of the mappings that can exist between representations.

Schwartz & Dreyfus (1993) examined how individuals integrated information between different representations designed to teach the concept of function. They used the TRM microworld which allows users to switch between algebraic, tabular and graphical modes. They defined two measures of students' performance with the software, a *convergence index* and a *passage index*. The former describes the efficiency with which a learner users available information to progress towards a solution. If a learner progress towards a right answer quickly, then they will have a high convergence index and will be assumed to have correctly interpreted the information at each stage in the solution. The passage index describes the extent to which a students keeps track of the available information when switching between representations. Thus, a student might be described as 'Pg = (4 2+ 2-)', which states that they switched representation four times, twice transferring all the available information successfully, and twice not. Using these measures they describe four prototypical students who differed in the success of their problem solving. For example, a student with high passage and convergence indexes was shown to be able to use the presented information successfully and keep track of it through the different representations. Another student who did not switch between representations 'Pg = (0 0+ 0-)' converged quickly on a solution through knowledge of algebraic representations alone. In contrast, less successful students had much lower convergence indices and did not pass information between representations successfully. Schwartz & Dreyfus conclude that such measures of representation use will provide useful insights into the design and use of learning environments.

One worry about the application of the passage index is whether it distinguishes between learners who understand how the representations relate to each other from those who understand how each representation relates to the domain. Transfer of information between two representations could be mediated through the domain or could occur by directly transferring information from one representation to another. It seems difficult to tell from the passage index which of these processes has occurred or, which seems more likely, how these processes were combined to translate information. (This issue is discussed further in section 7.4.3.)

3.7 A FRAMEWORK FOR LEARNING WITH MERS

A goal of this thesis is to identify some of the conditions under which MERs facilitate learning. It was argued that an integrative framework is currently missing and that one of the reasons for this was that not enough attention has been paid to the different purposes of MERs.

In this section, an initial set of dimensions will be proposed that are unique to learning with MERs. It is argued that these dimensions need to be addressed when designing multi-representational software for these different purposes. The different multi-representational learning environments discussed above vary along these dimensions, although descriptions of the systems often leave some of these dimensions unstated. No consideration is given to choices facing designers of all computer-based leaning environments in this framework (*e.g.* nature of help, whether representations should be constructed or given, *etc.*) as the focus is on the issues unique to learning with MERs.

These dimensions fall into two main classes: the nature of the representational system employed (points 1, 2 and 3 below) and the way these representations are implemented in an environment (points 4 and 5). These dimensions are only concerned with the nature of representations and how they are utilised and supported. For each particular learning goal, they would need to be applied in relation to tasks (*e.g.* learning, problem solving, communication) and to different users (experts, novices, children or adults).

(1) the amount of information per representation

(2) the similarity of representations

(3) how many representations should be used

(4) automatic translation between representations

(5) the ordering and sequencing of representations

Three broad types of use of MERs in learning environments were identified in section 3.1. These were (a) that MERs support different ideas and processes; (b) MERs can be used to constrain interpretations and (c) that MERs promote a deeper understanding of the domain. MERs may be used to fulfil one or a number of these different purposes within learning environments. It is argued that in order to achieve these different purposes, multi-representational systems should be analysed along the proposed dimensions. The learning environments discussed in earlier sections of this chapters did differ in terms of these dimensions. However, it is difficult to tell exactly how they could be defined as often these sorts of design decisions have not been made explicit. It is also difficult to avoid the conclusion that too often these decisions have not been made in a systematic, principled way. Each of the proposed dimensions will now be considered in turn.

One obvious difference between learning with one representation and learning with a multi-representational system. is in the way that information may be distributed in the MERs system. At one extreme, each representation could express the same information. Here, the only difference would be in their computational properties (Larkin & Simon, 1987). At the other extreme, each representation could convey completely different information. MERs may also be partially redundant, so that some of the information is constant across the representations. Thus, one important dimension to consider is the redundancy of information between representations.

A second difference with MERs is that they can also be presented in a wide variety of formats. The classic distinction is that of modality, but in section 3.2 a number of other ways of distinguishing between representations were discussed. Consequently, to achieve the identified purposes of MERs, they may be best served by different combinations of representations (*e.g.* graphs, tables, and equations; mathematical and non-mathematical representations).

A further necessary question facing designers of multi-representational learning environments is how many representations to employ. By definition, a multi-representational environment should use at least two representations, but many use more than that. A related issue is how many representations to use simultaneously? Many learning environments do not employ all the available representations at once.

In addition to decisions about the nature of the representational system, an additional consideration is how these representations are used and supported within a learning environment. For example, with the advent of computer technology, it is now possible to automatically link representations in a way that was not possible with pen and paper techniques. So, the fourth issue that should be considered is whether to provide automatic translation between representations such that a learner would act in one representation and see the results of these actions in another. As discussed above (sections 3.4.3, 3.5), learners have difficulty in translating between representations. However, it does not necessarily follow that we should provide this translation for users. It may be possible to over-automate and so not provide learners with the opportunity to construct knowledge of how to translate between representations themselves.

The final dimension is concerned with the ordering and sequencing of representations. If the MERs in a system are not presented simultaneously, two further issues arise. The first issue is the order in which representations should be presented. When an order has been determined, then decisions still have to made about when to add a new representation or switch between representations. Additionally, we need to consider whether these decisions should be under learner or system control.

These different dimensions will of course interact with each other. For example, without some degree of redundancy between representations, automatic translation is not possible. If all representations are co-present, then there is no need to consider the order that representations are presented in. Each decision taken about a dimension must take these interactions into account. In all cases, these issues should be considered in relation to the learning demands of MERs (discussed in section 3.3).

3.8 CONCLUSION

This chapter has reviewed research on learning with MERs. It has considered different approaches to classifying representations. These classifications are needed in order to describe representations used in a multi-representational systems and also to make comparisons between different multi-representational systems. The advantages that MERs can bring were outlined, but it was made clear that the use MERs within learning environments increase learning demands which must be met if these advantages are to be felt by learners. This chapter ended by proposing a set of dimensions which can be used to describe multi-representational software.

These dimensions will be applied to the descriptions of the design and evaluations of the two multi-representational learning environments which form the basis of the research conducted for this thesis. This research has focused on how the nature of the representational system may influence learning and has manipulated both the similarity of format and redundancy across representations. These have been considered in relation to the learning demands of representations, especially the demands of translating between representations. The number of representations has been restricted throughout the experiments to the simplest case of two representations. They are always co-present, so the issue of ordering and sequencing of representations does not apply. Chapter Four describes the first learning environment, COPPERS. A description of the goals and design of the system is given in relation to multiple solutions (reviewed in section 2.4) and learning with MERs.

CHAPTER FOUR

COPPERS: A Computer-Based Learning Environment for Multiple Solutions

This chapter describes COPPERS*, a learning environment that has been designed to teach children to produce multiple solutions to coin problems. Research reviewed in Chapter 2 described the general rationale for this instructional goal. One aspect of mathematical understanding identified by a number of researchers (*e.g.* Baroody, 1987; Schoenfeld, 1992) that is seen as a barrier to developing number sense was understanding that mathematical problems can involve multiple solutions. In this chapter, a more detailed description of how this research informed the design of COPPERS is given. The instructional goals of this system are considered and the features designed to support these goals discussed. Discussion will primarily avoid the implementational level - section 4.2, provides a brief description of these issues.

4.1 INSTRUCTIONAL GOALS

The domain taught by COPPERS involves arithmetic problems such as 'What is 3 × 20p + 4 × 10p ?' Users must answer this question by providing alternative decompositions of this total. One way to answer the problem is to calculate the total to this sum (*i.e.* (3 × 20p = 60p) + (4 × 10p = 40p) = £1.00) and then provide multiple decompositions to this total (*e.g.* £1.00 = '20p + 20p + 10p + 50p', or '10p + 2p + 2p + 1p + 5p + 10p + 10p + 5p + 5p + 50p', *etc.*). An alternative is to decompose the sub-totals (*e.g.* 3 × 20p = '20p + 10p + 5p + 5p + 10p', 4 × 10p = '1p + 2p + 2p + 5p + 5p + 5p + 20p'). To successfully solve these problems, children

---

* COPPERS was originally designed and implemented as partial requirement for an MSc in Knowledge Based Systems at the University of Sussex. A detailed description of the system implementation can be found in Ainsworth (1992). This chapter will provides an overview of the system in order to provide a background for the experiments based on COPPERS. Where additional implementation was conducted for this thesis, it will be indicated. $COPPERS_1$ is used to refer specifically to the MSc system, $COPPERS_2$ to the current version of the system.

must demonstrate a number of skills and have certain conceptual knowledge. They must (at least) know:

- the meaning of the symbols '×' and '+'
- how to perform the operations of multiplication and addition
- that the order in which the operations are performed is important
- that there are multiple decompositions to these problems
- how to calculate multiple decompositions

The educational goals of the system primarily concentrate upon the latter two aspects of these problems.

4.2 IMPLEMENTATION OF COPPERS

COPPERS was created for the Apple Macintosh computer using SuperCard™ 1.6, and written primarily in Supertalk. SuperCard™ is a variant of the more common HyperCard™ programming environment. To give an indication of the scope of the project, $COPPERS_1$ contains three windows, 15 cards, one dedicated menu, and many fields, graphics, buttons and icons. This is supported by around 2000 lines of code. $COPPERS_2$ includes three extra windows and 400 extra lines of code.

The design and implementation of $COPPERS_1$ included feedback from children and primary school teachers throughout the development phase. Children helped to develop appropriate wording of questions and contributed to the design of the interface. Teachers' advice was particularly sort in designing the feedback on answers. $COPPERS_2$ was developed after the first intervention study (sections 5.3), and was designed to make certain aspects of the system more salient (*e.g.* tabular feedback) and to inhibit poor interaction strategies by children (discussed in more detail later).

4.3 INSTRUCTIONAL APPROACH

Many of the educational principles underlying the design of COPPERS are based on a system for teaching multiplication in the classroom described by Lampert (1986a, 1986b). Two objectives of her scheme have been implemented in COPPERS. The first goal is to develop understanding that there can be multiple routes to the solution of mathematical problems. COPPERS serves this objective by providing questions for which there is one right total, but requires this total to be decomposed in a number of different ways. The second goal is the importance of allowing children's concrete and everyday knowledge to support the learning of other types of understanding such as principled and computational knowledge.

Both Lampert's approach and the COPPERS environment can be placed in a wider context of general mathematical pedagogy. Analysis of the nature of mathematical understanding (*e.g.* Schoenfeld, 1992; Lampert, 1990) and cross-cultural comparisons have prompted calls for revised approaches to mathematics instruction. For example, Fuson (1992) proposes that mathematics learning should involve:

- situations that are meaningful and interesting to children
- alternative solutions
- sustained engagement in mathematical situations, rather than on quickly finding answers
- analysis and acceptance of errors

Each of these dimensions will be discussed in relation to the design of COPPERS.

Meaningful Situations

The domain chosen by Lampert and adapted for use in COPPERS is that of coin problems. Lampert aimed to support the acquisition of principled and computational knowledge by emphasising connections to children's concrete and everyday knowledge. In order to do this, the mathematical situation must be both relevant and familiar to children. There is evidence that young children can benefit from presenting problems in which numbers refer to meaningful situations (*e.g.* Hughes, 1986), even if these situations are imaginary. However, it is not sufficient to simply present concrete manipulables (*e.g.* Schoenfeld, 1986).

It was proposed that money problems satisfy many of these criteria. They are familiar and meaningful to children. Furthermore, there is evidence that pre-literate children and unschooled adults can apply relevant concepts such as additive composition and place value when dealing with coin problems (Nunes & Bryant, 1996). Finally, computation is involved in dealing with money in everyday life. This should encourage children to frame the problem as a mathematical one and hence facilitate understanding (Kaput & Maxwell-West, 1994).

Alternative solutions

As discussed in section 2.3, children's beliefs about the nature of mathematics contrast strikingly with current views on the nature of mathematics. One aspect of this belief identified by, (amongst others), Phillip *et al.* (1994), Schoenfeld (1992) and Baroody (1987) is that there is only one correct way to solve a problem. Thus, the primary educational goal addressed in COPPERS is on finding alternative ways to solve money problems. Each problem has a single correct answer, but there are many different ways to produce this (*e.g.* £1.00 = '20p + 20p + 10p + 50p', or '10p + 2p + 2p + 1p + 5p + 10p + 10p + 10p + 50p', *etc.*). In section 2.4, understanding that one answer can be composed in different ways was identified as the simplest aspect of developing understanding of multiple solutions and strategies in mathematics. As such it seems appropriate to encourage children in their earliest formal education in mathematics to consider different solutions. This aspect of COPPERS' curriculum has been the subject of two experiments (see Chapter Five).

Sustained Engagement in Mathematical Situations

COPPERS is designed to ask children only few questions per session, so that the focus is upon generating many answers to a single problem. Lampert believed that such an approach to coin problems would encourage reflection upon strategy. In addition, COPPERS is designed so that there is no single best answer. Many comments made about the system during development suggested that the least number of coins solution should be presented as a model of the best way to solve a problem. However, this suggestion was resisted and a three coin solution is no more or less acceptable than a laboriously constructed 25 coin answer. When answers to problems are demonstrated by the system, they also deliberately avoid giving the quickest solution.

Analysis and acceptance of errors

In Lampert's classroom method, children were encouraged to record all of their answers in a 'summary table' to allow reflection upon why some answers were successful and some not. This has been implemented in COPPERS. After each question, a detailed breakdown of an answer is given in terms of all the number and types of coins used by children. In addition, a tabular representation is used to summarise this information for all the answers (right or wrong) given for each question. This has been the subject of one experiment with COPPERS (see section 5.2), and these representations are discussed in more detail later.

These four features of mathematics instruction were described by Fuson in the context of classroom teaching. However, they are obviously applicable to the design of computer-based learning environments. The following section discusses how COPPERS attempts to achieve these objectives by considering the design of the system in more detail.

4.4 SYSTEM DESIGN

This section will give a brief account of some of COPPERS' more important features. The following issues will be discussed: (a) problem representation and generation, (b) how problems are answered and, (c) feedback on answers. These areas are covered as they are key features of the environment and were selected for empirical analysis. Further details of the environment are discussed briefly in section 4.4.4, and at length in Ainsworth (1992).

4.4.1 Problem Representation and Generation

The common aspect to all problems is the addition, or addition and multiplication, of coins. However, the way this is represented can be varied, as can the complexity of the problem. In total, there are six stages of abstraction, with three levels of difficulty within each level.

The first two stages involve a single type of coin,. Level 1 requires addition, and level 2, multiplication.

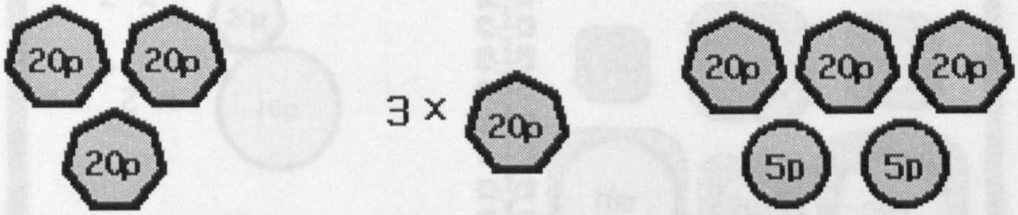


Figure 4.1 Example problems at level one, two and three

All the other stages include different types of coin. Level three presents pictures of coins that must be added to produce a total. Levels four to six involve multiplication and then addition of the partial products. The only difference between these levels is the way the problem is represented: at level four, the representation involves a mixture of text and graphics; at level five representation is solely textual and; at level six an algebraic notation is introduced.

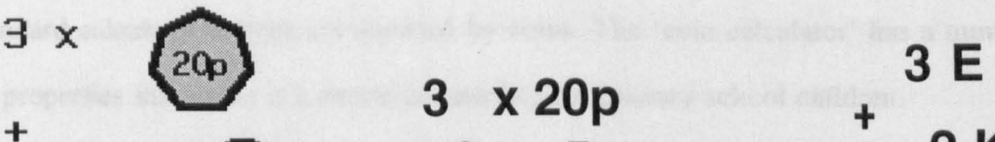


Figure 4.2 Example problems at level four, five and six

The difficulty of each problem is governed by three components. The first factor describes the number of different types of coins in a problem. If problems are represented using only pictures of coins, then they can be solved by addition alone. However, with two or more types of coin in multiplication problems, both multiplication of partial products and then addition of these subtotals is required. The second factor is the maximum number of each type of coin used (*i.e.* the multiplier).

Finally, the system can limit the range of coins used in problems (the multiplicand). In early stages, for example, the maximum value coin is 50p. All of these factors can be manipulated to meet a desired teaching objective.

4.4.2 Answering the Question

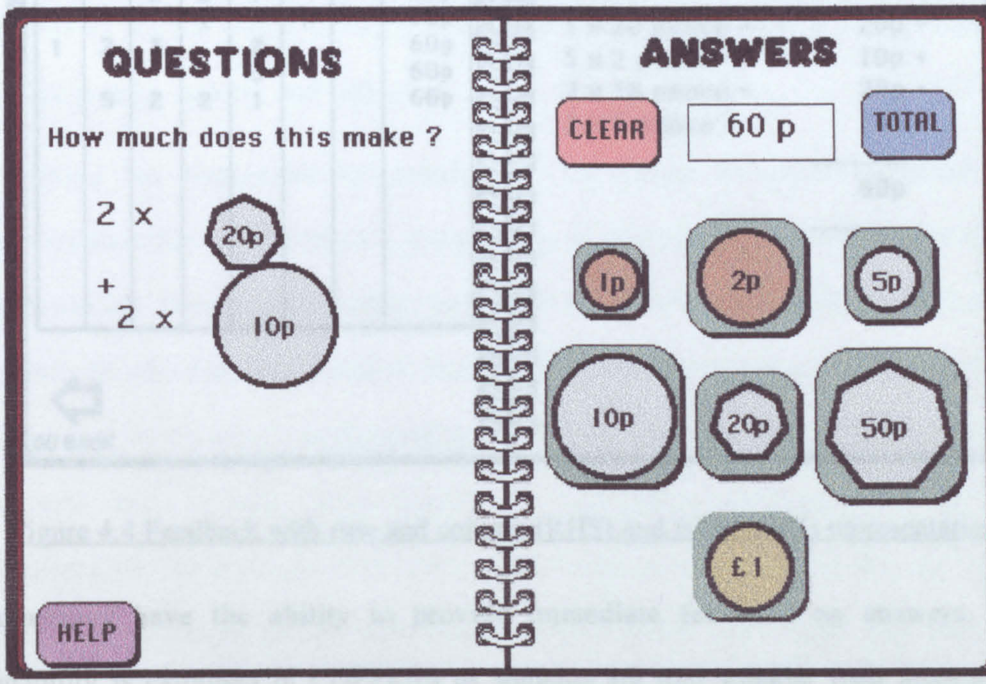


Figure 4.3 The 'coin' calculator (RHS) and example question (LHS)

In order to answer her questions, Lampert's pupils used pen and paper. The method chosen for COPPERS is the 'coin calculator', (illustrated in figure 4.3), where standard calculator buttons are replaced by coins. The 'coin calculator' has a number of properties that make it a desirable interface for primary school children:

- It is a simple to explain and use (pilot work found that even when the designer and user did not share a common language that its role was easily conveyed).
- It provides a familiar form of interface. The majority of children are now introduced to calculators during the primary schools years.
- It acts to reduce the burden of remembering number facts.
- Coins can be removed. For example, this can make the problem more difficult or to encourage more unusual decompositions.

4.4.3 Representations used for Feedback*

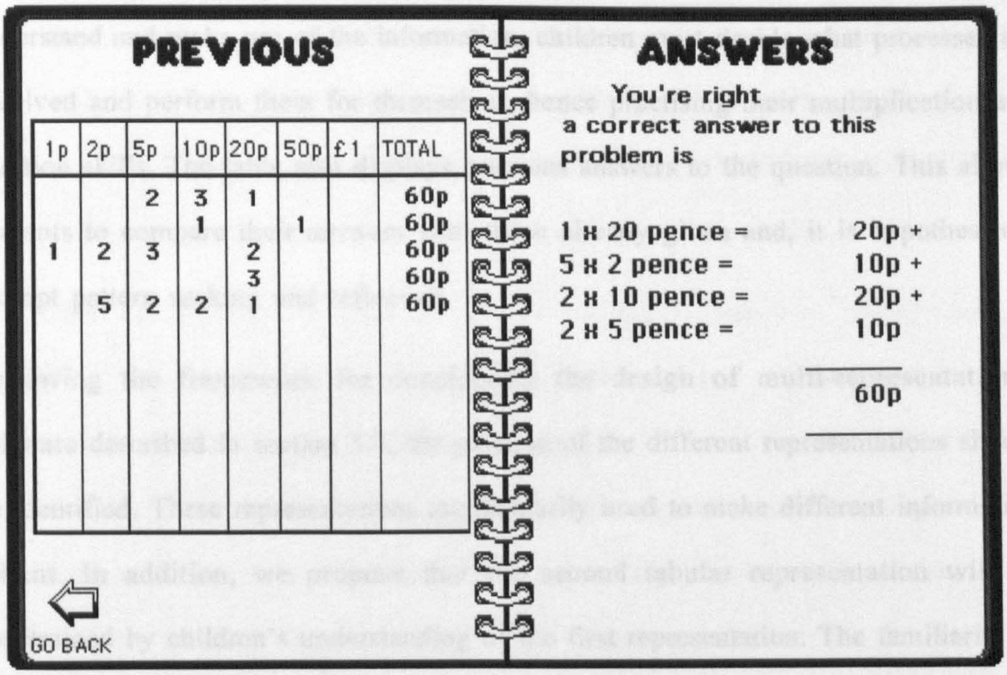


Figure 4.4 Feedback with row and column (RHS) and table (LHS) representations

Computers have the ability to provide immediate feedback on answers. This capability is exploited in COPPERS as students are told whether their answers are correct and are shown their answers broken down into partial products. This is performed in two ways and, by highlighting, the system encourages students to map between the different representations (Figure 4.4).

The first representation (RHS figure 4.4) is a common one in the primary classroom, and will be referred to as the column and row representation. The user is reminded of how many of each type of coin they used. The operations of multiplication and addition needed to produce the total are made very explicit in the column and row representation, making the arithmetical operations one of the most salient aspects of the representation.

The second representation (LHS figure 4.4) is tabular and similar to the one described by Lampert. In contrast to the row and column representation, the summary table is

* Re-implementation of feedback was performed before Experiment One to make the mapping between representations clearer.

less familiar to the children and the arithmetical operations are implicit. To understand and make use of the information, children must decide what processes are involved and perform them for themselves, hence practising their multiplication and addition skills. The table also displays previous answers to the question. This allows students to compare their answers with those already given and, it is hypothesised, prompt pattern seeking and reflection.

Following the framework for considering the design of multi-representational software described in section 3.7, the purpose of the different representations should be identified. These representations are primarily used to make different information salient. In addition, we propose that the second tabular representation will be constrained by children's understanding of the first representation. The familiarity of the row and column representation in COPPERS constrains the possible interpretations of the unfamiliar table representation by indicating the appropriate format and operators for the table representation. There is a second way that these two representations can constrain interpretation - by exploiting differences in the level of abstraction of the representations. Coin problems such as '5p, 10p, 5p, 10p' and '5p, 5p 10p, 10p' may appear very different to a young child if they do not understand commutativity. The tabular representation coin values used in COPPERS does not express ordering information. Therefore, if children translate between the representations, the equivalence of the two different orderings in the row and column representation is more likely to be recognised.

The classifications of representations introduced in section 3.2 can also be applied to these representations. The first approach identified was equivalence (*e.g.* Larkin & Simon, 1987). The two representations are informationally equivalent (when considering one answer) given the column and row notation, the table could be derived and vice versa. However, as described above, they differ in their computational properties. Both representations are mathematical according to Kaput's taxonomy. Applying Lohse *et al.*'s taxonomy, the row and column representation would be considered as a written symbol system and the tabular representation of coin values as a table. When considering the modality of the representations, it can be seen that the row and column representation is propositional; tabular representations are commonly referred to as semi-graphical. Thus, as described above these representations do not have identical formats. In COPPERS, both of the representations are used for display not action as they provide feedback on answers.

The other dimensions described in section 3.7 were number of representations, automatic translation and sequencing of representation. There are just two representations used for feedback and translation between the representations is signalled by means of highlighting. As both representations are presented simultaneously, the issue of ordering and sequencing of representations does not apply.

Learning demands were kept to a minimum with these representations. One representation is familiar (even if the place value concept it embodies is not) to children of the intended age range for COPPERS. The second representation is likely to be less familiar, but learning demands of interpreting the table may be supported by presenting it alongside the familiar representation. Signalling how the representations relate to each other should hopefully reduce the third translation demand. However, it is possible that the additional learning demands associated with the tabular representation will mean that it will not support the desired learning outcomes for children of this age. This question was addressed empirically in the Experiment One (Chapter Five).

4.4.4 Other System Details

The following section gives brief details of a number of other system features. These have not been evaluated in the thesis and are described simply to give readers a fuller summary of the design of the learning environment.

COPPERS include a very simple student model based on performance measures. This technique assumes that in order to describe students' knowledge, it is sufficient to measure how well they solve problems in that area - an obvious simplification.

However, the major advantage of this technique is that this information is readily available. Performance measures taken include number of right answers, number of wrong answers and the number of correct multiple solutions. These are used to govern such factors as problem difficulty and changes in representation. For example, if a student scored a number of wrong answers (this number can be set by anyone interacting with the computer who has 'teacher' status) then the system responds by offering help or by making the problem easier.

COPPERS has a limited number of teaching actions. One action it can take is to ask students whether they would like help. This strategy is a compromise between intrusive tutoring and help based purely on request, and is based on information in the student model. As the complexity of the problems is directly related to the notation used to display them, an obvious source of help is to rephrase the question (as many times as required or possible) in progressively more concrete terms. For example, if a student was working with a problem involving both addition and multiplication, then the question could be re-represented as a purely addition problem. Another form of help is to demonstrate a solution to a problem. This is the option taken when it is not possible to present the question at a lower level of abstraction. A further type of help offered is for questions presented with an algebraic notation, which simply reminds students of their letter to coin mappings.

The second teaching action that can be taken is to alter the difficulty of the problems in a domain contingent fashion. Poor performance will lead to users being given a problem that is easier than the one they have just completed, and good performance, harder. Again, how the parameters used to make these decisions can be adjusted by anyone with teacher status.

COPPERS teaching style is fairly directive. Students may only answer questions or quit the system. However, they retain some freedom of choice. Users share responsibility with the system for deciding whether the next question is harder or easier than the previous ones. Users are given this choice on roughly one in every three occasions.

4.5 COPPERS2

After a limited amount of formative evaluation and a larger scale experiment with 40 infant school children (Experiment One, Chapter Five), aspects of the interface underwent substantive re-implementation. In this section, the changes to the interface will be detailed, with only a limited explanation of these changes. Much fuller details of the experimental findings that motivated these changes can be found in Chapter Five.

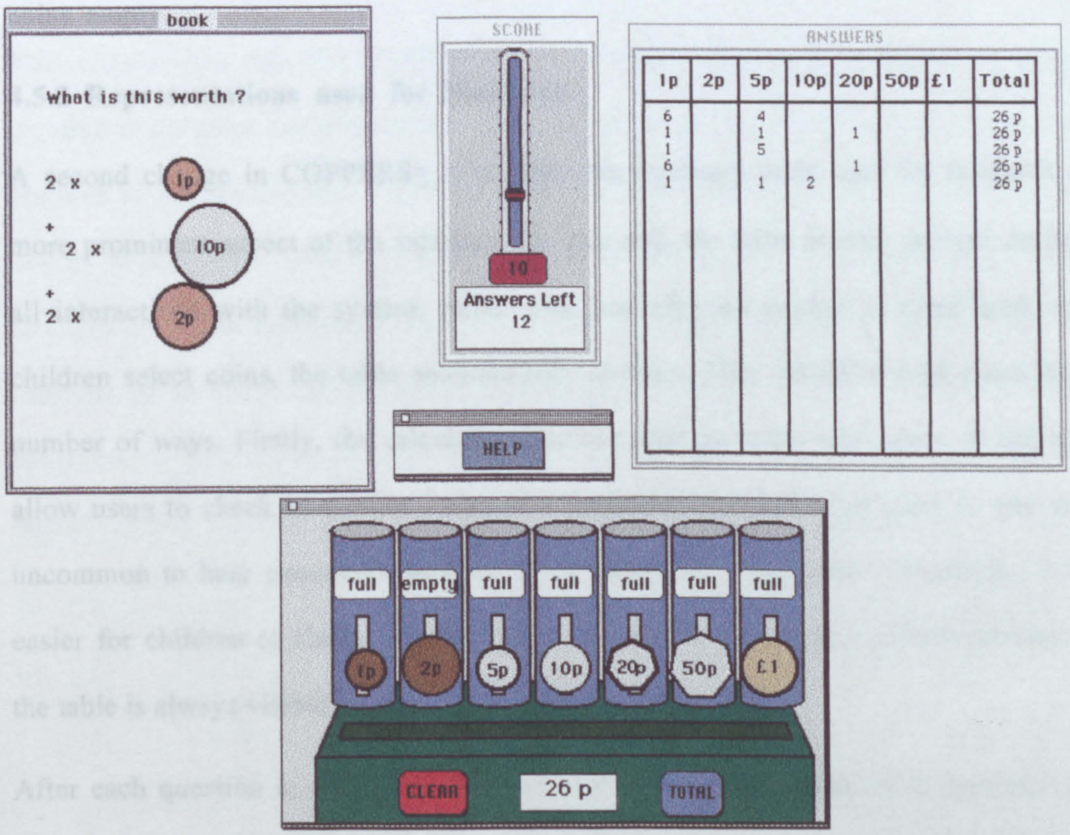


Figure 4.5 The new interface for COPPERS2

4.5.1 Answering the question

In the original system, users generated answers by pressing coins on the 'coin calculator'. Children of six to nine years had found this interface feature simple to use and understand. However, after the results of the first evaluation it was decided to modify the interaction metaphor to change some of the actions allowed to the user.

The principal motivation was to stop children producing an answer to a problem by simply copying the coins given in the question (*e.g.* if the question was '2 × 50p + 3 × 5p' pressing two 50p and then three 5ps). To inhibit this strategy, access to coins used in the question should be restricted. However, it was felt that the metaphor of calculator implies a limitless amount of coins and so therefore it was abandoned. To this end, the calculator was replaced by tubes of money that could either be full or empty. When children 'click' on a coin tube, a coin falls from the slot and into a money box that has a total indicator on the side (illustrated in figure 4.5). The new interface stops users from copying the question format by leaving one of the tubes of coins empty.

4.5.2 Representations used for Feedback

A second change in COPPERS2 is to make the summary table used for feedback a more prominent aspect of the interface. To this end, the table is now present during all interactions with the system, rather than just after an answer is completed. As children select coins, the table automatically updates. This should benefit users in a number of ways. Firstly, the calculator interface only gave the total score. It did not allow users to check how many coins of a particular type had been used. It was not uncommon to hear comments such as 'How many 10p's was that?'. Secondly, it is easier for children to check whether they have already given their current answer if the table is always visible.

After each question is completed, the row and column representation is still used to indicate whether an answer is correct. The mapping between this representation and the table continues to be made more explicit by means of highlighting the relevant portions of the table.

Another form of feedback was added to the interface to provide more information to the user about their performance. Specifically, children wanted to know how well they were doing. This was provided by adding a score window with a visual pointer and numerical score. The score reflects how many correct answers users had given. To encourage users to give multiple answers for each question, the score is biased so that a single answer on a question is worth 1 point, but a second answer on the same question is worth 2 points, *etc*. This window also displays how many answers users have left to give before they completed a session. These features allow learners to keep better track of their progress throughout a session.

A number of other minor changes were made to the interface including the provision of a navigation window and changes to the introduction that made the task demands more obvious and involved less reading.

4.6 CONCLUSION

This chapter has described COPPERS, a learning environment designed to teach children to consider multiple solutions to coin problems. It has described the principal features of the environment and how those features were re-implemented during the course of this thesis. Particular attention has been placed on the representations used in COPPERS. The following chapter describes two experiments that were designed to examine three questions: (a) whether children need to be taught to give multiple solutions to mathematical problems, (b) whether COPPERS met its educational objectives and, (c) what aspects of the system contributed to this goal.

CHAPTER FIVE

Experiments One and Two (COPPERS)

In this chapter, two experiments with COPPERS are discussed. These experiment had three main aims. The first aim was to explore children skills at producing multiple solutions to coin problems before intervention with COPPERS. The second goal of the experiments was to evaluate the effectiveness of COPPERS at supporting the development these skills. The final aim was to determine which aspects of the system design contributed to successful learning outcomes.

Experiment One

5.1 AIMS

5.1.1 Pedagogical Aims

There were two basic questions that needed to be addressed in relation to COPPERS: can young children easily produce multiple answers to coin problems? and if not, does the COPPERS environment provide appropriate support for them to develop the required knowledge and skills?

Learners' beliefs about the nature of mathematics have been identified as influencing the development of number sense. One belief that has concerned both researchers and teachers is children's belief that mathematical problems only have one correct answer (*e.g.* Baroody, 1987; Phillip, *et al.*, 1994; section 2.3). If these concerns also apply to English primary school children, it would be expected that the apparently simple task of producing multiple answers to coin problems would prove difficult. A review of the literature revealed little examination of these types of problems. Hence, one of the primary goals of the experiment was to examine how many solutions children could be expected to give before receiving teaching directly addressed at this issue.

The second goal was to examine whether COPPERS successfully met its objectives and could teach the skills and knowledge described in section 4.1 Consequently, children were pre-tested with types of problems set by the computer before the first intervention session. Two further multiple solutions tests were given following the intervention, one immediately after and one six weeks later.

5.1.2 Design Aims

The second goal of the evaluations was to treat COPPERS as a research laboratory to examine what features of a computer environment would help teach children to consider multiple solutions. Hence, the aim was not to compare a computer environment with others forms of teaching (*e.g.* teachers, pen and paper, real coins). Instead, detailed within system evaluations were conducted to identify which features contribute most to learning.

An analysis of the most important system features was undertaken to identify a series of predictions. For example, what would be the effect of the presence or absence of on-line help, immediate feedback, type of feedback, or learner control on a student's performance? As it is obviously not possible to evaluate all aspects of a system in a single study, a smaller list was selected for empirical testing. This focused upon two main issues that were considered central to the design of COPPERS: (a) how many solutions per question provides sufficient practice to develop this aspect of children's mathematical knowledge and; (b) how important are (multiple) external representations in supporting the development of this knowledge.

COPPERS' goal is to support multiple answers, hence the design of the system must consider how many solutions should be given for each answer. Thus, the first feature evaluated was the consequence of requiring multiple correct answers per question rather than just a single answer. Pilot work suggested that most users could be persuaded to give four answers per question, so users giving a single answer per question were contrasted with those giving four. It was proposed that practising multiple solutions would result in better learning outcomes.

The second aspect of the system that was evaluated was the representation of feedback. There is abundant evidence that the way information is presented affects how people reason (section 3.1). In addition, it has been claimed that presenting multiple representations can improve learning (section 3.3) either by supporting different processes, through constraining interpretation or by supporting abstraction.

COPPERS shows users whether their answers are correct and displays these answers in terms of partial products (Figure 4.4). This information is presented in two ways and, by using highlighting, the system encourages students to map between the different representations (described in detail in section 4.4.3). The first representation is a standard row and column representation. The second representation is a table similar to the one described by Lampert (section 4.4.3). The analysis of the role of these representations proposed in section 4.4.3, suggested that children who saw the additional tabular representation of their answer would have better learning outcomes than those who saw only the row and column representation. However, this will only be the case if learners can successfully meet the learning demands inherent in the addition of the second representation. It is known that young children have difficulty in working with tabular representations and often fail to use them successfully (*e.g.* Underwood & Underwood, 1987; Hoz & Harel, 1995). Chapter Three described in some detail the extra demands faced by a learner with more than one representation (sections 3.4, 3.5). Hence, it is proposed that these benefits will only be found if the children in the study could meet the learning demands of this additional, tabular representation.

5.2 METHOD

5.2.1 Design

The impact of these features (multiple answers and table feedback) was examined by producing several forms of the program which varied the presence of these elements. A three factor mixed design was used. The first factor was the presence of a summary table in addition to the row and column notation (table, notable). The second was the number of answers required for each question (multiple, single). Half the children were required to give four correct novel answers to four questions and half to give a single answer to 16 questions. The third factor was a within groups measure, time (pre-test, post-test, delayed-test). This resulted in four experimental groups, with ten subjects in each group. Subjects were assigned to conditions using a randomised block design by mathematical ability. Each had the same number of boys and girls and the mean age of the subjects did not differ.

A number of different measures were used to analyse children's performance. The first measure examined was the number of novel (*i.e.* no duplicate) correct solutions children gave to the three pen and paper tests. This measures both the children's skill at performing the arithmetic correctly and the accuracy and number of the decompositions. The second measure of performance is the number of solutions given irrespective of accuracy. This score includes incorrect answers created either because the initial calculation was in error or because of a mistake in the decomposition (*e.g.* a slip such as writing 41 rather than 42 one pence coins). The final performance measure examined was the percentage accuracy of the solutions, *i.e.* (total correct solutions / (total correct solutions + total errors)). Together, these variables permit analysis of whether any improvement in performance was due to increase in an accuracy, number of decompositions or a combination of both factors.

5.2.2 Subjects

Forty mixed ability year two pupils from a state infant school took part in the experiment. Their ages ranged from 6 years 10 months to 7 years 9 months; mean 7 years 3 months. All children were experienced with calculators and computers.

5.2.3 Materials

General mathematical Test

A general test of mathematical concepts and skills for seven to eleven year olds was given to all the subjects (Basic Number Screening Test - Gilham and Hesse, 1976).

Pre-test and Post-test Material

These tasks examined children's ability to give multiple solutions to the sorts of coin problems generated by COPPERS. The tests consisted of three problems, although it should be noted that there are many more possible answers. The problems were very similar to the ones the computer generates (see Figure 5.1 for an example and Appendix One for a complete test and all the possible solutions for one of these questions). Three parallel versions of the forms were created and children seated together were given different versions to prevent copying. In order to answer the question, the children were given blank pieces of paper and instructed to draw coins that would make the same total as the one in the question.

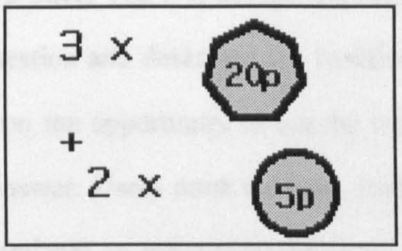


Figure 5.1 An example coin problem

5.2.4 Procedure

Pre-tests

Subjects were given the Basic Number Screening Test (BNST) in groups of five. The instructions for the test were given verbally following the recommendations for group administration. The test took approximately 20 minutes.

Multiple solutions pre-tests were given the following day. Again children were taken from the classroom in groups of five. In addition to the problem sheets and blank pieces of paper, children were given pictures of British coins to remind them of their values. Children were encouraged to give as many answers as they could for each question and not to start on the next question until they could not think of any more answers. The importance of working individually was stressed.

Computer Intervention

Subjects used the computer individually in a quiet corridor. The experimenter was present to help explain the instructions and how to use a mouse driven computer. To ensure sufficient practice with the system, each child used COPPERS twice (the time they spent on the computer was between 60 and 90 minutes), separated by approximately two weeks.

Although there were four different versions of the computer program, each one had the same basic structure. The user was greeted by a screen welcoming them to COPPERS. If necessary, subjects were given instructions by the experimenter about how to use a mouse. They were then given control of the mouse and worked through the instructions at their own pace. The instructions explained what the task was, how users should answer the question and described the functions of some of the interface features. Subjects were given the opportunity to see the computer generate a question and then demonstrate an answer. Users must read the instructions the first time they use the system, but this is optional on subsequent sessions.

COPPERS asks questions requiring addition, multiplication, or addition and multiplication of coins. All the problems set in this study required both addition and multiplication. All questions presented were generated dynamically, hence the problems each child was set were different. However, they were all generated according to the same rule: the highest value coin used was 50p, the largest allowed multiple was 3, and there were either 2 or 3 partial products.

Each subject was required to give sixteen correct answers. The subjects in the multiple answers condition gave four novel correct answers to each of four questions. The subjects in the other condition were required to give a single correct answer to each of the 16 problems they were set. All answers were generated using the 'coin calculator'.

Post-test

Two further multiple solutions pen and paper tests were administered to the subjects within a) ten days of their second computer trial and b) six weeks after that.

5.3 RESULTS

To examine the effects of the intervention, a number of [2 by 2 by 3] ANOVAs were carried out on the pre-test, post-test and delayed post-test data. The design for the analyses was 2 (table, no-table) by 2 (multiple, single) by 3 (pre-test, post-test, delayed post-test). The first two factors (feedback and practise) were between groups and the third, time, a within group repeated measure. The results from one subject were dropped. He was an extreme outlier scoring nearly six standard deviations above the mean at pre-test. Four children were unavailable to take the delayed post-test. All the data is presented per test. The average number of solutions per question can be found by dividing this total by three.

5.3.1 Multiple Solutions

A number of different measures were used to analyse the children's performance. The first measure examined was the number of novel (*i.e.* no duplicate answers) correct solutions children gave to the pen and paper tests. This can be seen in Table 5.1* .

Table 5.1. Number of correct novel multiple solutions by feedback, practice and time

	Table		No Table	
	Multiple	Single	Multiple	Single
Pre-test	2.60 (2.12)	1.78 (1.20)	1.88 (2.03)	4.38 (3.2)
Post-test	14.70 (5.54)	10.22 (6.20)	8.13 (6.31)	8.38 (4.87)
Delayed test	8.50 (5.21)	8.44 (5.88)	7.88 (7.02)	9.63 (10.10)

There was a significant main effect of time for the number of correct solutions ($F(2,62)=30.69$, $p<0.001$). The only significant interaction was for feedback and time ($F(2,62)=3.70$, $p<0.030$) (Figure 5.2). There was no significant interaction between practising multiple solutions (four versus one answer on the computer) and time ($F(2,62)=1.61$, $p=0.207$).

* Throughout the whole thesis, results are given as the average scores per test; figures in brackets are the standard deviations.

A simple main effects analysis showed that there were significant differences between the table and no-table at the post-test ($F(1,93)=5.479$, $p<0.0214$); subjects in the table condition produced significantly more correct solutions. This was the only occasion when the groups differed significantly. Tukey's unplanned comparisons showed that both the table and no tables groups improved from pre-test to post-test and from pre-test to delayed-test:

- table groups ($q=10.53$, $p<0.01$ & $q=6.36$, $p<0.01$)
- no-table groups ($q=4.78$, $p<0.01$ & $q=5.24$, $p<0.01$).

However, the table group's scores also decreased significantly from post-test to the delayed post-test ($q=4.17$, $p<0.05$), although remaining significantly above pre-test performance.

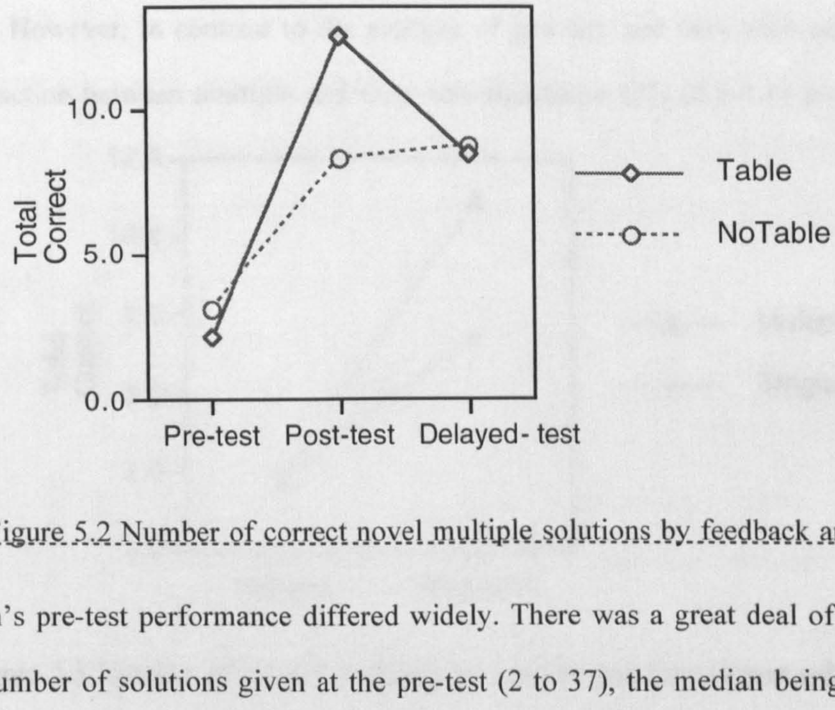


Figure 5.2 Number of correct novel multiple solutions by feedback and time

Children's pre-test performance differed widely. There was a great deal of variability in the number of solutions given at the pre-test (2 to 37), the median being 5. Hence, it was decided to perform an aptitude by treatment analysis, but due to small cell sizes it was not possible to split the data by the median. It was decided to look at those subjects who had most to learn and so higher performers at pre-test were removed from the sample and the results re-analysed. Twenty-nine of the subjects gave seven or less answers over the whole of the pre-test and the remaining eleven gave nine or more answers (this is displayed in table 5.2).

Table 5.2. Number of correct novel multiple solutions by feedback, practice and time

Lower performing subjects

	Table		No Table	
	Multiple	Single	Multiple	Single
Pre-test	2.43 (1.92)	1.87 (1.25)	2.00 (1.94)	3.00 (2.45)
Post-test	14.00 (6.30)	8.50 (5.78)	8.11 (4.65)	5.00 (1.87)

There were insufficient numbers of subjects at the delayed post-test, so only the pre-test and post-test scores were examined. When an [2 by 2 by 2] ANOVA was performed, the significant interaction between table and time remained ($F(1,25)=5.50$ $p<0.03$). However, in contrast to the analysis of practice and time with **all** subjects, the interaction between multiple and time was significant ($F(1,25)=4.44$ $p<0.045$).

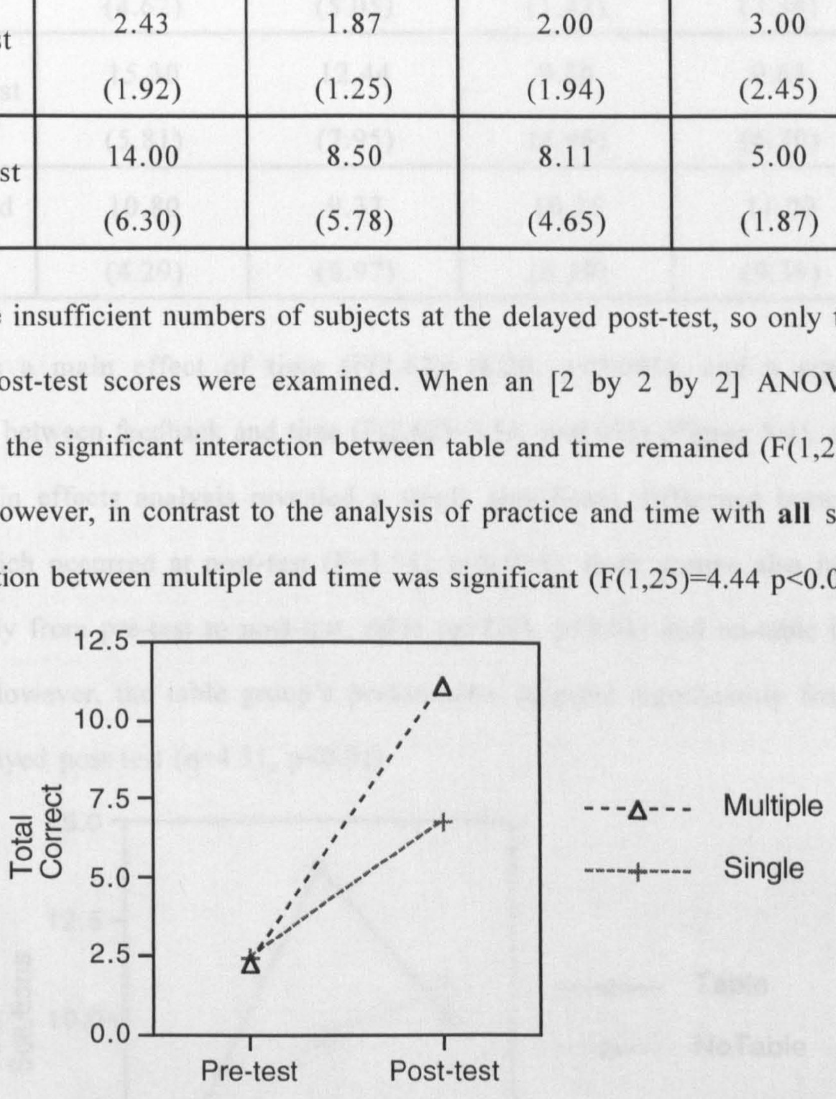


Figure 5.3 Number of correct solutions by practice and time (lower subjects)

Analysis showed that the groups only differed at post-test ($F(1,50)=5.912$ $p<0.019$), and that both conditions improved significantly: multiple ($q=8.51$, $p<0.01$) and single ($q=4.37$, $p<0.01$). This is illustrated in figure 5.3.

The second measure of performance is the number of solutions given irrespective of accuracy. Table 5.3 shows the results for all subjects expressed as the mean number of novel solutions for the three questions per test.

Table 5.3. Number of novel multiple solutions by feedback, practice and time

	Table		No Table	
	Multiple	Single	Multiple	Single
Pre-test	7.60 (4.62)	6.67 (5.05)	4.50 (1.42)	7.63 (3.34)
Post-test	15.30 (5.81)	12.44 (7.95)	9.50 (4.96)	9.63 (6.30)
Delayed test	10.80 (4.29)	9.33 (6.97)	10.25 (6.39)	11.00 (9.39)

There was a main effect of time ($F(2,62)=16.20$, $p<0.001$), and a significant interaction between feedback and time ($F(2,62)=3.54$, $p<0.035$) (Figure 5.4). Again, a simple main effects analysis revealed a single significant difference between the groups which occurred at post-test ($F=3.542$ $p<0.035$). Both groups also improved significantly from pre-test to post-test: table ($q=7.63$, $p<0.01$) and no-table ($q=4.71$, $p<0.05$). However, the table group's performance dropped significantly from post-test to delayed post-test ($q=4.31$, $p<0.01$).

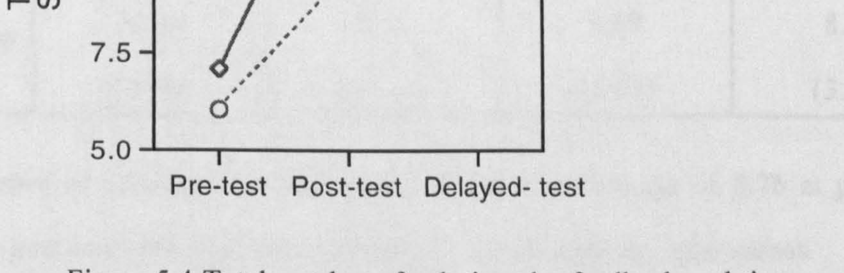


Figure 5.4 Total number of solutions by feedback and time

The final performance measured examined was the percentage accuracy of the solutions. This can be seen in Table 5.4

Table 5.4. Percentage accuracy by feedback, practice and time

	Table		No Table	
	Multiple	Single	Multiple	Single
Pre-test	41.56 (32.64)	35.96 (26.71)	35.51 (34.01)	57.32 (29.67)
Post-test	96.07 (5.52)	84.49 (16.61)	85.73 (13.23)	78.48 (18.32)
Delayed test	77.98 (29.47)	92.22 (10.42)	70.98 (32.81)	77.20 (29.96)

Analysis revealed a single significant effect, that of time ($F(2,62)=31.24$, $p<0.001$).

5.3.2 Types of Solution

A preliminary analysis of the types of solutions children produced was performed. The number of coins per solution, and the number of different types of coin per solution were examined at both pre-test and post-test.

Table 5.5. Number of coins per question by feedback, practice and time

	Table		No Table	
	Multiple	Single	Multiple	Single
Pre-test	7.29 (5.05)	6.05 (5.60)	4.41 (2.25)	4.63 (1.57)
Post-test	12.34 (4.99)	9.71 (5.13)	8.69 (5.63)	8.78 (3.87)

The number of coins per answer increased from an average of 5.76 at pre-test to 10.94 at post-test ($F(1,35)=22.46$, $p<0.001$). There were no interactions.

The final analysis was on the number of types of coins per question (Table 5.6). Minimally, there must be type of one coin per question. The maximum number of coins is seven (1p, 2p, 5p, 10p, 20p, 50p, £1.00).

Table 5.6. Number of types of coins per question by feedback, practice and time

	Table		No Table	
	Multiple	Single	Multiple	Single
Pre-test	2.3 (0.69)	2.13 (0.70)	1.99 (0.55)	2.10 (0.21)
Post-test	2.58 (0.27)	2.56 (0.31)	2.79 (0.55)	2.75 (0.32)

Again, there was a single main effect of time ($F(1,35)=31.84$, $p<0.001$). An average of 2.09 types of coins per question was given at pre-test which increased significantly to 2.67 by post-test.

5.3.3 Interaction Strategies

Subjects in all conditions had the task of adding and multiplying the coins in the problem regardless of whether they were then required to produce multiple depositions of this total. They commonly used two types of strategy to reach this total. The first strategy that the children employed was to use the 'coin calculator' to copy the format of the question. For example, if the question asked 'what is 2 × 20p + 3 × 10p?' they would enter two 20 pences and three 10 pences. This approach is unlikely to be associated with learning new flexible approaches to decomposition. The second type of strategy was to calculate either part or whole of the sum and then press different coins to reach this total. For example, users might say '2 × 20p = 40p', and then press four 10 pences to make the total.

The opportunity for using these strategies to calculate the total differs between the multiple and single conditions. In the single answer condition, subjects were set 16 different problems and so had 16 opportunities to use this strategy. In the multiple solutions condition, subjects gave four different answers per question and so had only to perform the initial calculation four times. To allow comparisons between the groups, the percentage of answers that were generated by the copying strategy has been expressed as a ratio of the number of different questions set. Only a few children used a single strategy throughout all of their interactions. The mean percentage of (first) answers generated by the copying strategy was 31%. The percentage of times that subjects used this strategy to obtain an answer to the multiplication problem was significantly negatively correlated with the total number of correct solutions at the post-test ($r=-0.3561$ $p<0.05$) and the total number of answers irrespective of accuracy ($r=-0.3472$, $p<0.05$). It would seem that the more times a student used this strategy on the computer, the poorer their subsequent performance on the pen and paper tests

5.3.4 General Mathematical Ability

Subjects had been given a general mathematics test (Basic Number Screening Test) at the beginning of the study. No significant correlation was found between the number of correct multiple solutions and the BNST at pre-test ($r=0.130$). However, the correlation between BNST and the total number of solutions (irrespective of accuracy) was significant ($r=0.334$, $p<0.05$). The BNST seems to predict the ability to produce multiple decompositions of an answer but not to ensure these decompositions are correct. The BNST scores did not correlate with any of the measures taken at the post-test.

The only other measure that correlated with the BNST was the subject's strategy for producing the initial total (*i.e.* percentage of the time the question format was copied) ($r=-0.4583$, $p<0.01$). The higher the BNST score, the less likely a subject was to copy the question format.

5.4 DISCUSSION

The aims of this first experiment were to a) examine children's base-line performance; (b) determine whether COPPERS could be used to successfully teach children to produce multiple solutions to coin problems and; c) to evaluate the role of various system components on learning outcomes. Specifically, this study examined the effects of practising giving multiple solutions for one problem and of presenting information about a user's performance in a summary table.

5.4.1. Multiple Solutions Performance

The first goal of the experiment was to investigate children's initial performance at producing multiple answers to coin problems. Given the research on the nature of children's beliefs about mathematics (*e.g.* Schoenfeld, 1992; section 2.3), it was proposed that children would initially perform poorly on this task. This hypothesis was supported by the experiments. At pre-test, the 6-7 year old children produced an average of 2.66 correct answers across three questions, *i.e.* less than 1 correct answer per problem. Even if errors are included, the children's scores do not improve substantially. Children produce an average of 2.2 answers per question. This experiment showed that primary school children do not easily produce multiple answers to these problems upon demand.

COPPERS appears to support the development of these skills. Subjects improved upon their pre-test scores by nearly 400% to produce an average of 3.5 correct novel solutions per question. Although there was a drop in performance to 2.9 answers per question at delayed post-test, scores remained significantly higher than pre-test. In addition, delayed testing took place in the last week of the summer term - less than ideal circumstances. Other measures of performance showed a similar effect: children increased the total number of solutions and increased the accuracy of the results. Thus, the improvement at post-test was not due to the development of a single aspect of the necessary skills and knowledge.

Due to limited number of subjects available, this study lacked a non-intervention control (subjects who took the pen and paper tests but did not experience the computer trials). Hence, it is impossible to state with complete certainty that the increase in the number of correct solutions was due to the computer intervention. It is conceivable that the subjects' continuing general mathematical education led to this improvement (although no classroom teaching specifically addressed this problem during the two months of the study), or that the subjects' progress was due to the effects of repeated testing.

Although COPPERS appears effective at teaching children to produce multiple solutions in this domain, replication with a non-intervention control is necessary before this claim can be made with complete certainty.

5.4.2. Types of solution

In addition to the increased number of solutions, the nature of these solutions may have changed after the intervention. However, it is difficult to perform detailed pre-test to post-test comparison. The number of correct solutions was very low at pre-test, and the main aim of the study was to test the effectiveness of the learning environment, not to probe children's strategies.

Both the average number of coins per question and the average number of different types of coins per question increased from pre-test to post-test. The vast majority of decompositions at pre-test were very routine. For example, the most common correct answer to '2 × 1p + 2 × 20p' was '10p + 10p + 10p + 10p + 2p'. Eleven out of the 18 children who gave a correct answer to this problem, generated this solution. The second most common was '20p + 20p+ 2p' given by nine of these children. Given the limited number of correct answers at pre-test, these two solutions accounted for the majority of correct answers. In total, only 15 different decompositions were identified for this question at pre-test from a possible total of 271.

At post-test, there was a completely different pattern of results. There was much greater variety of solutions, both within and between individuals, although there were still some preferred responses ('2 × 20p + 3 × 2p' = '1p + 5p + 10p + 10p + 20p', and '1p+ 5p + 10p + 10p + 10p + 10p'). These two decompositions were given by eleven and twelve children respectively out of a possible 39. Together they accounted for 15% of the solutions. From the 151 total answers generated for question one, 46 different decompositions were identified. In addition, solutions tended to be much less routine, such as '1p + 2p + 2p + 2p + 2p + 2p + 5p + 10p + 10p + 10p' or '1p + 1p + 1p + 1p + 1p + 1p + 1p + 1p + 1p + 2p + 5p + 10p + 20p'. The performance of the children at post-test suggests a much more flexible and inventive approach to decomposing numbers. Rather than learning a few common approaches to these problems from the computer, the range of solutions given suggests that children were generating their own decomposition strategies.

5.4.3. Representations Used for Feedback

The first aspect of COPPERS' design that was examined was the presence or absence of a summary table. This provided information about the current and, where relevant, previous answers to a question. It had been proposed that if children could learn to use the tabular representation, then its presence would improve performance. This hypothesis was supported by the study. The children who used the versions of COPPERS with the summary table produced significantly more novel correct solutions and more solutions in total at post-test compared to children who did not see the summary table.

There are a number of plausible explanations for the better performance of children in the table conditions. Lampert (1986a) proposed that such a table would be primarily useful for allowing subjects to compare their previous answers, especially those that had been in error, and to provide a record of work for their teacher to analyse. A further possibility was that the table served to remind students of answers they have already given. This may reduce repeat answers, either because of memory lapse or because of misunderstandings of commutativity of addition (the order that addition of the partial products is done is unimportant). However, these functions were proposed by Lampert and for the original computer system in the context of multiple solutions; the current study found that the presence of a table led to better performance regardless of the number of answers the children gave on the computer. While the table would allow students to compare wrong answers with right answers regardless of the multiple/single manipulation, only subjects in the multiple condition would be able to compare different right answers. Therefore it would seem that the table serves as more than a reminder of answers already given.

The table could serve a number of different functions in promoting the generation of multiple solutions. However, none of the records of computer use show any significant differences between the conditions (*e.g.* number of different coins used per question, numbers of buttons pressed, *etc.*) nor do subjects choose spend any longer looking at the feedback, although feedback does takes longer in the table condition. The table may have affected generation of solutions in many ways and a number of different approaches were observed: children tried to use as few columns as possible or as many as possible; they aimed to get high numbers in particular columns; made patterns across the columns, *etc.* For example, one subject noted his answer read like a palindrome across the table, 'its the same backwards as forwards' and tried to create another palindrome on his next go. It is therefore impossible with the granularity of information available from the computer records to distinguish these proposed different strategies unless a single one was used consistently. However, given recent emphasis on mathematics as 'the science of patterns', it is encouraging that children were beginning to seek and generate patterns in their answers.

Analysis of the different format and operators of the table and the row and column representations was conducted to explain the better performance of children in the table conditions. Tables tend to make information explicit, emphasise empty cells and hence direct attention to unexplored alternatives, highlight patterns and regularity, and represent variability (*e.g.* Cox & Brna, 1995). In this case, the table also emphasised order. Hence, the table representation makes different information salient. In addition, the table serves as a symbolic representation of the multiplication and addition procedures involved in finding solutions to the problems. Numbers in the columns must be multiplied by the column heading and then added together to get the total amount of money. The operators used to interpret a table therefore require children to practise multiplication and addition, skills that COPPERS attempts to teach.

The table and row and column representations simultaneously provide information on the same problem in different ways. Recently, a number of researchers have argued for the benefits of employing MERs (see section 3.3). Hence, the improved learning outcomes of the table groups may be due to the combination of row and column, and table representations used. It was proposed in section 4.4.3 that these two representations supported different inferences and could be used to constrain interpretations of the representation and the domain. Thus, the improved performance of the table group could be due to either of these reasons. Again, the granularity of information available from the computer records does not allow this question to be answered.

The framework proposed in Chapter Three suggested that children would only be able to take advantage of MERs if they could meet their learning demands. In addition, there is evidence that table representations may be hard for young children to understand and use (*e.g.* Underwood & Underwood, 1987). In this case, it did seem that children in this study were able to meet these additional learning demands. This success may be due to the way that the table was used within the system. A number of researchers (*e.g.* Kaput, 1992) have suggested that unfamiliar, abstract representations should be used for display before action and should be supported by more familiar or concrete representations. This suggestion matches the use to which the table is put within COPPERS. Translation between representations is known to be difficult (see section 3.4), again COPPERS attempts to support this learning demand by the use of highlighting to make clear the mapping between the representations. A number of researchers (*e.g.* Barwise & Etchemendy, 1992; Cox, 1996) have argued for mixed modality representations to support learning. As tabular representations are considered semi-graphical and the row and column representation is obviously propositional, this combination of representation comes close to achieving this objective.

Hence, although there are many explanations of the improved performance of the table group, the current experiment does not allow us to isolate which one(s) caused the observed improvement. These results do indicate the importance of (multiple) external representations for learning in this domain.

5.4.4. Practising Multiple Solutions

One of the central instructional issues in the design of COPPERS was the requirement to give more than one answer per question. This was examined by contrasting subjects who were required to give one correct novel answer per question with those who gave four answers per question. Preliminary analysis suggested that the hypothesis that practising multiple solutions would positively influence learning outcomes was not supported. Subjects who practised multiple solutions did not produce more correct solutions at post-test, nor was their performance better on any of the other measures taken such as accuracy of their solutions. However, when the (initially) lower performing subjects' scores were analysed separately, then practising multiple solutions was found to be important. For this group, children who had given four answers on the computer produced significantly more correct solutions at the post-test.

It would seem that children who already had some skills at producing multiple solutions did not show further improvement if they practised multiple decompositions but for children who initially had poorer performance, practising multiple solutions was important.

This obviously raises the question of why only the lower performing children were influenced by practice. One explanation considered is the nature of children's beliefs about mathematics. It is known that children of this age have difficulty accepting there can be multiple ways to solve problems (Baroody 1987; reviewed in section 2.3). The conditions in which subjects practised multiple solutions to one problem may not only have given them skills to perform multiple decompositions but also have legitimised the concept of multiple correct answers to a question. Higher performing subjects who gave some multiple solutions at pre-test demonstrated an understanding of this concept, although their skills could still improve. Lower performing subjects needed support to develop both the concept and the skills

Another potential explanation which could be tested empirically, is that the number of solutions that the computer requested was not sufficiently stretching for the high performing subjects. The computer asks for four answers per question; these subjects had produced an average of four answers initially. It would be possible to test this prediction by setting the number of solutions required to a higher value (*e.g.* eight). If four answers had not been sufficiently stretching, then this higher value should positively affect learning outcomes.

5.4.5. Interaction strategies

The way in which the children used the computer was related to the number of correct multiple solutions that they gave at the post-test, regardless of the experimental condition. As described above, there are a number of ways in which an answer to the initial calculation can be reached. The first way is to calculate the answer to the whole problem and then choose which buttons to press to reach this total. This provides practice in addition and multiplication, and decomposing a total. The second strategy involves multiplying each partial product and then decomposing that subtotal by pressing coin buttons. This strategy still involves practising decompositions and multiplication, but reduces the demands of the problem. A third strategy is to simply press the coins that are in the question. This means that not only are no decompositions made upon a total (and in the single answer condition no decompositions made at all), but also that subjects neither practice multiplication nor addition. It might therefore be expected that children who used this strategy would not show the same amount of improvement as children who commonly used either of the other two strategies.

The percentage of times that children copied the question when calculating the initial total was significantly negatively correlated with their scores at the post-test. They produced less correct novel multiple solutions and fewer solutions in total. This indicated that these subjects were not simply less accurate (as they had not practised multiplication and addition on the computer) but produced fewer decompositions overall. Subjects in the multiple conditions would still have practised decompositions but those in the single answer condition would have had no opportunities to do so.

It appears from these results that the 'copying' strategy may be detrimental to the process of learning to produce correct multiple solutions. It would therefore seem wise to ensure that students are not able to use this strategy. This concern motivated the new design for the interface discussed in section 4.5.

5.4.6 General Mathematics Ability

The primary motivation for testing children's more general mathematical knowledge and skills was to ensure an even distribution of these skills across the different conditions. However, it was possible to use these results to speculate about the relation between pre-existing mathematical knowledge and abilities to produce multiple solutions to coin problems. There were few significant correlations between scores on the BNST and either pre-test or post-test measures. The only significant correlation at pre-test was with the number of total solutions. This may suggest that children with higher maths scores were more likely to accept that there could be multiple answers for problems without necessarily having the mathematical skills to produce multiple correct decompositions. There was no correlation between post-test scores and the BNST. The intervention seems to have weakened this relation. However, the BNST did correlate with the percentage of time that children used the 'copying' strategy. This strategy was related to poorer post-test performance. Thus, there is no simple relation between mathematics skills as measured by the BNST and multiple solutions performance. This is consistent with the informal evidence of Price and Forman whose work on 'how many different ways' problems such as the Witch's Spell (reviewed in section 2.4) suggested that children who are not normally considered good at mathematics found these sorts of problems interesting.

5.5 CONCLUSION

Children's pre-test performance on multiple solutions was consistent with the research on children's mathematical beliefs. On average, children gave less than one right answer per question before receiving direct teaching. The improvement in children's scores from pre-test to post-test suggested that COPPERS could successfully teach children to give multiple solutions to coin problems. However, in order to claim that COPPERS was responsible for the improvement in performance comparison with a non-intervention control is necessary.

Providing learners with an extra tabular representation of their answers during the intervention improved post-test scores. This suggested that children had overcome the learning demands of this extra representation and were able to use it successfully. There was no overall relation between giving multiple rather than single answers during the computer intervention and learning outcomes. However, children with low initial scores performed better at post-test when they had practised multiple solutions on the computer. This issue motivated the design of the second experiment.

Experiment Two

5.6 AIMS

This experiment was designed to address some of the issues raised in Experiment One. In particular the question investigated was how to decide upon the right number of answers per question to improve learning outcomes.

5.6.1 Pedagogical Aims

The primary pedagogical aim of the experiment was to determine whether the effectiveness of COPPERS at supporting the development of multiple solutions understanding could be replicated. Experiment One had found nearly 400% improvement from pre-test to post-test. However, without a non-intervention control, it was impossible to determine how much of this improvement was due to the effects of repeated testing and how much from the intervention. Hence, the design of this experiment introduced a non-intervention control. Additionally, the children in this experiment were older than those in Experiment One. Experiment One used children from year two (six to seven years). The subjects in this experiment were from year four (eight to nine years). This provides the opportunity to examine whether the results of the first study were only applicable to infant school children.

5.6.2 Design Aims

The previous study had found only lower performing subjects benefited from giving four answers per question. One of the proposed explanations for this was that the higher performing children had not been stretched sufficiently on the computer. Consequently, this experiment further examined the issue of how to set children the 'right' number of solutions per question.

This was approached in two ways. It was hypothesised that in order to improve the performance of high performers, they needed to give more answers per question. Hence, an eight answer per question condition was included. This was far beyond the average number that children gave during the pre-test. However, this selection was still arbitrary. There was little to motivate this decision rather than say nine or twelve solutions. A second possible solution to this problem was to examine if users of the system would be able to set sensible numbers of multiple solutions for themselves. Hence, a second condition was introduced to provide learners with this degree of control.

The issue of learner control is a difficult one. There has been little theoretical basis for the use of high learner control, although some attempt has been made (Milheim & Martin, 1991). Although most researchers agree that control and perceived control is important for motivation (*e.g.* Lepper, Woolverton, Mumme, & Gurtner, 1993), it is difficult to draw robust conclusions from the learner control literature. Reeves (1993) states that control can mean very different things in different systems. For example, does it affect pace, content, representation or sequencing of instruction, with or without advice? Steinberg (1989) suggests that giving learners control will only be effective to the extent that they then chose a successful learning strategy. Results from the previous study suggested that an effective strategy would maximise the number of answers per question. Hence, it was proposed to examine what strategies learners choose and any effects on learning outcomes.

5.7 METHOD

5.7.1 Design

A two factor mixed design was used. There were four levels to first factor (condition) which varied the number of answers children were required to give to each problem they were given on the computer. The first group were required to give four answers to four question (four), the second group eight answers to two questions (eight), the third group could choose how many answers they gave per question (limited to 16 answers) (autonomous) and the fourth group were a no-treatment control group. This resulted in four experimental groups, with ten subjects in three groups and 20 in the autonomous group. This condition had more subjects as the style of interaction with the computer was examined to see if it was affected by previous ability or was related to learning outcomes. The second factor, time, was within groups. Each group had similar number of boys and girls. The mean age of the subjects and their scores on a maths test did not differ significantly. Dependent variables were identical to those used in the previous experiment.

5.7.2 Subjects

Fifty mixed ability year four pupils from a state junior school took part in the experiment. They ranged in age from 8 years 4 months to 9 years 2 months; the mean age was 8 years 9 months. All the children were experienced with calculators and mouse driven computers were present in their classrooms.

5.7.3 Materials

General Mathematical Tests

A general test of mathematical concepts and skills for seven and 8 year olds was given to all the participants; the Y1, Young (1979). This replaced the BNST used in Experiment One as it was deemed more appropriate to the age group.

Pre and Post-test Material

These were identical to those described for Experiment One.

5.7.4 Procedure

Pre-tests

Subjects were given the Y1 in groups of a ten and a second experimenter was present. The test took approximately 30 minutes. The multiple solutions pre-tests were identical to those used in Experiment One.

Computer Intervention

COPPERS was re-implemented before Experiment Two (see section 4.5 for full details). To recap, the following changes had been made:

- The coin calculator was replaced by coin tubes. One tube is empty, corresponding to part of the question. This is designed to prevent the 'copying' strategy.
- The table is now visible continuously and is updated as each coin is selected.
- A 'score' indicator was added which consisted of a pointer and a numerical score. This gives one point for the first answer, two for a second answer, *etc.*

The procedure of the computer intervention was identical to Experiment One. The only difference was that children in the autonomous group were told that they had to give 16 answers to the computer, but that they could give as many answers as they liked to as many questions as they chose and the scoring mechanism was explained to them.

Post-test

Two further multiple solutions pen and paper task were administered to the subjects within ten days of their second computer trial and again five weeks after that.

5.8 RESULTS

To examine the effects of the intervention a number of [4,3] ANOVAs were carried out on the pre-test, post-test and delayed post-test data. The design of the analyses was of the form 4 (control, four answers, eight answer, autonomous) by 3 (pre-test, post-test, delayed post-test). The first factor (practice) is between groups and the second a within group measure. The results from two subjects have been excluded from the analysis. One child was recognised as having special educational needs and the task seemed outside her abilities. A second was discounted as he scored significantly higher than all the other subjects at pre-test.

5.8.1 Multiple Solutions

The first measure examined was the number of novel correct solutions given for the three item pen and paper tests (see Figure 5.5 and Table 5.6).

Table 5.6. Number of correct novel multiple solutions by practice and time

	Control	Autonomous	Four	Eight
Pre-test	3.20 (3.23)	4.32 (3.85)	3.22 (3.11)	3.90 (3.67)
Post-test	3.10 (3.70)	8.05 (3.10)	6.67 (2.92)	11.80 (3.52)
Delayed test	3.50 (4.12)	7.42 (4.21)	7.44 (3.84)	12.00 (3.56)

There were significant main effects of time ($F(2,88)=33.03$, $p<0.001$) and practice ($F(3,44)=9.08$, $p<0.001$). The interaction between practice and time was also significant ($F(6,88)=3.75$, $p<0.002$) (Figure 5.5). Simple main effects showed significant differences at post-test ($F(3,132)=5.14$, $p<0.002$) and delayed post-test

(F(3,132)=5.44, p<0.002). All but one of the experimental groups showed significant increase in performance from pre-test to post-test and from pre-test to delayed-test:

- autonomous (q=5.23, p<0.01 & q=4.34, p<0.01)
- four (q=3.32 & q=3.37, p<0.05)
- eight (q=8.02, p<0.01 & q=8.22, p<0.01)

There was no significant change from post-test to delayed post-test in any condition. Although the differences in the means seems large, there were a limited numbers of subjects and high variances, therefore the only significant differences found were between the control & eight group at post-test (q=6.44, p<0.01) and four and eight groups (q=3.80, p<0.05). At delayed post-test the only significant difference was between the eight and control group (q=6.29, p<0.01).

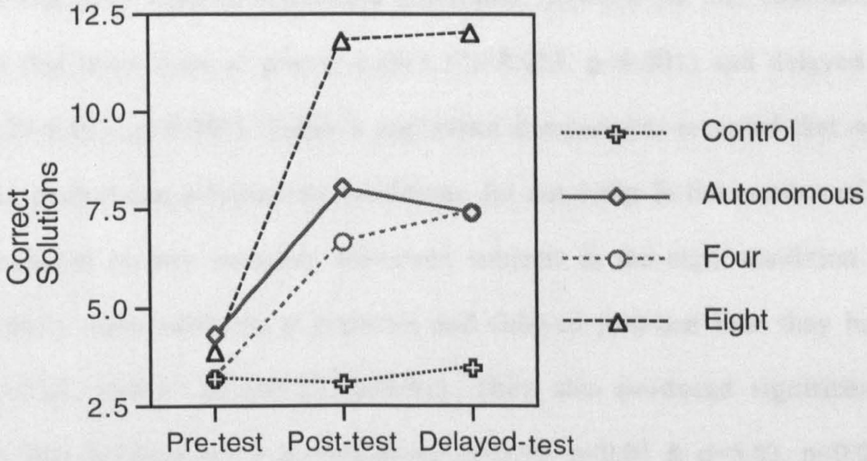


Figure 5.5 Number of correct novel multiple solutions by practice and time

Another measure of performance examined was the number of solutions subjects produced irrespective of accuracy of the solutions (see Table 5.7 and Figure 5.6).

Table 5.7. Number of novel multiple solutions by practice and time

	Control	Autonomous	Four	Eight
Pre-test	7.80 (3.49)	6.74 (3.75)	8.89 (4.14)	8.70 (3.60)
Post-test	6.10 (2.84)	8.48 (2.92)	8.56 (5.22)	13.30 (2.41)
Delayed test	6.70 (3.10)	8.79 (4.03)	8.56 (2.41)	13.3 (3.16)

There was a trend for a main effect of time ($F(1,45)=3.63$, $p<0.063$) and a significant main effect of practice ($F(3,44)=7.23$, $p<0.001$). There was a significant interaction between practice and time ($F(6,88)=3.08$ $p<0.009$). Simple main effects analysis showed that there were no significant differences between the four conditions at pre-test but that there were at post-test ($F(3,32)=8.023$, $p<0.001$) and delayed post-test ($F(3,132)=8.023$, $p<0.001$). Tukey's unplanned comparisons revealed that subjects in the four, control and autonomous conditions did not differ in the number of solutions they produced on any occasion. However, subjects in the eight condition produced significantly more solutions at post-test and delayed post-test than they had at pre-test ($q=5.23$, $p<0.01$ & $q=5.23$, $p<0.01$). They also produced significantly more answers than students in the control group ($q=5.49$, $p<0.01$ & $q=5.03$, $p<0.01$).

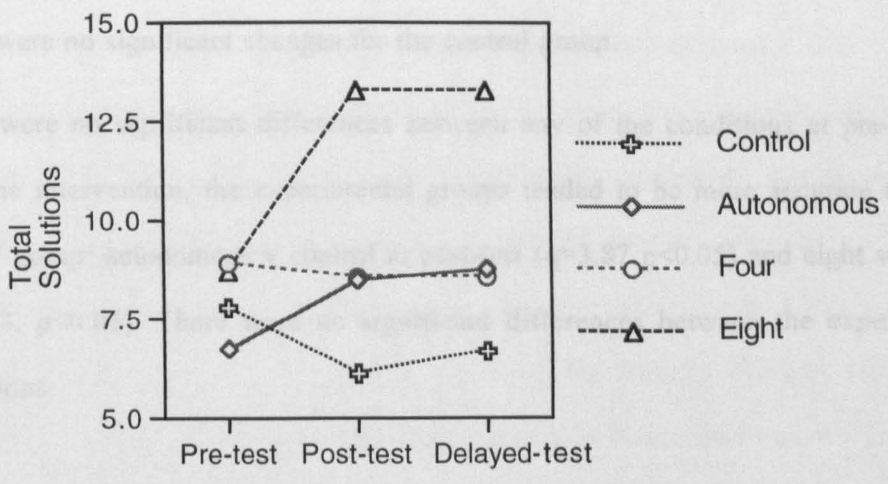


Figure 5.6 Novel multiple solutions by practice and time

The final dependent variable discussed is percentage accuracy (total correct solutions/(total correct solutions + total errors) $\times$ 100) . This was independent of the number of answers that were given (Table 5.8).

Table 5.8. Percentage accuracy by practice and time

	Control	Autonomous	Four	Eight
Pre-test	44.98 (42.34)	56.84 (40.16)	42.16 (37.14)	48.81 (35.25)
Post-test	51.30 (48.50)	93.81 (9.84)	81.57 (19.87)	87.07 (14.77)
Delayed test	45.55 (48.13)	81.41 (27.81)	84.72 (34.11)	89.78 (14.55)

Analysis identified two significant main effects; time ($F(2,88)=10.833$, $p<0.001$) and practice ($F(3,44)=6.176$, $p<0.0013$). Unplanned comparisons revealed that all experimental groups performed significantly more accurately at post-test and at delayed post-test compared to their pre-test performance:

- autonomous ($q=5.24$, $p<0.05$ & $q=3.38$, $p<0.05$)
- four ($q=3.84$, $p<0.05$ & $q=4.15$, $p<0.05$)
- eight ($q=3.93$, $p<0.05$ & $q=4.21$, $p<0.05$)

There were no significant changes for the control group.

There were no significant differences between any of the conditions at pre-test but after the intervention, the experimental groups tended to be more accurate than the control group; autonomous v control at post-test ($q=3.87$ $p<0.05$) and eight v control ($q=4.03$, $p<0.05$). There were no significant differences between the experimental conditions.

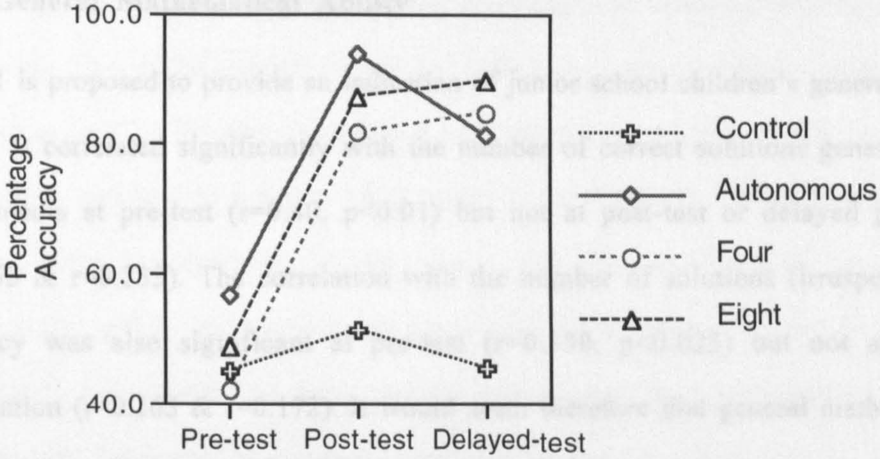


Figure 5.7 Percentage accuracy by practice and time

5.8.2 Interaction strategies

The number of answers per question that children in the autonomous condition chose to give during the intervention sessions was examined. Children in the autonomous group answered an average of 9.6 questions during their interaction with the computer and therefore produced an average of 1.66 answers per question. However, the variability was very large: some children answered 16 different questions and one gave 16 different answers to the same question. The children were very consistent in their strategy of interacting with the computer. The correlation between the number of questions answered on their first and second interaction was significant ($r=0.76$, $p<0.001$).

The number of questions the autonomous subjects chose to answer (and therefore number of answers per question) was not related to general maths ability, as measured by the Y1 ($r=-0.063$). There was also no correlation between the strategy used and the number of correct solutions at pre-test ($r=-0.078$), post-test ($r=-0.108$) or delayed post-test ($r=-0.382$). Nor was there significant correlations with total number of solutions at either pre-test or post-test. However, the strategy children used on the computer did significantly correlate with delayed post-test performance ($r=-0.391$, $p<0.05$). It would seem if there is any relation between the strategy children in the autonomous group chose to use on the computer and their performance, it is weak.

5.8.3 General Mathematical Ability

The Y1 is proposed to provide an indication of junior school children's general maths ability. It correlated significantly with the number of correct solutions generated by the subjects at pre-test ($r=0.40$, $p<0.01$) but not at post-test or delayed post-test ($r=0.253$ & $r=0.235$). The correlation with the number of solutions (irrespective of accuracy was also significant at pre-test ($r=0.330$, $p<0.025$) but not after the intervention ($r=0.165$ & $r=0.172$). It would seem therefore that general maths ability was related to children's initial ability to produce multiple solutions for coin problems but that intervention with COPPERS reduced this relation.

5.9 DISCUSSION

5.9.1. Multiple Solutions Performance

One of the motivations for this study was to examine whether the improvements in performance found in Experiment One could be replicated. Experiment Two found significant improvement from pre-test to post-test for the experimental groups. At pre-test, these subjects were producing an average of 1.27 correct answers per question and at post-test 2.95 answers per question, an improvement of 231%. This can be compared to the control group who produced 1.07 correct answers per question at pre-test and 1.03 correct answers at post-test. This difference in performance is evidence that it is the computer intervention that leads to improved performance and not simply the effects of repeated testing. Additionally, there were no significant differences from post-test to delayed post-test for experimental or control groups. Therefore, the improved performance for experimental groups seems reasonably robust.

5.9.2 Practising Multiple Solutions

It had been predicted that there might be a relationship between the number of answers per question students were required to give on the computer, and those they chose to give subsequently at post-test. However, Experiment One showed that this relationship was not as strong or as simple as had originally been predicted. The only children to benefit from practising multiple solutions were lower performing children. Experiment Two examined this further by adding an extra condition requiring eight answers, as well as the four answer and the autonomous condition. The hypothesis that getting children to produce eight answers per question would lead to better performance was supported by the study. All the experimental groups had significantly better performance at post-test but the eight answer group had significantly more correct answers than the four answer and control groups at post-test, and than the control group at delayed post-test.

The ability to give a number of correct solutions is based on the two different skills of calculation and correct multiple decompositions. To examine whether improvement in performance is due to increased competence in either or both of these skills, the total number of solutions (irrespective of accuracy) was examined. For this measure, the prediction that children in the experimental groups would produce more answers than children in the control group was not supported; only children in the eight group improved. It would appear that the significant improvement for total correct solutions observed for all the experimental groups had different causes. The four and autonomous groups improved for the most part because the accuracy of their calculations increased. However, the increase for the eight group was also due to the increase in the total number of decompositions.

The accuracy of the solutions produced by all experimental group was significantly better at both post-test and delayed post-test. There was no significant differences amongst these groups; they all approached ceiling. All the experimental groups were significantly better than the control group after the intervention.

These results suggest that if the goal of the computer use is to encourage children's skills at addition and multiplication or accuracy of calculation and decomposition then any of the experimental conditions will be sufficient. However, children will only produce *more* correct decompositions if in the eight answer condition.

It is suggested that the reason only eight answers (as opposed to four) proved effective is related to the zone of proximal development, (Vygotsky 1978). This is the region of activity in which learners can perform successfully given the aid of supporting context, in this case that of the computer. Taking this view, it is necessary to set problems on the computer that would be out of reach for children without support. However, to diagnose the dimensions of the zone of proximal development is a difficult task. Nevertheless, it should be possible to identify its lower boundary by analysing the child's unaided performance. With this information, problems could be set that are out of reach for the unsupported child and which therefore fall within their zone of proximal development.

5.9.3 Computer v Learner Control

The decision to give learners limited control over aspects of their interaction with the computer was examined in the current study. One group of children were given control over the number of answers they had to give per question while the others were given predetermined limits. The decision to give learners some control over their choice of numbers of solutions was motivated by the difficulty in deciding upon how to determine the 'right' number of solutions. If children chose to maximise the number of answers per question themselves, then the decision could be made by each user in the context of their own knowledge. However, by drawing on the research on children's beliefs about mathematics (see section 2.3), it might be predicted that they would not chose to do this.

The most immediately striking result is the small number of answers per question that autonomous children gave while on the computer; an average of 1.66 answer per question. This is perhaps surprising given that the children knew that they would receive more points if they gave more answers per question. The simple explanation that they either did not understood the points system, or else did, but were not 'falling for it' seems unlikely given their comments. They were concerned to know how well they were doing (and how well their friends were doing!). In fact many of the children seemed to be caught between the lure of the points and that of the new question making audible bargains with themselves (*e.g.* "I'll have a new one, but I'll answer three on the next one"). It would be interesting to know quite why they wanted to answer a new question so much. Some of the reasons could be:

- that they still believed that more than one answer on a question was 'cheating'
- that they thought that answering a new question would be harder
- that answering a new question would be easier
- that answering a new question is more interesting

Unprompted comments to the experimenter indicated that the children might be choosing their strategy for any of these reasons but more in-depth interviews are needed to tell for certain. It was also interesting to observe that even children who continuously chose to answer new questions rather than re-answer old ones would spontaneously re-answer a question if they had got it wrong.

The range of answers per question was very large, ranging between the maximum and minimum possible values. One child gave 16 answers on one question on both interactions with the computer, while four others answered 16 different questions. There was a high correlation between children's behaviour on their first and second time on the computer. This would suggest that children had some deliberate strategy and were not just randomly pressing for new questions. However this strategy was not related either to general maths ability or to measures taken at pre-test or after intervention. There was no significant correlation either between strategy and general maths ability, total number of correct solutions or total solutions at pre-test.

There was also almost no relation between how many questions the subjects answered on the computer and post-tests measures. The hypothesis that children in the autonomous condition who had given more answers per question on the computer would have better learning outcomes was only supported at delayed post-test for total number of solutions. However, the majority of children gave very few answers per question; 79% gave on average less than two answers per questions. Only two of the children consistently gave four or more answers per questions (the fewest number of solutions that the computer demanded in this experiment). The one subject who gave 16 answers to one question, however, showed the greatest improvement of all the subjects in the autonomous group. Given the similarity of behaviour on the computer, it does not seem surprising that there were no differences between the users' learning outcomes.

The results for giving learner's control over the choice of number of solutions suggests that this is not an effective approach to teaching children to give multiple solutions. These results are consistent with Steinberg's (1989) view that learners will benefit from more control only if they are capable of selecting an appropriate learning strategy. In this case, they did not do so. This is not surprising given the research already reviewed on how children's mathematical beliefs lead them to expect a single correct answer for mathematical problems (Baroody, 1987). If learners can not abandon this belief, they are highly unlikely to chose an effective strategy.

5.9.4 General Mathematics Ability

Again no simple pattern was found of relations between general mathematical aptitude and the production of multiple solutions. It was not the case that children who scored better on the general mathematics tests were better at this task. There were significant correlations at pre-test with accuracy and correct solutions but none at post-test. Nor did ability predict interaction strategy with the computer. There was no relation between mathematical ability and strategy in the autonomous group.

It is tempting to speculate that one of the reasons for these results is that children with better mathematics skills have already stabilised their beliefs about the nature of mathematics (*e.g.* one correct answer per question as quickly as possible). These children have been exposed to a primary mathematics curriculum where the goal is to answer many problems in a short time using a single solution (see Fuson, 1992). Giving multiple solutions to a single problem will therefore be an uncommon experience and it easy to see why these children would not want to abandon a strategy which has previously led to success. Thus, the mathematical skills and knowledge that they could use on these problems are under-utilised.

5.10 CONCLUSION

This results of this Experiment Two confirmed and extended the finding that COPPERS could successfully teach children to give multiple solutions to coin problems. Relative to a non-intervention control, experimental subjects improved significantly and this enhanced performance remained stable to delayed post-test. It was argued that if the goal of using a system such as COPPERS is to support the accuracy of mathematical calculation, then any number of answers of question may be sufficient. However, in order to develop knowledge of multiple solutions then it is necessary to set this number to be beyond that given without the aid of supporting context. Furthermore, learners in this domain are unlikely to choose this strategy unless the computer requires them to do so.

5.11 GENERAL CONCLUSION

These experiments have evaluated a computer-based learning environment that support children's skill and understanding for producing multiple solutions for a single problem. Based on research that described the relation between children's number sense and mathematical beliefs, it was predicted that children would find producing multiple solutions difficult. In line with this prediction, children were found to produce a very low number of solutions, but, with a limited amount of teaching, they show impressive and sustained improvement. Two aspects of the computer system were found to be strongly positively related to learning outcomes. The first required children to produce (with support) more solutions than they would naturally give. The second was the benefit of providing a extra, tabular representation of users' answers.

In order to further explore how to support children's understanding of alternative ways to answer mathematical problems and the role of (multiple) external representations in supporting such learning, a new system was designed and implemented. It was created to address a further aspect of the 'right answer' misconception by focusing on inexact answers and procedures - estimates. The proposed users of the system were older children allowing a richer repertoire of representations to be used. The design of the system is discussed in the next chapter.

CHAPTER SIX

CENTS: A Computer-Based Environment for Computational Estimation

Computational estimation is seen as dependent upon good number sense and it is also proposed that developing children's estimations skills will lead to better number sense (Sowder, 1992a) (reviewed in section 2.3). In this chapter, a detailed description of how research in this area informed the design of CENTS is given. The instructional goals of this system are considered. The execution of these goals is discussed in terms of the general instructional approach and specific support. In particular, a detailed description of the representations available in CENTS is provided. Discussion will primarily avoid the implementational level. Section 6.2 provides a brief description of these issues.

6.1 INSTRUCTIONAL GOALS

To successfully perform computational estimation a wide range of mathematical knowledge is required. LeFevre *et al.* (1993) proposed that three types of knowledge are necessary for computational estimation - conceptual, procedural and factual. The factual knowledge required for mental calculation, for example, would include knowledge of place-value and memorised number facts. Conceptual knowledge is needed to choose an estimation strategy that will produce approximate numbers to facilitate computation. Procedural knowledge is required to perform the approximation. CENTS was primarily designed to support the development of the procedural and conceptual aspects of estimation. No attempt is made to teach aspects of mental calculation such as place value. This is considered to be essential pre-requisite knowledge.

The educational goals of the system are to

- teach children strategies that they can use to estimate problems
- encourage children's understanding of how transforming numbers to produce an intermediate solution affects subsequent accuracy

- support the development of the required underlying conceptual knowledge
- encourage users to consider estimation in terms of LeFevre *et al.*'s conceptual principles of simplicity and proximity.

The following sections describe how each of these issues is addressed in the system.

6.2 IMPLEMENTATION OF CENTS

CENTS was created for the Apple Macintosh computer using SuperCard™ 2.0, and written primarily in Supertalk. The system (presently) contains 15 windows, three dedicated menus, 30 text fields, and 40 graphics. Numerous buttons, backgrounds, icons, cursors and sound resources were created. CENTS is run by around 4,500 lines of code.

Teachers and children were involved in all stages of the design and implementation of CENTS. Many of their suggestions have been incorporated, although not all were considered appropriate (particularly those that included monsters, street fighters, *etc.*!). Changes to the system after the initial development phase were commonly based upon children's comments. In total, programming and development took around nine (part-time) months.

6.3 INSTRUCTIONAL APPROACH

6.3.1 General Instructional Method

The general pedagogical approach taken by CENTS (figure 6.1) is to encourage the children to consider estimation in a flexible and thoughtful way. To this end, the metaphor of an experiment is used. Users make predictions about a particular estimate, perform the estimation, and then have the opportunity to examine the results of the estimation process in the light of their predictions. After each problem, children log the results of (at least) two different estimation strategies in an on-line work book. They describe how they transformed the numbers, how accurate each estimate was, and how difficult they found each estimate. At the end of a session, children are encouraged to review the log book to investigate patterns in their estimates. For example, that truncation will always give you an estimate that is lower than the exact answer, rounding using intermediate compensation will generally be the most accurate strategy, *etc.*

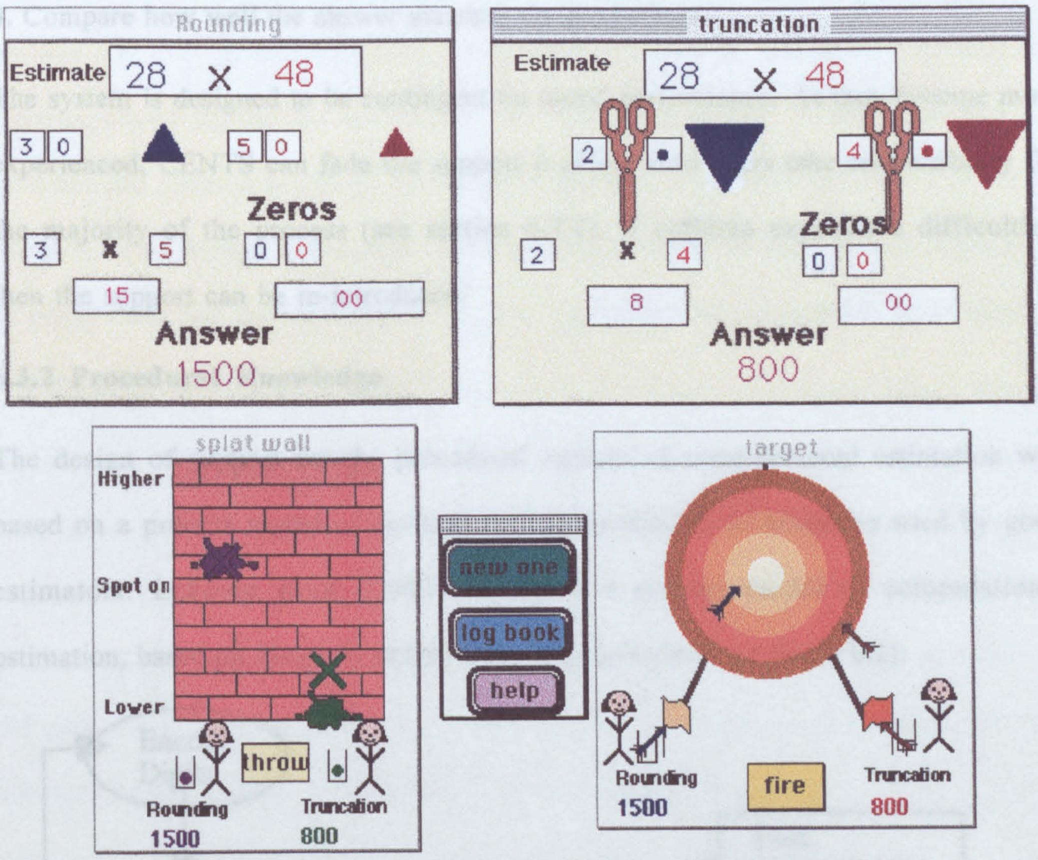


Figure 6.1 An illustration of a completed problem with CENTS

This produces the predict-test-explain cycle that has been found to promote understanding in science education (*e.g.* Howe, Rodgers & Tolmie, 1990). In CENTS, the prediction and analysis stages are supported by multiple representations of the underlying conceptual principles. Hence, a session involves the following stages:

Given the problem - estimate 387 × 123

1. Produce the intermediate solution.

round to 400 × 100

2. Predict the accuracy of your estimate based on the intermediate solution

not very close to the exact answer

lower than the exact answer

3. Produce the estimate

40000

4. Compare how well the answer matched the predictions.

The system is designed to be contingent on users' performance. As they become more experienced, CENTS can fade the support it offers until users take responsibility for the majority of the process (see section 6.3.2). If children experience difficulties, then the support can be re-introduced.

6.3.2 Procedural Knowledge

The design of support for the procedural aspects of computational estimation was based on a process model of estimation and descriptions of strategies used by good estimators. LeFevre *et al.* (1993) described a process model of computational estimation, based on Siegler's model of strategy selection (see figure 6.2).

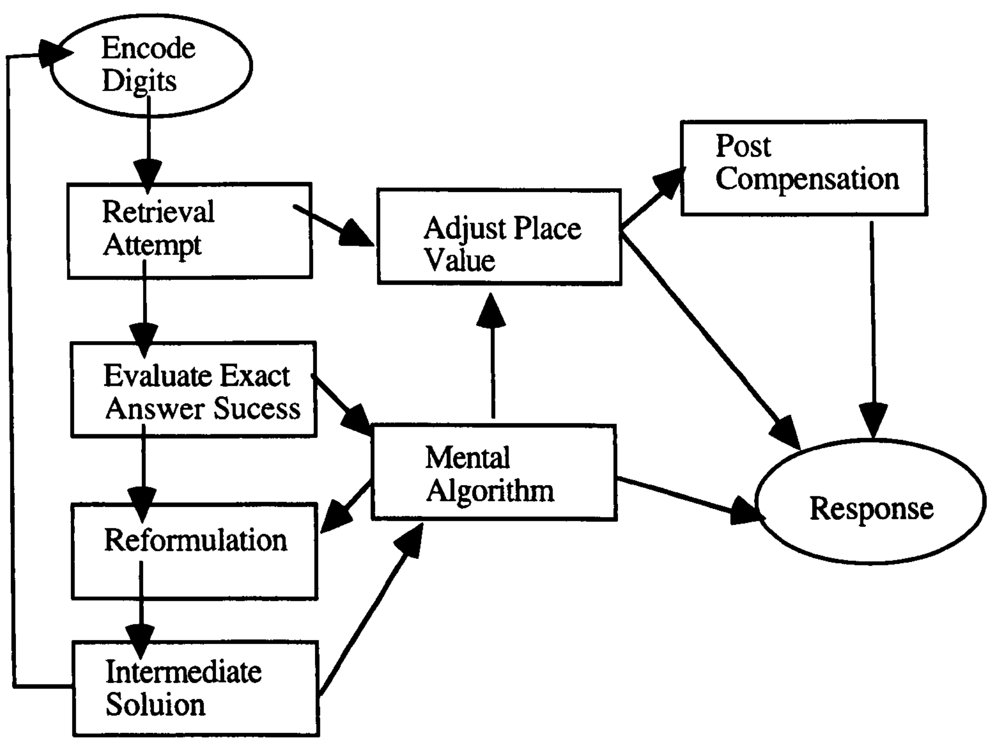


Figure 6.2 LeFevre *et al.*'s (1993) process model of estimation (p 123)

The first step in the model is an attempt at retrieval. A familiar pair of numbers may automatically activate a potential solution. The next stage would be to consider whether one should calculate an exact answer solution. This will involve knowledge both of the complexity of the sum and knowledge about your own mathematical capabilities. Note that this process does not insist on formal algorithms. If an exact answer is not produced, then reformulation (which includes intermediate compensation) should be attempted. The next step in the model is to use the intermediate solution produced by the reformulation to generate an answer either by retrieval or by mental calculation. The final stages in the process include adjustment of place-value and post-compensation which should occur if the earlier reformulation was tagged as needing some adjustment.

Reys *et al.* (1982) identified three types of estimation strategy: reformulation, translation and compensation. Reformulation involves altering numerical data to produce a more mentally manageable form without altering the structure of the problem. Translation refers to the action of changing the mathematical structure of the problem to a more mentally manageable form. Finally, compensation involves adjustments to numbers either during or after computation.

CENTS teaches children strategies for performing estimation of multiplication sums. A number of different strategies were prototyped and two fully implemented. These two strategies are rounding and truncation, both examples of reformulation strategies. No attempt was made to support translation strategies. These are highly idiosyncratic and require reformulation as a subprocess. The final kind of strategy, compensation, is supported by encouraging children to consider intermediate compensation when rounding. Post compensation is predicated upon informed insight into the proximity of an estimate. Thus, CENTS attempts to support the initial development of this skill by encouraging reflection upon the accuracy of an estimate.

<u>Rounding</u>

Rounding was the most common strategy found in Reys *et al.*'s study. Interviews with teachers during the design and implementation phase of CENTS suggested that it is also the one most likely to be taught in British primary schools. An example of rounding is given below.

Rounding: the factor is transformed to the nearest multiple of 5, 10, 100, *etc.*

e.g. estimate 323 × 48

323 is closest to 300, 48 is closest to 50

so 300 × 50 = 15000

In order to consider how to support the acquisition of rounding, LeFevre *et al.*'s (1993) process model was used to describe the steps necessary for successful rounding.

For example: estimate '323 × 48' by rounding.

(1) - try to retrieve answer NO

(2) - round larger number to 300 (tag 00)

(3) - try to retrieve 3 × 48 NO

(4) - round smaller number to 50 (tag 0)

(5) - try to retrieve 3 × 5 YES = 15

(6) - add three tagged zeros = 15,000

(7) - respond

However, neither LeFevre's process model or Reys *et al.* descriptions of successful estimators describe how the numbers are rounded to create the intermediate solution. Hence, a small informal study based on observation and interviews with successful estimators was conducted in order to examine this process. This suggested for rounding that the following steps must be conducted (note, steps a and b are not necessarily performed in this order).

For each number to be rounded to create an intermediate solution:

(a) decide whether to round to the nearest 5, 10, or 100 *etc.*

(b) decide whether to round up or down

(c) round number

(d) tag in working memory direction and magnitude of the rounding.

Thus, these two accounts were combined to produce a description of the stages involved in rounding. These form the basis of the support that CENTS provides (see figure 6.3). The system provides a lot of structure when children are inexperienced with estimation. When users improve, the support fades leaving children with more decisions and responsibility. The following system description will concentrate on the entry level support.

Stage 1. Rounding the number. Sowder & Wheeler (1989) found that children often do not round to a 'right' number, for example, rounding 461 to 300, or rounding 461 to 460 without the ability to multiply by 46. In problems such as '63 × 42', children have been known to round the first number to 100 and the second to 0. This stage is supported in CENTS by a 'slot and fill' approach. The required number of zeros are already in place, children must enter the front end digits. They are given a hint about which direction might be the best to round (an arrow pointing in the suggested direction). Initially, this is to the closest answer without considering the principle of intermediate compensation. It is possible to ignore this hint and round to a different number. However, if children choose a solution that is further than away than any of the computer's preferred solutions, (*e.g.* rounding 448 to 300 rather than 400 or 500), then this is corrected. As users' experience grows, more choice is made available (*e.g.* they could choose to round 448 to 450). To help children keep track on the numbers in the problem, the first factor is represented with red text and graphics and the second in blue.

Stage 2. Noting the direction and magnitude of the transformation. Case & Sowder (1990) proposed that primary school children would not be able to keep track of 'how far off' their estimate was due to overloaded working memory. The system supports memory load by providing a simple representation of the direction and magnitude of the transformation in the form of a proportionally sized triangle pointing either up or down. It should be stressed that this in based on relative not absolute transformation. Hence, transforming 18 to 20 would result in a much larger arrow than 88 to 90.

Stage 3. Front end extraction After transforming the numbers, the next stage is to extract the digits to be multiplied. Initially, CENTS performs this stage for the student.

Stage 4. Place value tagging When the digits are extracted, the 'zeros' are collected together and stored for subsequent place value correction. As children become more experienced, they take responsibility for tagging the number of zeros they will need. The original factor from which the 'zeros' come is indicated by their colour, either red or blue.

Stage 5. Multiplying the extracted digits. Users must enter the product of the extracted digits. In the event of a wrong answer, CENTS either displays a number square or, for solutions with factors that are greater than 12, suggests that users begin their answer again. This is to ensure that children don't fail at the task because they can't recall their 'timetables', but also serves to discourage them from rounding to numbers that they subsequently cannot multiply. The text colour for stages five and six now changes to purple to indicate that the two factors have been combined.

Stage 6. Final Answer. The product of the front end extraction is combined with the tagged zeros in order to correct for place value and the estimate is displayed.

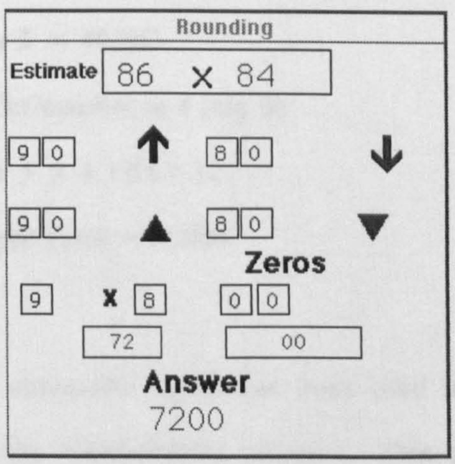


Figure 6.3 A solution created by rounding

Truncation

Truncation normally produces a less accurate estimate than rounding. Nonetheless, it is an important approach to teach children as it is a particularly easy strategy. Sowder & Wheeler found younger children showed a preference for rounding, but did not have adequate skills to carry it out. An example of truncation is:

Truncation where the right-most digits are ignored

e.g. estimate 323 × 48

323 is changed 300, 48 is changed to 40

so 300 × 40 = 12000

Truncation simply requires extraction of the front digits. Consequently, children do not need to work out which is the closest 'nice number'. In addition, load on working memory should be reduced as is it not necessary to remember the direction of the transformation.

For example: estimate 323 × 48 by truncation.

(1) - try to retrieve answer NO

(2) - truncate larger number to 3 (tag 00)

(3) - try to retrieve 3 × 48 NO

(4) - truncate smaller number to 4 (tag 0)

(5) - try to retrieve 3 × 4 YES = 12

(6) - add three tagged zeros = 12,000

(7) - respond

Again, interviews with successful estimators were used in order to examine this process of producing the intermediate solution. This provided the following description for truncation.

To truncate numbers to produce the intermediate solution

(a) decide whether to truncate to 1 digits, 2 digits, *etc.*

(b) truncate number

(c) tag magnitude of truncation

Again, these two levels of description were combined to produce the structure and support provided by CENTS. The support for truncation is necessarily very similar to that provided for rounding (see figure 6.4). The only difference between the strategies that is predicted by the process model occurs when producing the intermediate solution.

Stage 1. Truncating the number. Although superficially the actions of rounding down and truncation appear similar, they result from different processes. In order to emphasise this, CENTS employs a different metaphor for truncation to that of rounding. For truncation, children are encouraged to consider how to remove the needed digits from the unwanted ones. Users must place scissors in the correct position in order to 'chop off' the non-essential digits. Again, as users experience grows, they are given more choice over how many digits to extract.

Stage 2. Noting magnitude of the transformation. This is represented using the triangles to record the change. In the case of truncation the transformation is always down. Again, it is the proportional change on the number that is represented. This can result in very striking differences. For example, truncating 17 to 10 is a far greater proportional change that 87 to 80.

Stages 3 to 6 are identical to those described for rounding.

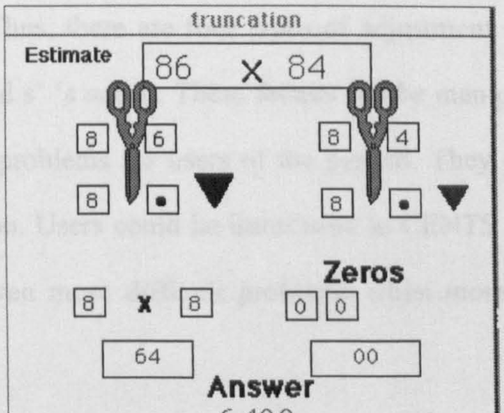


Figure 6.4 A solution created by truncation

Compensation

The use of intermediate compensation is indirectly supported by CENTS. The problem generation routine is designed to provide some problems which would most accurately be solved by intermediate compensation. It was also hoped that when children review the logbook for patterns in their estimates, that they will see that intermediate compensation normally results in a very accurate answers. As discussed above, post-compensation requires children to have sense of how their estimate differs from the exact answer. The development of this skill receives a great deal of attention in CENTS and is discussed in section 6.3.4.

6.3.3 Problem Generation

The estimation problems given by CENTS are dynamically generated. A number of factors can be manipulated to determine problem complexity. The first is the number of digits in the problem. CENTS can set two (digit) by two problems, two by three, three by two problems and three by three problems. A wider range of problems could easily be supported but would require some reprogramming of the interface. A second dimension of problem complexity is the size of adjustment to be made to each factor. LeFevre *et al.* used two (absolute) definitions, small and large. A small adjustment on a 2 digit problems would be 2 or less, *i.e.* 12 to 10 or 18 to 20. A large adjustment involves changing the number by 4 of more. These proportions are scaled up for the three digit problems. Thus, there are four types of adjustment available, 's(mall) and s', 'l(arge) and l', 'l and s' 's and l'. These factors can be manipulated in order to alter the complexity of the problems for users of the system. They can be programmed to alter through the session. Users could be introduced to CENTS with '2 by 2' 's and s' problems and then given more difficult problems when more experienced with the system.

An alternative to the adjustment size aspect of problem generation is also provided. It is designed to provide questions which, when rounding, would be best solved by either rounding up, rounding down or intermediate compensation. This was included to ensure that children could be given the opportunity to explore the whole space of possibilities for rounding. This might be particularly desirable if the aim was to expose children to differences between the supported strategies.

6.3.4 Insight into Accuracy

In order to become flexible and accurate estimators, children need to develop understanding of how their estimate relates to an exact answer (Trafton, 1986). This means that they need to consider how transforming numbers to create an intermediate solution affects the accuracy of the final outcomes. This is necessary if post-compensation is to be used. The sense of 'reasonableness' of answer is a fundamental component of the appropriateness of an estimate. Under-estimating the price of goods in a shopping trolley could prove much more embarrassing at the checkout than over-estimating!

A review of the literature revealed little research in this area. Hence, an analysis of this knowledge was undertaken. What follows is therefore speculative and will require further research to be undertaken.

There are two aspects of insight into accuracy. The first component is the direction of the estimate - is it under or over the exact answer? The second is the magnitude of the difference between the estimate and the exact answer. For example, is your estimate close, far away, within 10%, 30%, *etc.* of the right answer. Further complications are introduced by the requirement to reason about relative rather than absolute transformations. Children need to understand that transforming 25 to 20 (a change of 20%) is much greater than changing 95 to 90 (5.2%).

It seems plausible that there are a number of different levels for understanding accuracy. Some examples for deciding whether a solution is an over or underestimate are given below, but this is not meant to be interpreted as a stage model of insight into accuracy:

- If both factors have been transformed down then the estimate must be lower
- If both factors have been transformed up then the estimate must be higher
- If one factor has been transformed down and the other up then consider the absolute differences in the transformation to decide if the estimate is higher or lower than the exact answer
- If one factor has been transformed down and the other up then consider the relative differences in the transformation to decide if the estimate is higher or lower than the exact answer

Accuracy of estimates therefore receives considerable attention in CENTS. It forms the basis of predict-test-explain cycle discussed above. Children must predict the accuracy of an estimate and then test and compare this with the actual accuracy and other estimates. This is performed upon multiple representations of accuracy (see section 6.5). In addition, support for interpreting and remembering transformation of numbers is provided by the proportionally sized arrows which represent this process (described above).

6.3.5 Conceptual Knowledge

The system attempts to support the development of the three areas of conceptual knowledge outlined by Sowder & Wheeler (1989) (this was first discussed in section 2.5).

The Role of Approximate Numbers

The knowledge that the process of estimation necessitates approximate numbers is supported by requiring children to perform this approximation. It is not possible to continue with the estimation until numbers have been transformed into an intermediate solution. As a result, the answer to the problem will also be approximate. Additionally, the exact answer is never seen in the system. The representations used to display the accuracy of estimates are designed such that there is no 'right' answer to compare against. For example, the archery target (discussed in more detail below) deliberately has a central area representing '0-10%' away rather than a 'bullseye'.

The Role of Multiple Processes/Outcomes

Students have difficulty accepting that you can use a variety of procedures and that different answers can be correct (Baroody, 1987; Sowder & Wheeler, 1989; Chapter Five). This knowledge is supported in the system by requiring children to answer problems using (at least) two different strategies. This often (but not necessarily) results in two different estimates. Requiring multiple solutions allows for subsequent discussion of why different processes are acceptable and a comparison of the accuracy of different answers to the same problem.

The Role of Appropriateness

This involves recognition that the appropriateness of an estimate depends upon the context or desired accuracy. Of the three principles, this is probably the least well addressed in CENTS. The system was designed to set problems simply as computational sums without context. Research on the benefits of context is mixed and there is evidence that different contexts affect children's strategies (Forrester, Latham & Shire, 1990). It was decided that too little is known about the affects of context to implement this.

As the appropriateness of an estimate is dependant primarily on its context, this obviously limits the system. However, children were encouraged to consider their estimate in terms of its accuracy and simplicity. This may help to provide children with a language with which to consider appropriateness, something that Sowder and Wheeler found to be lacking in primary aged children (Sowder, personal communication).

Conceptual Principles

LeFevre *et al.* (1993) proposed that the conceptual knowledge necessary to perform estimation could be summarised by the two conceptual principles of proximity and simplicity. CENTS places considerable emphasis on these principles as it is predicted that by doing so children will come to understand more about the conceptual knowledge underpinning estimation.

Proximity receives the most attention in CENTS as it is the one most implicated in insight into accuracy. LeFevre *et al.* found that children rarely mention this principle - adults on the other hand seem guided by it. Proximity is supported in the system by the use of representations which describe children's estimates in terms of percentage deviation from the exact answer. Children predict what they think the accuracy of their estimate will be. Subsequently, they are shown the actual accuracy of their estimate and can compare their predictions to the results. This is designed to encourage children to think about estimation as entailing proximity, and should help them start to develop the skills needed to decide how to best reformulate numbers in order to achieve an accurate estimate (see above).

Simplicity is the knowledge of the best way to modify a problem to produce a solvable intermediate solution. Conceptual understanding of this principle was found even in the youngest children in the LeFevre *et al.* study, yet to operationalise this principle is far from easy. Firstly, simplicity depends upon a child's knowledge and their judgement about such knowledge. A reasonable heuristic for reformulation strategies is to assume that the more numbers left after rounding the more complicated the problem. However, knowledge of simplicity must be strongly situated. If you know that 7 packets of your favourite sweets which cost 35p can be bought for £2.45, then you can directly retrieve this solution ($7 \times 35 = 245$) and hence making it is very simple solution. In addition, the different strategies themselves are more or less difficult. Some strategies involve less steps and less demand on working memory than others. As discussed above, truncation requires less decisions and working memory demands than rounding. Finally, some strategies are more familiar to the children than others. For example, rounding is the most commonly taught strategy in the United Kingdom. This will serve to make rounding comparatively easier for children until they have had considerable experience with other strategies.

Given these difficulties of defining and operationalising simplicity, it was decided to let the children judge for themselves the relative difficulty of producing each estimate. CENTS currently gets children to describe how difficult they found each estimate when completing the log book. This obviously results in a less than balanced approach to the two principles, since more emphasis is placed on accuracy. However, for the age group of the proposed users of the system, LeFevre *et al.* found that it is the proximity principle that is undeveloped, while simplicity is already understood.

6.4 ANALYSIS OF REPRESENTATIONS

Many advantages have been proposed for learning with MERs (discussed in section 3.3). However, the learning demands associated with MERs were reviewed at length (section 3.4) to show that using MERs is not unproblematic. CENTS has been designed to address how different combinations of representations affect the process and outcomes of learning. As described above (section 6.3.3), their role in CENTS is to express the accuracy of estimation in relation to an (hidden) exact answer. As CENTS is designed to investigate issues in the use of MERs, a number of different representations of proximity are available to an experimenter or teacher.

All representations are based on the percentage deviation of the estimate from the exact answer ((exact answer - estimate / exact answer) $\times$ 100). This is a commonly used measure to analyse the accuracy of an estimate (Levine, 1982; Dowker, 1992). No matter how the nature of surface features of the representations differ, the deep structure is always based on this relationship. They are used both for display (how accurate the estimate was) and also for action (children's prediction of how accurate the estimate will be, given the intermediate solution).

A number of features of the representations can be manipulated. In section 3.2, various approaches to describing the different properties of representations were introduced. These included taxonomic approaches and attempts to define fundamental properties (equivalence is obviously not relevant until comparing two or more representations). The following sections will review the nature of the representations based on these approaches. The first dimensions considered are based on two of Palmer's (1978) classification for describing information presented in a representation - the type of information ('amount of information' will be used in the thesis for clarity) and the resolution of the information. Secondly, the modality of the representation is described as this remains the most common way of classifying representations. Finally, Kaput's distinction between mathematical and non-mathematical representations is used to categorise the representations used in two broad classes.

In addition, these features can differ across the representations that are presented together. This allows properties of multi-representational systems to be varied along the dimensions of redundancy between representations and the similarity of format. Each of these dimensions will be discussed in turn.

Amount of Information

It was proposed (section 6.3.3) that there are two different dimensions to accuracy of estimates - direction and magnitude. Hence, the representations used in CENTS can be chosen either to display direction or magnitude separately or can display both dimensions simultaneously. This is particularly interesting when considering multi-representation systems because it allows for different levels of (informational) redundancy across representations. In the case of CENTS, three levels of redundancy are possible - no redundancy, partial redundancy and full redundancy.

In no redundancy situations, each representation expresses a different dimension of accuracy. Thus, one representation is used to display direction (either higher or lower than the exact answer) and one to express magnitude (either continuously or categorically). When MERs are fully redundant, then the same information is derivable from both of them. For example, both representations could express direction and (continuous) magnitude. Finally, MERs could be partially redundant. In this case, there is some overlap between the information derivable in the representations. For example, one representation could express magnitude only, while the other both magnitude and direction. This flexibility allows predictions to be tested about how redundancy between representations affects learning.

Resolution

Two types of resolution of proximity information were created for use in CENTS. The first is to present the accuracy of estimates in a categorical system. Thus, two estimates, one 12% away from the right answer, the other 16% away might be considered to fall into the same category. Descriptions such as 'close', '10-20%', '2nd band of the target' might all be labels for this category. In the case of the representations used in CENTS, the categories depicted represented 10% deviation bands. Evidently, this choice is, to some extent, arbitrary.

The second granularity of information chosen for use in CENTS was a continuous one. For the given example, it would be possible to discriminate between the two different estimates of 12% and 16% away, (*e.g.* higher up the splatwall). It should be noted that the exact percentage deviation was not expressed. For example, if the estimate was 11.85% away from the answer, the system displays 12% away.

These two different resolutions were to chosen to express different views on accuracy. It is often convenient to think of estimates in a categorical manner. There can be few occasions when even sophisticated estimators would need to discriminate between 12% and 16% inaccurate. However, children tend to view categories as having hard boundaries and to consider themselves as wrong if they predict a category adjacent to the 'right' one.

CENTS can therefore offer two different views on the proximity of representations. Hence, when providing MERs these can either be at the same level of granularity or at different ones.

Modality of Representations

Perhaps the most common distinction between representations is whether they are graphical or propositional (*e.g.* Barwise & Etchemendy, 1992). Many researchers have found that in different situations representation of one modality may be more effective than another, (*e.g.* Larkin & Simon, 1987; see section 3.1). Although, this was not a primary aspect of this thesis, the representations used were constructed to differ in modality (about 70% were graphical and 30% propositional).

Type of Representations.

There are a large number of taxonomies of representations (for a review see section 3.2). Each field concerned with the role of external representation seems to have created at least one (but normally many) of their own. One very useful one for the design of CENTS (although not sufficient to describe all the features of the representations) was the distinction proposed by Kaput (1987) between ambient symbol systems such as pictures and natural language and other, normally school taught, representations such as graphs, tables, schematic diagrams (referred to as mathematical representations). The intended age group of CENTS (late primary school children) are considerably more familiar with pictures than they are with mathematical representations. Given the research into the role of expertise in understanding external representations (*e.g.* Petre & Green, 1993), this may be a crucial dimension.

6.4.1 Representation Descriptions

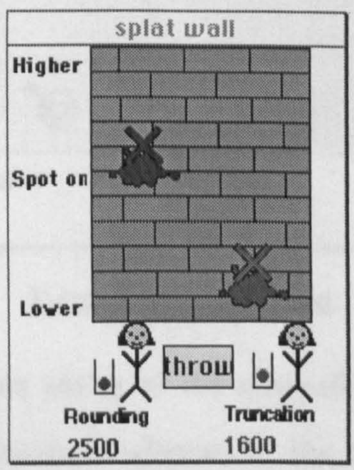
If each single representation differed along the dimensions of amount of information, resolution, modality and type, then 24 different representations of proximity would be needed. However, some of these cells may be empty. It may not be possible to have a representation that is pictorial but not graphical. Not all 24 proposed possibilities were created, altogether a total of eight different representations have been used in CENTS to date (Table 6.1). The choice of representation implemented was primarily dictated by the experimental questions that the evaluations with CENTS addressed (see Chapters Seven and Eight). Each representation will be briefly described in turn.

Table 6.1. Representations currently available in CENTS

	Type	Modality	Resolution	Amount
Splatwall	pictures	graphical	continuous	D & M
Archery Field	pictures	graphical	continuous	D & M
Target	pictures	graphical	categorical	M
Hoops	pictures	graphical	categorical	D
Marbles	pictures	graphical	continuous	M
Numerical D & M	maths	propositional	continuous	D & M
Histogram	maths	graphical	categorical	M
Numerical M	maths	propositional	continuous	M

key: D = direction M= magnitude

<u>Splatwall</u>



<u>Figure 6.5 Splatwall</u>

The underlying analogy for all the pictorial representations that express magnitude information is based on distance from a goal. The 'splatwall' operates by throwing 'paintballs' at a wall to indicate percentage deviation of the estimate in continuous terms. It expresses both magnitude and direction. Thus, in terms of Cheng's functional roles of diagrams, (Cheng, 1996a; see section 3.2) it depicts both states and values and allows for comparisons to be made. The states in Cheng's terms for this representation are either under or over-estimates. The values as these are continuous representations is the deviation away from the right answer.

This representation was created to emphasise that being in the middle is best and is therefore unlike most common representations of distance which are biased so higher is better. Children act upon the representation by marking on the wall how close they believe their final estimate will be to the exact answer. Hence, this representation supports a direct manipulation interface. The accuracy of the estimate is indicated by throwing a paint ball at the wall to leave a 'splat'. A compromise was made between space and precision, so that the wall represents accurately deviations of ±50%. Deviations of above this amount (and there are very few) are represented by placing the 'splat' on the limit.

Archery field

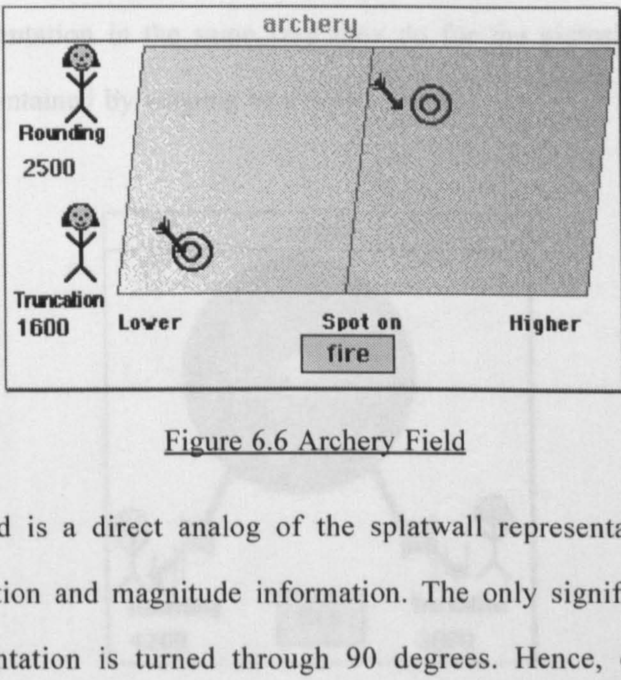


Figure 6.6 Archery Field

The archery field is a direct analog of the splatwall representation. It expresses continuous direction and magnitude information. The only significant difference is that the representation is turned through 90 degrees. Hence, estimates that are 'higher than the exact answer' must be mapped onto to the right of the centre and 'lower than the exact answer' to the left of the centre.

Numerical display - direction and magnitude

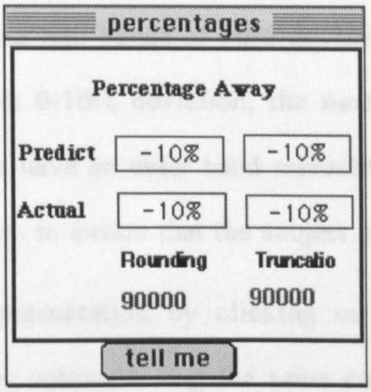


Figure 6.10 Numerical Display

Both prediction and display with the numerical representation express percentage deviation in digits. The direction of the deviation is given by '+' and '-' signs. The function roles assigned to this representation are identical to those proposed for the splatwall and archery field. Children act upon the representation using the keyboard. Accuracy is given to the closest 1%. Although space and precision constraints do not affect this representation in the same way they do for the pictorial representations, consistency is maintained by keeping to a ±50% range.

Archery target

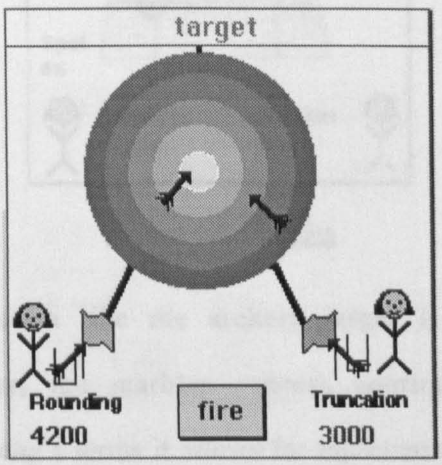


Figure 6.7 Archery Target

The design for the archery target representation of accuracy is based on real archery target. Hence, it represents magnitude information in a ordered categorical fashion. Cheng's functional roles analysis would suggest that the archery target depicts states as it expresses ordered categorical data. Accuracy can be read off in terms of distance from the centre. Comparable to a real archery target, the high/low dimension is meaningless. It is categorised into bands of 10% deviation represented by different tones. The centre represents 0-10% deviation, the next 10-20%, *etc.* and the last above 40%. The decision to have an outer band represent all values greater than 40 rather than 40-50% was taken to ensure that the subject never 'missed' the target.

Children act upon the representation by clicking on a circle to indicate their prediction of accuracy. This colours a flag the same colour as the band they have selected, so that users can compare their prediction to the answer. The accuracy of the estimate is indicated by firing an arrow at the target. Note, that there is a central wide area rather than a 'bulls eye' on the target as these representations are designed to de-emphasise the 'right' answer.

Marbles

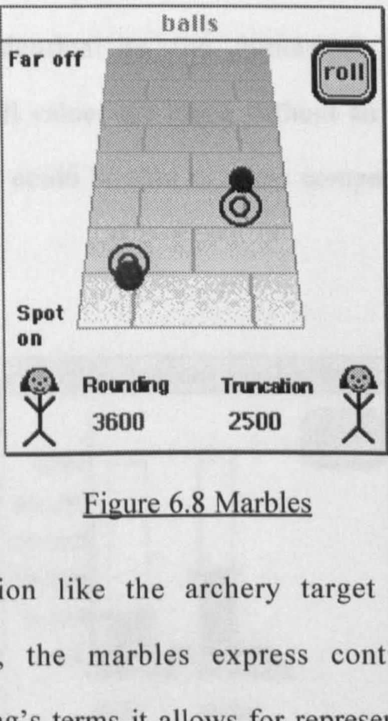


Figure 6.8 Marbles

The marbles representation like the archery target is a pictorial magnitude representation. However, the marbles express continuous rather categorical information. Thus, in Cheng's terms it allows for representing and comparing values. The metaphor used is that of rolling balls along a road. The further a marble rolls, the greater the magnitude of the deviation. This representation was particularly complicated to design as it was difficult to devise a magnitude representation which was not contaminated by direction interpretations.

Children act upon the representation by clicking on the road to indicate their prediction of accuracy. Hence, like all the pictorial representations, it is accessed via a direct manipulation interface. A ball is rolled along the road to express the accuracy. The representation is sensitive to deviations of up to 50%.

Numerical display - magnitude

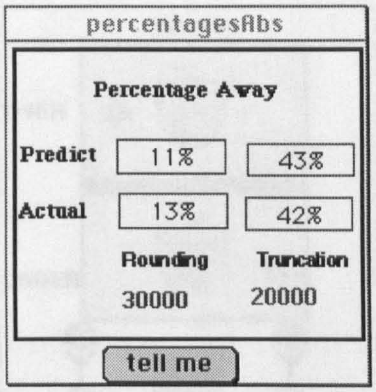


Figure 6.11 Numerical Display - magnitude only

This representation is identical to the numerical direction and magnitude representation except that all values are given without an indication of direction, *i.e.*, without a '+' or '-' sign. It could be said to allow comparison of the different values of magnitude information.

Histogram

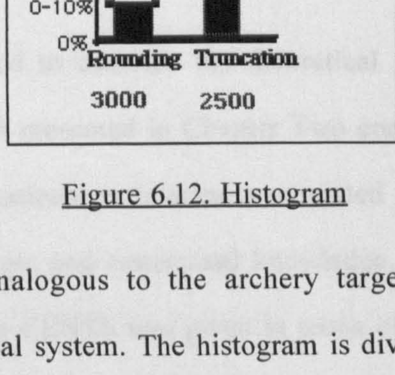


Figure 6.12. Histogram

This representation is analogous to the archery target. It expresses magnitude information in a categorical system. The histogram is divided in bands of 10% from 0% up to 40+% (figure 6.12). The numerical values are available from the representation. The functional roles provided to this representation are the same as those assigned to archery target. To predict how accurate their estimate will be, children simply click on the graph to mark it. Accuracy is represented by shading the histogram. This representation therefore differs from the other mathematical representations as it exploits perceptual processes to a greater extent.

Hoops

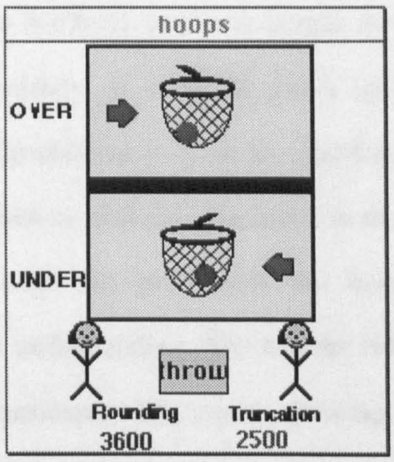


Figure 6.9 Hoops

The hoops representation uses a similar metaphor to the magnitude representations by providing distance feedback by throwing balls into hoops. These simply represent whether the estimate is higher or lower than the exact answer. Therefore, even 0.1% above an exact answer would be represented as higher than the right answer. Consequently, this representation could be said to depict just two states. Children select a hoop by clicking on it. The computer throws a ball into the hoop to indicate the correct direction.

6.5 CONCLUSIONS

This chapter has attempted to describe the theoretical rationale for the design of CENTS. The research first presented in Chapter Two concerning the knowledge and skills required for computational estimation was related to the support provided for the development of strategic and conceptual knowledge. Secondly, a description of the representations used in CENTS was given in terms of the four of the dimensions they differed upon:

- the amount of information expressed
- the resolution of information
- the modality of the representation
- the type of representation

The primary purpose of the MERs is therefore simply to support different ideas and processes by allowing a variety of different views on the task to be displayed. However, they may also help children to come to understand the domain more fully if children can translate between or abstract over them in the ways discussed in section 3.4.3 and 3.5. At this stage, no prediction has been made about how each representation may support understanding. Nor has the different learning demands of the representations been considered. The argument being developed in this thesis is that when considering MERs, it may not be sufficient to examine the properties of individual representations. Thus, these issues are addressed in the next two chapters which describe three evaluation studies with CENTS. These concentrate upon three primary questions: (a) what is the baseline performance of children who have not been taught computational estimation; (b) whether CENTS is a useful tool for learning computational estimation and, (c) how combining different representations contributes to the development of this understanding?

CHAPTER SEVEN

Experiment Three (CENTS)

The first experiment with CENTS had three main goals. The first aim was to examine strategies and knowledge involved in computational estimation. The second goal was assess how effective CENTS was at teaching children this knowledge. The final aim was to explore whether different combinations of MERs affected what children learnt.

7.1 AIMS

7.1.1 Pedagogical Aims

Two basic questions were addressed: (a) what was the nature of children's (untaught) computational estimation performance and; (b) could CENTS successfully support learning of the aspects of computational estimation that it was designed to teach ? These were:

- to teach children strategies that they can use to estimate solutions to problems
- to encourage children's understanding of how transforming numbers to produce an intermediate solution affects subsequent accuracy
- to support the development of the required underlying conceptual knowledge

This evaluation study concentrated upon the first two aspects of computational estimation. Therefore, data were collected to examine:

- the accuracy of estimates
- the appropriateness estimation strategies
- insight into how close an estimate was to the exact product of the factors (*e.g.* a little lower, a lot higher).

7.1.2 MERs Aims

CENTS uses MERs of proximity. As discussed in section 6.4, these can be displayed in many ways. One dimension of the design of multi-representational software that was introduced in section 3.7 was the similarity of format between representations. The aim of this experiment was to examine how children used MERs which varied the format of representations.

For each multi-representational system, two representations were used to emphasise different aspects of proximity. The first representation was categorical and conveyed only magnitude information (referred to as the categorical representation). The second representation was continuous and expressed both magnitude and direction information (continuous representation). Thus, the MERs for all the different formats were partially redundant. This level of redundancy was chosen so that the representations provided different views on the phenomenon, but still had some information in common. The representations were presented simultaneously and so the issue of how to sequence representations was not relevant. No automatic translation was provided by the system as the experiment aimed to analyse how difficult children would find mapping across representations that differed in format.

This experiment employed two different representational formats - pictorial and mathematical. Thus, three different types of MERs are available, two pictorial representations (picts, Figure 7.1) two mathematical representations (maths, Figure 7.2) or one pictorial and one mathematical (mixed, Figure 7.3). Each pictorial representation was graphical, but the maths and mixed systems employed one graphical and one propositional representation.

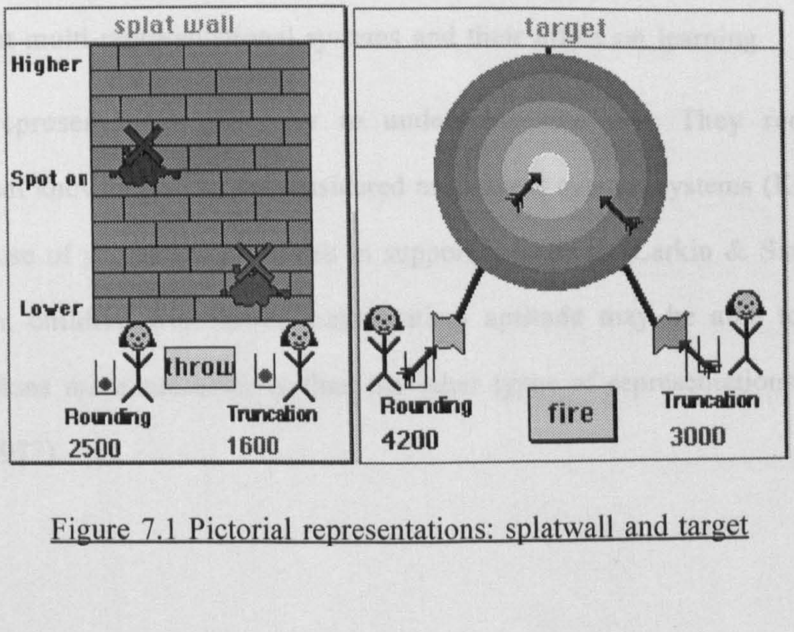


Figure 7.1 Pictorial representations: splatwall and target

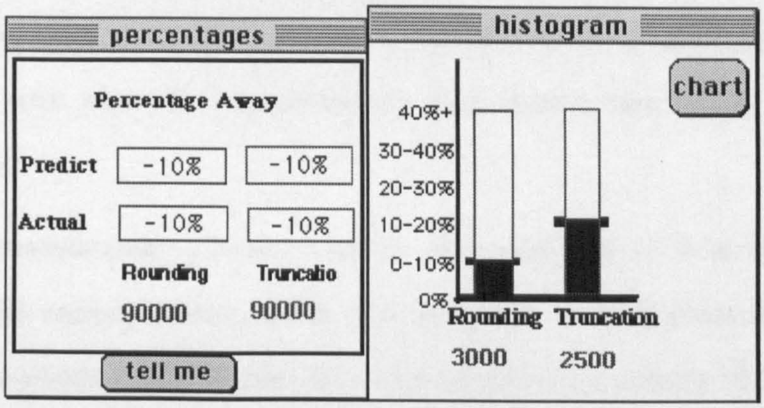


Figure 7.2 Mathematical representations: numerical display and histogram

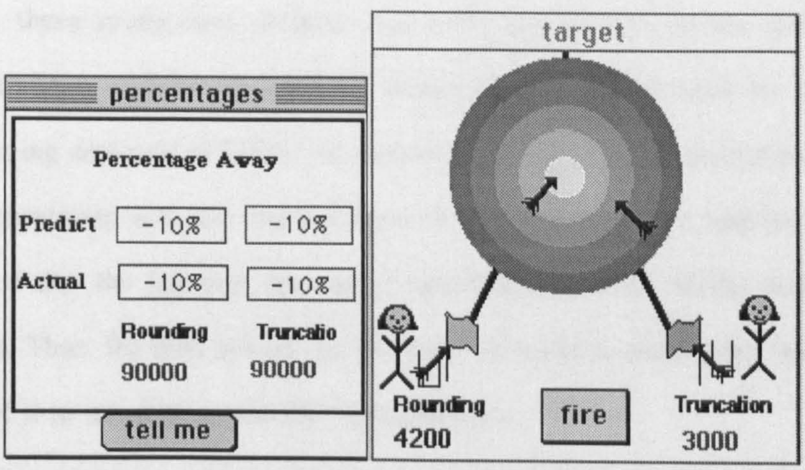


Figure 7.3 Mixed representations: numerical display and target

The existing literature on the properties of individual mathematical and pictorial representations was examined to derive a series of predictions about the properties of the different multi-representational systems and their affect on learning.

Pictorial representations are easy to understand and use. They require little mathematical knowledge, can be considered as ambient symbol systems (Kaput, 1987) and make use of perceptual processes to support inferences (Larkin & Simon, 1987). In addition, children with lower mathematical aptitude may be able to use these representations more successfully than the other types of representations (Cronbach & Snow, 1977).

By the same token, mathematical representations will be less easy to understand. They require more specialist knowledge and make less use of perceptual processes. Compared with pictorial representations, they should take longer to be used successfully.

Mixed representational systems combine the properties of both pictorial and mathematical representations. Hence, the advantages of both these representation should be available in this system. In addition, Dienes (*e.g.* Dienes 1973) argues for the multiple embodiments in mathematics education - the linking of imagery and symbolism. The mixed representations come the closest to achieving this.

However, these predictions assume that when representations are combined the effects are simply additive. In addition, these proposals only account for the first two of the learning demands of MERs (discussed in section 3.4), the format and operators of a representation and the relation between the representation and the domain. It was argued that the learning demand of translating between MERs should also be considered. Thus, for each system, an attempt was made to analyse how easy children would find it to translate across the representations.

The pictorial representations are based on the same analogy, 'distance from target' and are of the same type (classified according to Lesh *et al*,. 1987 and Lohse *et al.*'s, 1994, typologies). Prediction on both representations involves 'clicking' to select some of the representation. Feedback is given by identifying a part of the representation. Therefore, the format and operators of the representations are similar. Translation between the pictorial representations should be easily learnt.

The mathematical representations are of different types (Lohse *et al.*, 1994) and, in addition, mix modality as the graph exploits perceptual processes whilst the numerical display is propositional. Thus, the format and operators of the representations are quite different. However, both representations employ numbers. Children find it easy to recognise the similarities between representations if they contain the same numbers (DuFour-Janvier *et al.,* 1987). Although learning each representations may be difficult, it should be relatively easy to translate across representations.

Mixed representations are often used in the hope that pictures will act as a bridge to the less easily understood mathematical representations. However, these representations mix modality, use different format and operators and have no numbers in common. This situation is also the only one of the three that Lesh's model of translation (described in section 3.5) would describe as requiring a translation across representation types, rather than within types. Translation across these representations was therefore expected to be difficult.

This experiment was designed to examine the effects of these combinations of representations in light of these different predictions.

7.2 METHOD

7.2.1 Design

A two factor mixed design was used. The first factor varied representations of accuracy of the estimates. This resulted in four groups of 12 subjects consisting of subjects who received 'picts' (target and splatwall), 'maths' (histogram and numerical) and 'mixed' (target and numerical) representations. The final group were simply a no-intervention control who just took the pen and paper tests. The second factor, time, was within groups. A randomised block design was used and children were assigned to the different conditions on the basis of their scores on a mental maths test. Each group had similar numbers of boys and girls and the mean age of the subjects did not differ significantly.

A number of measures of children's performance were examined. These can be divided into two main groups: pen and paper tests given at pre-test and post-test, and computer traces which examine the users' behaviour during their sessions with CENTS. Paper measures examined the accuracy of children's estimates (answer accuracy), the prediction of the accuracy of an estimate (prediction accuracy) and the strategies they use. Prediction and answer accuracy are, in principle, independent.

Trace measures assessed behaviour with the representations. Answer accuracy was constrained by the system as it would not allow 'wrong' intermediate solutions, but prediction accuracy with the two different representations can be measured. This gave an indication of children's developing domain knowledge and their understanding of the representations. To examine whether children were learning to translate across representations, the similarity of users' behaviour across the different representations was measured (representational co-ordination). If children see the relation between representations, then their prediction on each representation should be the same, even if this prediction is incorrect. Thus, over a time, a trend towards increasing convergence should be observed. Finally, as predictions about representation use call for differential affects of ability, aptitude by treatment interaction were examined.

7.2.2 Subjects

48 mixed ability year five pupils from a state junior school took part in the experiment. They ranged in age from 9:9 to 10:8 years. All the children were experienced with mouse driven computers.

7.2.3 Materials

Mental Maths Test

A general test of mental mathematics was devised by combining exercises from books two and three of 'Think and Solve Mental Maths' (Clarke and Shepherd). It was piloted with a parallel class which was not taking part in the experiment.

Pre-test and Post-test Material

The task required children to estimate an answer to a multiplication problem. There were 20 questions, eight 3 digit by 3 digit problems (*e.g.* 213 × 789) and twelve 2 digit by 2 digit problems (*e.g.* 21 × 78). Given the strategies taught, five of the problems would be most accurately solved by rounding down, five by rounding up and ten problems by intermediate compensation. To probe the depth of insight that children had into the accuracy of their estimate, they were required to state how they thought their estimate differed from the exact answer (see figure 7.4, Appendix 2).

1. Estimate: 64 × 56 my estimate is **3000**

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

Figure 7.4 An example question (with answer) from the pen and paper test

Categories were labelled in both percentages and natural language. It was recognised that knowledge of percentages is not very developed in this age group. However, it was felt important to label the categories in such a way that proportional reasoning was emphasised. In addition, it also provided a definition for the category, one person's 'just less' may be another person's 'less'. This is particularly true when the numbers involved different orders of magnitude.

7.2.4 Procedure

Pre-tests

Children were given mental maths tests in their classroom. The class teacher read the items to the children and allowed them to query items if they had not understood a question. Children were allowed a short break after each block of ten items. In total, the test took about 30 minutes to complete.

The estimation tests were given the following day. Again, testing took place in the classroom with the teacher present who helped explain the task to the children. Instruction stressed that exact answers were not required, encouraged guessing rather than leaving an answer blank and explained how to use the insight measure. Subjects were allowed to proceed at their own pace through the test and generally took between 20 and 40 minutes to complete it. One child was stopped after an hour. Three parallel versions of each test had been created and, to prevent copying, children seated together were given different versions.

Computer Intervention

Subjects used the computer individually in a quiet corridor or classroom. The experimenter was present to help explain the instructions. To ensure sufficient practice with the system, each child used CENTS twice (the total time spent on the computer was between 80 and 100 minutes), separated by approximately two weeks. The three different versions of the program had the same basic structure. The user was greeted by a screen welcoming them to CENTS. Instructions explained what the task was and how they should answer the question. The experimenter demonstrated the task to the children and then stayed to provide support if they became confused about how to operate the system (but did not provide direct teaching).

Children were set eight questions which they had to answer by truncation and by rounding. All questions presented were generated dynamically, hence the problems each child was set were different. Each child started with a two (digit) by two problem and gradually included larger problems (two by three and three by two) and ended with a three by three problem. After each problem, children filled in the log book recording details of their estimates.

Post-test

Children received a parallel version of the estimation test within 10 days of their second computer period. Interestingly, application of the test took longer than at pre-test, requiring between 20 and 80 minutes.

7.3 RESULTS

To examine the effects of the intervention, a number of [4 by 2] ANOVAs were performed on the pre-test and post-test data. The design for the analyses was 4 (control, maths, mixed picts) by 2 (pre-test, post-test). The first factor, format. was between groups and the second, time, a within group repeated measure. In addition, trace logs from the two intervention sessions were analysed using [3 by 2 by 2] ANOVAs. Children gave two answers for each problem, hence the strategy used was included as a factor. The design was 3 (maths, mixed, picts) by 2 (rounding, truncation) by 2 (time 1, time 2). The first factor was between groups and the last two (time and strategy) are within groups factors.

7.3.1 Answer Accuracy

A commonly used measure of estimation performance is the percentage deviation of the estimate from the exact answer. This was examined using an [4 by 2] ANOVA on the pen and paper data (Table 7.1). The results from one subject have been dropped. She was an extreme outlier scoring 10 standard deviations above the mean at pre-test.

Table 7.1. Percentage deviation of estimate by format and time

	Control	Mixed	Maths	Picts
Pre-test	89.83% (16.9)	88.99% (9.5)	101.25% (62.5)	102.95% (55.7)
Post-test	82.66% (13.9)	60.71% (24.6)	55.29% (45.1)	57.56% (34.1)

As can be seen from Table 7.1, the pre-test performance of the children was very poor. The average percentage deviation from the correct answer was 96%. This created two problems. Firstly, the data were non-homogenous and no transform could solve the problem. Secondly, this measure has traditionally only been used on deviations of up to 40%. Consequently, other measures of performance were designed.

One problem with using a percentage deviation is that a large number of children performed appropriate transformations, correct front-end extraction and multiplication, but failed at place value correction. To distinguish those children who only failed at the final step from those who used incorrect strategies or just guessed answers, the estimates were corrected for order of magnitude. A child answering 1200 to '221 × 610' would therefore be corrected from 99% to 11% inaccurate by this measure. However, a guess of 2500 would remain 80% inaccurate. This should identify which children were generating plausible estimates, only failing at order of magnitude.

Table 7.2. Percentage deviation of corrected estimate by format and time

	Control	Mixed	Maths	Picts
Pre-test	38.58% (10.6)	38.32% (18.1)	38.08% (15.6)	40.74% (12.16)
Post-test	42.08% (10.4)	27.08% (14.5)	24.02% (17.2)	27.84% (19.1)

Analysis using an [4 by 2] ANOVA showed a significant main effect of time ($F(1,44)=10.84$, $p<0.002$). The interaction between format and time was also significant ($F(3,44)=3.006$, $p<0.040$) (Figure 7.5). Simple main effects analysis found no significant differences between the groups at pre-test ($F(3,88)=0.114$), but there were differences at post-test ($F(3,88)=4.57$, $p<0.02$). The control groups performance did not change, but all three experimental groups improved significantly:

- control ($F(1,44)=0.84$)
- mixed ($F(1,44)=4.58$, $p<0.038$)
- maths ($F(1,44)=7.42$, $p<0.009$)
- picts ($F(1,44)=7.025$, $p<0.011$)

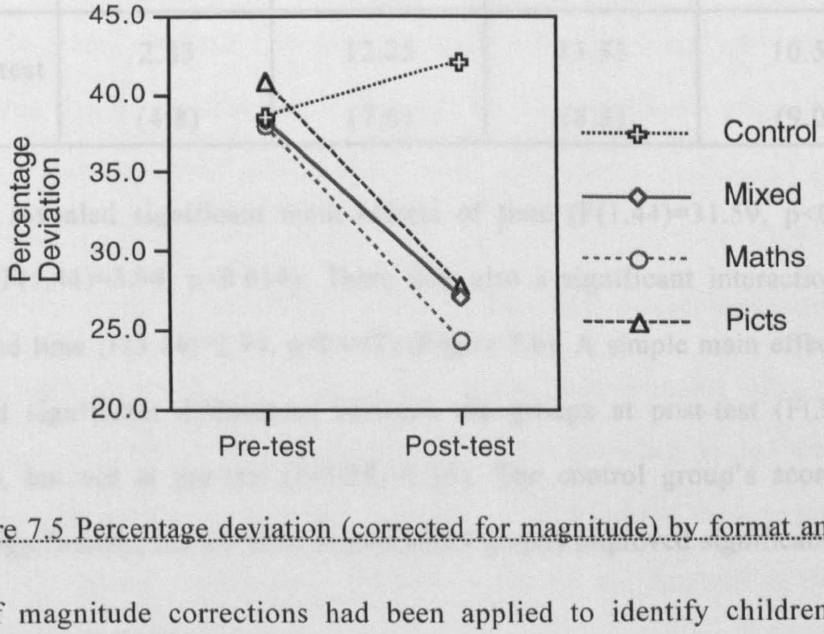


Figure 7.5 Percentage deviation (corrected for magnitude) by format and time

Order of magnitude corrections had been applied to identify children who had correctly produced an intermediate solution. However, when examining how many orders of magnitude were needed to correct the estimate, the degree of correction significantly decreased from pre-test to post-test ($F(1,44)=22.014$, $p<0.001$). Children did gain some understanding of place value correction during the study.

7.3.2 Strategic Knowledge

A second way of examining subjects' estimation skills was to code whether their estimates were produced using a recognised strategy. Estimates were identified as rounding up, rounding intermediate compensation, rounding down (or truncation), exact answer (or attempt to produce one), addition, or unknown. (The other strategies identified by Reys *et al.* (1982) were not found in these studies). Hence, strategies that involved front end extraction were coded as appropriate and all others as inappropriate*. The number of estimates generated by a recognised strategy was examined (Table 7.3).

Table 7.3. Numbers of estimates generated by a recognised strategy by format and time (out of 20)

	Control	Mixed	Maths	Picts
Pre-test	1.91 (3.8)	6.58 (8.0)	4.33 (5.1)	2.75 (3.5)
Post-test	2.83 (4.8)	12.25 (7.6)	13.33 (8.3)	10.58 (9.0)

Analysis revealed significant main effects of time ($F(1,44)=31.59$, $p<0.001$) and format, ($F(1,44)=3.98$, $p<0.014$). There was also a significant interaction between format and time ($F(3,44)=2.94$, $p<0.043$) (Figure 7.6). A simple main effects analysis identified significant differences between the groups at post-test ($F(3,88)=6.19$, $p<0.001$), but not at pre-test ($F(3,88)=1.16$). The control group's scores did not change significantly, but all three experimental groups improved significantly:

---

* A second coder examined 10% of the scripts. No formal inter-rater reliability was performed as over 97% of codes agreed.

• control (F(1,44)=0.19)

• mixed (F(1,44)=7.44, p<0.01)

• maths (F(1,44)=7.97, p<0.001)

• picts (F(1,44)=4.67, p<0.001)

Tukey tests showed that the picts group did not perform significantly better than the control group at post-test, although the other experimental groups did: mixed v control (q=4.17, p<0.05) and maths v control (q=4.64, p<0.05) (see figure 7.6).

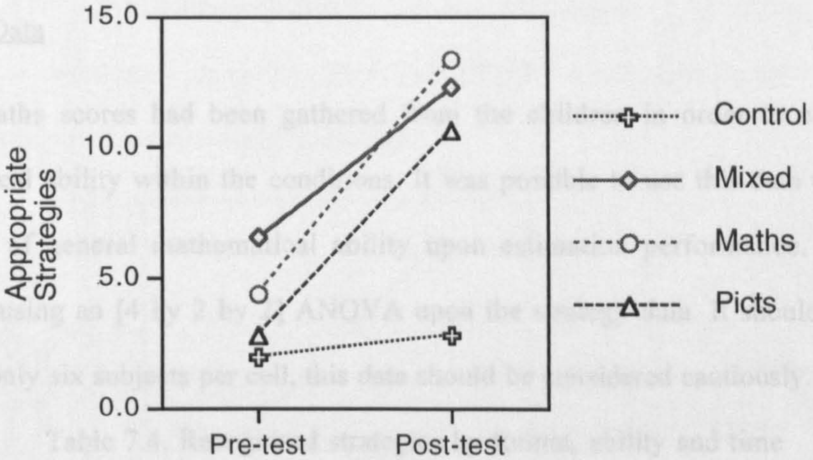


Figure 7.6 Number of appropriate estimation strategies by format and time

The results for the analyses of accuracy and appropriate strategies can therefore be seen to be very similar. The control group did not significantly alter their performance and all the experimental groups showed significant improvement from pre-test to post-test. The performance of the different experimental groups was almost identical. It seems that children can learn to estimate with CENTS and that improvements in performance were not due to the effects of repeated testing.

The tests were constructed such that in 25% of the cases rounding down (truncating) was the most accurate (taught) strategy to use, 25% rounding up and the remaining 50% intermediate compensation. If children were picking the most accurate known strategy, then this pattern should be reflected in the scores. Analysis of the estimates at post-test found that intermediate compensation is the most common strategy accounting for 44.7% of appropriate estimates. However, there is also a very high incidence of rounding down/truncation (42.8%). This indicates that some answers were generated by truncating as opposed to rounding down. Inspection of individual children's results suggested that some children invariably truncated. This is not necessarily to be discouraged. As it was suggested that truncation is an easier strategy than rounding, then this strategy can provide success where an attempt at rounding might lead to failure. Only 12.5% of answers were generated using a rounding up strategy. This is much less than would be expected if children were choosing the most accurate way to solve the problem.

Aptitude Data

Mental maths scores had been gathered from the children in order to control for mathematical ability within the conditions. It was possible to use this data to explore the effect of general mathematical ability upon estimation performance. This was examined using an [4 by 2 by 2] ANOVA upon the strategy data. It should be noted that with only six subjects per cell, this data should be considered cautiously.

Table 7.4. Recognised strategies by format, ability and time

(Higher mental maths scores)

	Control	Mixed	Maths	Picts
Pre-test	2.33 (4.8)	9.83 (8.3)	6.17 (6.3)	5.0 (3.75)
Post-test	1.5 (5.3)	17.17 (2.76)	17.83 (2.6)	15.17 (7.7)

(Lower mental maths scores)

	Control	Mixed	Maths	Picts
Pre-test	3.17 (2.7)	3.33 (6.8)	2.51 (2.8)	0.05 (.8)
Post-test	2.50 (4.7)	6.17 (7.89)	8.83 (9.8)	6.01 (8.3)

As before, there were main effects of format and time. There was also a main effect of ability (F(1,40)=16.49, p<0.0002) (Tables 7.4, Figure 7.7). There were no significant interactions with ability. Therefore, CENTS seems suitable for children of with wide ranges of mathematical knowledge.

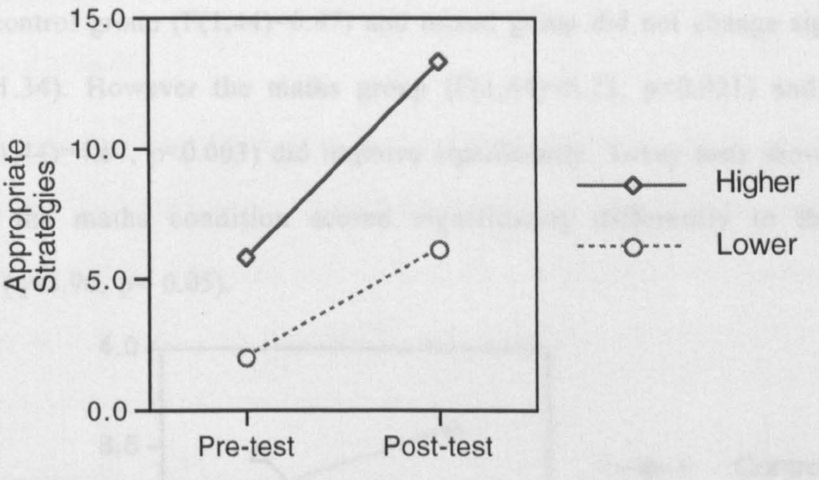


Figure 7.7 Estimation strategies by ability and time (collapsed across format)

7.3.3 Prediction Accuracy

The above measures explored the improvement in the use and application of estimation strategies. However, they did not examine insights the children may have into the process of estimation and how the estimates differ from the exact answer. This was assessed using the tick boxes which subjects filled in to indicate how far away an estimate was from the exact answer. The responses were coded as the difference between the category that they should have selected given their estimate and those that they did which provides a score between 0 and 8 per answer. This was examined using an [4 by 2] ANOVA.

Table 7.5. Difference between prediction and estimate by format and time

	Control	Mixed	Maths	Picts
Pre-test	3.26 (0.91)	3.06 (1.74)	3.06 (0.91)	3.46 (0.62)
Post-test	3.58 (1.18)	2.67 (0.97)	2.28 (0.81)	2.44 (1.32)

Analysis revealed a significant main effect of time ($F(1,44)=8.25$, $p<0.007$) and a significant interaction between format and time ($F(3,44)=3.28$, $p<0.03$) (Table 7.5 and Figure 7.8). There were no significant differences between groups at pre-test ($F(3,88)=0.456$), only at post-test ($F(3,88)=4.14$, $p<0.008$). The performance of both the control group ($F(1,44)=0.97$) and mixed group did not change significantly ($F(1,44)=1.34$). However the maths group ($F(1,44)=5.73$, $p<0.021$) and the picts group ($F(1,44)=4.67$, $p<0.003$) did improve significantly. Tukey tests showed that at post-test, the maths condition scored significantly differently to the control condition ($q=3.90$, $p< 0.05$).

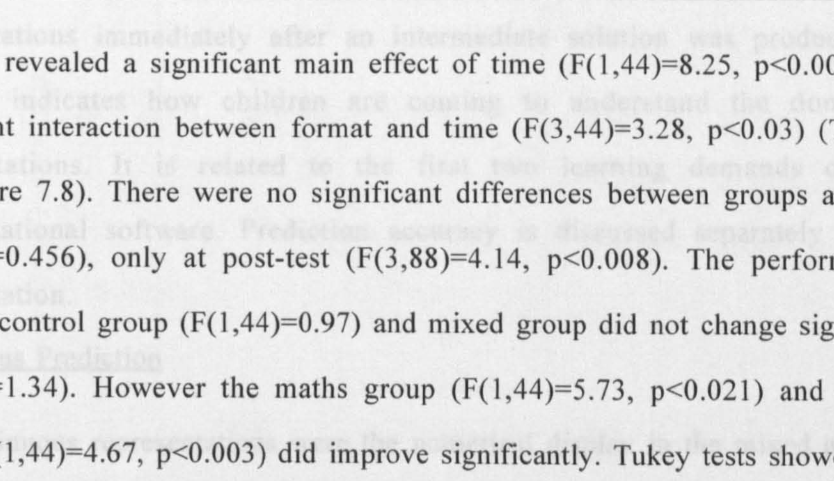


Figure 7.8 Prediction accuracy by format and time

This is the only pen and paper measure where there is any difference between the experimental conditions. It is also the measure that is most directly related to the representations. In order to examine more closely how the different MERs may have affected learning outcomes, the computer logs generated during the intervention session were examined.

7.3.4 Process Data

The first measure of performance examined was prediction accuracy. This is similar to the paper tests of prediction as children were asked to predict how far their estimate will be from the exact answer. Prediction was performed using different representations immediately after an intermediate solution was produced. This measure indicates how children are coming to understand the domain and representations. It is related to the first two learning demands of multi-representational software. Prediction accuracy is discussed separately for each representation.

Continuous Prediction

The continuous representations were the numerical display in the mixed and maths conditions and the 'splatwall' in the picts condition. The data from the splatwall were recoded as percentage deviation scores using the underlying model which drives the representation. An [3 by 2 by 2] ANOVA was conduced upon the on-line data from the two trials with CENTS (see Table 7.6).

Table 7.6. Difference between prediction and estimate by format, time and strategy

Percentage deviation

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	18.85 (15.1)	19.92 (10.4)	20.56 (16.1)	21.39 (11.0)	13.94 (5.7)	13.75 (5.6)
Time 2	16.85 (8.5)	19.17 (8.4)	10.18 (5.5)	13.32 (7.3)	11.82 (4.02)	11.65 (5.22)

These data did not pass homogeneity of variance tests, so were transformed using a natural log function. There were significant main effects of time ($F(1,33)=9.02$, p<0.005) and strategy (F(1,33)=4.29, p<0.046); answers generated by rounding were predicted more accurately. There was also significant interaction between time and format (F(2,33)=3.81, p<0.032) (Figure 7.9). Simple main effects showed no significant differences between the groups at time one, but there were at time two (F(2,66)=3.73, p<0.029). The maths condition demonstrated significant improvement in performance (F(1,33)=14.67, p<0.001).

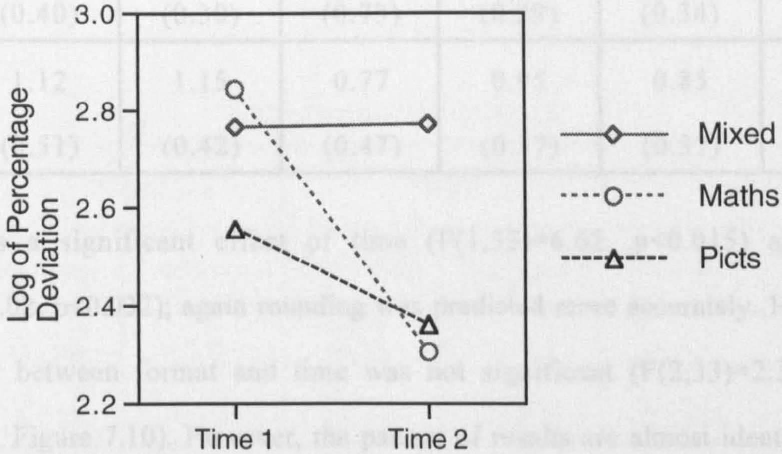


Figure 7.9 Predictive accuracy by format and time (collapsed across strategy).

Categorical Prediction

The second type representation employed in this experiment was a categorical representation which represented magnitude (*i.e.* either histogram or archery target). The system logs which category the user predicted and this can be compared to the one they should have predicted given their estimate. This gave a difference score (between 0 and 4) that was analysed by an [3 by 2 by 2] ANOVA.

Table 7.7. Difference between prediction and estimate by format, time and strategy

Category differences

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	1.05 (0.40)	1.29 (0.30)	1.19 (0.73)	1.29 (0.29)	0.99 (0.34)	1.09 (0.46)
Time 2	1.12 (0.51)	1.15 (0.42)	0.77 (0.47)	0.95 (0.37)	0.85 (0.37)	1.01 (0.50)

There was a significant effect of time ($F(1,33)=6.62$, $p<0.015$) and strategy ($F(1,33)=5.02$, $p<0.032$); again rounding was predicted more accurately. However, the interaction between format and time was not significant ($F(2,33)=2.36$, $p<0.11$). (Table 7.7, Figure 7.10). However, the pattern of results are almost identical to those for the continuous representations described above.

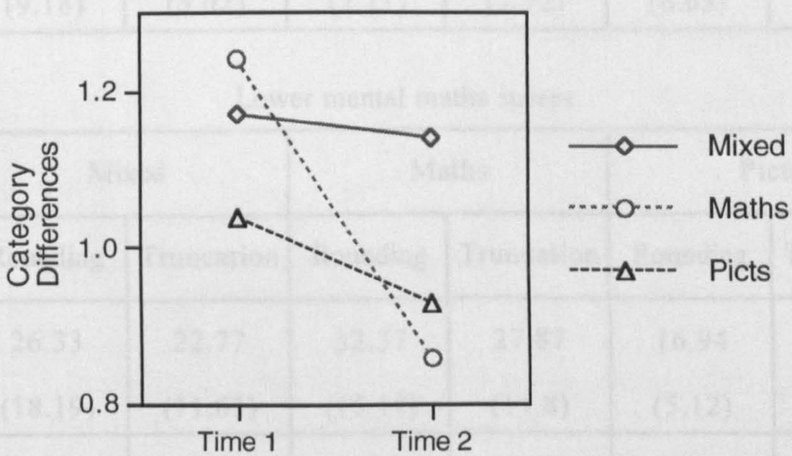


Figure 7.10 Predictive accuracy by format and time (collapsed across strategy).

Thus, for both representations, the poorer understanding demonstrated by children from the mixed condition was apparent by the second intervention session.

Aptitude Measures

The representations differed in terms of mathematical knowledge required to interpret them. Hence, it was predicted that there may be an effect of children's mathematical ability. The continuous representations were examined and a median split by the mental maths scores was performed upon the data. This was then analysed using an [3 by 2 by 2 by 2] ANOVA. The design was 3 (mixed, maths, picts) by 2 (high, low scores) by 2 (time 1, time 2) by 2 (rounding, truncation). The first two factors are between groups and the last two are within groups.

Table 7.8. Difference between prediction and estimate by format, ability, time and strategy

Higher mental maths scores

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	11.36 (6.22)	17.07 (9.23)	8.75 (2.91)	14.91 (5.32)	10.93 (4.91)	12.00 (2.08)
Time 2	15.91 (9.18)	12.06 (5.02)	6.54 (2.21)	7.58 (2.72)	10.54 (6.03)	12.27 (6.71)

Lower mental maths scores

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	26.33 (18.19)	22.77 (11.63)	32.37 (15.17)	27.87 (11.8)	16.94 (5.12)	15.27 (4.99)
Time 2	17.79 (8.63)	25.10 (6.91)	13.84 (5.59)	19.04 (5.76)	13.04 (5.53)	12.23 (4.00)

These data did not pass homogeneity of variance tests and so were transformed using a natural log function. As before, analysis revealed main effects of time and strategy and a trend for a main effect of format ($F(1,30)=2.61$, $p<0.089$) (Table 7.8). There was also a main effect of ability ($F(1,30)=27.60$, $p<0.001$). The students who had been judged to have greater mathematical ability (by mental maths scores) were significantly better at predicting the accuracy of their estimates than those with lower scores.

There proved to be a strong trend towards an interaction between format and ability ($F(2,30)=2.980$, $p<0.066$). A simple main effects analysis found the only significant differences between representation use were for children with lower mental maths scores ($F(2,30)=3.462$, $p<0.044$). The three different representations were differentially affected by ability. The mixed and maths conditions demonstrated a significant effect of ability on representation, mixed ($F(2,30)=7.69$ $p<0.01$) and maths ($F(2,30)=23.65$, $p<0.001$). Children with higher mental maths scores use these representations more successfully. However, there were no significant difference between higher and lower mathematical ability children for the picts condition ($F(2,30)=2.18$) (Figure 7.11).

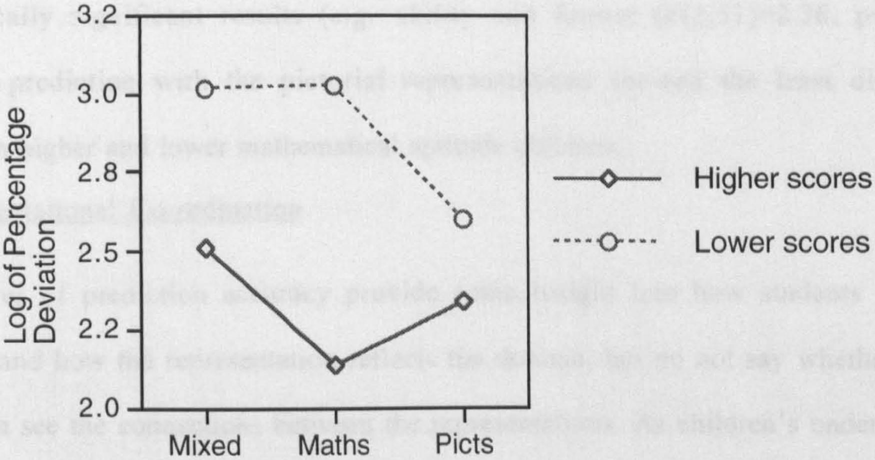


Figure 7.11 Predictive accuracy by format and ability (collapsed across strategy).

All groups were significantly worse at predicting truncation estimates but there was also a three way interaction between time, task and ability, ($F(1,30)=5.91$, $p<0.021$) (see Figure 7.12). Simple simple main effects analysis showed that the high ability group significantly improved performance in truncation ($F(1,30)=4.72$ $p<0.038$), but not in rounding ($F(1,30)=.03$). The lower ability group improved at rounding ($F(1,30)=7.78$, $p<0.009$) but not truncation ($F(1,30)=0.687$).

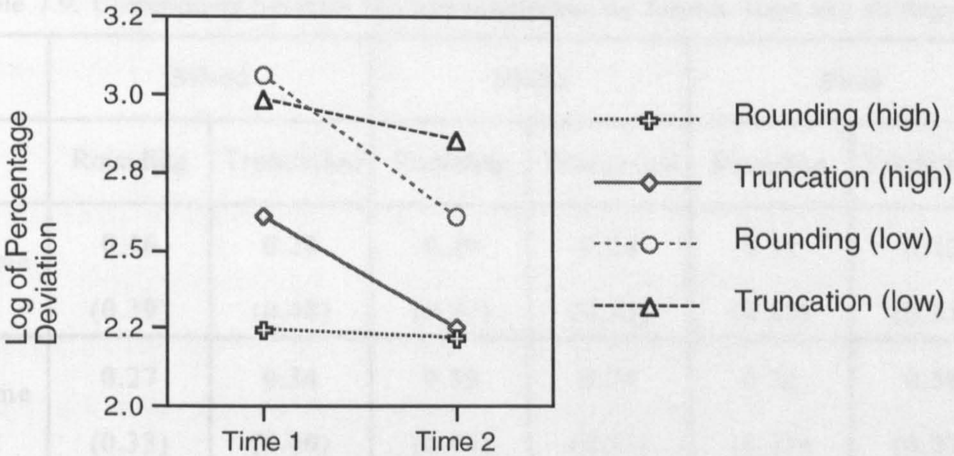


Figure 7.12 Predictive accuracy by ability, task and time (collapsed across format).

It is also possible to repeat this analysis for the categorical representations (target and histogram). The pattern of results is very similar, although as for the previous analysis of categorical representations, it manifests as trends rather than as statistically significant results (*e.g.* ability and format ($F(2,51)=2.36$, $p<0.098$). Again, prediction with the pictorial representations showed the least difference between higher and lower mathematical aptitude children.

Representational Co-ordination

Measures of prediction accuracy provide some insight into how students come to understand how the representation reflects the domain, but do not say whether or not children see the connections between the representations. As children's understanding of the representational system improves, their behaviour should become similar across both representations, even if this behaviour is still flawed with respect to the domain. For example, if a prediction of very close is made on the first representation, it should be made on the second representation as well.

The first analysis correlated the predictions on the two different representations across each session. This was examined by an [3 by 2 by 2] ANOVA.

Table 7.9. Correlations between the representations by format, time and strategy

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	0.46 (0.39)	0.28 (0.48)	0.49 (0.39)	0.44 (0.46)	0.32 (0.42)	0.42 (0.31)
Time 2	0.27 (0.33)	0.34 (0.40)	0.59 (0.44)	0.74 (0.33)	0.56 (0.37)	0.58 (0.27)

There is a noticeable trend for the correlations to be higher on the second use of the system although this difference is not significant ($F(1,33)=3.629$, $p<0.065$) (Table 7.9, Figure 7.13). It was predicted that different conditions would differentially improve in co-ordination. Simple main effects showed improvement for the maths group ($F(1,33)=3.73$, $p<0.062$), and the picts group ($F(1,33)=3.824$, $p<0.059$). However, the mixed group showed no evidence of improved co-ordination ($F(1,33)=0.345$).

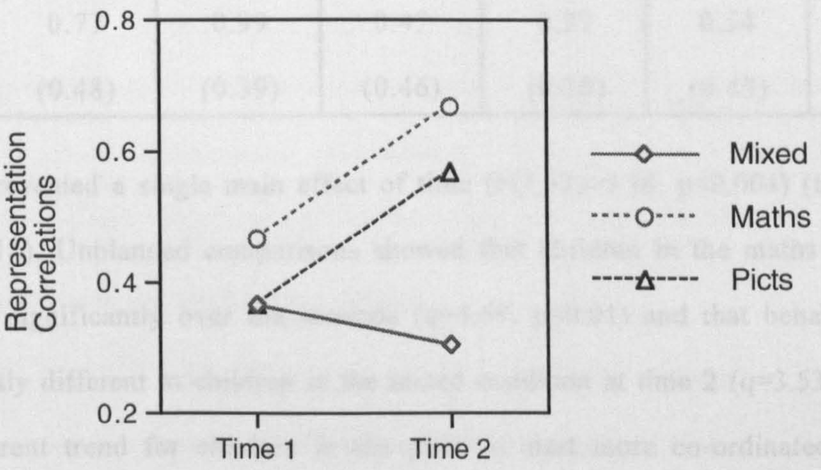


Figure 7.13 Correlation of behaviour on the representations by format and time

Correlations are insensitive to any rescaling of the representations by the children. For example, if children in the mixed condition had predicted 'band a' (target) with 1% (numerical), followed by 'band b' (target) with 2% (numerical), 'band c' (target) with 50% (numerical), they would be perfectly co-ordinated, but this obviously represent rescalings by the children. It is also likely that some multi-representational systems are more likely to be rescaled than others. Both the mathematical representations contain numbers which may inhibit rescaling.

An alternative way to examine similarity of behaviour was to convert the continuous representation into the appropriate absolute category so that the two predictions could be compared (prediction A - prediction B). If both representations were used to predict the same answer, then the sum of the differences between the two groups should be 0. If the answers were maximally discrepant, the maximum average value is 4 (as there were 5 categories in total).

Table 7.10: Difference in prediction between the representations by format, time and strategy

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	1.05 (0.50)	0.93 (0.48)	0.91 (0.56)	0.90 (0.54)	0.74 (0.38)	0.81 (0.43)
Time 2	0.77 (0.48)	0.99 (0.39)	0.47 (0.46)	0.37 (0.38)	0.54 (0.43)	0.64 (0.34)

Analysis revealed a single main effect of time ($F(1,33)=9.36$, $p<0.004$) (table 7.10, Figure 7.14). Unplanned comparisons showed that children in the maths condition improved significantly over the sessions ($q=4.69$, $p<0.01$) and that behaviour was significantly different to children in the mixed condition at time 2 ($q=3.53$, $p<0.05$). The apparent trend for children in the picts to start more co-ordinated was not significant.

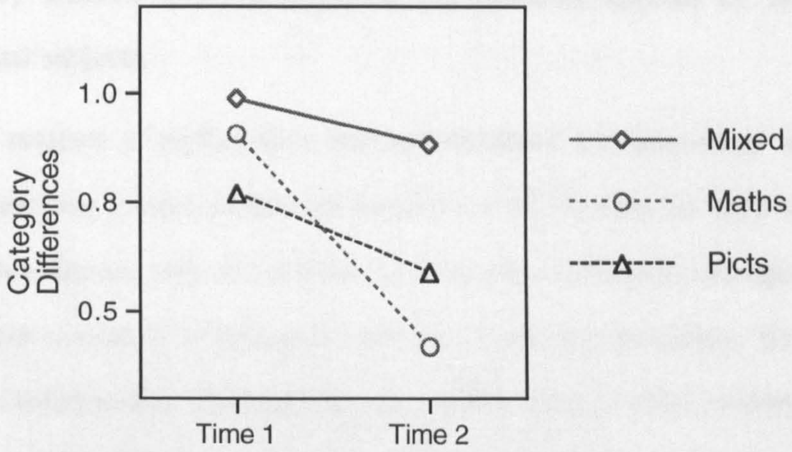


Figure 7.14 Differences between prediction behaviour on the two representations by format and time

The two different representational co-ordination measures provide a very similar account of children's ability to translate over representations. It would seem that children given mixed representations find it much more difficult to translate between the representations as they show no improvement in co-ordination over the two sessions.

7.4 DISCUSSION

The two primary goals of this evaluation were to examine whether CENTS could successfully teach computational estimation and to explore the effects of different combinations of representations on children's understanding of estimation.

7.4.1 Computational Estimation Skills

Children's estimation skills were examined in a number of different ways. The first analysis assessed whether an estimate had been generated using an appropriate strategy (in this case, one that involved front-end extraction). The children's knowledge of estimation at pre-test was generally low. An average 22% of answers were generated using an appropriate strategy. At post-test, children in the control group did not use more appropriate strategies (14%). Children in all experimental groups had improved significantly - 60% of post-test estimates were generated using an appropriate strategy. Hence, it would seem that the strategies that were taught to children by CENTS were remembered and correctly applied by many of the experimental subjects.

The other measure of performance that was examined was percentage accuracy. At pre-test, there was a mean percentage deviation of 95.5% from the right answer. This demonstrates that not only did children not know any estimation strategies, they also did not have a sense of a 'ball park' estimate. It was not uncommon for children to estimate a solution that was less than one of the factors of the problem (much to their teacher's horror!)

The model of estimation proposed by LeFevre *et al.* (1993) suggests that estimation involves three main stages: production of an appropriate intermediate solution, then calculation of this value, and finally place value correction. Given the generally poor performance of children at pre-test, it was not possible with the percentage deviation measure to discriminate between children who had no idea about how to estimate and those who performed all stages but the final one correctly. The children's answers were corrected for order of magnitude (*i.e.* to include the final step) and the results re-examined. The three experimental groups' modified percentage deviation scores show a significant improvement from 39% to 26% after the intervention. The control group did not improve with 38% at pre-test and 42% at post-test. This measure confirms the strategy analysis which showed that children were becoming significantly better at estimating after using CENTS. In addition, although children remain relatively poor at producing the right place value correction, this did improve significantly over time.

Further analysis had investigated whether general mental mathematical aptitude was related to learning outcomes. No strong claim is made about these data, as the mental mathematics was not measured with a standardised test. Children with higher mental maths scores were found to have performed better on the estimations tasks at both pre-test and post-test. Encouragingly, there was no interaction between aptitude and time. This suggests that CENTS provides an appropriate learning environment for children of all abilities and is not only suited for higher ability students.

One area of concern was the relatively poor order of magnitude correction performed by the experimental subjects at post-test. Although, performance had significantly improved from pre-test, children were often incorrect, especially on the three by three digit sums. As discussed in section 6.3, CENTS provides a variety of help settings - the higher the level of help, the more support provided by the computer. During the intervention, this parameter had been set relatively high for place value support. It was decided that in subsequent studies with CENTS, that users should be given responsibility for place value correction much earlier in the intervention. This would allow a longer period of time for children to practise these skills.

The strategic support provided by CENTS was constant across all three conditions. Consequently, it had not been proposed that there would be any differences between the experimental groups on these measures. The prediction was supported by the data. The only differences found between the experimental groups were related to the use of representations and how these affected children's understanding of the accuracy of an estimate.

7.4.2 Prediction Accuracy

In order to become a flexible, accurate estimator, children need to understand how transforming numbers to produce an intermediate solution affects the accuracy of the subsequent estimate. The development of this skill is supported in CENTS by asking children to predict the accuracy of their estimate when they have produced an intermediate solution. They perform this action and receive feedback by using the different representations.

Pen and paper measures of prediction accuracy showed that children in the maths and picts conditions made significant improvements in this skill. However, children in the mixed group became significantly more accurate estimators (answer accuracy) without becoming better at knowing how accurate their answers were (prediction accuracy).

This skill is the one most directly supported by the MERS. Hence, any differences in what children learnt from the different combinations of representations would be expected to manifest in this measure. Predictions based on analysis of the individual representations had suggested that the mixed representations should provide the best rather than the worst conditions for learning. However, analysis of the learning demands of mixed representations had identified translation across these representations as particularly difficult. Hence, the results from the pen and paper tests suggested that there was some affect of learning translation between MERS. This was further examined by analysing the intervention logs.

One interesting analysis of the insight measure was to examine the number of times children indicated they thought their estimate was exactly right. This prediction was made equally across all three experimental conditions at post-test, it accounted between 11.5% and 14% of predictions in each condition. However, this prediction seemed to be associated with very different levels of understanding across children. Some children predicted exactly right on a high proportion of estimates and seemed to believe that an estimation provided the right answer. Other children, however used it in a very specific circumstance, for example, the problem '18 × 92' when transformed to '20 × 90'. These children appeared to reason in absolute terms and believed if one number was reduced by two and the other increased by two, then the changes cancelled out. This represents sophisticated, if flawed, reasoning on the part of these children. Hence, this particular prediction seems associated with children with the both the least and most knowledge of prediction accuracy.

7.4.3 Representation Use

This experiment was designed to examine how learning with MERs which differed in similarity of format influences learning. Two aspects of multiple representation usage were examined. The first analysis assessed how each of the representations was used with respect to the domain, *i.e.* prediction accuracy. The second analysis concerned the similarity of children's behaviour across both representations. The analysis of similarity of behaviour was proposed to test representational co-ordination. These two analyses are related but independent. It is perfectly possible to fully understand how two representations relate to each other, but still have incomplete knowledge of the domain. Thus, the prediction accuracy is related to the first two learning demands of MERs - learning the format and operators of a representation and learning the relation between the representation and domain. Representational co-ordination is proposed to measure the third learning demand of MERs - translation between representations.

These measures are similar in kind to the measures proposed by Schwartz & Dreyfus (1993) (described initially in section 3.6). They both attempt to distinguish domain and interpretation knowledge from knowledge of translation between representation. The most fundamental distinction between these measures and those of Schwartz & Dreyfus is that these researchers were interested in differences between how individual students used representations. Whereas this research aimed to explore how different types of representations influenced translation. Thus, the measures used in the thesis needed to be sensitive to the degree of similarity of use. Schwartz & Dreyfus's measure used a categorical system to describe whether all of the information was correctly transferred from one representation to another (+1 or -1). Another important difference is that by using the representations in the mixed condition that were also present in other conditions, it is possible to begin to separate out the two different processes of translation between different representations occurring directly, or through mapping onto domain knowledge as a mediating agent (first raised in section 3.6). If no direct translation was occurring between representation, then performance with an representation should not be affected by a second representation as interpretation of the representation should only be through the domain. If, however, translation between representation does occur, then it would be expected that the same representation would be used differently depending on other representations it was paired with.

Incidental differences between the two approaches are based more upon the nature of the domains. The function problems of Schwartz & Dreyfus require representations to be used at different stages of the problem as a solution is slowly converged upon. The estimation task uses representations once in each (much shorter) problem, and these representations are co-present rather than switched between. Consequently, while there is obvious similarity between representational co-ordination as defined in this thesis and the passage index of Schwartz and Dreyfus, it would not have been appropriate to use their approach to address the questions posed by this research.

The two measures (prediction accuracy and representational co-ordination) used to examine children's performance are related to the benefits claimed for MERS discussed in section 3.3. The first measure relates to the proposals of a number of researchers (*e.g.* Tabachneck *et al.*, 1995; Cox & Brna, 1995; Kaput, 1987) that one of the advantages of MERs is that they allow for different ideas and processes to be represented and supported. On this view of MERs, it is less important that users of MERs recognise the similarity of representations; instead the emphasis is on understanding how each representation reflects the domain.

A second suggested advantage of MERs is that promote deeper understanding by allowing learners to abstract across representations to uncover invariances in a domain (Kaput, 1989). For this use of MERs, it is crucial that users are able to translate between the different representations. This was examined using representational co-ordination.

The analysis of how the representations were used with respect to the domain showed that the two types of representation, (continuous and categorical) demonstrate a strikingly similar pattern of results across children's two interactions with the system. However, the only statistically significant interactions were for the continuous representation. The maths group became significantly better at predicting the accuracy of their estimates over time. It would seem (unsurprisingly) that there was a significant cost associated with learning how to use mathematical representations.

Once understood these representations were used successfully. Children in the picts group did not improve at predicting. However, at time one there was a demonstrable trend for better prediction accuracy than the other groups. At time two, the picts condition demonstrated almost identical performance to the maths groups. The mixed group did not improve and were significantly worse than the maths group at time two. Hence, it would seem that, relative to the other groups, the mixed group were worse at predicting accuracy using the representations.

However, both representations that were used by the mixed group were also common to one of the other groups; the target was used by the picts group and the numerical representation was used by the maths group. When the representations were employed in these conditions, they were used successfully. Hence, it was proposed that the explanation of the poorer performance of the mixed group lies in the combination of the representations rather than in the individual representations.

It was proposed that if users were able to translate across the different representations then their representation usage should be essentially identical, even if imperfect with respect to the domain. Hence, it was predicted that as experience with the system increases, there should be a trend towards increasing convergence. This convergence is seen with both the maths and picts group, but not with the mixed group. This failure to converge suggests that children are not successfully able to translate between the two mixed representations and hence don't construct the same degree of domain knowledge as the other groups. This is then reflected on their post-test performance.

7.4.4 Properties of the MERS

The properties of each multi-representational system were examined in order to explain why co-ordination occurred in picts and maths cases but not in the mixed condition. Many explanations could be provided for why the picts representations were successfully co-ordinated. In this particular case of target and 'splatwall', both representations were based upon the same metaphor. Each represents proximity as physical distance from a goal. Children selected part of the representation to indicate the accuracy of a representation. Feedback was provided by the computer 'throwing a missile' at the representation. Thus, both the format and the operators for these representations were almost identical. Additionally, the interaction supported by the representations was very similar. A direct manipulation interface was used to act upon both representations. Obviously, these similarities need not necessarily apply to all combinations of pictorial representations.

In addition to these factors, pictures (and natural language) are 'ambient symbol systems' (Kaput, 1987). It is known that expertise is needed to successfully use external representations (*e.g.* Petre & Green, 1993). Children of this age will have had considerable opportunity to interpret language and pictures, but relatively little experience with other representations. Hence, it would be expected that translation between two familiar types of representations would be more easily achieved.

Translation between the different mathematical representations also occurred successfully. This was initially more surprising. The histogram representation is graphical and exploits perceptual processes. By contrast, the numerical display is propositional. The interface to the representations is also different. The histogram was acted upon by direct manipulation and the numerical display via the keyboard. These representations are also relatively unfamiliar to children of this age.

The explanation proposed is that mapping between the representation was facilitated as both representations use numbers. DuFour-Janvier *et al.* (1987) suggested that children only believed that two representations were equivalent if they both used the same numbers. Thus, the numbers could be used to help learners translate across the representations.

The mixed representations differed in terms of modality - the archery target representation is graphical and the numerical display is propositional. The interface to the representations also mixed direct manipulation and the keyboard. This multi-representational system also combined mathematical and non-mathematical representations. Amongst others, Kaput has made a strong distinction between these types of representation. Research on multimodal functioning when children are acquiring new mathematical concepts (*e.g.* Watson, Campbell & Collis, 1993) and research on word algebra problems (*e.g.* Tabachneck *et al.*, 1994) suggest that different types of representation may also lead to completely different strategies. Finally, research on novice-expert differences (*e.g.* Chi *et al.*, 1981) would predict that learners would find it more difficult to recognise the similarity between representations when their surface features differ. Thus, it can be seen for mixed representations that failure of overlap occurred at all levels.

7.4.5 Mathematical Aptitude and Representation

It was proposed that the representations require differential amounts of mathematical knowledge in order to be used successfully, hence there may be aptitude by treatment interactions in the children's use of representations. It was found that the measure of mathematical aptitude used (the mental maths scores) did interact with representation use. Prediction accuracy for the maths and mixed group was significantly affected by ability. Children with higher mental maths scores were significantly better at predicting than children with lower scores. This relation was not demonstrated for children in the picts group. There were no significant differences in prediction between high and low scoring children in this condition. Hence, for this task, it would seem children identified as have better existing mathematical knowledge and skills were not affected by the type of representation. Children measured as lower mathematical aptitude were found to benefit from pictorial representations.

Previous research on the aptitude by treatment interactions and representations has been inconclusive. Although, many studies how found effects that lower aptitude subjects benefited from pictorial representations, others have found no effect (see Snow & Yalow, 1982; for a review). However, we might expect these differences to be particularly acute for young children. Pictures and natural language are by far the most commonly experienced representations for children of this age. They would have had only limited experience with mathematical representations and diagrams. This may serve to increase any aptitude by treatment interaction.

7.4.6 Strategies

Two analyses found differences between the two different strategies available on the computer (rounding and truncation). For both the categorical and continuous representations, the accuracy of rounding estimates was predicted significantly better than the accuracy of the truncation estimates. In addition, there was a three way interaction between condition, ability and strategy for prediction accuracy. The high ability group significantly improved performance in truncation, but not in rounding, whereas the low ability group improved at rounding but not in truncation. Neither of these effects had been anticipated before the analysis.

Observation of the children's behaviour provides an explanation of why truncation was predicted less accurately than rounding. The majority of children tended to underestimate rather than overestimate the inaccuracy of an estimate, especially on smaller numbers (*e.g.* 14 × 16). In addition, it is truncation which tends to produce the most inaccurate estimates, again, particularly on smaller numbers (*e.g.* the problem above solved by truncation to 10 × 10 is a massive 55% inaccurate, but by rounding to 10 × 20 is only 10% inaccurate). It was obvious by their comments to the experimenter that many children had difficulty with the concept that a procedure performed correctly could result in such an inaccurate result.

This may provide an explanation of the observed three way interaction. Higher scoring children are better at rounding at time 1. At time 2, their prediction on truncation answers has improved to the same level as rounding. This suggests that predicting rounding is easier than predicting truncation, and that the higher ability children's performance on rounding is nearly at ceiling at time one. The same explanation serves for the lower ability children's performance. These children improve significantly at rounding but not at truncation. As rounding is the easier strategy, they are able to learn how to predict this strategy first. This explanation would predict that if given more time, lower scoring children's performance on truncation would improve to be similar to that on rounding.

7.5 CONCLUSION

This experiment designed to examine: (a) children's untaught estimation performance, (b) whether using CENTS improves children's understanding of estimation and (c) how multi-representational systems that differed in similarity of format affect learning. In line with Case & Sowder's (1990) model of the development of estimation, pre-test performance was low. Children rarely knew (or invented) any appropriate estimation strategies and this was reflected in the inaccuracy of their estimates. However, with only limited teaching with CENTS (two session) children in all experimental groups improved significantly at performing estimation. They used more appropriate strategies and were more accurate.

However, on the measure most strongly related to the representations, only children in the mathematical and pictorial groups improved at predicting how accurate an estimate would be. Children in the mixed group improved at estimating without improving at understanding the relation between their estimate and the right answer. Examination of computer records found systematic differences between how the MERs were used to predict the accuracy of estimates. In particular, it was argued that as children in the mixed condition did not converge their behaviour across the different representations over time, they were unable to recognise the similarities between the representations.

Thus, this experiment found that the degree of similarity of format between representations did influence what children learned. Representations that were similar in format (pictures and mathematical representations) were associated with better learning outcomes for prediction accuracy than those that were dissimilar (mixed representations). However, it would be premature to conclude that mixed representations should be avoided in learning environments. One problem with this experiment is that due to time limitations it was not possible to fully balance the conditions. A further mixed condition consisting of the histogram and the 'splatwall' would have ruled out the possibility that these results were due to the individual representations used. It seems unlikely that the particular representations used would have been uniquely difficult to co-ordinate, but there will almost certainly be variations in this effect.

Additionally, it is possible that if children in the mixed condition had been given more time and experience, they would have shown similar improvements to those in the other conditions. It could be the case that these representations take longer to understand, but eventually performance with them will reach or surpass that of the other representations. Experiment Four was designed to address this issue.

CHAPTER EIGHT

Experiments Four and Five (CENTS)

In this chapter two further experiments with CENTS are reviewed. Each addressed the concern raised by Experiment Three in regard to how users of multi-representational software learn to integrate information from representations. This issue was explored over longer periods of time (Experiment Four) and under conditions of varying information redundancy (Experiment Five).

Experiment Four

8.1 AIMS

8.1.1 Pedagogical Aims

In Experiment Three, CENTS was shown to be effective at teaching children computational estimation. After the intervention, the experimental subjects produced more accurate estimates and used more appropriate strategies. However, children's performance was not at ceiling. Only 60% of estimates were generated by an appropriate strategy and, in addition, place value correction remained relatively poor. As discussed in section 7.4.1, the support provided by CENTS for place value correction was altered in an attempt to improve learning outcomes. Additionally, this experiment also included two extra intervention sessions. This was also expected to improve learning outcomes.

Thus, the pedagogical aim of this experiment was to see if children's improved estimation performance was replicated or even enhanced.

8.1.2 MERs Aims

Experiment Three showed that mixed representations resulted in poorer learning outcomes. It was argued in section 7.4.3, that this was because users had not been able to translate across the representations. However, both the task (computational estimation) and the learning environment were new to the children. This placed especially heavy learning and working memory demands upon children. This is consistent with Sweller's cognitive load approaches to describing learning (*e.g.* Sweller, 1988; Chandler & Sweller, 1992). Cognitive load accounts suggest that the task demands are initially very high when learners are introduced to a problem. However, with practice aspects of the task become automated which frees resources for other aspects of the task. Therefore, one possible explanation of the results from Experiment Three was that co-ordinating mixed representations was only likely to be a short-term problem. When children become more experienced with the learning environments and with estimation problems, then mixed representations may be more easily co-ordinated.

This hypothesis was tested by adding two further intervention sessions to the experiment, producing a total of four CENTS trials in all. Four sessions (about 200 minutes in total) was chosen as a likely number of sessions over which children should become familiar with the environment and task. If convergence has still not occurred by the fourth session, then it is plausible to argue that mixed representations are more than just only an initial problem. In addition, for systems such as CENTS, it is highly unlikely that in normal classroom use children would be allowed more time than this to use a computer-based learning environment aimed at a single aspect of the curriculum.

8.2 METHOD

8.2.1 Design

This experiment employed the same representations and design as Experiment Three. A two factor mixed design was used. The first factor had three levels which varied representations of accuracy of the estimates. This resulted in four groups of 12 subjects consisting of subjects who received 'picts' (target and splatwall), 'maths' (histogram and numerical) and 'mixed' (target and numerical) representations. The final group was a no-intervention control who took the pen and paper tests. A second factor, time, was within groups. A randomised block design was used and children were assigned to the different condition on the basis of their scores on a mental maths tests. Each group had similar numbers of boys and girls and the mean age of the subjects did not differ significantly.

8.2.2 Subjects

48 year five and six pupils from a state junior school took part in the experiment. They had been selected by their teachers to be the best at mathematics in their (vertically grouped) classes, (the top 10 from 30). The children ranged in age from 9:5 to 11:2 years. All the children were experienced with mouse driven computers.

8.2.3 Measures

The same measures of performance were used as for Experiment Three. Paper measures examined the accuracy of children's estimates (answer accuracy), the prediction of accuracy of the estimate (prediction accuracy) and the strategies used. Trace measures were used to examine use of the representations, both with respect to the domain (prediction accuracy) and with respect to other representations (representational co-ordination). As the children in this study were chosen from the highest performers in each class, no effects of ability were examined.

8.2.4 Materials

The materials were used were the same as those in Experiment Three. Wording of the mental maths tests was agreed with the school's maths co-ordinator.

8.2.5 Procedure

Pre-tests

Children were given mental maths tests in groups of ten. The experimenter read the items to the children and allowed children to query items they had not understood. Children were allowed a short break after each block of ten items. In total, the test took about 30 minutes to complete.

The estimation tests were given in class groups. Instruction stressed that exact answers were not required, encouraged guessing rather than leaving an answer blank and explained how to use the insight measure. Subjects were allowed to proceed at their own pace through the test and took between 15 and 45 minutes to complete it. Three parallel versions of each test were used. Children seated close together were given different versions.

Computer Intervention

Two computers were set up in an spare classroom. Each child used CENTS a total four times, each session was separated by approximately one week. The total time they spent on the computer was between 150 to 220 minutes. Slightly different versions of the computer program were used across the session. For example, for the first two sessions included an introductory question that did not require prediction. The latter versions of the program handed more responsibility and freedom to the user. For example, they allowed three digit problems to be rounded to two significant figures (*e.g.* 132 to 120). The experimenter provided support if users became confused about how to operate the system, but did not provide direct mathematical teaching.

Children were set eight questions which they had to answer by truncating and by rounding. All questions presented were generated on line, hence the problems each child was set were different. Each started with a 2 (digit) by 2 problem, gradually moved on to larger problems (2 by 3 and 3 by 2) and ended with a 3 by 3 problem. After each problem, the children filled in the on-line log book recording details of their estimates.

Post-test

Children received a parallel version of the estimation test within 7 days of their final computer period.

8.3 RESULTS

To examine the effects of the intervention, a number of [4 by 2] ANOVAs were carried out on the pre-test and post-test. The design for the analyses was 4 (control, maths, mixed, picts) by 2 (pre-test, post-test). The first factor was between groups and the second a within group repeated measure. Trace logs from the intervention sessions were analysed. The first and last session were examined when the numbers of sessions was not the focus of the analysis. For representational co-ordination measures, where changes over the intervention were of particular interest, all four sessions were examined using [3 by 4 by 2] ANOVAs. The design was 3 (maths, mixed, picts) by 4 (time 1, time 2, time 3, time 4) by 2 (rounding, truncation). The first factor was between groups and the others within groups factors.

8.3.1 Answer Accuracy

Pen and paper measures were taken to examine whether the computer intervention successfully taught children to become accurate estimators. As before, both the accuracy of their estimates (uncorrected and corrected for place value) and the appropriateness of their estimation strategies were examined.

The percentage deviation of the estimate from the exact answer was used to examine the accuracy of the estimates.

Table 8.1. Percentage deviation of estimate by format and time

	Control	Mixed	Maths	Picts
Pre-test	89.20 (14.83)	83.59 (6.81)	92.28 (24.22)	84.26 (7.33)
Post-test	85.78 (11.87)	27.79 (18.98)	20.10 (15.42)	36.74 (45.11)

As can be seen from Table 8.1, the pre-test performance of the children was poor. There was an average 87% deviation from the correct answer. At post-test, the experimental groups were much closer with an average 28% deviation. Again no analysis is performed as the data were extremely non-homogenous.

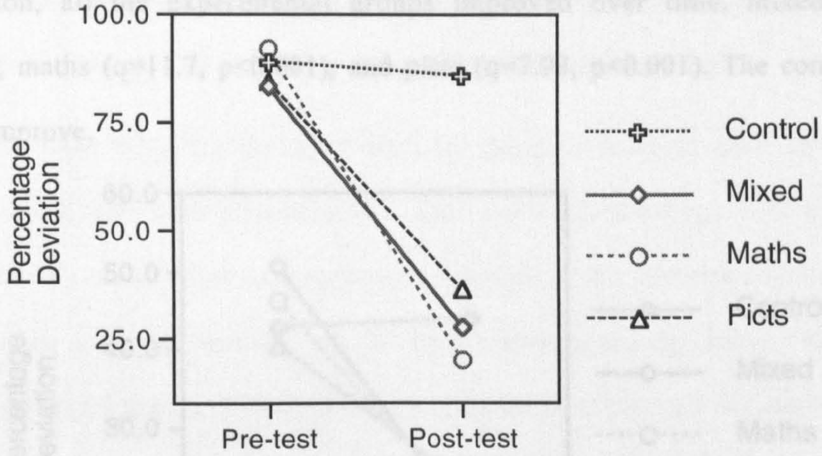


Figure 8.1 Percentage deviation by format and time

A second measure of accuracy was the corrected percentage deviation. This adjusted children's answers to the correct order of magnitude and hence distinguished between children who performed appropriate transformations, but failed at final place value correction, from those who used inappropriate strategies or simply guessed an answer.

Table 8.2. Percentage deviation of corrected estimate by format and time

	Control	Mixed	Maths	Picts
Pre-test	42.68 (4.48)	50.83 (7.11)	46.24 (8.85)	40.74 (11.69)
Post-test	43.59 (11.07)	17.16 (12.13)	18.71 (7.73)	22.01 (6.89)

Analysis by an [4 by 2] ANOVA yielded significant main effects of format ($F(1,42)=6.28$, $p<0.002$) and time ($F(1,42)=147.33$, $p<0.001$). There was a significant interaction between format and time ($F(3,42)=21.38$, $p<0.001$). Tukey tests found that the significant differences between the experimental groups and the control group was at post-test:

- mixed v control ($q=8.72$, $p<0.001$)
- maths v control ($q=8.21$, $p<0.001$)
- picts v control ($q=7.12$, $p<0.001$)

In addition, all the experimental groups improved over time, mixed (q=14.7, p<0.001), maths (q=11.7, p<0.001), and picts (q=7.98, p<0.001). The control group did not improve.

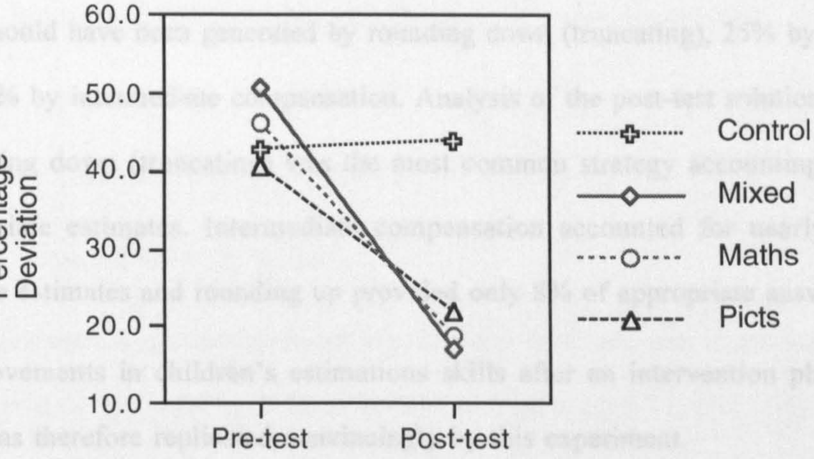


Figure 8.2 Percentage deviation (corrected for magnitude) by format and time

8.3.2 Strategic Knowledge

Subjects' estimates were coded into appropriate (rounding up, rounding down or truncation and intermediate compensation) and inappropriate strategies (as for Experiment Three). The number of estimates generated by a recognised strategy was examined (Table 8.3).

Table 8.3. Numbers of estimates generated by an appropriate strategy by format and time (from 20)

	Control	Mixed	Maths	Picts
Pre-test	0.75 (3.75)	0.25 (1.25)	0.63 (3.18)	0.09 (0.46)
Post-test	0 (0)	18.63 (1.80)	18.08 (3.85)	19.58 (0.68)

As can be seen from the Table 8.3, the pre-test performance of all the children was very low. Only 2.15% of all answers were generated by an appropriate strategy. By post-test, the experimental groups scored an average of 93.8%, effectively moving from floor to ceiling. The control group performance remained static.

If children were selecting the most accurate (taught) strategy, then 25% of their answers should have been generated by rounding down (truncating), 25% by rounding up and 50% by intermediate compensation. Analysis of the post-test solutions showed that rounding down (truncating) was the most common strategy accounting for 62% of appropriate estimates. Intermediate compensation accounted for nearly 30% of appropriate estimates and rounding up provided only 8% of appropriate answers.

The improvements in children's estimations skills after an intervention phase using CENTS was therefore replicated convincingly by this experiment.

8.3.3 Prediction Accuracy

Measures of answer accuracy allowed analysis of the improvement in the use and application of estimation strategies. However, they did not permit assessment of any insights that children may have into either the process of estimation or how an estimate differ from an exact answer. This was examined using the tick boxes which subjects filled in to indicate how far off they thought an estimate was from the exact answer. The responses were coded as the difference between the prediction and the category that should have been selected given the estimate. This was then examined using an [4 by 2] ANOVA.

Table 8.4. Difference between prediction and estimate by format and time

	Control	Mixed	Maths	Picts
Pre-test	4.81 (1.09)	4.601 (0.87)	4.53 (0.68)	3.82 (0.94)
Post-test	4.04 (0.92)	1.99 (0.79)	2.07 (1.4)	2.27 (1.1)

There were main effects of format ($F(1,42)=6,8$, $p<0.001$) and time ($F(1,42)=110.38$, $p<0.0001$) and a significant interaction between time and format

(F(3,42)=6.04, p<0.002) (Figure 8.3). Simple main effects showed no differences between the conditions and pre-test but that there were differences at post-test (F(3,84)=11.37, p<0.001). The experimental groups were significantly better at predicting the accuracy of estimates than the control group at post-test:

- mixed v control (q=6.17, p<0.001)
- maths v control (q=5.93, p<0.001)
- picts v control (q=5.33, p<0.01)

All the experimental groups improved significantly from pre-test to post-test: mixed (q=10.75, p<0.001), maths (q=9.70, p<0.001), and picts (q=6.11, p<0.001).

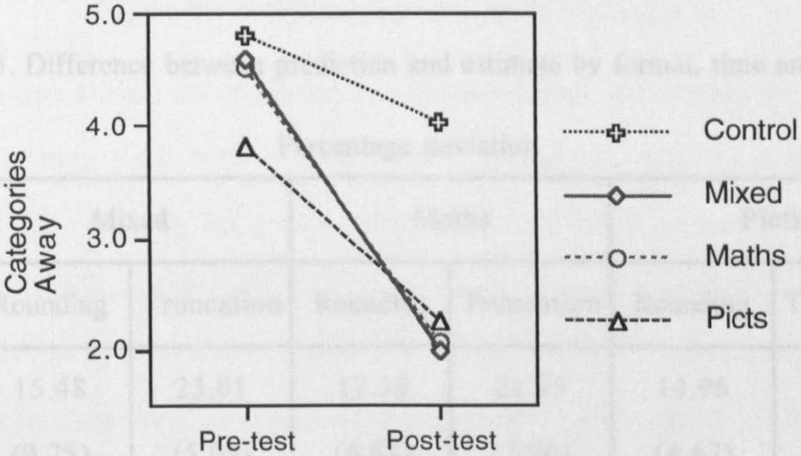


Figure 8.3 Prediction accuracy by format and time

Analysis of prediction accuracy therefore differs from Experiment Three. Here, children in all the experimental conditions improved their performance over time. Such a result was consistent with the proposal that mixed representation are only problematic for short periods of time. In order to examine more closely how the different MERs may have affected learning, the computer logs generated during the intervention session were examined.

8.3.4 Process Data

To examine how the children's performance changed with experience on CENTS and the effects of the different conditions, a number of analyses were performed. Two types of measures were examined: those that analyse how the children's understanding of the domain is reflected in their use of representations and those that measure children's understanding of how the representations relate to each other.

Continuous Prediction

This examined accuracy of prediction using the continuous representations which provide the percentage deviation of the estimate from the exact answer. The predictions were represented as numbers for mixed and maths and as a 'splatwall' for picts condition. An [3 by 2 by 2] ANOVA was conducted with the on-line data from the subjects' first and last trials with CENTS. The design was 3 (mixed, maths, pictures) by 2 (time 1, time 4) by 2 (rounding, truncation) (see Table 8.5).

Table 8.5. Difference between prediction and estimate by format, time and strategy

Percentage deviation

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	15.48 (9.75)	23.01 (5.05)	17.36 (6.68)	21.95 (7.96)	14.96 (4.67)	13.02 (4.83)
Time 4	10.08 (4.75)	11.67 (6.06)	6.57 (4.67)	9.75 (4.91)	9.92 (4.41)	9.63 (4.15)

Analysis revealed a main effect of time ($F(1,31)=44.1$, $p<0.0001$) and of strategy ($F(1,31)=5.31$, $p<0.028$); rounding solutions were predicted significantly more accurately than truncation solutions. There was also a trend towards a main effect of format ($F(2,31)=2.99$, $p<0.065$). A trend for an interaction between time and format ($F(2,31)=2.87$, $p<0.07$) was also observed (Figure 8.4). There were significant differences between the conditions after the first session on the computer ($F(2,62)=4.79$, $p<0.012$) but not after all four sessions ($F(2,62)=0.98$). At time one, the picts group were performing significantly better than the other groups: picts v maths ($q=4.10$, $p<0.05$) and picts v mixed ($q=4.02$, $p<0.05$). However by time four, the other experimental groups had improved significantly, the picts group had not improved further (mixed (q=5.89, p<0.05) and maths (q=7.55, p<0.05).

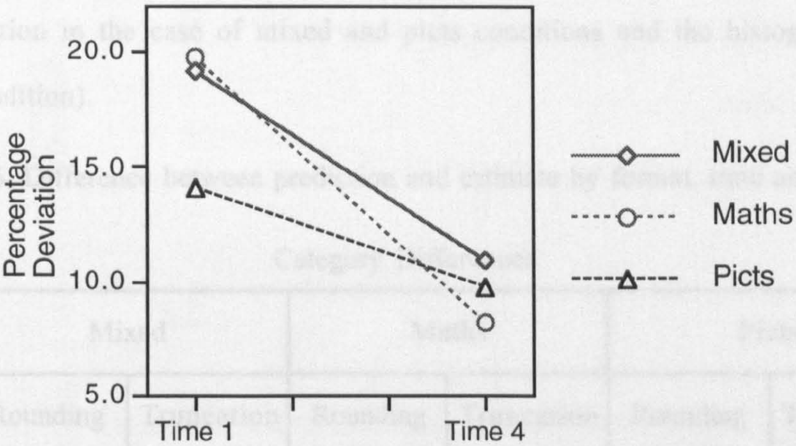


Figure 8.4 Prediction accuracy by format and time

There was also a trend for an interaction between format and strategy ($F(2,31)=2.99$, $p<0.064$) (Figure 8.5). Simple Main effects showed no differences between the conditions for rounding, but showed that there were for truncation ($F(1,31)=5.91$, $p<0.01$). The mixed and maths conditions predicted less accurately on truncation problems ($F(1,31)=7.23$, $p<0.01$), and ($F(1,31)=3.99$, $p<0.05$) respectively. However, there were no differences in prediction accuracy between truncation and rounding for the picts group.

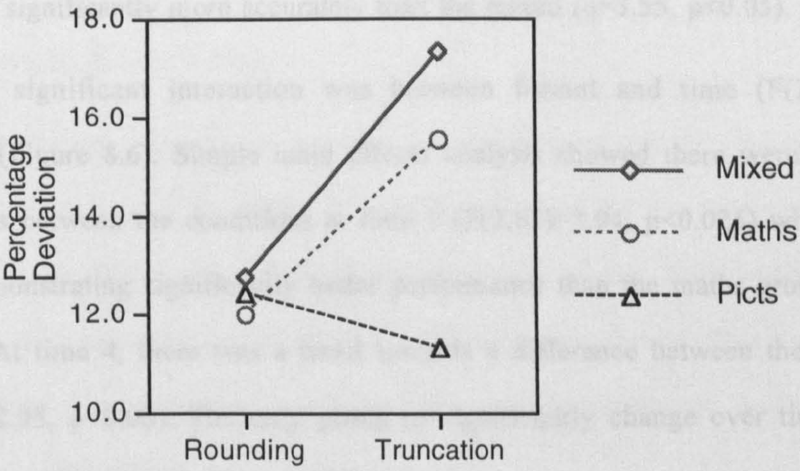


Figure 8.5 Prediction accuracy by format and strategy

Categorical Representations

This analysis was repeated for the categorical representations (the target representation in the case of mixed and picts conditions and the histogram in the maths condition).

Table 8.6. Difference between prediction and estimate by format, time and strategy

Category Differences

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	0.92 (0.31)	1.37 (0.36)	1.27 (0.48)	1.36 (0.47)	0.85 (0.28)	0.96 (0.38)
Time 4	1.01 (0.31)	1.17 (0.44)	0.58 (0.25)	1.00 (0.50)	0.75 (0.40)	0.93 (0.63)

As with the continuous representations, there were main effects of time ($F(1,31)=7.29$, $p<0.012$) and strategy ($F(1,31)=15.93$, $p<0.001$); again, rounding was predicted significantly more accurately than truncation. There was also a trend towards a main effect of format ($F(2,31)=3.27$, $p<0.051$) with the picts group predicting significantly more accurately than the mixed ($q=3.55$, $p<0.05$).

The only significant interaction was between format and time ($F(2,31)=3.65$, $p<0.038$) (Figure 8.6). Simple main effects analysis showed there were significant differences between the conditions at time 1 ($F(2,62)=3.94$, $p<0.025$) with the picts group demonstrating significantly better performance than the maths group ($q=4.01$, $p<0.05$). At time 4, there was a trend towards a difference between the conditions ($F(2,62)=2.95$, $p<0.06$). The only group to significantly change over time was the maths group ($F(1,31)=13.78$, $p<0.001$).

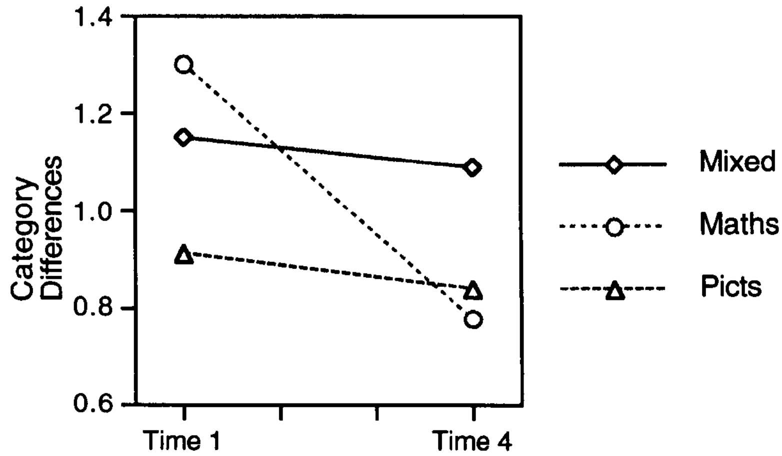


Figure 8.6 Prediction accuracy by format and time

Unlike Experiment Three, differences were found in the use of categorical and continuous representations. In this experiment, both of the maths representations were associated with poorer performance initially, but improved significantly over time. Prediction with the picts representations demonstrated a tendency for better initial performance and by Time 4 had very similar performance to the maths representations. However, prediction with the continuous mixed representation showed improvement over time, whilst the categorical representation did not. Again, the way that individual representations are used seems to be affected by the other representation presented with it.

Representational Co-ordination

These analyses were designed to examine the similarity of subject's behaviour across the two representations. As students' understanding of the representational system improves, their behaviour should become similar across both representations. Although, their understanding of the domain could remain flawed. This was examined by correlating the predictions on the two different representations. Analysis was by an [3 by 4 by 2] ANOVA.

Table 8.7. Correlations between the representations by format, time and strategy

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	0.06 (0.41)	0.27 (0.35)	0.39 (0.38)	0.36 (0.36)	0.26 (0.38)	0.28 (0.47)
Time 2	-0.01 (0.35)	0.27 (0.33)	0.35 (0.41)	0.51 (0.42)	0.44 (0.44)	0.49 (0.38)
Time 3	0.17 (0.38)	0.21 (0.42)	0.65 (0.26)	0.43 (0.53)	0.58 (0.35)	0.66 (0.29)
Time 4	0.05 (0.45)	0.16 (0.35)	0.77 (0.25)	0.66 (0.43)	0.51 (0.43)	0.81 (0.20)

There were main effects of format ($F(2,31)=9.45$, $p<0.001$), and time ($F(3,31)=5.78$, $p<0.0011$). There was a slight trend towards a main effect of strategy ($F(2,31)=3.23$, $p<0.08$). There was a significant interaction between format and time ($F(6,31)=2.27$, $p<0.043$) (Figure 8.7). Simple main effects showed significant differences between the conditions at times two, three and four ($F(2,124)=3.99$, $p<0.021$; $F(2,124)=6.32$, $p<0.0024$; $F(2,124)=13.499$, $p<0.0001$). At time four, both maths and picts were significantly more co-ordinated than the mixed group ($q=4.47$, $p<0.01$; $q=4.07$, $p<0.05$). Both the maths and picts group improved over time ($F(3,93)=4.05$, $p<0.01$, & $F(3,93)=5.76$, $p<0.002$). The mixed group showed no evidence of improved co-ordination even after four trials on the computer.

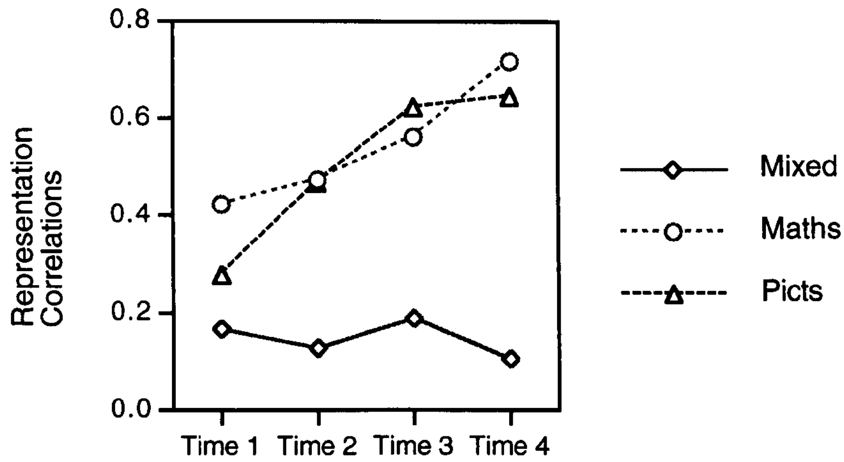


Figure 8.7 Correlation of behaviour on the representations by format and time

An alternative way to examine similarity of behaviour was to convert the continuous representation into the appropriate absolute category so that the two predictions could be compared (prediction A - prediction B).

Table 8.8. Difference in prediction between the representations by format, time and strategy

	Mixed		Maths		Picts	
	Rounding	Truncation	Rounding	Truncation	Rounding	Truncation
Time 1	1.06 (0.63)	0.98 (0.48)	0.79 (0.44)	0.94 (0.49)	0.73 (0.29)	0.83 (0.44)
Time 2	0.87 (0.42)	0.74 (0.26)	0.57 (0.34)	0.68 (0.31)	0.60 (0.31)	0.72 (0.35)
Time 3	0.78 (0.31)	0.97 (0.31)	0.39 (0.31)	0.47 (0.32)	0.48 (0.23)	0.49 (0.25)
Time 4	1.02 (0.41)	0.97 (0.31)	0.22 (0.26)	0.42 (0.37)	0.54 (0.41)	0.56 (0.25)

This analysis revealed a very similar pattern of results to the other co-ordination measure with main effects of format ($F(2,31)=10.668$, $p<0.001$), time ($F(2,31)=7.94$, $p<0.0001$) and a trend towards a main effect of strategy ($F(2,31)=3.33$, $p<0.078$).

There was a significant interaction between format and time (F(6,31)=2.28, p<0.042) (Figure 8.7). Simple main effects showed that there were differences between the conditions at time 3 and time 4 (F(2,124)=6.28, p<0.0025; F(2,124)=13.989, p<0.0001). At time 3, the maths scores were significantly lower than the mixed scores (q=3.55, p<0.05), and at time 4, both maths and picts differed from mixed (q=5,74, p<0.001; q=3.80, p<0.05). The scores in the maths group improved significantly over time (F(3,93)=8.19, p<0.0001) and there was a trend for the picts group to improve (F(3,93)=2.43, p<0.07). The mixed group scores did not change over time.

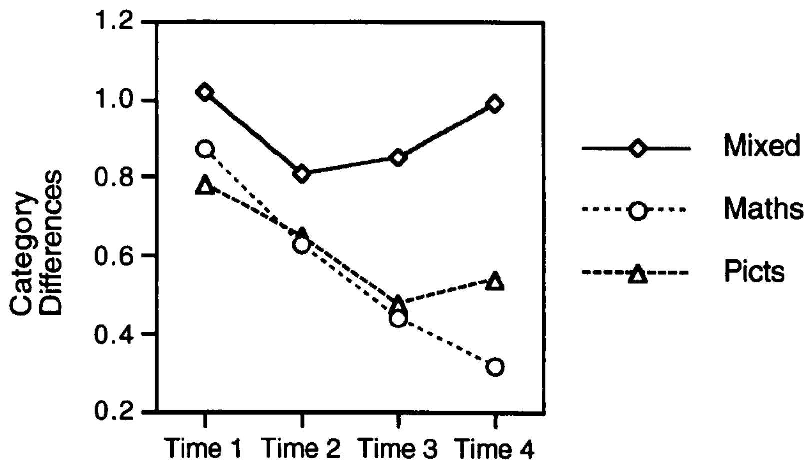


Figure 8.7 Differences between prediction behaviour on the two representations by format and time

This measure of representational co-ordination told a very similar story to that of the correlation data. Predictions by children in the mixed group on the two representations were no more similar to each other by the final session than they were at the first session.

8.4 DISCUSSION

The first goal of this experiment was to examine whether the improvement found in children's estimation skills after an intervention involving CENTS could be replicated. The second goal was to explore the effects of combining different representations over a longer period of time than that examined in Experiment Three.

8.4.1 Computational Estimation Skills

The first measure of estimation skill examined was estimation accuracy. At pre-test, children's estimates were very inaccurate. The mean average percentage deviation was 87.3% away from the correct answer. Again, children showed little ability to produce even 'ball park' estimates. This was particularly striking as school maths lessons before and during the time of the intervention were focusing on informal approaches to mental calculation of 'large' numbers.

The experimental groups demonstrated a significant increase in the accuracy of their estimates. At post-test, the average percentage deviation was 28.2%. In contrast, the control group's performance remained stable at 85.6% inaccurate.

The estimates were corrected for order of magnitude and the results re-examined. The control groups performance did not change over the study, but the three experimental groups' modified percentage deviation scores show a significant improvement from 46% to 19% after the intervention.

Minor re-implementation of CENTS between Experiments Three and Four was aimed addressing the poor order of magnitude correction identified as an area of concern. No direct comparison was performed as the children in this experiment differed along a number of dimensions from children in the original study (*e.g.* age, catchment area). However, the post-test percentage deviation measure in this study seemed to indicate that the experimental subjects were better at producing the right place value correction than previously; post-test percentage deviation was 58% in Experiment Three compared to 28% in Experiment Four. It was proposed that this represented a combination in the change of the support and the additional numbers of sessions in this experiment. The final measure of strategic knowledge of estimation examined was whether the strategy used to estimate the answer was appropriate. At pre-test, very few recognised strategies could be determined. The vast majority of answers appeared to simply be guesses. There was also almost no attempt to calculate an exact answer. A total of 2% of pre-test answers were recognised as employing an estimation strategy. At post-test, no control subject generated an answer using an appropriate strategy. However, experimental children produced an average of 94% of their answers by using a strategy that involved front-end extraction. This effectively represents a floor to ceiling improvement over the intervention.

The improvement in estimation skills found in Estimation Three were convincingly replicated by this experiment. Therefore, it was concluded that CENTS provides appropriate support to learn strategies for computational estimation problems.

8.4.2 Prediction Accuracy

CENTS requires children to predict the accuracy of an estimate to support the development of insight into how transforming numbers is related to the answer. This is performed using multiple representations.

Pen and paper measures of prediction accuracy showed that children in all experimental conditions improved at predicting the accuracy of their estimates. The control group did not improve significantly. This contrasted with Experiment Three which found that children in the mixed condition did not improve at prediction accuracy. The improvement in children's prediction accuracy in the mixed conditions was consistent with the hypothesis that mixed representations are only problematic for a short period of time when the initial task demands are great. Such an account would be predicted by cognitive load accounts of learning which would propose that with practice additional resources become available as aspects of the task become automated. The intervention logs were examined in order to assess whether this explanation was correct.

8.4.3 Representation Use

As in Experiment Three, two types of analyses were performed upon the data. Prediction accuracy on both representations was examined to identify children's developing understanding of the domain and the representations used. Representational co-ordination measures were used to assess children's knowledge of the relation between representations.

Prediction accuracy with the categorical representations showed a strikingly similar pattern of results to Experiment Three. Children in the picts condition were more accurate than children in the other conditions at Time 1. By Time 4, the maths group had significantly improved their predictions. The mixed group showed no improvement with this representation.

However, the use of the continuous representations did not match Experiment Three so exactly. Again, the picts group were identified with better initial performance and the maths group significantly improved performance over time. However, in contrast to Experiment Three the mixed representations group also improved significantly.

The picts and maths groups showed a very consistent use of representations, across both the representations in this experiment and with Experiment Three. The mixed group, however, used the representations differently. The numerical representation used as the continuous representation in both the maths and mixed conditions, was used similarly in both conditions. However, the archery target (the categorical representation in both mixed and picts cases) was used differently depending upon condition. Again, there was evidence that the way a representation was used was related to the other representation it was presented alongside.

Finally, the two measures of representational co-ordination were examined. It was argued that if mixed representations were only problematic due to initial task demands, four sessions should have provided sufficient experience for co-ordination to occur. However, if disparate representations remain harder to co-ordinate, even when children were experienced with the learning environment, then by the fourth session on the computer children still may not co-ordinate representation use.

It was shown that the maths and picts groups became significantly more converged over time. However, even after four sessions on the computer (a total of over three hours experience) the mixed group behaviour did not become more co-ordinated. The poor representational co-ordination demonstrated by children in the mixed condition in Experiment Three was therefore replicated in this experiment. Furthermore, it was shown that this occurs even when children had gained considerable experience with the representations. Thus, even when children in the mixed condition had been given considerable practice with computational estimation problems and with the representations, they still failed to co-ordinate their representation use. This suggests that failure to co-ordinate representations is not solely due to the heavy demands of initial learning.

The analysis of both types of representation use (with respect to the domain and to each other) provide an explanation of the learning outcome measures for the mixed condition. Unlike children in the maths and picts conditions, children in the mixed condition did not learn to translate across the representations. This led them to abandon their attempts to work with one of the representations (categorical) and to concentrate on the other representations (continuous). The second representation contains both the direction and the magnitude information and so provides all the information in the first representation plus more. Therefore, these children appeared to have made a highly sensible decision. It was not proposed that children would find it impossible to learn about proximity from one well chosen representation, so the mixed group improvement was perfectly consistent with their use of the representations.

8.5 CONCLUSION

This study replicated the finding that CENTS could be used to teach children aspects of computational estimation. The improvement in using appropriate estimation strategies went from floor to ceiling and the other measures of estimation performance also showed considerable improvement. Therefore, CENTS is proposed to have met its pedagogical objectives.

The second aim of the experiment was to explore the effect of mixing representations when children were given much longer to use the learning environment. The results of Experiments Four replicated the finding that mixed representations were considerably more difficult to co-ordinate than either pictorial and mathematical representations by themselves. However, this did not result in poorer performance overall as it did in Experiments Three. Unlike Experiment Three, where neither of the representations were used successfully by children in the mixed condition, in this experiment they appeared to concentrate upon a single representation and learnt to use it effectively. Thus, as the continuous representation contains all the necessary information, then children could learn to understand prediction accuracy without referring to the other representation.

In this configuration of the system (partially redundant representations), there is no particular necessity to map across representations. However, if learners concentrate upon one representation and ignore both a second representation or the translation across representations, many of the proposed benefits of multiple representations will not occur (see section 3.3). Consequently, MERs could not be used to constrain interpretations or to support abstraction.

In addition, one of the most common justifications for the use of MERs is that one representation may be insufficient to display all the needed information. Each representation in a MERs may be used to convey a different aspect of a domain. Thus, in this case there is no (informational) redundancy between representations.

This issue of redundancy between MERs was explored in Experiment Five. CENTS allow designers three levels of redundancy of representation - full, partial and none. Experiments Three and Four addressed combinations of partial redundant representations. Experiment Five was designed to explore whether learners can integrate information from representations when they were either fully redundant or when they shared no information in common.

Experiment Five

8.6 AIMS

8.6.1 Pedagogical Aims

This experiment had no new pedagogical aims. CENTS had been shown in two previous experiments to successfully teach aspects of computational estimation. It was predicted that this should be replicated by Experiment Five.

8.6.2 MERs Aims

The MERs used in Experiments Three and Four have some overlap as each conveyed information about the magnitude of an estimate. However, the resolution of magnitude information was different. The categorical representations displayed information to the nearest 10% whilst the continuous representations provided detail to the nearest 1%. The continuous representation also expressed the direction of the estimate (under or over-estimate). The continuous representation contained all the information present in the categorical representation plus the additional direction information. Consequently, the continuous representation contained all the information that was required to help learners refine their understanding.

However, MERs are often employed when a single representation can not display all of the required information for a domain. Thus, each representation in the multi-representational system may convey a different part of the concept. For example, the MoLE learning environments for model logic, Oliver & O'Shea (1996) (described in section 3.3) presents users with two different representations - one of the worlds and one of the relation between worlds. In this situation, it is assumed that learners will be able to integrate information from all the MERs. However, the previous experiments had shown that in certain circumstances, children were unable to translate across representations. This resulted either in impoverished learning outcomes (Experiment Three) or on concentration upon a single representation (Experiment Four). These effects may have different consequences for learning depending upon how much unique information is conveyed by each representation.

Experiment Five therefore employed two different levels of redundancy. A fully redundant system allows the same information to be derived from both representations. In this case, both representations expressed direction and magnitude in a continuous fashion. Where there is no redundancy across representations, completely different information is derivable from each representation. Consequently, one representation was used to convey direction information and the other (continuous) magnitude.

It was proposed that when it was difficult to integrate information between representations (*i.e.* mixed representations), then providing no redundancy MERs will result in poorer learning outcomes. When mixed representations are fully redundant, even if translation across representations does not occur, learning outcomes should be unaffected as each representation is sufficient to develop the required understanding. In addition, it was also expected that full redundancy would aid translation. It should be easier to see how two representations relate to each other if both convey exactly the same information.

The mixed representations were contrasted with pictorial representations. The previous experiments had found that providing two pictures facilitated co-ordination. Hence, for both full and no redundant representations, it should be easier for learners to integrate information from pictorial representation.

The last issue addressed was whether there would be any differences in how well prediction accuracy would be learnt across the different levels of redundancy. When each representation expresses a different aspect of the situation, initial learning may be facilitated as each dimension could be considered separately. Ultimately, learners will need to integrate these dimensions to build a full understanding of the domain.

8.7 METHOD

8.7.1 Design

A three factor mixed design was used which varied both the format of representations (either picts or mixed) and the redundancy across the representations (either full or none). This resulted in four groups of twelve subjects who received one of 'picts - full' (splatwall and archery field see Figure 8.8), 'mixed - full' (splatwall and numerical display, Figure 8.9), 'picts - none' (hoops and marbles Figure 8.10) and 'mixed - none' (hoops and numerical display, see Figure 8.11) (see section 6.4 for fuller descriptions of the representations). A final group was a no-intervention control who took the pen and paper tests. A third factor, time, was within groups. A randomised block design was used and children were assigned to the different condition on the basis of their scores on a mental maths tests. Each group had similar numbers of boys and girls and the mean age of the subjects did not differ significantly.

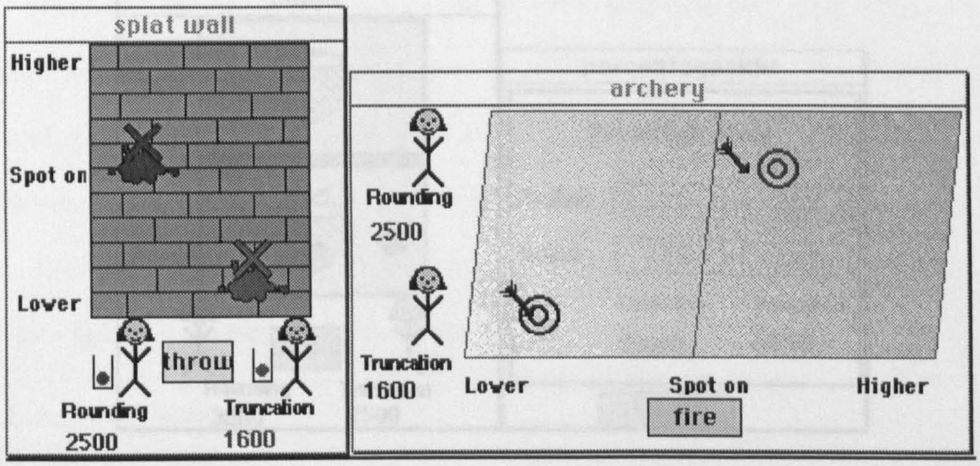


Figure 8.8 Picts - full: splatwall and archery field

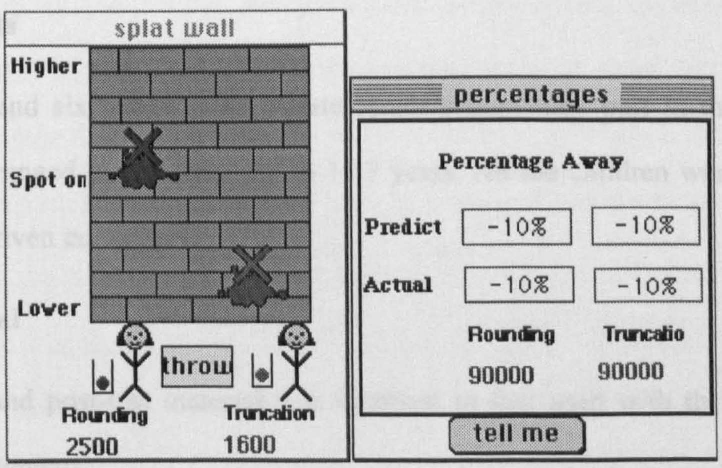


Figure 8.9 Mixed - full: splatwall and numerical display

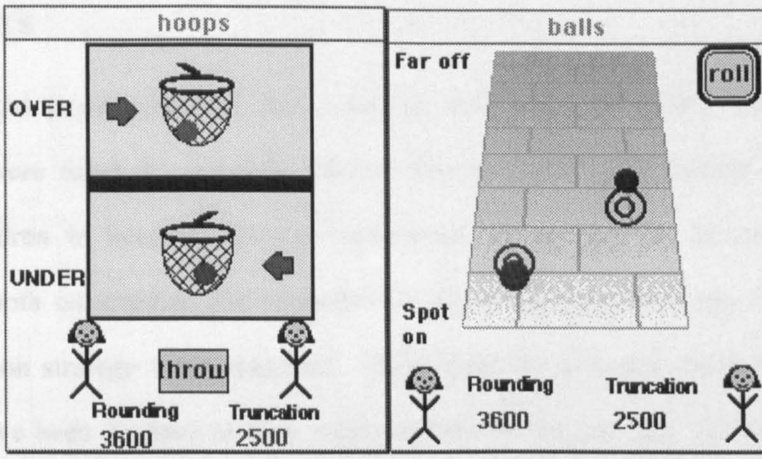


Figure 8.10 Picts - none: hoops and marbles

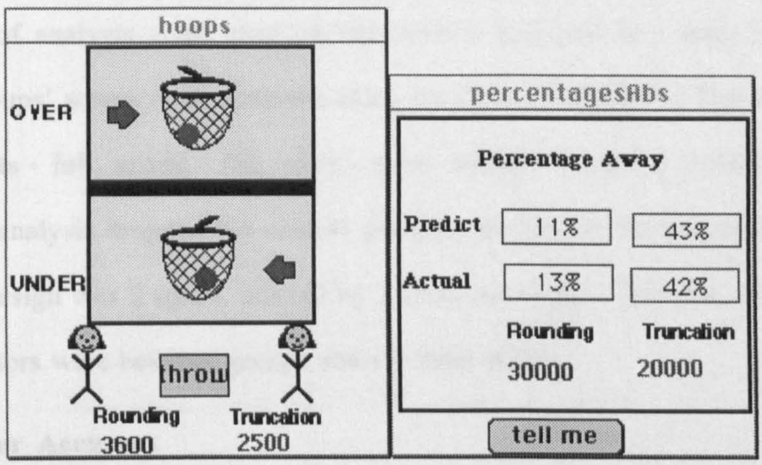


Figure 8.11 Mixed - none: hoops and numerical display

8.7.2 Subjects

60 year five and six pupils from a state junior school took part in the experiment. The children ranged in age from 9:9 to 11:7 years. All the children were experienced with mouse driven computers.

8.7.3 Material

The pre-test and post-test material was identical to that used with the two previous CENTS experiments

8.7.4 Procedure

The procedure was identical to that used for Experiment Four, except that children received only two computer interventions sessions.

8.8 RESULTS

Both pen and paper measures and computer logs were examined. Pen and paper measures were taken to examine whether the computer intervention successfully taught children to become accurate estimators. As before, the accuracy of their estimates (both uncorrected and corrected for place value) and the appropriateness of the estimation strategy were examined. Throughout the analyses, the scores for three children have been dropped as they were unavailable for one part of the intervention or testing phase (one each from the control, picts-full and mixed-full groups).

Two types of analysis were used on the pre-test and post-test data. First, all the different groups' scores were analysed using an [5 by 2] ANOVA. The design was 5 (control, picts - full, mixed - full, picts - none, mixed - none) by 2 (time 1, time 2). Subsequent analysis dropped the control group to analyse by format, redundancy and time. This design was 2 (picts, mixed) by 2 (full, none) by 2 (pre-test, post-test). The first two factors were between groups and the third within.

8.8.1 Answer Accuracy

The percentage deviation of the estimate from the exact answer was used to compare estimates given by the control group to those given by the experimental groups. Three additional extreme outlying subjects were removed from this analysis (the average percentage deviation for one of these subjects was 3696343.73% !)

Table 8.8. Percentage deviation of estimate by condition and time (collapsed across experimental groups)

	Control group	Experimental groups
Pre-test	92.21 (46.25)	89.03 (33.32)
Post-test	92.42 (24.04)	39.22 (39.28)

There were main effects of condition ($F(1,50)=9.16$, $p<0.004$), and time ($F(1,50)=12.69$, $p<0.001$). The interaction between time and condition was also significant ($F(1,50)=12.91$, $p<0.001$). As can be seen from Table 8.8, the experimental groups improved significantly ($F(1,50)=60.53$, $p<0.001$), the control groups performance did not improve. Further analysis was conducted upon the experimental groups alone (Table 8.9).

Table 8.9. Percentage deviation of estimate by format, redundancy and time

	Mixed		Picts	
	Full	None	Full	None
Pre-test	93.76 (25.57)	86.10 (27.45)	87.93 (26.82)	87.46 (17.26)
Post-test	51.26 (56.87)	33.45 (27.30)	30.52 (25.63)	39.31 (37.54)

The only significant effect was a that of time ($F(1,38)=87.81$, $p<0.001$). All experimental groups improved equally after the intervention.

These analyses were repeated for the accuracy of the estimation after it had been corrected for order of magnitude.

Table 8.10. Percentage deviation of estimate by condition and time (corrected for order of magnitude) (collapsed across experimental groups)

	Control group	Experimental groups
Pre-test	40.42 (8.53)	44.34 (8.38)
Post-test	40.40 (10.31)	20.70 (11.69)

As can be seen from Table 8.10, the experimental groups and the control group behaved very differently. When an [5 by 2] ANOVA was applied to the data, there were main effects of condition ($F(1,4)=4.42$, $p<0.004$), time ($F(1,4)=109.6$, $p<0.001$) and an interaction between time and condition ($F(1,4)=7.43$, $p<0.001$). Simple Main effects analysis demonstrated that there were no significant differences between the conditions at pre-test ($F(4,104)=1.41$), but that there were at post-test ($F(4,104)=9.92$, $p<0.001$). The only group that failed to improve their scores were the control group, ($F(1,52)=0.003$). Subsequently, further analysis was performed upon the experimental groups alone.

Table 8.11. Percentage deviation of estimate by format, redundancy and time (corrected for order of magnitude)

	Mixed		Picts	
	Full	None	Full	None
Pre-test	42.59 (11.77)	40.23 (11.82)	46.21 (6.75)	48.34 (9.23)
Post-test	21.72 (11.22)	16.58 (9.35)	17.46 (8.96)	26.90 (14.58)

There was a main effect of time ($F(1,42)=121.49$, $p<0.001$), and a trend for a main effect of format ($F(1,42)=3.62$, $p<0.065$). The only significant interaction was for format and redundancy ($F(1,42)=4.21$, $p<0.05$). Simple Main effects analysis revealed that the only difference between the levels of format was for no redundancy ($F(1,42)=8.18$, $p<0.007$).

Analysis of the individual children's performance provided an explanation of this anomalous event. Altogether seven children demonstrably failed to learn how to estimate. At post-test, they produced less than 30% of their answers using an appropriate estimation strategy. These children were removed from the analysis (one from mixed-full, one from mixed-none, one from picts-full and four from picts-none), and the results re-analysed (Table 8.12). The only significant effect for this modified data was one of time ($F(1,35)=185.45$, $p<0.001$).

Table 8.12. Percentage deviation of estimate by format, redundancy and time (without non-learners)

	Mixed		Picts	
	Full	None	Full	None
Pre-test	41.73 (12.04)	38.61 (10.89)	46.62 (6.93)	47.22 (5.56)
Post-test	19.25 (8.10)	14.86 (7.54)	15.85 (7.70)	18.99 (8.10)

8.8.2 Strategic Knowledge

As before, subjects answers were coded to see if they were produced by an appropriate strategy (*i.e.* one which involved front-end extraction). To this end, the subjects' estimates were coded into appropriate (rounding up, rounding down or truncation and intermediate compensation) and unrecognised strategies. The number of estimates generated by a recognised strategy was examined (Table 8.13).

Table 8.13. Numbers of estimates generated by an appropriate strategy by condition and time (from 20) (collapsed across experimental groups)

	Control group	Experimental groups
Pre-test	0.09 (0.30)	0.17 (0.49)
Post-test	3.64 (6.23)	16.44 (6.34)

Initial analysis compared the control group to the experimental groups. An [5 by 1] ANOVA on the post-test data found a main effect of condition ($F(4,52)=10.17$, $p<0.001$). Tukey tests showed that all the experimental groups differed from the control group. There were no differences between the experimental groups.

- mixed-full v control ($q=7.0$, $p<0.001$)
- mixed-none v control ($q=7.41$, $p<0.001$)

- picts-full v control ($q=7.63$, $p<0.001$)
- picts-none v control ($q=5.41$, $p<0.01$)

Further analysis on the experimental groups alone confirmed that there were no significant differences between the conditions.

Table 8.14. Numbers of estimates generated by an appropriate strategy by format, redundancy and time (from 20)

	Mixed		Picts	
	Full	None	Full	None
Pre-test	0.0 (0.0)	0.17 (0.38)	0.27 (0.64)	0.25 (0.63)
Post-test	16.72 (6.25)	17.50 (5.6)	17.91 (4.39)	13.75 (8.25)

There were no main effects of either format or redundancy and no significant interactions between any of the variables (Table 8.14).

As noted in the previous experiments, the tests were constructed such that in 25% of the cases rounding down (truncating) is the most accurate strategy, 25% rounding up and the remaining 50% intermediate compensation. Analysis of the types of solution (at post-test) confirmed that rounding down (truncating) was the most common strategy accounting for 55% of appropriate estimates. This indicated that some answers were generated by truncating as opposed to rounding down. Inspection of individual children's results suggested that some children invariably truncated. Intermediate compensation accounted for nearly 35% of appropriate estimates and as with the previous experiments rounding up was rare, providing only 10% of answers

8.8.3 Prediction Accuracy

The prediction accuracy data was also analysed. This was determined by calculating the difference between the prediction given by the subject using the tick boxes and the one they should have predicted.

Table 8.15. Prediction accuracy by condition and time (collapsed across experimental groups)

	Control group	Experimental groups
Pre-test	3.42 (0.72)	3.55 (0.97)
Post-test	3.48 (0.78)	2.21 1.15)

A [5 by 2] ANOVA showed a main effect of time ($F(1,52)=42.14$, $p<0.001$), and an interaction between time and condition ($F(4,52)=3.29$, $p<0.02$). Simple Main effects analysis showed no differences between the conditions at pre-test ($F(4,104)=1.20$) but showed differences at post-test ($F(4,104)=3.74$, $p<0.01$). Again, the control group was the only condition where scores did not improve significantly ($F(1,52)=0.01$)

Table 8.16. Prediction accuracy by format, redundancy and time

	Mixed		Picts	
	Full	None	Full	None
Pre-test	3.31 (0.63)	3.39 (1.19)	4.12 (0.81)	3.40 (1.02)
Post-test	2.35 (1.14)	2.05 (0.97)	2.47 (1.62)	2.03 (0.86)

Subsequent analysis of the experimental groups showed no differences between the conditions. Analysis by [2 by 2 by 2] ANOVA identified a main effect of time ($F(1,42)=45.73$, $p<0.001$) but yielded no main effects of either format or redundancy and no interaction between these variables.

In this experiment, the redundancy factor manipulated whether the subjects' predicted magnitude and direction separately or together. Hence, the prediction of these measures was examined separately. Correctly predicted direction was scored one point each, resulting in a maximum possible value of 20.

Table 8.17. Direction accuracy by format, redundancy and time (out of 20)

	Mixed		Picts	
	Full	None	Full	None
Pre-test	9.00 (3.28)	8.92 (4.56)	8.45 (3.26)	10.08 (3.58)
Post-test	9.91 (7.34)	10.17 (5.41)	7.45 (4.54)	11.75 (6.19)

Analysis using a [2 by 2 by 2] ANOVA found no main effects or interaction. Hence according to this measure, children did not increase their skills at predicting the direction of an estimate after the intervention (Table 8.17).

The magnitude of the prediction accuracy without direction was also examined. Again, this was scored by examining the difference between the category chosen and the one that should have been selected, ignoring direction. For example, a prediction of "very much lower" and "very much higher" were coded with the same score.

Table 8.18. Magnitude accuracy by format, redundancy and time (category differences)

	Mixed		Picts	
	Full	None	Full	None
Pre-test	1.74 (0.56)	1.63 (0.70)	1.44 (0.34)	1.66 (0.58)
Post-test	1.48 (0.57)	1.15 (0.58)	1.06 (0.58)	1.36 (0.60)

Analysis revealed a single significant effect, that of time ($F(1,42)=10.74$, $p<0.002$), (Table 11). Hence, the improvements seen in the general prediction measure must be primarily reflecting improvement in magnitude prediction. There was a trend for an interaction between format and redundancy ($F(1,42)=3.56$, $p<0.066$). Simple Main effects identified a trend for a difference between the levels of format for the full redundancy condition, ($F(1,42)=3.46$, $p<0.056$). There was no interaction with time.

8.8.4 Process Measures

These analyses examined how the children used CENTS' representations. As before, two types of analyses were performed. First, the domain and representation knowledge as expressed through representation use (prediction accuracy) was considered. Secondly, the similarity of behaviour across the two representations (representational co-ordination) was analysed.

Prediction Accuracy

The analysis of prediction accuracy was more complicated in this experiment since redundancy as well as format was manipulated. Each level of redundancy must first be analysed separately. The no redundancy presentation provided a separate representation for direction and magnitude, so each these dimensions must be examined separately. Direction was measured by scoring a 1 when the subject correctly identifies the direction of prediction. Hence, there was a maximum score of 9 per session (Table 8.19) Magnitude was scored by percentage deviation without direction (Table 8.20). These data were examined using two [2 by 2 by 2] ANOVAs. The design for the analysis was 2 (mixed, picts) by 2 (time 1, time 2) by 2 (rounding, truncation). The first factor was between groups, the others within.

No redundancy

The first analysis concerned the direction scores for the no -redundancy representations.

Table 8.19. Direction accuracy for no redundancy representations by format, time and strategy

	Mixed (hoops)		Picts (hoops)	
	Rounding	Truncation	Rounding	Truncation
Time 1	7.58 (0.99)	7.75 (0.75)	7.58 (1.16)	7.92 (1.24)
Time 2	7.83 (0.72)	8.83 (0.39)	7.42 (0.90)	8.25 (1.21)

There was a main effect of strategy ($F(1,22)=12.12$, $p<0.003$); truncation was predicted more accurately than rounding. There was also an interaction between time and strategy ($F(1,22)=4.82$, $p<0.04$) (Figure 8.12). Simple Main effects indicated no significant differences between the conditions at time 1 ($F(1,44)=0.75$), though there were at time 2 ($F(1,44)=10.08$, $p<0.001$). Subjects did not improve over the two sessions at predicting rounding ($F(1,44)=0.24$) but did improve at predicting truncation ($F(1,44)=6.87$, $p<0.02$).

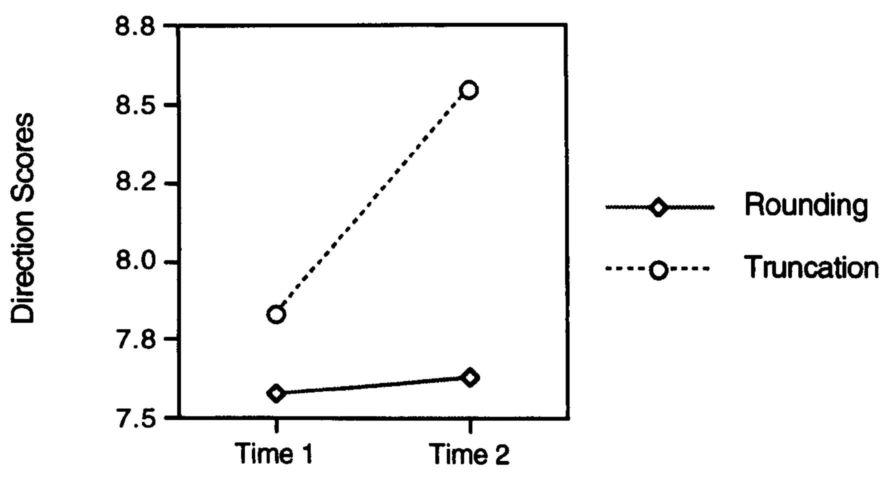


Figure 8.12 Direction accuracy for no-redundancy representations by time and strategy.

The analysis for the magnitude representations revealed a different pattern of results (Table 8.20).

Table 8.20. Magnitude accuracy for no redundancy representations by format, time and strategy

	Mixed (numbers)		Picts (marbles)	
	Rounding	Truncation	Rounding	Truncation
Time 1	8.48 (3.71)	15.98 (3.92)	10.92 (4.33)	14.72 (5.19)
Time 2	7.26 (2.92)	10.98 (6.53)	8.14 (3.18)	11.83 (3.57)

Both time ($F(1,22)=12.75$, $p<0.002$) and strategy ($F(1,22)=54.23$, $p<0.001$) yielded significant main effects. In contrast to the direction measure, however, it was rounding that was predicted more accurately. There were no interactions.

Full redundancy

The fully redundant representations, (*i.e.* the splatwall, and archery target for the picts format and the splatwall and numerical representations in the mixed condition) provide information about the percentage deviation of the estimate from the exact answer. They combine direction and magnitude information.

These data were examined using an [2 by 2 by 2 by 2] ANOVA. The design for the analysis was 2 (mixed, picts) by 2 (time 1, time 2) by 2 (rounding, truncation) by 2 (representation 1 - splatwall, representation 2 - archery target or numerical display). The first factor was between groups, the others within.

Table 8.21. Percentage accuracy for fully redundant representations by format, time, strategy and representation

Rounding

	Mixed		Picts	
	Splatwall	Percentages	Splatwall	Archery
Time 1	12.40 (2.39)	16.34 (7.72)	11.74 (2.15)	12.48 (3.15)
Time 2	9.20 (3.84)	8.06 (4.86)	9.06 (3.78)	9.97 (4.94)

Truncation

	Mixed		Picts	
	Splatwall	Percentages	Splatwall	Archery
Time 1	20.51 (5.82)	20.87 (7.20)	15.84 (4.48)	17.10 (3.83)
Time 2	13.63 (3.22)	14.64 (7.53)	13.88 (3.75)	14.51 (3.92)

There were main effects of time ($F(1,20)=34.43$, $p<0.001$) and strategy ($F(1,20)=50.19$, $p<0.001$); rounding was predicted more accurately than truncation. There was also an interaction between format and time ($F(1,20)=6.44$, $p<0.02$) (Figure 8.13). Simple Main effects showed that at time 1, there were differences between the conditions ($F(1,40)=4.13$, $p<0.05$), but found none at time 2 ($F(1,40)=0.09$). This confirms the earlier studies that have found an initial advantage for pictorial representations. There was also a noticeable trend for a 4 way interaction between condition, time, strategy and representation ($F(1,20)=3.98$, $p<0.06$) (see Table 8.23).

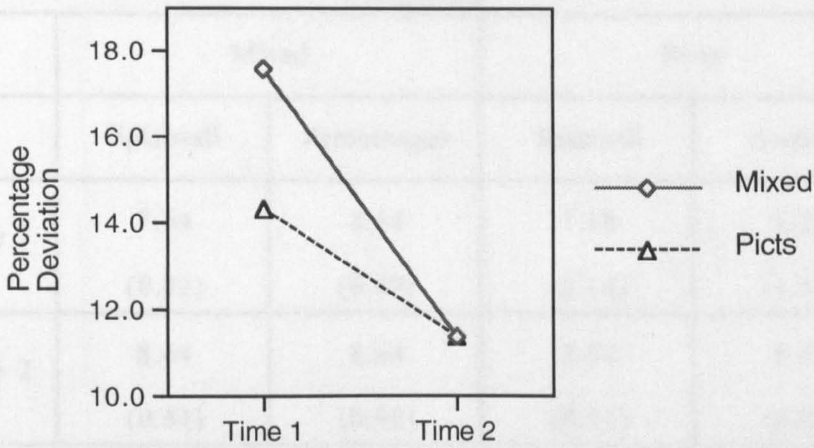


Figure 8.13 Percentage deviation of no-redundancy representations by format and time

In order to compare the no redundancy and the full redundancy conditions more directly, the performance on each of the representations in the full redundancy conditions can be recoded to give separate direction and magnitude information. This was then analysed using two [2 by 2 by 2 by 2] ANOVAs on format, time, strategy and representation.

Table 8.22. Direction accuracy for fully redundant representations by format, time and representation

Rounding

	Mixed		Picts	
	Splatwall	Percentages	Splatwall	Archery
Time 1	6.45 (0.93)	6.9 (1.13)	6.18 (1.08)	5.45 (1.86)
Time 2	7.18 (1.77)	7.73 (1.79)	7.74 (1.10)	7.73 (1.27)

Truncation				
	Mixed		Picts	
	Splatwall	Percentages	Splatwall	Archery
Time 1	7.54 (0.82)	8.54 (0.69)	7.18 (2.14)	7.18 (1.47)
Time 2	8.64 (0.51)	8.64 (0.92)	8.64 (0.51)	8.55 (0.82)

Analysis revealed main effects of time ($F(1,20)=26.02$, $p<0.001$) and strategy ($F(1,20)=41.48$, $p<0.001$); truncation was predicted more accurately than rounding. There was an interaction between format and time ($F(1,20)=4.53$, $p<0.05$) (Figure 8.14).

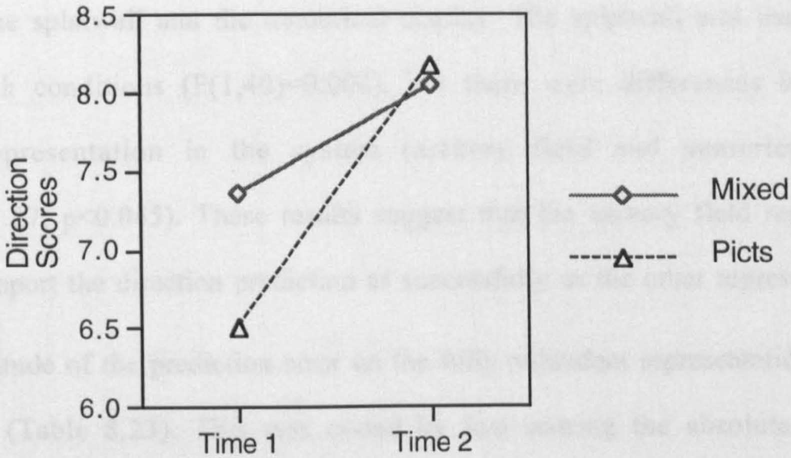


Figure 8.14 Direction scores by format and time

Simple Main effects identified differences between the levels of format at time 1 ($F(1,40)=4.84$, $p<0.04$), but no differences at time 2 ($F(1,40)=0.084$). Initially, children in the picts condition predicted less well than mixed, but both levels of formats improved over time; mixed ($F(1,20)=4.14$, $p<0.049$) and picts ($F(1,20)=26.14$, $p<0.001$). There was also an interaction between format and representation ($F(1,20)=5.45$, $p<0.03$). (see figure 8.15)

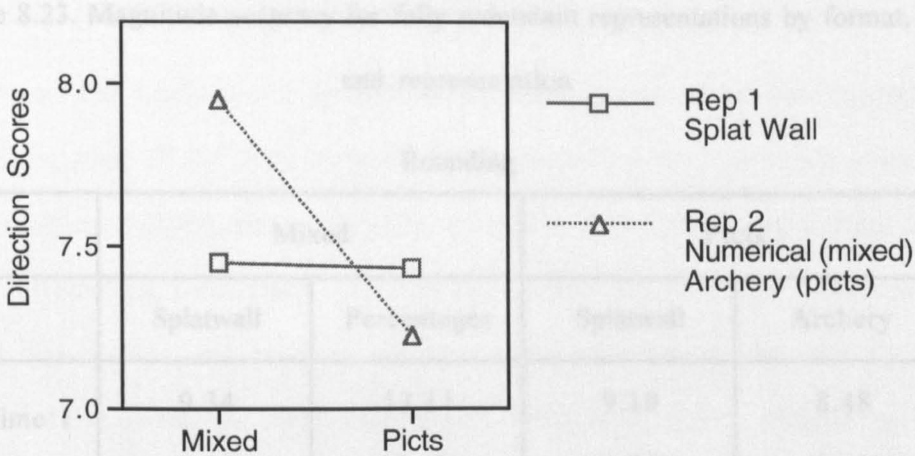


Figure 8.15 Direction scores by format and representation

Simple Main effects demonstrated that for the picts condition there was no difference between the representations (splatwall) ($F(1,20)=0.93$)). There was a difference between the representations for the mixed condition, ($F(1,20)=5.58$, $p<0.03$), *i.e.* between the splatwall and the numerical display. The splatwall was used similarly across both conditions ($F(1,40)=0.004$), but there were differences between the second representation in the system (archery field and numerical display) ($F(1,40)=4.27$, $p<0.045$). These results suggest that the archery field representation did not support the direction prediction as successfully as the other representations.

The magnitude of the prediction error on the fully redundant representations was also examined (Table 8.23). This was coded by just scoring the absolute percentage deviation of the estimate. For example +25% and -25% would be coded as the same.

Table 8.23. Magnitude accuracy for fully redundant representations by format, time and representation

Rounding

	Mixed		Picts	
	Splatwall	Percentages	Splatwall	Archery
Time 1	9.34 (2.95)	13.11 (6.45)	9.10 (1.86)	8.48 (2.41)
Time 2	7.73 (3.33)	7.86 (4.61)	7.12 (3.00)	7.23 (3.69)

Truncation

	Mixed		Picts	
	Splatwall	Percentages	Splatwall	Archery
Time 1	16.55 (5.54)1	17.25 (4.77)	13.61 (2.68)	14.31 (3.28)
Time 2	12.68 (2.88)	12.97 (4.42)	12.89 (3.52)	11.85 (4.08)

There were main effects of time ($F(1,20)=14.28$, $p<0.002$) and strategy ($F(1,20)=68.82$, $p<0.001$); rounding was predicted more accurately than truncation. There was also a four way interaction between format, time, strategy and representation. This was examined further by splitting the data by strategy which indicated a three way interaction between format, time and representation for rounding ($F(1,20)=5.20$, $p<0.040$) but not for truncation ($F(1,20)=0.33$). Simple simple main effects analysis showed that the only representation on which prediction improved significantly over time was the mixed, numerical representation ($F,(1,20)=13.78$, $p<0.002$). This appears to be due to its initial disadvantage.

Full and no redundancy representations

To compare the direction or magnitude of the prediction across all four experimental conditions, each of the fully redundant representations was examined separately. These analyses took the form of 2 (picts, mixed) by 2 (full, none) by 2 (time 1, time 2) by 2 (rounding, truncation). The first two factors were between groups and the second two within. This analysis will discuss the direction dimension for the two different fully redundant representations (Tables 8.24, 8.25). Magnitude is not reported as further analysis simply confirmed the main effects of time and strategy reported above.

Table 8.24. Direction accuracy for representation 1(splatwall) by format, redundancy and time Rounding

	Mixed		Picts	
	Full (Splatwall)	None (Hoops)	Full (Splatwall)	None (Hoops)
Time 1	6.45 (0.93)	7.58 (0.96)	6.18 (1.08)	7.58 (1.16)
Time 2	7.18 (1.78)	7.83 (0.72)	7.73 (1.10)	7.42 (0.90)

Truncation

	Mixed		Picts	
	Full (Splatwall)	None (Hoops)	Full (Splatwall)	None (Hoops)
Time 1	7.55 (0.82)	7.75 (0.75)	7.18 (2.14)	7.92 (1.24)
Time 2	8.64 (0.51)	8.83 (0.39)	8.64 (0.51)	8.25 (1.21)

As before, there were main effects of time (F(1,42)=33.26, p<0.001), and strategy (F(1,42)=20.92, p<0.001). There was also a main effect of redundancy (F(1,42)=5.62, p<0.022). Prediction accuracy was higher for the no redundancy representations

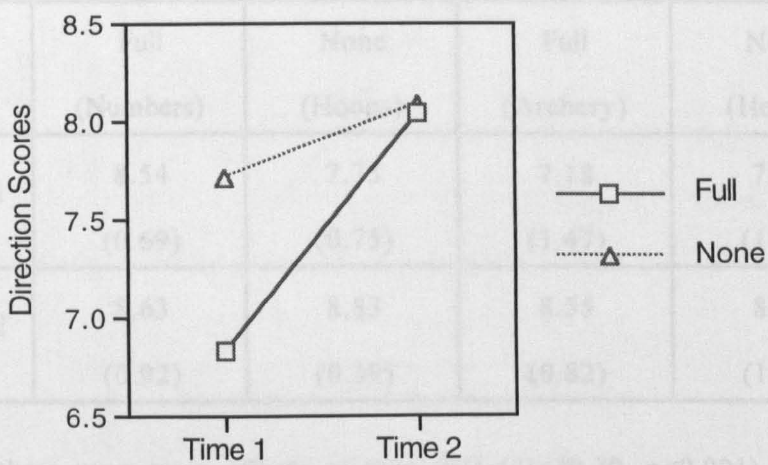


Figure 8.16 Direction scores by redundancy and time

The interaction between time and redundancy (F(1,42)=5.77, p<0.021) also proved to be significant (figure 8.16). Simple Main effects found that the only difference between the levels of redundancy was at time 1 (F(1,84)=11.35, p<0.002). The full redundancy group were the only ones to improve over time (F(1,84)=23.32 p<0.001).

This analysis was then repeated to compare the second representation in fully redundant conditions with the representations in the no-redundancy representations.

Table 8.25. Direction for representation 2 (numbers, archery) by format, redundancy and time

Rounding

	Mixed		Picts	
	Full (Numbers)	None (Hoops)	Full (Archery)	None (Hoops)
Time 1	6.91 (1.14)	7.58 (0.96)	5.45 (1.86)	7.58 (1.16)
Time 2	7.73 (1.79)	7.83 (0.72)	7.72 (1.27)	7.42 (0.90)

Truncation

	Mixed		Picts	
	Full (Numbers)	None (Hoops)	Full (Archery)	None (Hoops)
Time 1	8.54 (0.69)	7.75 (0.75)	7.18 (1.47)	7.92 (1.24)
Time 2	8.63 (0.92)	8.83 (0.39)	8.55 (0.82)	8.25 (1.21)

As above there were main effects of time ($F(1,42)=20.79$, $p<0.001$) and strategy ($F(1,42)=48.56$, $p<0.001$). In contrast to the analysis for representation 1, however, there was no main effect of redundancy, but a main effect of format ($F(1,42)=4.46$, $p<0.041$). The mixed representation was predicted more accurately than the picts. This appears to be due to the archery field representation (discussed above).

There was an interaction between redundancy and time ($F(1,42)=5.28$, $p<0.027$). Simple Main effects showed that the levels of redundancy differed at time 1 ($F(1,84)=6.15$, $p<0.015$), and that the full redundancy group improved over time ($F(1,42)=22.53$, $p<0.001$). (see Figure 8.17)

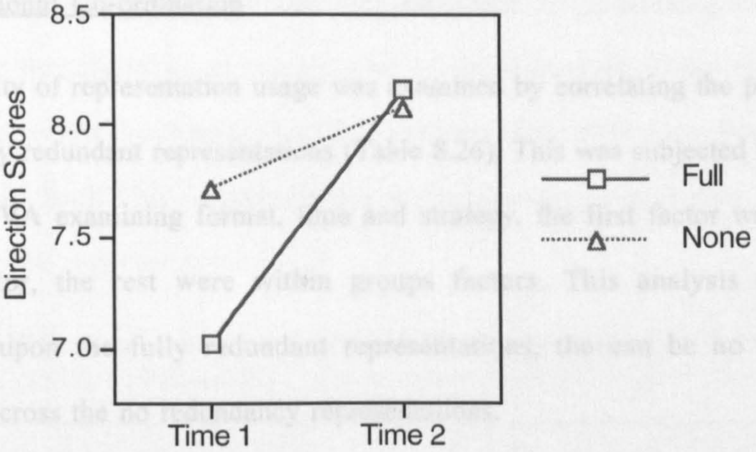


Figure 8.17 Direction scores by redundancy and time

A further interaction between redundancy and strategy ($F(1,42)=6.71$, $p<0.013$) (Figure 8.18) was also found. Simple Main effects identified rounding as the only strategy to be influenced by redundancy ($F(1,84)=6.31$, $p<0.014$).

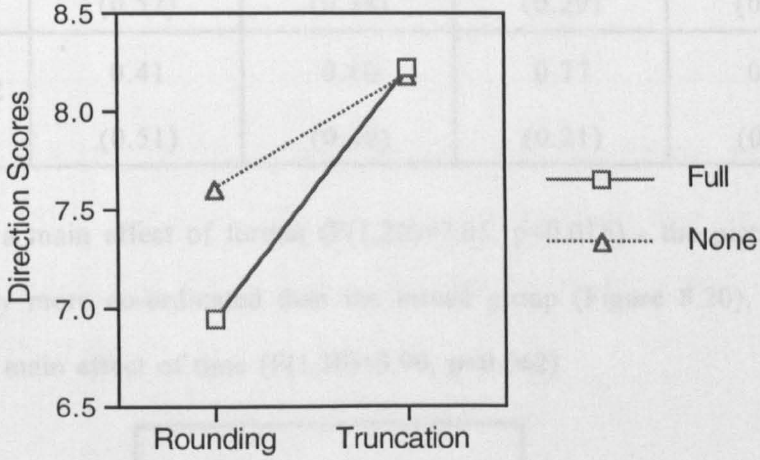


Figure 8.18 Direction scores by redundancy and strategy

There were two three way interactions. The first was between format, redundancy and time ($F(1,42)=8.63$, $p<0.006$). Simple Main effects indicated that the only difference between levels of redundancy was for the picts format at time one ($F(1,42)=10.43$, $p<0.003$) where the fully redundant representation was associated with poorer performance. The second three way interaction was between redundancy, time and strategy ($F(1,42)=7.09$, $p<0.011$). At time 1, when performing rounding, the level of redundancy was important ($F(1,84)=12.84$, $p<0.001$). Again, it was the fully redundant representation that led to poorer performance.

Representational Co-ordination

The similarity of representation usage was examined by correlating the predictions on the two fully redundant representations (Table 8.26). This was subjected to an [2 by 2 by 2] ANOVA examining format, time and strategy, the first factor was a between groups factor, the rest were within groups factors. This analysis can only be performed upon the fully redundant representations, the can be no similarity of behaviour across the no redundancy representations.

Table 8.26. Correlations between the representations by format, time and strategy

	Mixed		Picts	
	Rounding	Truncation	Rounding	Truncation
Time 1	0.27 (0.52)	0.32 (0.38)	0.65 (0.29)	0.54 (0.35)
Time 2	0.41 (0.51)	0.40 (0.40)	0.77 (0.21)	0.77 (0.31)

There was a main effect of format (F(1,20)=7.05, p<0.015) - the picts group were significantly more co-ordinated than the mixed group (Figure 8.20). There was a trend for a main effect of time (F(1,20)=3.90, p<0.062).

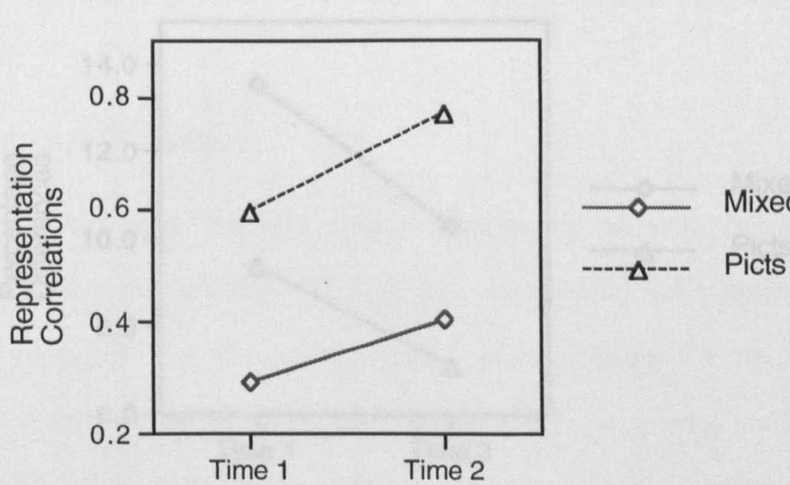


Figure 8.20 Correlation of predictions on the fully redundant representations by format and time

Representational co-ordination was also measured by examining the difference between the two measures, *i.e.* (prediction A - prediction B). If the representations were used to predict the same answer, then the sum of the differences between the two groups should be 0. This measure differs to the correlations as it does not allow for rescaling on the representations.

Table 8.27 Difference in prediction between the representations by format, time and strategy

	Mixed		Picts	
	Rounding	Truncation	Rounding	Truncation
Time 1	13.91 (5.91)	13.15 (4.72)	9.64 (6.35)	9.19 (6.59)
Time 2	9.66 (7.32)	10.97 (6.08)	6.62 (5.28)	7.73 (6.07)

This produced a very similar result to the other analysis, although both the effects manifest as trends. The picts group were more co-ordinated than the mixed group ($F(1,20)=4.08$, $p<0.057$) (Figure 8.21). There was a trend for a main effect of time ($F(1,20)=3.70$, $p<0.069$).

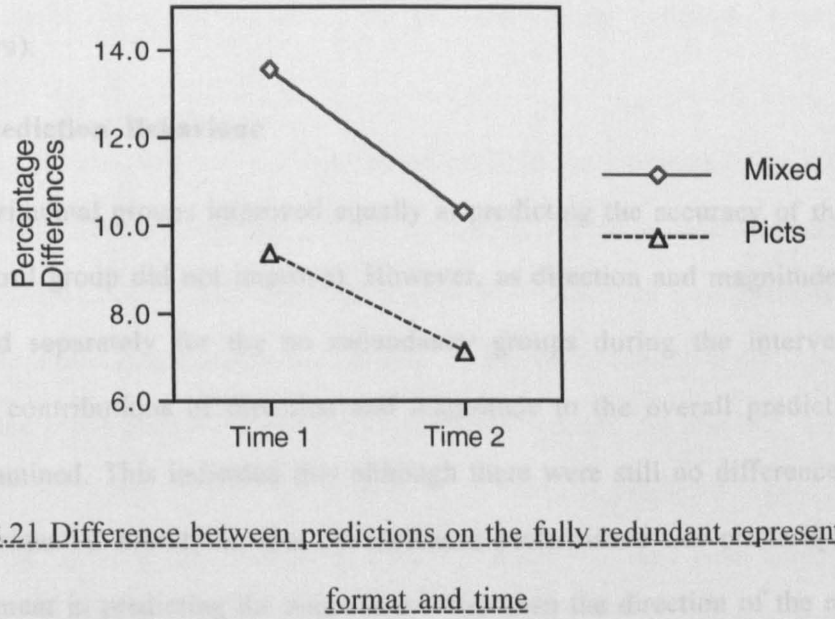


Figure 8.21 Difference between predictions on the fully redundant representations by format and time

Both these analyses therefore confirm the results of Experiments Three and Four which found that pictorial representations were more co-ordinated than a mixed system combining of mathematical and pictorial representations.

8.9 DISCUSSION

8.9.1 Computational Estimation Skills

The results of this experiment for the most part confirmed previous studies. Experimental groups produced more accurate estimates at post-test and used more appropriate strategies. The control group failed to improve upon any measure of performance. One result which was an exception to this general pattern of replication was that for percentage deviation (corrected for order of magnitude), where there was an interaction between format and redundancy, with the picts-none group performing worse than the other conditions. It is possible that this may be a genuine effect of representation type upon estimation performance. This interpretation is unlikely since this result would contradict the previous experiments in which representation had no affect upon strategic knowledge. Secondly, in all the experiments, a small proportion of children failed to learn, and they were equally balanced across the conditions. In this case, it seems that due to inadvertent sampling bias, more children who failed to learn were in the 'picts-none' condition (four as opposed to one in all the others).

8.9.2 Prediction Behaviour

All experimental groups improved equally at predicting the accuracy of the estimate (the control group did not improve). However, as direction and magnitude had been examined separately for the no redundancy groups during the intervention, the separate contributions of direction and magnitude to the overall prediction scores were examined. This indicated that although there were still no differences between the experimental conditions, that the increased performance was primarily due to an improvement in predicting the magnitude rather than the direction of the estimate.

This was a surprising result. It was thought that direction would be easier to predict that magnitude. Children found the direction predictions easier than magnitude on the computer. By the second time on CENTS, the direction predictions were almost at ceiling, an average of 8.1 correct predictions from a total of 9. Examining the children's scores on the post-test in more detail, it can be seen that there was huge variability in the number of correct direction predictions (0% to 100%). Ten children got 75% or more of the prediction right, seven got none right (five of these seven were those who had failed to learn any estimation strategies).

Strategies for predicting varied widely. Three children always predicted 'above', while three others always predicted 'below'. The former was rarely associated with high scores. However, invariable prediction of 'below' was commonly very successful, as often these children were producing their answers by truncation. The majority of children did not stick to one type of prediction throughout but adjusted their prediction on each question, again with varying degrees of success. There were few answers generated by rounding up, but these were generally predicted very accurately. Predicting the direction of intermediate compensation is often very difficult as it rests on understanding that relative (not absolute) differences govern accuracy. In total, 57% of rounding down predictions were correct, 47% of intermediate compensations were correct and 71% of rounding up solutions were predicted accurately.

One cause for concern was the number of children that indicated that they thought their estimate was exactly right. This, of course, serves to decrease the direction scores (but incidentally, especially in the case of intermediate compensation, will lead to a fairly accurate magnitude prediction). At pre-test, 'exactly right' accounted for 10% of the predictions given by the experimental groups. At post-test, however this accounted for nearly 25% of the answers. Two subjects predicted 'exactly right' for all answers. These children acted as if they believed they had been taught a method for exact multiplication during the intervention (it should be noted that these children did not know a long multiplication algorithm). They may have based their judgements upon introspection about the procedure rather than the outcome. If they were confident they had performed the process correctly, then the outcome should be 'right'.

A further proportion of children predicted 'exactly right' whenever they performed intermediate compensation, but predicted higher or lower for other solutions. A small number of children predicted 'exactly right' whenever their intermediate solutions appeared in absolute terms to cancel out, *i.e.*, $18 \times 42 = 20 \times 40$. This is strategy that is commonly demonstrated by adults and represents quite sophisticated, if flawed, reasoning on the parts of the children.

These results suggest that both the intervention and the pen and paper measures could be improved. Given the high performance at direction prediction upon the computer, it might be that children had become complacent about this skill and were failing to reflect upon their prediction. There may be over reliance upon the support provided by the computer; perhaps the system needs to fade this support. However, it may be that this result represents working memory failure upon the children's part. Case & Sowder's (1990) model would certainly predict that due to limited capacity children of this age would perform poorly on this task. Children may fail to remember how they changed the numbers once the intermediate solution has been reached. There is some evidence for this interpretation since in at least one subject's results, the predictions would have been nearly perfect *if* the answers had been generated according to the most accurate strategy. However, the actual estimation strategy used was truncation. It would appear that this subject performed the prediction separate to the estimate and without remembering the strategy that he had used.

A second cause for concern is a worry that children's beliefs about the nature of mathematics were interfering with successful performance on this task. It has been argued by many researchers and teachers that children beliefs about mathematics can negatively affect performance (*e.g.* Phillip *et al.* 1994; Schoenfeld, 1992; see section 2.3 for a review). In this case, children are likely to view producing an answer as the most important part of the task. However, by far the most complex part of the tests is to calculate the relation between the estimate and the right answer. If children viewed the estimation as the only important mathematical goal, and they may have paid only limited attention to the prediction tasks.

Given these difficulties, it may be wise to adapt the tests. There is abundant evidence from the previous experiments that CENTS can teach estimation skills. Hence, there is less need to test this directly. Instead, children could be given semi-completed tests, where intermediate solution and final estimate would be given. Their sole task would be to calculate how accurate the given estimates were. This would both provide support for working memory and remove difficulties in comparison between children who used different types of estimation strategy. Only limited research as examined what children know about estimations that they have not produced (*e.g.* Dowker, 1993). It would seem interesting to expand this work to older children using more complex estimation tasks than Dowker's children used.

8.9.3 Representations

Although there were no differences between the experimental conditions in the final outcomes, there were differences in how the representations were used during the intervention. Analysis of prediction behaviour with the representations produced a fairly complex pattern of results, many of which confirm the earlier studies, however there were some differences. The following discussion is structured to around each of the factors in the experiment: Format, redundancy, time and strategy.

<u>Format</u>

In this experiment, just two of the formats that have been used in previous experiments were selected; pictures, and a mix of pictures and mathematics. Some of the results confirmed those of earlier studies. For example, analysis of prediction accuracy on the fully redundant representations confirmed that pictures were initially used more successfully than mixed representations. The no redundancy representations did not show this effect, either for direction or magnitude. It will be argued later that this was because the task of predicting accuracy was easier when each decision was made separately. When direction for the full redundancy representations was examined, the interaction between format and time was significant. In contrast to previous experiments and the above analysis, for the first intervention session, children in the picts condition performed worse than mixed representations. There was also an interaction between format and representation. The splatwall was used similarly across both formats, but the archery field produced significantly poorer performance than the numerical display. Thus, the archery field seemed to be successful for expressing magnitude, but supported direction poorly. This representation is effectively the splatwall turned through 90 degrees and emphasises that apparently small changes in presentation may have significant effects upon how well the representation supports learning.

This point is interesting to consider in relation to the different approaches to classifying representations first discussed in section 3.2. Palmer's (1978) analysis could employ the principle of uniqueness (whether the representation directly supports the required inference or whether additional information is needed to construct the right interpretation) to explain the difference between the representations. In the case of splatwall, the direction of the estimate maps directly onto above and below on the splatwall. However, the archery field requires users to understand that left represents 'under the exact answer' and right, 'above the exact answer'. Thus, the additional inference seems likely to be the reason why this representation was used less successfully than the other pictorial representations. The other approaches to classifying representations such as modality and the taxonomic approaches, (*e.g.* Lohse *et al.*, 1994) would not be able to make any distinction between these two representations.

The numerical display seemed to support direction predictions well. One possible explanation is that this representation, as opposed to the others in the full redundancy groups, requires users to first state direction and then state magnitude. This turned prediction of the accuracy of an estimate into two separate decisions, similar to the no redundancy representations where direction has been predicted more successfully than with full redundancy. This is not a fundamental property of the representation. It could have been implemented so that user's must select direction and magnitude simultaneously. This interpretation again suggests that caution should be taken about generalising from representations in too global terms.

The numerical display representation was associated with poorer performance for magnitude predictions. It was also found that for composite predictions (*i.e.* those that collapse direction and magnitude into one score) mixed representations were worse for prediction accuracy at time 1 than pictorial representations. This finding replicates the previous two experiments with CENTS. The four way interaction between format, redundancy, time and strategy reported in section 8.8.4 was explained by the numerical representation supporting predictions of magnitude prediction poorly at time 1 when producing answers by rounding although not when estimating using truncation. In other words, when task demands were particularly hard, this representation was used less successfully.

Redundancy

This experiment manipulated redundancy over representations. It was proposed in section 3.7, that one dimension that is important in considering multi-representational software is how the information is shared over the representations. In this case two levels of redundancy were used, full redundancy where all the information derivable in one representation is also present in the other. Under conditions of no redundancy, each representation presents unique information. It was suggested in section 8.6.2, that limiting redundancy between representations may aid initial learning.

An analysis of direction predictions which compared each of the fully redundant representations with the no redundancy representations was performed. The first representation in the fully redundant MERs was the splatwall, the second either the archery field or a numerical display. There was a fairly consistent effect of redundancy. No redundancy was either generally better (compared to the splatwall) or better at time 1 (compared to the archery field and numerical display). Thus, the hypothesis that when each representation presented only a single dimension, it may be easier to learn about the effects of different dimensions, received some support. It was also found that the fully redundant representations were associated with poorer performance when the task demands were more complex, *i.e.* when predicting rounding, or using the archery field or at time 1. Thus, although much more research is needed, there does seem to be evidence that limiting redundancy over representations may be beneficial (at least in the short term).

Strategy

For fully redundant representations, prediction accuracy was higher for estimates produced by rounding. This replicates the effects of strategy in Experiments Three and Four which had found that prediction accuracy was higher for rounding than for truncation. When examining direction and magnitude predictions separately, both were influenced by strategy, but with different effects. Truncation estimates were predicted more accurately for direction, and rounding estimates were predicted more accurately when just considering magnitude. One explanation for these results was that as truncation invariably results in a lower answer, it was much simpler to predict the direction (lower!) than it was for rounding, which can either be higher or lower. This result was discussed in section 7.4.6, when it was suggested that predicting magnitude may be easier for rounding solutions than truncation solutions as rounding tends to produce more exact answers. Children have difficulty believing that mathematical procedures that are 'correct' can lead to inexact solutions. Again, this seems consistent with the 'right answer hypothesis' proposed by Baroody (1987).

It is particular difficult to predict direction on answers produced by intermediate compensation. Truncation, however, can produce highly inaccurate answers, in contrast to rounding which tends to produce more accurate answers. Hence, as children have a tendency to overestimate the accuracy of their answer, rounding will normally be predicted more accurately.

Time

The vast majority of measures (*e.g.* composite predictions, predictions of the magnitude of estimates) showed simply a main effect of time, *i.e.* children were more successful on their second use of the system. The only exception to this was that direction predictions for the no redundancy representations showed no improvement. However, as an average of eight out of the nine answers given during the first intervention session were correct, this indicates a ceiling effect. This does not mean that learning did not occur within a session. For example, 68% of predictions for the first two questions were correct but 92% of the final two questions were answered correctly.

8.9.4 Representation Co-ordination

Representational co-ordination was examined across the fully redundant representations (this analysis obviously is not applicable to no redundancy representations). The results of this study replicated the previous ones in that predictions across pictorial representations were significantly more co-ordinated than across mixed representations. This confirms that the previous effects of representational co-ordination are unlikely to be due to the particular combination of representations.

However, this time there was a trend for both groups to improve over time. In previous experiments with partially redundant representations, pictorial and mathematical representations had showed increased convergence with time, but mixed representations had not. The results of Experiment Five are in line with the prediction that full redundancy would help children learn to translate across the representations as it would increase the similarity between representations.

8.10 CONCLUSION

Again, CENTS was shown to be successful at teaching primary school children aspects of computational estimation. Investigation of the effects of MERs showed that when the relation between co-ordination and learning outcomes was analysed, the results were in line with previous experiments - pictorial representations were more successfully co-ordinated than mixed representations. In addition, it had been predicted that a fully redundant system should aid representational co-ordination. Even for children who don't learn to co-ordinate, concentrating on one representation should provide them with all the information they need to know. Hence, we would not expect to see a difference in learning outcomes between the different formats for fully redundancy.

However, it was predicted children in the mixed, no redundancy group should perform less well at post-test. It was argued that children in the mixed condition would be impeded in integrating information across the representations. Hence, as correct prediction requires both direction and magnitude to be related, then these children should be at a disadvantage. This prediction was not supported by the data.

Two of the possible explanations for the performance of the children are as follows. The first explanation is based on the evidence that limiting the redundancy between representations may make aspects of the task easier. This allowed children to focus their attention on integrating information from both representations. Hence, the task of integrating information from mixed representations was supported by making others aspects of the learning goal easier.

The second explanation concerns the nature of all subjects post-test performance. It was found that few of the children improved at both the magnitude and direction components of prediction accuracy, - the only significant improvement was for magnitude. As this was the case, less redundancy would not be implicated in children failing to integrate understanding of both dimensions. There is no evidence that any of the multi-representational systems supported this level of understanding.

Thus this experiment replicated the previous ones which found a) that CENTS can successfully teach estimation and; b) that mixed representations were less well co-ordinated than other representations. There was tentative support for the hypothesis that representations dedicated to a single aspect of a situation support initial concept acquisition more effectively. However, as none of the experimental conditions were associated with integration of these concepts, longer term consequences of less redundancy across representations remain unexplored.

8.11 GENERAL CONCLUSION

This chapter reported two further experiments aimed at exploring CENTS effectiveness at teaching computational estimation and at examining the effects of MERs on learning. In both experiments, all experimental subjects significantly improved their estimation performance. Different combinations of representations were manipulated to explore the effects of format and redundancy on learning. In all cases, mixed representations were associated with less co-ordination during the computer intervention. In Experiment Four, with four intervention sessions on the computer, children in the mixed condition appeared to concentrate upon a single representation. This is in contrast to children in pictures and mathematical representation conditions who learnt to effectively use both representations. Thus, it was suggested that in conditions of partial redundancy it is possible for children to learn about the domain if they concentrate upon a single complete representation.

Experiment Five examined the effects of different levels of redundancy, either full redundancy where the same information is displayed in both representations or no redundancy where each representation presents different information. The influence of different levels of redundancy on learning with MERs is less clear cut, however it appears that there may be an initial advantage to be gained by limiting redundancy.

The final chapter will discuss in more detail the implications of the results of the experiments with CENTS and COPPERS for the design of multi-representational software and primary mathematics teaching.

CHAPTER NINE

Discussion and Synthesis

This final chapter presents a summary and synthesis of the thesis. The research objectives are reviewed and the design of the learning environments are considered with respect to multiple representations and number sense. A brief summary of the more important experimental findings is given. Then, the limitations of the systems and the evaluation studies are considered, along with the potential for improvements in evaluation and design. The techniques and results of the empirical studies are used to address the question of design of multi-representational software. The final sections present suggestions for future work and provide a general summary of the thesis.

9.1 AIMS AND OBJECTIVES

This thesis has addressed the design of computer-based learning environments for primary mathematics. Evaluation of the systems focused on how different representations of the domain affect learning. Examination of the goals of primary mathematics teaching identified the 'right answer' misconception as an important and under-researched domain. Children who hold this view believe that all problems must have a correct answer, that there is only one correct way to solve a problem and that inexact answers and procedures are undesirable. It was argued (Chapter 2) that such a misconception is implicated in children's failure to develop good number sense. Thus, one principal goal of this thesis was to design, implement and evaluate learning environments that successfully address these aspects of number sense. The first system, COPPERS, was designed to teach children that there can be many correct solutions to problems. The second system, CENTS, was designed to teach computational estimation. It supports the development of strategic and conceptual knowledge, and places considerable emphasis upon the development of insight into accuracy in estimation.

Furthermore, inspired by the possibilities inherent in computer-based environments for novel and interesting uses of representations, the second goal of this thesis was to explore issues in the successful use of multiple external representations (MERs). It was argued that comparatively little is known about how children use multi-representational learning environments. Consequently, little is known about how to achieve successful multi-representational software. To inform design, the research aimed to:

- uncover and analyse different uses of MERS
- examine the various learning demands of MERS
- evaluate the influence of different combinations of representations on learning outcomes
- consider the instructional implications of supporting such learning

To this end, a literature review was performed upon many different domains. The issue of external representations in learning and problem solving has been considered in (at least) the following areas: Artificial Intelligence, Cognitive science, Cognitive Psychology, Developmental Psychology, Mathematics and Science Education, Human Computer Interaction, Instructional Science, Intelligent Tutoring Systems, and Visual Programming Languages. These areas were reviewed (although not all received equal attention) in order to develop a conceptual framework with which to consider the use of MERs in computer-based learning environments.

Some of the issues raised by this analysis were selected for further empirical study, and systems were designed to allow exploration of these areas. The primary focus of this aspect of the research was on how nature of the multi-representational system may influence learning. This was considered in terms of the similarity of format and redundancy across representations.

Both COPPERS and CENTS were designed to use MERs. COPPERS employed multiple representations (a) sequentially, to display questions in increasingly abstract notations (from concrete to algebraic) and (b) in parallel to describe answers in two complementary representations. The role of the latter use of MERs was addressed empirically.

CENTS was designed to assess how combining different types of representations influenced learning. Estimation accuracy was presented in MERS, and learners used them to predict the accuracy of their estimates. Combinations of representations were manipulated to alter redundancy of information and format of representations. Consequently, predictions concerning how to best combine representations to support learning could be tested.

Thus, this thesis has been concerned with four main issues:

- assessment of children's performance at computational estimation and multiple solutions tasks
- the development and evaluation of learning environments that teach these areas
- empirical analysis of how system features affected learning outcomes, specifically in relation to MERs
- the creation of a framework and methodology to help inform design of multi-representational learning environments

The following sections reviews the first three of these areas, which are directly based on the empirical research. The final, more general issue of the design of learning environments, is considered in section 9.6.

9.2. NUMBER SENSE - MULTIPLE SOLUTIONS AND COMPUTATIONAL ESTIMATION

Before each intervention study, data was collected on children's performance upon multiple solutions and computational estimation tasks. This served two functions: (a) it served as a baseline to compare with post-intervention performance and; (b) it justified the pedagogical goals of the systems by demonstrating that children in this age group are unlikely to possess these mathematical skills without direct teaching.

It was argued in Chapter Two that children aged six to eleven would demonstrate mathematical performance consistent with the 'right answer' hypothesis (Baroody, 1987; Lampert 1990). It was predicted that primary school children would demonstrate little proficiency at producing multiple answers for (apparently) simple mathematical problems and would have difficulty with inexact mathematical procedures such as estimation. This prediction was supported by the research conducted with both COPPERS and CENTS (described in Chapters Five, Seven and Eight).

9.2.1 Multiple Solutions

Experiment one examined the six to seven year old subjects' performance on the multiple solutions pre-test. These children produced an average of less than one correct solution per question. The eight to nine year old children demonstrated only slightly better performance with 1.2 correct solutions per question. In addition, Experiment Two showed that, without direct teaching, merely re-testing children (and thus it might be argued beginning to legitimise multiple solutions as a mathematical goal) did not improve performance. A non-intervention control group's performance remained constant across three pen and paper tests.

The results of these experiments showed that primary school children do not easily produce multiple solutions to coin problems. This might appear counter-intuitive, as the mathematical skills to achieve this should be within the competence of children of this age (and indeed are after the intervention). This is consistent with worries about the nature of children's mathematical beliefs negatively influencing performance where multiple solutions are required. This was also supported by the weak relation between more general mathematical ability (as measured by the BNST and Y1) and multiple solutions performance. This poor performance provided support for COPPERS' pedagogical goal of teaching children to produce multiple answers for a single mathematical problem.

9.2.2 Computational Estimation

Previous research on children's (and adult's) estimation abilities has examined both the skills and concepts of computational estimation (Reys *et al.*, 1982; Sowder & Wheeler, 1989) and has included a developmental model (Case & Sowder, 1990). Such research predicts that children in the age group examined (10-11 years) will demonstrate only limited estimation skills. This prediction was supported by the evaluation studies (Chapters Seven and Eight).

Three measures of children's estimation performance were taken: (a) analysis of the strategy used to produce the estimate, (b) assessment of the accuracy of the estimate, and (c) analysis of the insight children had into the accuracy of an estimate.

At pre-test, children rarely applied an appropriate strategy (such as rounding or truncation). Averaged across all three experiments, only 7% of pre-test responses used an appropriate strategy. This was also reflected in the accuracy of the estimates. At pre-test, subjects' estimates were an average of 91.4% away from the right answer. For example, an estimate to '25 × 55' might be given as 118 when the exact answer is 1375. Frequently, the estimate was lower than one of the factors or less than the sum of the two factors.

The final measure taken was insight into the estimate - *i.e.* whether children could recognise whether an estimate was higher or lower, close to or far from the exact answer. In all of the experiments, children demonstrated only a limited understanding of accuracy - for example, their prediction of the direction of an estimate was at chance.

Children showed little understanding of either skills for producing estimates or the ability to generate or recognise 'ball park' responses. In addition, once again, non-intervention controls did not improve after repeated testing. It would seem that the skills and knowledge that CENTS was designed to support are not present in primary children before direct teaching.

Results of the investigations of children's understanding of multiple solutions to coin problems and computational estimation were consistent with previous research on the nature of children's beliefs about mathematics and computational estimation. These results show that primary school children have difficulties performing successfully on these sorts of problem and are unlikely to develop these skills without some teaching. This motivates the implementation of the two systems designed to support this type of mathematical understanding.

9.3 DESIGN AND EVALUATION OF THE LEARNING ENVIRONMENTS

COPPERS and CENTS were designed to support different aspects of number sense and to explore how different combinations of representations influenced the development of this understanding. This section will give a brief summary of the motivation for building these systems, and will assess whether the systems met their educational objectives.

9.3.1 Motivation

The arguments for teaching children to give multiple solutions to single questions and to accurately estimate the answers to multiplication problems were reviewed in Chapter Two. To briefly recap, these areas have been associated with the development of number sense. It was proposed that successful performance in these areas depends heavily on number sense and that be helping children to develop these skills they also gain a deeper sense of number. It was argued that children commonly hold beliefs that can be characterised as the 'right answer' hypothesis. Furthermore, these beliefs are not compatible with current approaches to mathematics education which emphasise pattern seeking, number sense, hypothesis testing and active search for solutions (Schoenfeld, 1992). Thus, the systems were designed to address and challenge these beliefs by providing situations where multiple solutions and inexact procedures were required to solve mathematical problems, and by supporting the development of skills and knowledge fundamental to achievement of these mathematical goals.

9.3.2 Multiple External Representations

There is considerable evidence that appropriate external representations can aid learning and problem solving. Recently, given the opportunities presented by computers as tools for learning, there have been many advantages proposed for providing MERs. Three broad classes of claims were identified in section 3.3: (a) that MERs support different ideas and processes, (b) that MERs constrain interpretation, and, (c) that MERs promote a deeper understanding of the domain. The evidence to support these proposals was reviewed in Chapter Three. Analysis of the learning demands of MERs with respect to these benefits was undertaken (section 3.4). One of aims of this thesis was to examine under what conditions MERs supported learning.

The types of representations used by the systems varies. COPPERS uses different representations of coins - either pictorial, numerical or algebraic. It also employs two representations for feedback - a tabular representation and a row and column representation. CENTS offers users (and researchers) an even wider choice of representations. It can employ pictorial representations (such as the 'splatwall' or archery target), or mathematical representations (histogram, numerical values) which can be displayed in any combination. The redundancy of information across displayed representations can also be varied - both representations can provide the same information, can provide completely different information, or have partial overlap. Thus, both COPPERS and CENTS were designed to exploit and explore the different inferential capabilities of external representations, either singly or in combination.

9.3.3 Evaluation of the Learning Environments

<u>COPPERS</u>

The results of the evaluation studies conducted with COPPERS suggest that it can be used to teach young children to produce multiple solutions to coins problems. As discussed above, initial performance on this task was low - the majority of subjects in Experiment One and Two did not produce correct multiple solutions to the coin problems. After the intervention phase, (two sessions using one of the COPPERS versions), all experimental subjects significantly improved the number of correct multiple solutions by the post-test. Averaged over the two experiments, children increased the number of correct solutions given by over 300%.

In addition, a simple analysis of the types of solutions was performed. It was found that both the average number of coins per question and the average number of different types of coins per question increased from pre-test to post-test. An informal post-hoc analysis suggested that children were demonstrating less routine solutions after the intervention.

COPPERS was shown to fulfil its desired educational function. Children taught with the system learnt to provide multiple solutions to coin problems and produced more complex and less canonical solutions after the intervention.

CENTS

CENTS was designed to teach the concepts, strategies and insight required for computational estimation. Only the latter two goals have been evaluated. Subjects were asked to estimate answers to twenty multiplication problems. Pre-test scores suggested that subjects effectively lacked the skills and knowledge necessary to perform this difficult mathematical task. The intervention studies showed that CENTS can be considered to be an effective environment for teaching some aspects of computational estimation.

The accuracy of estimates is the most commonly used measure of performance (*e.g.* LeFevre *et al.,* 1993; Rubenstein, 1985). All the studies with CENTS showed that experimental subjects significantly improved the accuracy of their estimates (answer accuracy). The mean percentage deviation was 91% at pre-test and 42% at post-test across the three experiments. However, although the pre-test means across the three experiments were similar (ranging from 87% to 98%), the post-test scores for the experimental groups were quite different (mean post-test scores: Experiment Three - 58%, Experiment Four - 28% and Experiment Five - 39%). This range is primarily due to differences in order of magnitude correction, and may reflect a change in the system after Experiment Three to emphasise place value correction. It should also be noted that in Experiments Three and Five, experimental subjects received two training sessions, whilst in Experiment Four, each subject used CENTS four times.

Estimation performance was also scored for percentage deviation from the exact answer after an order of magnitude correction had been applied. It was argued in Chapter Seven that this represents a more useful measure of subjects' performance. It distinguishes children who guess answers from those that apply a correct strategy but fail at place value correction. For the three experiments, the mean percentage deviation (corrected for magnitude) was 91.2% at pre-test and 22.0% at post-test. Again, there was little difference between the experiments at pre-test averages, but for this measure the difference at post-test was also smaller (mean post-test scores: Experiment Three - 26%, Experiment Four - 19% and Experiment Five - 21%).

Both performance measures showed that experimental groups became significantly more accurate after the intervention and provided support for the claim that CENTS can teach children to become more accurate estimators.

An examination of the strategies used by the subjects to estimate their solutions was performed. Estimates were identified as rounding up, rounding intermediate compensation, rounding down (or truncation) or unknown. (The other strategies identified by Reys *et al.,* 1982 were not found in these studies). Hence, strategies that involved front end extraction were coded as appropriate and all others as inappropriate. CENTS attempts to teach children rounding and truncation. Unsurprisingly, we can see that after the intervention, (experimental) subjects estimate problems using more appropriate strategies. At pre-test, the average score was 1.7 (from 20) and 15.8. at post-test. Again, there was some variation of outcomes across the experiments (mean percentage of appropriate estimates: Experiment Three - 61%, Experiment Four - 94% and Experiment Five - 82%). However, it can be seen that the majority of subjects were able to learn and correctly apply the strategies supported by CENTS.

Considerable emphasis was placed upon supporting insight into the accuracy of an estimate in CENTS (prediction accuracy). This skill is slow to develop (Case & Sowder, 1990), yet represents an important part of a sophisticated estimator's skills (Trafton, 1986). Hence, the system provides MERs of estimation accuracy and children used these representations for both prediction and display.

Pen and paper post-tests of prediction accuracy revealed that these intervention goals received qualified support. The majority of experimental subjects improved significantly (differences between conditions will be discussed in section 9.4). Children were asked to indicate how accurate they thought their estimates were by indicating a category that described each estimate. These were labelled in both natural language and percentage deviations. A typical category was, 'much less than: 20-30% below'. Thus, it was possible to score prediction accuracy independently of the accuracy of an estimate. At pre-test, the average difference between a child's prediction and the right prediction was 3.7 categories. At post-test, this had improved to 2.3 categories difference.

However, closer analysis in Experiment Five revealed that whilst subjects were better at predicting the magnitude of the estimate, they did not improve at predicting its direction. Explanations that were considered in section 8.9.2 included the difficulty in predicting direction when answering using intermediate compensation, and children's complacency about their knowledge given their performance on the computer (ceiling). It was also suggested that a post-test which solely tested prediction accuracy might provide a cleaner result.

In summary, CENTS and COPPERS have been shown to be able to teach primary school children much of the relevant skills and knowledge involved in multiple solutions and computational estimation.

9.4 SYSTEM FEATURES

The final aim of the experiments was to use the systems as laboratories to test more general predictions about how to support learning. The principal focus was on testing claims for (multiple) external representations in computer-based learning environments.

Two aspects of COPPERS design were examined: (a) the number of answers per questions required from users, and (b) the importance of providing an additional tabular representation. Experiments with CENTS built on a fundamental issue that arose from the COPPERS experiments and concentrated upon learning with multiple representations. Two aspects of MERs were examined: the format of the combined representations and the redundancy of information across the representations.

9.4.1 Multiple Solutions

In Experiment One, subjects who were required to give four rather than one solution per question did not generally produce more correct solutions at post-test. However, for the lower performing children, giving multiple answers did lead to significantly better performance. In Experiment Two, this was examined further by adding an eight answer condition and an autonomous condition where users could choose their own number of answers per question. A four answer condition was retained. It was found that all the experimental groups had significantly better performance at post-test. In addition, the eight answer group had significantly more correct answers than the four answer and control groups at post-test, and the control group at delayed post-test. However, closer analysis revealed that the four and autonomous groups improved solely because the accuracy of their answers improved. The eight group, in addition to becoming more accurate, also increased the total number of decompositions.

This result was interpreted in terms of the zone of proximal development (Vygotsky 1978). It was suggested (section 5.9.2) that the strategy employed by the computer should be to require a number of answers that was just beyond what a child could produce without support.

Given the difficulty of diagnosing the region of the zone of proximal development, and the motivating effects of selecting one's own goals, it might be argued that children should be given the opportunity to select the number of answers per question for themselves. However, children given this choice in the autonomous condition provided an average of 1.66 answer per question and did not increase their number of decompositions after the intervention. It was argued that this behaviour is in line with children's mathematical beliefs (reviewed in section 2.3). Thus, it would seem that the computer should at least decide upon a minimum number of solutions. Children could then give additional solutions if they wished.

9.4.2 External Representations

The representations used by COPPERS were examined in Experiment One. The system employs two types of representation for describing the answers that children give. A standard row and column representation is used to describe the partial products of an answer. In addition, a summary table of these partial products is employed to provide information about the current and (where relevant) previous answers to a question. The hypothesis that the table would improve the children's performance was supported. The children who were provided with the summary table in addition to the row and column representation produced significantly more solutions at post-test than children who were only given the row and column representation.

A number of explanations were proposed (5.4.3) that could account for this result. Lampert (1986a) had suggested that such tables would be useful for allowing subjects to compare their previous answers, especially those that had been in error. The research on the computational properties of representations suggests that tables encourage pattern seeking and reflection on unexplored alternatives. Informal assessment of children's behaviour suggests that this was occurring, but given the granularity of information available from the computer records this can not be proved for certain.

In addition, the table serves as a symbolic representation of the multiplication and addition procedures involved in finding solutions to the problems. Numbers in the columns must be multiplied by the numbers in the column heading and then added together to get the total amount of money. The operators used to interpret a table therefore require children to practise the skills that COPPERS is attempting to teach.

Finally, the table and row and column representations simultaneously provide information on the same problem whilst presenting it in different ways. Thus, the advantages proposed for translating across multiple representations may account for the improved learning seen in the tabular representation condition.

This experiment does not isolate which one of these explanations account for the improved performance of the table group. Further research must be conducted to establish this. One possible future experiment would be to examine the table presented without the row and column representation. This would allow the role of multiple representations to be separated from the cognitive properties of the table representations *per se*. Further experiments could examine whether the table was promoting deeper mathematical exploration by gathering much more detailed accounts of children's behaviour.

9.4.3 Combining Representations

One dimension that was proposed as unique to learning with MERs was the need to consider the similarity of format across representations. Experiment Three examined this issue. The research literature was reviewed to derive a series of predictions about the relative advantages of different types of representations. This was firstly done by considering the properties of individual representations. For example, pictures may be beneficial initially and with lower performing children. Mathematical representations take longer to learn, but ultimately prove to be more effective. A mix of representations may offer the best solution, in that pictorial representations can be used to bridge understanding to the more symbolic ones.

However, it was argued in section 3.4, that multi-representational software that aims to exploit abstraction across representations requires learners to understand three aspects of MERs:

- the format and operators of a representation

- the relation between a representation and a domain
- the relation between the representations

Thus, in section 7.1 a second set of predictions about the relative advantages of different representations was articulated, based upon considering the learning demands of MERs. In order to test these predictions, three different representational systems were created: pictorial (splatwall and target), mathematical (numerical and histogram) and mixed (numerical and target). Hence, each representation in the mixed representational system was also present in either the mathematical or pictorial system. In this way, the first two learning demands were constant for each representation, but the final learning demand was varied.

To briefly summarise the argument in section 7.4., subjects in the pictures and mathematical conditions improved on measures of prediction accuracy from pre-test to post-test; the mixed group did not. An explanation was proposed by examining the process data. Learners' knowledge of the representations and their domain knowledge was reflected in their use of representations for prediction. By the second session on the computer, the maths and pictures conditions were performing significantly better than children with mixed representations. Thus the poorer understanding manifested on the post-test was also present during the computer interaction.

To measure the final learning demand, the similarity of children's actions on both representations was assessed (representational co-ordination). It was argued that if learners understood the relation between the representations, their actions should be identical over both representations, even if their prediction is wrong with respect to the domain. It was found that over time the mathematical group and pictures group converged their behaviour across the representations, but that the mixed group did not.

Consequently, it was argued that the learning demands of representational translation in the mixed condition were so great that children could not benefit from the support provided by these representations. Predictions based upon considering the learning demands of MERs were supported over those that simply considered the properties of each representation in isolation.

In order to explain this effect, a comparison of the properties of each representational system was performed. The pictorial representations are based on the same metaphor and have similar formats and operators. Pictures are also very familiar representations to children of this age. The mathematical representations have different formats and operators, as one representation is graphical and one is propositional. However, it was proposed that mapping between the representation was facilitated by the fact that both representations use numbers. The mixed representations differ in terms of formats and operators and have no obvious mechanism to support the mapping. In addition, they also mix mathematical and non-mathematical representations. Failure of overlap therefore occurs at all levels.

However, before recommending that representations that are *too* difficult to co-ordinate are avoided in computer based learning environments, it was obviously necessary to replicate this effect. In addition, it might be argued that the negative effect of mixed representations is a temporary one caused by the excessive memory and learning demands at the initial stage of a task. Hence, this experiment was repeated with a further two intervention sessions.

9.4.4 Representational Co-ordination over Time

If failure to co-ordinate mixed representations is only short term problem, it was predicted that representational co-ordination should be seen to converge by the 4th session of the intervention. In addition, it would be expected that there would be no differences in learning outcomes between the experimental conditions. It might even be argued that the extra work needed to build links across the mixed representations would have led to a better understanding of the domain.

The post-test results from Experiment Four showed that children in all experimental conditions had improved at both answer accuracy and prediction accuracy. Such a result was consistent with the prediction that mixed representation were only problematic over short periods of time. However, this simple explanation was rejected when the process measures of prediction accuracy and representational co-ordination were examined.

Prediction accuracy on the categorical representation showed a strikingly similar pattern of results to that of Experiment Three. Subject's use of the categorical representation in the mixed condition did not improve over time. However, prediction with the continuous representations was dissimilar to Experiment Three. In this case, pictures were initially associated with better performance as before, but prediction accuracy improved in all conditions.

When representational co-ordination was examined, it was found that mathematical and pictorial representations converged over time. However, the mixed group again showed no evidence of co-ordination, even after four sessions. Therefore, the hypothesis that failure to co-ordinate in mixed conditions only occurs during a limited initial period was rejected.

It was argued that children's continued failure to translate across the mixed representations ultimately lead them to abandon their attempts to work with one of the representations (categorical) and to concentrate on the other representations (continuous). This decision was highly strategic as the continuous representation contains both the direction and the magnitude information. Reliance solely on this one complete representation would account for the mixed group's improvement on the test.

The results of the Experiments Three and Four suggested that mixed representations are considerably more difficult to co-ordinate than pictorial and mathematical representations. This may either result in poorer performance overall (Experiments Three) or concentration upon a single representation (Experiments Four). However, concentrating upon one representation and ignoring both a second representation and the translation across representations, means that many of the proposed benefits of multiple representations cannot occur. Multiple representations could not be used to constrain interpretations or to support abstraction.

In addition, one of the most common claims for the use of multiple representations is that one representation may be insufficient to display all the needed information. Consequently, MERs may be used to convey different aspects of a domain. The hidden assumption in this approach is that a learner can then integrate information from all of these representations in order to fully understand the domain.

Experiment Five was therefore designed to address the affects of representational co-ordination under conditions of varying redundancy.

9.4.5 Redundancy

Integrating information across MERs is particularly important if the representations express different concepts. In the situations described above there was some overlap between the representations, as each involved magnitude. However, the granularity of magnitude information was different and the continuous representation also expressed direction. Hence, failure to integrate information across these representations will not necessarily result in impoverished learning.

Two further levels of redundancy were employed to address this issue. For full redundancy, the same information is derivable from both representations. In this case, both representations expressed direction and magnitude in a continuous fashion. Where there in no redundancy, completely different information is derivable from each representation. Consequently, one representation was designed to convey direction information and the other (continuous) magnitude.

The hypothesis tested was that when it is difficult to integrate information across representations (*i.e.* mixed representations), having no redundancy will lead to poorer learning outcomes. If mixed representations are fully redundant then translation should be facilitated. Even without translation however, learning should still occur as one representation expressed all the needed information. Full and none redundancy pictorial systems were included for comparison as an example of MERs where co-ordination is facilitated. In addition, it was proposed to examine whether no-redundancy made the concepts easier to learn initially.

There were differences in how the representations were used during the intervention phase. There was a fairly consistent effect of redundancy. For direction predictions, no redundancy was either generally better (compared to the splatwall) or better initially (compared to the archery field and numerical display). This would suggest that it may be easier to learn about the effects of different dimensions when they are presented separately. The fully redundant representations were particularly associated with poorer performance when the task demands were more complex, *i.e.* when predicting rounding or during the first intervention session.

Analysis of the effects of format for the most part confirmed the results of earlier studies. For example, analysis of prediction accuracy on the fully redundant representations confirmed that pictures were initially used more successfully than mixed representations. The no redundancy representations did not show this effect. It was argued that this occurred as the task of predicting accuracy is easier when each decision is made separately.

Analysis of representational co-ordination across the fully redundant representations replicates the results of previous studies. Pictures were significantly more co-ordinated than mixed representations. There was a trend for both groups to improve over time. This was in line with prediction that full redundancy should help children map across the representations, but that mixed representations will still hamper co-ordination.

None of these differences in representation use on the computer were reflected in learning outcomes. Counter to the prediction, all experimental groups improved equally at prediction accuracy. It had been expected that there would be no differences between pictures and mixed representations under conditions of full redundancy. However, the results of Experiment Five contradicted the prediction that children in the 'mixed - no redundancy' condition would be disadvantaged in their attempts to integrate information across the representations and so perform less well at post-test.

Two possible explanations of this effect were proposed. Firstly, no redundancy appeared to aid understanding. This may have aided children attempting to integrate information from the mixed MERs. Secondly, post-test measures showed that few of the children improved at both magnitude and direction components of prediction accuracy. The only significant improvement was for magnitude. Hence, no redundancy mixed MERs could not be particularly associated with lack of integration of both dimensions as there is no evidence that any of the MERs supported this.

Further research is needed to clarify the effects of manipulating redundancy. Some suggestions for interesting directions in which to take this research are discussed in section 9.7

9.4.6 Summary

In brief, this research has identified a number of consistent effects. Primary school children do exhibit performance consistent with the 'right answer' hypothesis. However, short intervention sessions with CENTS and COPPERS can impact upon these beliefs and their behavioural corollary. All children in an experimental condition (regardless of what condition they were in) improved their performance on some of the outcomes measures. No non-intervention control group was observed to improve on any measure of performance. These findings seem robust. All the experiments reported in the thesis support this claim.

There is also considerable evidence that the external representations employed in the learning environments affect the process and outcome of learning. In COPPERS, providing an additional tabular representation provoked better learning outcomes. In CENTS, the way children used MERs was affected by the format and to some extent the redundancy of the representations. Pictures were normally associated with better initial performance and may aid lower ability children. Mathematical representations took longer to learn but were then used successfully during later computer sessions. Representational co-ordination was high in both these conditions. Mixed MERs were invariably associated with low levels of representational co-ordination. This was predictive of poor outcomes in Experiment Three. This effect was replicated over a longer time (Experiment Four) and with different representations (Experiment Five). There was some tentative evidence to show that distributing information over representations made certain concepts easier to learn. The generality of these findings in considered in section 9.7.

The results of these experiments show that when considering learning environments that use MERs, it is not sufficient to analyse each representation separately. The effects of representations will vary depending upon the way they are combined.

9.5 CURRENT LIMITATIONS AND SUGGESTED IMPROVEMENTS

9.5.1 Systems

Both COPPERS and CENTS are prototype learning environments. A number of additional features would be required if these systems were to be used in an everyday classroom. Trivially, such features would include documentation, re-implementation on a platform supported in schools, teacher controls and the removal of many features which supported these systems as experimental devices (*e.g.* the mouse-click level logging).

In addition, the experiments with children identified a number of aspects of system design which could be improved. Some of these were addressed by re-implementation during the course of the research, others remain to be undertaken. It was not possible during this research to evaluate all of the design features of the systems (*e.g.* the help provided has never been analysed). Consequently, the main focus of this brief review will be on aspects of the systems that were assessed.

<u>COPPERS</u>

One aspect of COPPERS underwent substantial modification during the course of the research - the way children select coins to answer the questions. In $COPPERS_2$, the 'coin calculator' was replaced by 'coin tubes'. This was prompted by the finding in Experiment One, that some children simply copied the coins in the question in order to provide an answer. This strategy was associated with poor learning outcomes. The 'tubes' prevented the application of this strategy by allowing one of the 'coin tubes' corresponding to an element of the question to be empty.

This new metaphor also allows COPPERS to ask more interesting questions. 'Coin tubes' could be presented which only contain a limited number of coins. COPPERS could set questions which would require users to provide as many solutions as possible within these restrictions. It would also be possible to implement an explicit 'exchange facility', allowing users, for example, to change a 5p coin into five 1p coins. Thus COPPERS could be used to set harder and potentially more interesting questions - making it a more flexible classroom tool.

The results of Experiments One and Two suggested that to improve performance, children should give more solutions per question on the computer than they produced before the intervention. This could affect the design of COPPERS in two ways: teachers could set minimum numbers of solutions, either on a per child or a per class basis or, more ideally, the system should monitor children's performance to adapt the minimum number of solutions as a user's expertise increases.

COPPERS' instructional goal was to teach children to consider alternative ways to answer problems. However, for many types of maths problems, it is useful to consider how the solutions differ (*e.g.* more effective, elegant, simpler, interesting, *etc.*). Indeed, as part of discussions with children after the studies, the experimenter often asked them questions about their solutions "What's your favourite? Which would be the teachers favourite? Which is the prettiest and ugliest? Which is the most fun? Which one would be best in a hurry" The majority of children responded with puzzlement to the possibility of judging solutions of any dimension other than right or wrong. To get students to recognise that there are many solutions to a problem is a necessary first step; the second is to encourage students to strategically choose solutions based upon reflection of their specific learning objectives. To encourage such reflection it will be necessary to involve peers and teachers in the way recommended by Lampert's approach to collaborative teaching.

Further issues in the design of COPPERS which have not been addressed by the evaluations studies include when to change the difficulty of a problem (currently governed by a simple performance measure student model and user choice), and what the major factors are that determine difficulty in these problems (currently defined as maximum number of partial products, maximum value of coin allowed and the maximum value of multiplier). These were primarily based on intuitive analysis of the domain. It seems likely that children would benefit from using COPPERS as a collaborative tool. This would allow them to share and discuss different solutions. There are some intuitively plausible ways to redesign it to support this approach to learning. For example, children could be given different money boxes and would need to exchange coins with one another. They could set challenges for each other. Such approaches would obviously require substantive re-evaluation.

CENTS

Fewer aspects of CENTS were changed during the research. The only change made as a result of an experiment was to adjust how much experience children received before the system required them to make the place value correction. Initially, this was required only after substantive experience with the system. However, given the poor order of magnitude results of the children in Experiment Three, the level of help was adjusted so that children were required to make this judgement much more often. This would appear to have been successful, as order of magnitude correction was much better in Experiments Four and Five.

One aspect of CENTS that could be further addressed is the role of the log book. Although children completed this (*e.g.* describing problems and estimates, their views of proximity and simplicity, *etc.*), no exercises were set which involved it directly. Thus, the 'explain' aspect of the 'predict-test-explain' cycle was under-utilised within CENTS.

Early in the design phase, it had been proposed to set questions using realistic contexts. However, this decision was not implemented as it was felt that too little was known about how the difficulty of estimates was affected by context. Further research would be needed before such a feature could be added.

One obvious expansion to CENTS would be to support more estimation strategies. During prototyping, five different reformulation strategies were created for addition, and these could be integrated into the system. If more idiosyncratic strategies (*e.g.* translation strategies) were to be supported, a different approach would be needed. In this case, it might be more appropriate to get the system to occasionally demonstrate a different way to solve the problem. It seems both difficult and undesirable to step children in a rule based fashion through strategies such as 'nice numbers' and translation to a more appropriate form - numbers are only nice if one sees the relation, and the form is only more appropriate if you are comfortable with it. The granularity of description of children's knowledge needed in a student model would introduce CENTS to all the well-known problems of student modelling in ITS (*e.g.* Self, 1990).

Introducing compensation strategies would be easier. Effective compensation is based upon accurate insight into the estimate - a skill much emphasised in CENTS. It would be possible to encourage children to post-compensate after they had seen the feedback from the computer in order to make their estimate more accurate. This feature could be reserved for older children whom Case & Sowder would predict would have the competencies to understand and use post-compensation.

General Issues

Key improvements to CENTS and COPPERS have been discussed. Even if these proposed new features were implemented, much further study would be required to assess whether the effects of these components had the desired effect upon learning. It should be re-stressed that COPPERS and CENTS are not intended as a stand-alone systems. Much of the important conceptual understanding would come from teaching and discussion with teachers and peers. Ideally, to test the impact of these systems on children's understanding in a more realistic manner, the systems would be introduced as part of normal classroom teaching. Although follow up exercises were discussed with teachers, full integration was not attempted within this research due to the need to tightly control how the system was used. In addition, as no attempt was made to compare the system to other forms of teaching (each system acting as its own control), the aims of these experiments did not require such a comparison.

9.5.2 Evaluation studies

The evaluation studies were designed to assess the impact of system components upon learning outcomes. Thus, the granularity of information collected was appropriate to that level of analysis. The current data do not provide answers to questions about how children's conceptual understanding has changed as a result of the intervention. This task would be more easily achieved for CENTS than for COPPERS. Sowder and Wheeler (1989) identified the conceptual knowledge associated with successful performance of computational estimation. They also provide examples of measurement instruments.

Assessing conceptual knowledge would be more difficult for coin problems. There does not appear to be an existing framework to describe this type of knowledge. It would be interesting to attempt a systematic analysis of the types of solutions children produce to these problems and the strategies they develop. A few strategies were obvious from the children's answers. A pre-test to post-test analysis was not possible as there were simply not enough right answers at pre-test. Examples of possible strategies include (illustrated for 46p):

- a least coin strategy - use the minimum number of coins. *e.g.* '20p 20p 5p 1p'
- a trading strategy - trade a coin which was part of a previous answers for others, *e.g.* '10p 10p 20p 5p 1p'
- a least types of coin strategy - use the minimum number of types of coin *e.g.* '23 x 2p'

- a factor approach - use a coin which is a factor of a close number *e.g.* '9 x 5p, 1p'

Further interview based research would be required to identify the strategies that children were using. It would be interesting to see whether types of answer are related to intervention measures and/or learning outcomes.

The CENTS experiments used a fairly complicated pen and paper test which attempted to measure both estimation and prediction accuracy. Given the importance for this research of accurately measuring children's understanding of prediction accuracy and the problems highlighted by Experiment Five (see section 8.9.2), it was argued that a new post-test measure might be desirable. One possibility would be to provide children with the estimates (and intermediate solutions to demonstrate the process) and ask them to assess the accuracy of the estimates. This would have two advantages: variations in estimation accuracy would be eliminated and children could focus on the prediction task which is both more difficult than the original estimation and which also tends not to be viewed as a 'real' mathematical task. A further possibility might be to initially only teach children the estimation strategies using CENTS. Subjects' estimation and prediction accuracy could then be measured as a baseline before further intervention which stressed insight into accuracy.

9.6 THE DESIGN OF MULTI-REPRESENTATIONAL LEARNING ENVIRONMENTS

This research has shown that learning with MERs will not always be effective. Consequently, in order to achieve the desired learning outcomes, design of multi-representational learning environments requires careful consideration. Designers of such software must first consider the same issues that any environment which intends to exploit external representations should evaluate (*e.g.* what representation should be used, what role it should play, how abstract it should be, whether should representations be constructed by users or provided, *etc.*). In addition, there are additional design issues unique to MERs.

In section 3.7, a framework of questions that designers should address when creating multi-representational learning environments was proposed. Although research in this field is not sufficiently mature to provide general answers for specific cases, it is argued that by using the techniques developed within this thesis, these issues can begin to be addressed. The following issues will be considered in this section:

- the purpose of MERS
- the similarity of representations
- the amount of information per representation
- automatic translation between representations
- how many representations should be used
- the ordering and sequencing of representations

9.6.1 Purpose of Multiple Representations

An initial question that should be asked is, what are the goals of employing MERs? Three broad claims for the advantages of MERs were identified in section 3.3: that MERs can be used to support different ideas or process, they can be used to constrain interpretation and they can promote a deeper understanding of the domain. Each have different implications for the design of learning environments. This section will describe examples from each of these different uses of MERs in turn.

<u>Different Ideas and Processes</u>

Two main uses of MERs were identified in this category. The first was that MERs can be designed so that each representation in the multi-representational system conveys some different information. This use is common when one representation is insufficient to carry all the required information or would be too complicated if it did so. Sometimes, information may be partially redundant between representations, in other cases, there is no redundancy between the information expressed by each representation.

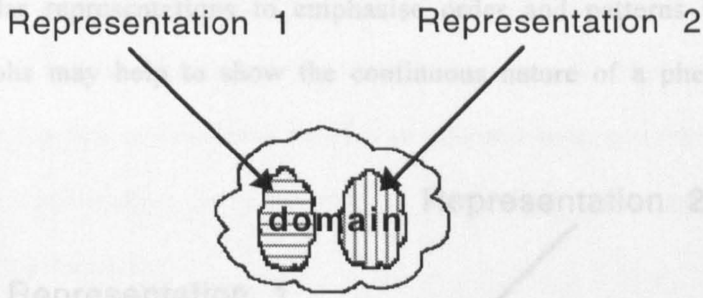


Figure 9.1 Using MERs to convey (completely) different information*

Figure 9.1 shows an abstract illustration of a learning environment that supports this form of MERs. Each representation in the system describes a different aspect of the domain. Note that there is no translation between the representations. The distance between the representation and the domain is intended to indicate the cognitive effort required to successfully use the representation.

This design is based on analysis of the experiments with CENTS (especially Experiment Three). It was demonstrated that when one representation was sufficient to learn the desired aspects of a domain, that presenting it alongside a second representation could interfere with successful learning. It was argued that this was due to the learning demands of translating between representations. Therefore in cases where representations are used to convey different aspects of the domain and no translation between representations is required (see later examples on constraints and abstraction for contrast), then encouraging learners to co-ordinate representations may in fact decrease learning outcomes. Thus, the additional learning demand of translation could be reduced either by only presenting one representation at a time or by letting the computer do any translation that is needed.

The second aspect of employing MERs to support different ideas and processes is when a designer aims to exploit the different computational properties of the alternative representations. For example, in some situations it may appropriate to

* length of the lines suggests the amount work needed to map between the representations or the representation and the domain. The shorter the line, the less work needed to make the mapping.

exploit tabular representations to emphasise order and patterns in numbers; in another, graphs may help to show the continuous nature of a phenomenon being examined.

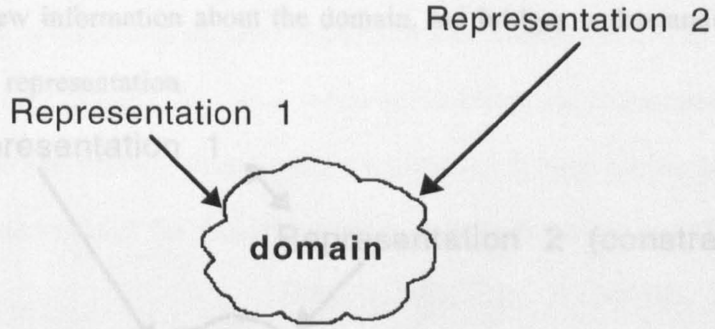


Figure 9.2 Using MERs with different computational properties

Again, based on the experiments with CENTS it is argued that the additional learning demand of translating between the representations could provide unnecessary load which would interfere with the designers aims (figure 9.2). Thus, if translation could be automatically provided by the system (*e.g.* tables automatically updating when graphs manipulated or vice versa) or if representations were presented sequentially rather than in a co-present fashion, learners should be less likely to be overburdened by the learning demands of translation.

Generally, when using MERs to support different information or computational properties, it would seem wise to reduce the third learning demand of translation to a minimum. The properties of the individual representations can then be analysed with respect to the first two learning demands of representations (format and operators of the representation and relation between the representation and the domain).

Constraining Interpretation

The second broad class of purposes of MERs identified in the thesis is to constrain interpretations of a situation. One way this may be achieved is to use a second representation to support interpretation of a more complicated, abstract or less familiar representation. For example, microworlds such as DM^3 (Henessey *et al.*, 1995) provide a simulation of a skater alongside a velocity-time graph. In such a situation, a common misunderstanding is that a straight line means no motion. This interpretation is not possible when the simulation shows the skater still moving. In cases such as this, the second more familiar or concrete representation is not intended to provide new information about the domain, but bridges understanding of the more complicated representation.

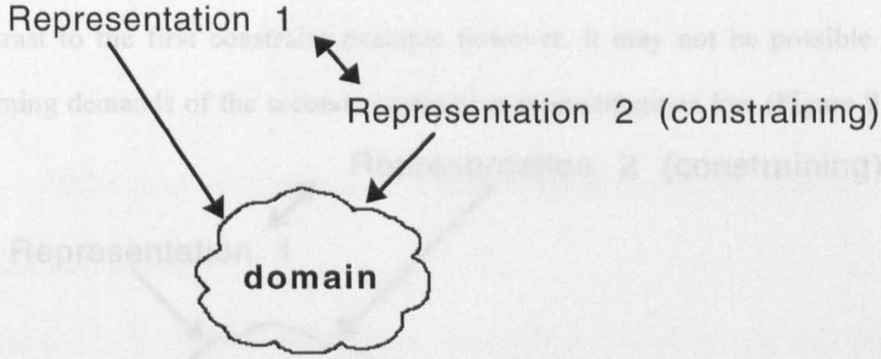


Figure 9.3 Using MERs to constrain interpretation of a less familiar representation

In contrast to the use of MERs to support different ideas and processes, it is crucial in this case that learners can co-ordinate the presented representations. Consequently, in order to achieve constraint on interpretation of a less familiar representation, designers need to ensure that complementary representations which aid translation are chosen. Again, there may be a case for the computer to support or to perform the translation between the representations in order to reduce the learning demands of translation. In addition the properties of the individual representations must be considered as the second representation should be easily understood in order to keep the learning demands of this representation as low as possible.

COPPERS uses the combination of the tabular representation and the row and column representation in this way. Experiment One demonstrated that children had improved learning outcomes when the tabular representation was presented to learners in addition to the row and column representation. The row and column representation is familiar to children of the intended age-group, however, the tabular representation is less familiar and requires children to make explicit the implicit arithmetical operations that are needed to make a correct interpretation of the table. Thus, the second constraining representation was chosen to be as familiar and easy to understand as possible. Translation between the representations was supported by the use of highlighting to signal the correspondences between the representations.

A further use of constraint between representations first introduced in section 3.3.2 is when constraints inherent in one representation affect the interpretation of another. In contrast to the first constraint example however, it may not be possible to keep the learning demands of the second constraining representations low (Figure 9.4).

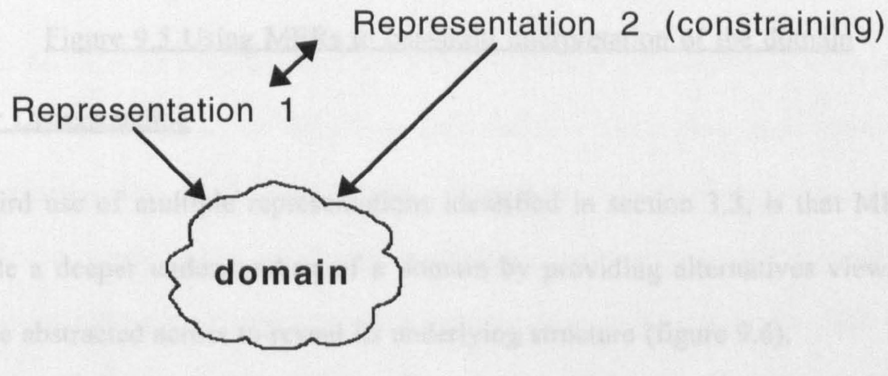


Figure 9.4 Using MERs to constrain interpretation by exploiting the properties of representations

COPPERS again provides an example of this type of design goal. A property of the less familiar tabular representation (order irrelevance) could be said to constrain interpretation of the row and column representations (which is order sensitive).

If learners are to be able to take advantage of this intended use, translation between the representations is crucial. Given the research with CENTS which showed that learners can have persistent difficulties in co-ordinating representations, then designers must consider how to support translation between representations. For example, this could be achieved by automatically translating between representations or by providing cues to help learners construct the appropriate mapping.

The final use of MERs for constraining interpretation is when the MERs are partially redundant. Thus information presented in each representation, when integrated, mutually constrain interpretation. In this case, as the constraint exists in the domain rather than in the representations, translation between the representation is not necessarily required. However, each representation must be fully understood: *i.e.* the first two learning demands must be met successfully (Figure 9.5).

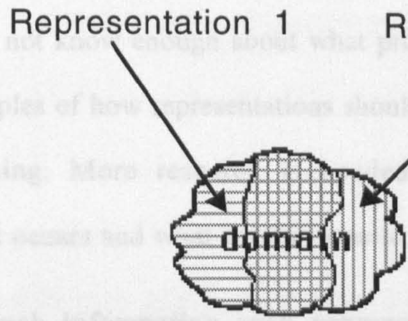


Figure 9.5 Using MERs to constrain interpretation of the domain

Deeper Understanding

The third use of multiple representations identified in section 3.3, is that MERs can promote a deeper understanding of a domain by providing alternatives views which must be abstracted across to reveal its underlying structure (figure 9.6).

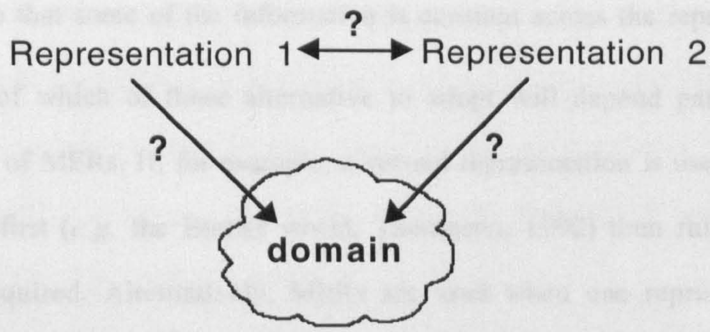


Figure 9.6 Using MERs to support abstraction

This goal provides designers with hard choices. If users fail to translate across representations, then abstraction can not occur. It was shown in Experiments 3 to 5 that learners find translating over representations which are superficially dissimilar to be difficult. This effect was found over long periods of time and even when the information in one representation was completely derivable from the other representation. However, in contrast to the other cases when translation between representations is desired, in this case translation between representations should not be made to easy. If alternatives representations do not provide sufficiently different views on a domain, then abstraction of invariances can not occur. Additionally, if the system performs all the translation for students, then they may not learn to translate for themselves.

As yet, we do not know enough about what properties affect translation to be able to produce principles of how representations should be combined in order to support this type of learning. More research is needed to examine how abstraction over representations occurs and what can be done to support it.

9.6.2 How much information each representation should express

Another factor to be considered in the design of multi-representational software is the degree of redundancy across the representations. At one extreme, each representation could express the same information. Here, the only difference would be in computational properties. At the other extreme, each representation could convey completely different information. Multiple representations may also be partially redundant, so that some of the information is constant across the representations.

The choice of which of these alternative to adopt will depend partially upon the intended use of MERs. If, for example, a second representation is used to display the actions of a first (*e.g.* the Blocks world; Thompson, 1992) then full redundancy is obviously required. Alternatively, MERs are used when one representation is not sufficient to clearly convey all of the desired information. In this case, there will almost always be less than full redundancy.

However, when MERs are used to display aspects of complicated situation to learners, it would be possible to vary the amount of redundancy. For example, CENTS allows for full, partial and no redundancy. In this case, designers must decide which situation will best support learning.

The level of redundancy which best supports learning seems to be an open research question. One possibility is that it is easier to learn complex ideas when each part is represented separately. Alternatively, it may be harder to learn with MERS that do not allow redundancy as the relation between representations (and therefore concepts) may be less obvious. The results of Experiment Five tentatively suggest that the initial acquisition of concepts may be facilitated when each representation expresses a different aspect of the situation. However, this result requires replication and extension, especially as no information about integration of information could be gathered.

It may also be the case that the information redundancy should change as a learner's expertise increases. For example, it may be better to use unique representations when first introducing them to concepts to allow learners to concentrate upon aspects of these concepts. Subsequently, we may wish to combine aspects of the domain within one representation so that interactions between the variables become more obvious.

9.6.3 Similarity between Representations

Multi-representational software can present representations in a variety of formats. Experiments reported in this thesis have involved representations that differ in terms of modality, amount and granularity of information. It was shown that when users were exposed to MERs with different formats, that behaviour with these representations changed. In particular, it was shown in Experiment Three that when failure of overlap between representations occurs at many different levels, then learning outcomes can be diminished.

The choice for a designer of the degree of similarity between representations is obviously related to the different purposes of representations. Thus, when representations are needed because of their distinct computational properties, there are bound to be differences in the format of these representations. The two feedback representations in COPPERS are a case in point. Equally, when choosing representations in order to encourage abstraction over representations, again, some differences in format will be required.

The primary consideration for a designer of multi-representation software is to balance the learning demands required by using representations of different format with the desired learning outcomes. Research on the way that different external representations support learning of specific tasks will obviously help to answer these questions (*e.g.* Bibby & Payne, 1993). The experiments presented in this thesis that aimed to address what factors influenced the ease of translation between representations are also relevant to this debate and the paradigm could be extended to other types of representations, tasks and learners. In section 9.7.2 an initial set of factors that may affect ease of translation between representations is proposed.

9.6.4 How many representations?

All the experiments reported in this thesis kept to the simplest case of MERs - two representations at a time. However, multi-representational software has been created which uses many more. The 'Visual Calculator' (Fox, 1988), for example, supplies five different representations simultaneously. A key question facing designers is how many representation should be employed in order to achieve a balance between the learning demands of the representations and the benefits they bring.

Using empirical techniques similar to those employed in this thesis may help resolve this question during the formative evaluation of a system. If the measured learning demands on the intended users for a particular representation (or combination of representations) are high, then as few representations as possible should be used. Thus in many cases it may not be appropriate to use MERs, since one representation may be sufficient. For example, Experiment Four showed that when children were given partially redundant representations, a highly effective strategy was to concentrate upon a single useful representation.

Another factor that must be considered is the purpose of the MERs. If MERs are used to constrain interpretation, one extra easily understood representation may be sufficient (see figure 9.3 above). Again, the more complicated situation arises when MERs are used to expose the structure of the underlying domain. Further domain specific research would be needed in order to analyse how MERs allow learners to begin to expose the domain invariances. This problem is reminiscent of the problem facing any learning situation where the aim is to support abstraction and generalisation, and as with these situations, finding out how learners come to understand the structure of the domain is likely to prove difficult. One possible way to begin to answer this question would be to exploit measures of representational co-ordination. When learners faced with new representations of some aspect of the domain quickly converge their behaviour, this may be evidence they have abstracted over the representations to build a model of the domain.

A further issue to consider is how many representations should be used simultaneously. For example, if the aim is to use different representation to support distinct ideas, learning demands would be reduced by using one representation at a time (Figure 9.1). Thus, each representation should be understood before introducing another representation. In addition, design in this area will be particularly affected by practical constraints. The monitors of most school machines are considerably smaller than those available to designers, and their resolution may not be as high.

9.6.5 Automatic Translation

Another question facing designers is whether to provide automatic linking between representations. Here, one acts in one representation and sees the results of these actions in another (*e.g.* the Blocks Microworld, Thompson 1992). This is currently an open research question. Kaput (1992) proposes that many benefits follow from dynamic linking as the computer reduces the cognitive load for the user. However, it may also be the case that over-automation does not encourage a user to actively translate across representations.

Experiments measuring representational co-ordination may help to answer this question. One possibility would be to present two representations - either linked or unlinked. Then, children would be given a new representation to use. The children who have better understanding of how the core features of a domain could be expressed would be expected to converge their behaviour faster on this the new representation.

When representations are not dynamically linked, then we may still want the system to support a user's mapping across the representations. COPPERS, for example, uses highlighting to show users how two informationally redundant representations relate to each other. Mapping could either be automatic or under the learner's control. For example, given a table of values, users may wish to select a row and then be shown the equivalent location on a graph.

Conditions for dynamic linking and mapping also vary with different degrees of redundancy. If there is less than full redundancy between representations, users would either have to work in the representation with the most information or provide the computer with extra information to disambiguate their intended action.

Designing computer support for translation across multiple representations requires researchers to consider many different issues. The answers to some of these questions could be provided if designers knew how difficult users would find co-ordinating and integrating information across the representations. Thus, during formative evaluation of software, representational co-ordination could be measured. This could then be used to determine the degree of support given within the finished system.

9.6.6 Ordering and Sequencing Representations

In systems where all the MERs are not presented simultaneously (unlike COPPERS and CENTS), two further issues arise - in what order to present representations and when to add a new representation. Decisions about ordering representations may be based upon analysis of the domain. For example, Kaput (1994) uses MathsCar to support the development of calculus. He argues that understanding is best supported by introducing integration before differentiation, and hence proposes representations such as velocity-time graphs should be introduced before position-time graphs. Another common approach is to move from concrete representations to increasingly symbolic representations, mimicking Bruner's modes of representation (*i.e.* from enactive through iconic to symbolic). This approach has been often been taken literally, although Bruner did not intend it to be interpreted in this way (Behr, Harel, Post & Lesh, 1992).

A related issue is whether the redundancy of information should be increased or decreased as more representations are made available. Experiment Five provided tentative evidence to suggest that limiting redundancy may help when learning the initial aspects of a task or concept. Thus, it may be beneficial to increase the redundancy over representations as a learner's expertise grows. It could be the case that as long as children come to see fully understanding the MERs (format, operators and links between each other and the domain), the order in which they were introduced may be irrelevant.

Even when this issue of how to order representations has been addressed, we are still faced with the question of when to change a representation or introduce a new one. One possible solution is to allow learners to make this choice. For example, the switchER system (Cox, 1996) allows users to move at will between their self-created representations. Cox argues that this can be beneficial as it can help learners to resolve impasses. However, there also is evidence to suggest that switching between representations can also be symptomatic of less understanding.

Another possibility is that learners should switch when they have exhausted all of the information available in the representation they are currently using for problem solving - Graphs and Tracks (Trowbridge, 1989) exploits this technique to good effect. For example, it suggests that users should switch from a velocity-time to a distance-time graph in order to gain information about the represented object's starting position.

Finally, the system may take responsibility for this decision. In this case, the task for the system is to determine when users have learnt all they can about the domain with the given representations, but not switch so soon (or so often) that the learning demands of the new representations overburden the user. One suggestion is to provide a new representation when the learner's behaviour is still flawed with respect to the domain but has converged over the current representations. This suggests that a new representation might be useful (to help debug or introduce domain knowledge), and would not over burden users as they have already learnt about the representations. The empirical methods that have been used to examine effectiveness of learning environments in this thesis could also be used to monitor this understanding dynamically within multi-representational software.

9.7 FUTURE WORK

Many suggestions have already been given in the previous section for immediate directions in which to take this research. Expansions to the systems were also have also been considered (9.5). This section will concentrate upon some additional longer term research issues raised by this thesis.

9.7.1 Multiple Solutions and Computational Estimation

In section 9.5.2, it was proposed that one extension to the research with COPPERS would be to examine children's solution to coin problems in more detail. It was found that children used less simplistic solutions to problems after the intervention. Typically, they moved away from a least coins strategy. This is interesting in light of Resnick's suggestion that the place value understanding develops though understanding non-canonical decompositions (*e.g.* 22 = 1 ten and 12 ones) (Resnick, 1983). Thus, it would be interesting to see whether children who created more complicated coin problems, showed better understanding of place value. It might also be the case that experience with COPPERS helped children gain this understanding.

As a consequence of collecting pre-test and post-test data to examine how CENTS influenced learning outcomes, a large amount of data was gathered on primary school children's understanding of computational estimation. This raised questions that would be interesting to follow up but which were beyond the remit of this thesis. For example, in all the experiments, children showed a disinclination to produce an estimate by rounding up the numbers to produce an intermediate solution. In addition, the author is not aware of any research which has examined children's insight into accuracy in estimation. The relation between answer and prediction accuracy deserves further study.

These studies also raised the question of what it means to understand computational estimation. The vast majority of research on computational estimation has simply examined estimation and mathematical skills - for example, what strategies are used, what mathematical skills are related to computational estimation. Sowder and Wheeler's (1989) assessed the conceptual and affective components of estimation. What children know about an estimate they have produced is unclear.

9.7.2 Multiple Representations

One of the major issues to emerge from this thesis has been the need to consider the factors that affect translation across representations. It was proposed in section 7.4 that the more the format and operators vary across representations, then the more difficult learners would find translation over the representations. In the cases described in this thesis, the most important difference was between mathematical and pictorial representations. For any given task, other plausible candidates that may strongly influence co-ordination include:

- the modality of the representations - propositional v graphical
- whether representations are static or dynamic
- including representations that differ in levels of abstraction
- the degree of redundancy across representations
- whether the representations encourage different strategies
- any differences in labelling and symbols
- alternative uses of representations *e.g.* display v action
- variance between the resolutions of presented information

However, a definitive statement of these factors would need to be predicated upon an integrative taxonomy of representations. As discussed in section 3.2, although there are many candidates for classifying representations but no one approach is as yet completely satisfactory.

In addition to the nature of the representations, the style of interface to the representation may also affect co-ordination. Recent research has demonstrated that different interfaces can influence what users learn. Consequently, some researchers are now arguing for a move from direct manipulation interfaces in educational technology (*e.g.* Gilmore, 1996) For example, Svendsen (1991) found that direct manipulation interfaces resulted in poorer performance that command lines interfaces for solving Tower of Hanoi problems. Churchill & Ainsworth (1995) argued that designers of computer-based learning environments often do not give sufficient attention to the way actions on representations are supported.

In addition to properties of the representations, learner characteristics may influence co-ordination. For example:

- a learner's familiarity with the representations
- a learner's familiarity with the domain
- a learner's cognitive style
- a learner's general aptitude in that domain
- a learner's age

It seems probable that if learners know the format and operators of each of the representations they are given, then learning to translate across the representations will occur more rapidly. This is also true of their domain knowledge. Consequently, it is argued that lower the learning demands are on other parts of the task, the more attention can be focused on translation.

Although this thesis studied children's understanding of MERs, it was not a developmental thesis in that it did not examine age related changes in this understanding. It is argued that children's performance can be seen as a characteristic of novices in a domain. However, it does not seem implausible that in general younger children will find co-ordinating representations particularly difficult.

Finally, the issue of cognitive style may well be relevant. Oberlander *et al.* (1996) suggest that one distinguishing characteristic of people who were classified as diagrammatic reasoners may in fact be able to translate information across representations more successfully.

Thus, to predict how easy it will be for someone to understand the relation between a set of presented representations for a given task, both the individual and the representation's characteristics will need to be considered.

Another question of particular concern is to identify what is different about learning with MERs. In particular, Kaput (1989, 1992) makes the strong claim that learning with multiple representations generates a robust, flexible, *deeper* understanding of a subject. However, it is difficult to know how to evaluate whether children have developed a deeper understanding of the domain. In practise it may be very hard to separate this effect from when learning has occurred because of one 'perfect' representation. Equally, children may learn the relation between each representation and the domain without learning to translate across the representations. For example, children could learn in which situations to use a velocity-time graph rather than a distance-time graph to reason about motion, but they might never understand that if the distance graph was differentiated you would get the velocity graph.

Identifying how this depth of understanding manifests itself and designing measurement instruments that examine this would allow us to both test these claims more precisely and provide insight into the development of expertise in an area. One approach may to examine when learners are able to map their knowledge onto a new representation, or even create a new representation to express their knowledge of a domain. This might show whether there were key features of the domain that children were missing. For example, in the case of CENTS, children might create a representation which expresses direction and magnitude but which is based on absolute rather than relative understanding.

A further significant issue for this thesis has been the importance of assessing children's understanding of the various learning demands they have faced when using multiple representations. Knowledge of the domain and of the representations have been considered separately to the knowledge of the relation among representations (representational co-ordination). These measures have been used to predict and explain learning outcomes. It has also been proposed as way of uncovering design principles for multi-representational software. Consequently, the generality of this approach needs to be evaluated. Using this method with learning environments which operate in different domains (such as chemistry and science) would extend this work. In addition, whilst it has been argued that children represent a general instance of novice performance, it would be useful to extend this framework to adult learning.

Two different methods of assessing representational co-ordination were developed in this thesis. They were based on asking users to perform the same actions on two co-present representations and then measuring the similarity of their behaviour. Schwartz & Dreyfus (1993) proposed a third measure of integration which was to score when users translated all of the available information from one representation to a new one that replaced it (the passage index). However, all of these three techniques only work when there is some redundancy across representations and when the representations are used for action rather than display. Further ways of analysing co-ordination and integration of information from representations would be needed in order to cover alternative uses of MERs. Schoenfeld *et al.* (1993) examined this issue using detailed microgenetic analysis of one learner's understanding of the connection between representations. If research is interested in finding out how learners translate between representations rather than under what conditions translation can successfully be achieved, then this level of analysis will be crucial.

A final fundamental issue raised by this thesis was how people learn the relationship between different representations. Experiments in this thesis have demonstrated conditions under which it is more or less easy to translate across representations. It has been argued that without any additional support it is the similarity between two representations' format and operators that will influence ease of co-ordination. It is also expected that co-ordination will be affected by the learners' familiarity with the representations. However, these experiments have not allowed us to understand the processes children (and adults) use when learning the relation between MERs. A number of possible bases for exploring this question exist, for example, models of analogical reasoning (*e.g.* Gentner, 1989), Plötzner's SEPIA model of integrating information from qualitative and quantitative multiple (internal) representations (Plötzner, 1995). One possible future direction of this research is to use multi-representational learning environments as a base to collect protocol data which could subsequently inform the development of a computational model.

9.8 GENERAL SUMMARY

This thesis has reported the development and evaluation of two mathematical learning environments that are related to the development of number sense. The first concerned multiple solutions to mathematical problems and the second computational estimation. Both learning environments were shown to effectively teach children these aspects of primary mathematics. Experience with COPPERS significantly improved children's performance at producing multiple solutions to coin problems. CENTS was shown to improve children's computational estimation strategies and the accuracy of their estimates.

The learning environments were developed to explore theories of instruction. Key design features were systematically altered, and children's computer use and learning outcomes measures were analysed. In particular, the effects of multiple external representations was examined. Initial research suggested that children as young as six could benefit from learning with multiple representations. Further research developed empirical techniques to analyse representation use which predicted and explained learning outcomes. Analysis of the roles and learning demands of multiple representations combined with these experimental findings generated a framework in which to consider the design of effective multi-representational software.

References

Adelson, B. (1981). Problem solving and the development of abstract categories in programming languages. Memory and Cognition, 9, 422-433.

Ainsworth, S. E. (1992) COPPERS: A mathematical microworld for multiplication. M.Sc. dissertation, University of Sussex.

Anzai, Y. (1991). Learning and use of representations for physics expertise. In K. Anders-Ericsson & J. Smith (Eds.), Towards a general theory of expertise: Prospects and limits. Cambridge: Cambridge University Press.

Baroody, A. J. (1987). Children's mathematical thinking: A developmental framework for pre-school, primary and special education teachers. New York: Teachers College Press.

Barwise, J., & Etchemendy, J. (1992). The language of first order logic: including the program Tarski's world 4.0. Center for the Study of Language and Information: CSLI lectures notes no. 34.

Basic Number Screening Test (1976), SevenOaks: Hodder and Stoughton.

Behr, M. J., Harel, G., Post, T., & Lesh, R. (1992). Rational number, ratio and proportion. In D. A. Grouws (Eds.), Handbook of teaching and learning mathematics (pp. 296-333). New York: Macmillan Publishing Company.

Bell, A., Greer, B., Grimison, L., & Mangan, C. (1989). Children's performance on multiplicative word problems: Elements of descriptive theory. Journal for Research in Mathematics Education, 20(5), 434-449.

Bestgen, B. J., Reys, R. E., Rybolt, J. F., & Wyatt, J. W. (1980). Effectiveness of systematic instruction on attitudes and computational estimation skills of preservice elementary teachers. Journal for Research in Mathematics Education, 11, 124-136.

Bibby, P. A., & Payne, S. J. (1993). Internalization and the use specificity of device knowledge. Human-Computer Interaction, 8(1), 25-56.

Bibby, P., & Payne, S. (1996). Instruction and practice in learning to use a device. Cognitive Science, 20(4), 539-578.

Borba, M. C. (1994). A model for student's understanding in a multi-representational software environment. In Proceedings of the 18th International Conference for the Psychology of Mathematics Education, vol. 2 (pp. 104-111). Lisbon.

Boulton-Lewis, G. M., & Halford, G. S. (1990). An analysis of the value and limitation of mathematical representations used by teachers and young children. In Proceedings of the 14th International Conference for the Psychology of Mathematics Education, (pp. 199-206).

Brna, P. (1996). Can't see the words for the tree: Interpretation and graphical representation. In I.E.E Colloquium on Thinking with Diagrams Digest no 96/010. London.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. Cognitive Science, 2, 155-192.

Bruner, J. (1966). Towards a theory of instruction. New York: W.W. Norwood & Company inc.

Case, R. (1985). Intellectual development. Orlando: FL: Academic Press.

Case, R., & Sowder, J. T. (1990). The development of computational estimation: A neo-piagetian analysis. Cognition and Instruction, 7(2), 79-104.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. British Journal of Educational Psychology, 62(2), 233-246.

Cheng, P. C.-H. (1996a). Functional roles for the cognitive analysis of diagrams in problem solving. In G. W. Cottrell (Ed.), Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society (pp. 207-212). Hillsdale, NJ: LEA.

Cheng, P. C.-H. (1996b). Law encoding diagrams for instructional systems. Journal of Artificial Intelligence in Education, 7(1), 33-74.

Cheng, P. C.-H. (1996c). Scientific discovery with law encoding diagrams. Creativity Research Journal, 9(2&3), 145-162.

Chi, M., T. H., Feltovich, P. J., & R., G. (1981). Categorization and representation in physics problems by experts and novices. Cognitive Science, 5, 121-152.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self explanations: How students study and use examples in learning to solve problems. Cognitive Science, 5, 145-182.

Churchill, E. F., & Ainsworth, S. E. (1995). Making claims about teaching systems. In K. Nordy, P. Helmersen, D. J. Gilmore, & S. Arnesen (Ed.), Proceedings of Interact 95, (pp. 415-418). Lillehammer.

Confrey, J. (1992). Function Probe©. Santa Barbara, CA: Intellimation library for the Macintosh.

Confrey, J. (1994). Six approaches to transformations of functions using multi-representational software. In Proceedings of the 18th International Conference for the Psychology of Mathematics Education, vol. 2 (pp. 217-224). Lisbon.

Confrey, J., & Smith, E. (1992). Revised accounts of the function concept using multi-representational software, contextual problems and students paths. In Proceedings of the 16th International Conference for the Psychology of Mathematics Education, vol. 1 (pp. 153-160). New Hampshire.

Cox, R. (1996) Analytical reasoning with multiple external representations. unpublished Ph.D. thesis, University of Edinburgh.

Cox, R., & Brna, P. (1995). Supporting the use of external representations in problem solving: The need for flexible learning environments. Journal of Artificial Intelligence in Education, 6(2/3), 239-302.

The CSMS Mathematics Team (1981). Children's understanding of mathematics: 11-16. London: John Murray (Publishers) Ltd.

Cronbach, L.J. & Snow, R.E. (1977) Aptitudes and Instructional Methods. New York: Irvington.

De Corte, E., & Verschaffel, L. (1987). The effect of semantic structure on first graders's strategies for solving addition and subtraction word problems. Journal for Research in Mathematics Education, 18(5), 363-381.

Dienes, Z. (1973). The six stages in the process of learning mathematics. Slough, UK: NFER-Nelson.

Dowker, A. (1989). Computational estimation by young children. In Paper presented at the British Society for Research into Learning Mathematics. Brighton Polytechnic.

Dowker, A. (1992). Computational estimation strategies of professional mathematicians. Journal for Research in Mathematics Education, 23(1), 45-55.

Dowker, A. (1993). Young children's views of other people's estimates. In Paper presented at the BPS Developmental Section Conference. Birmingham.

Dowker, A. (1996). Arithmetical estimation: When does it begin. Poster presented at BPS developmental section conference. Oxford.

Dufour-Janvier, B., Bednarz, N., Belanger, M. & (1987). Pedagogical considerations concerning the problem of representation. In C. Janvier (Eds.), Problems of representation in the teaching and learning of mathematics (pp. 109-122). Hillsdale, NJ: LEA.

Dugdale, S. (1982). Green globs: A micro-computer application for graphing of equations. Mathematics Teacher, 75, 208-214.

Duncker, K. (1945). On problem solving. Psychological Monographs, 58(5).

Ehrlich, K., & Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. Journal of Verbal Learning and Verbal Behaviour, 21, 296-306.

Elsom-Cook, M. (1990). Guided discovery tutoring. London: Paul Chapman Publishing Ltd.

Fischbein, E., Drei, M., Nello, M. S., & Merino, M. S. (1985). The role of implicit models in solving verbal problems in multiplication and division. Journal for Research in Mathematics Education, 16(1), 3-17.

Forrester, M. A., Latham, J., & Shire, B. (1990). Exploring estimation in young primary school children. Educational Psychology, 10(4), 283-300.

Fox, M. (1988) Theory and design for a visual calculator for arithmetic, CITE technical report no. 32, Institute of Educational Technology, Open University.

Fuson, K. C. (1992). Research on whole number addition and subtraction. In D. Grouws (Eds.), Handbook of teaching and learning mathematics (pp. 243-275). New York: Macmillan.

Genter, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), Similarity and analogical reasoning (pp 199-242). Cambridge: Cambridge University Press.

Gilmore, D. J., & Green, T. R. G. (1984). Comprehension and recall of miniature programs. International Journal of Man-Machine Studies, 21, 31-48.

Gilmore, D. J. (1996). The relevance of HCI guidelines for educational interfaces. Machine-Mediated Learning, 5(2), 119-133.

Green, T. R. G. (1989). Cognitive dimensions of notations. In A. Sutcliffe & L. Macaulay (Ed.), People and Computers V. (pp 443-460). Cambridge: Cambridge University Press.

Green, T. R. G. (1990). The cognitive dimension of viscosity: A sticky problem for HCI. In D. Diaper, D. Gilmore, G. Cockton, & B. Shackel (Ed.), Human-Computer Interaction - INTERACT 90. (pp 79-86). Amsterdam: Elsevier Science.

Green, T. R. G., Petre, M., & Bellamy, R. K. E. (1991). Comprehensibility of visual and textual programs: a test of superlativism against the 'match-mismatch' conjecture. In Proceedings of Empirical Studies of Programmers: Fourth Workshop (pp. 121-146).

Green, T. R. G., & Petre, P. (1996). Usability analysis of visual programming environments: a `cognitive dimensions' framework. Journal of Visual Languages and Visual Computing, 7, 134-174.

Greer, B. (1992). Multiplication and division as models of situations. In D. Grouws (Eds.), Handbook of teaching and learning mathematics (pp 276-295) New York: Macmillan Publishing Company.

Hennessy, S., O'Shea, T., Evertsz, R., & Floyd, A. (1989). An intelligent tutoring system approach to teaching primary mathematics (CITE Report No. 93). Institute of Educational Technology, Open University.

Hennessy, S., Twigger, D., Driver, R., O'Shea, T., O'Malley, C., Byard, M., Draper, S., Hartley, R., Mohamed, R. & Scanlon, E. (1995) Design of a computer-augmented curriculum for mechanics. International Journal of Science Education, 17(1), 75-92.

Hofstadter, D. R. (1985). Metamagical themas: Questing for the essence of mind and pattern. London: Penguin Books.

Howe, C.J., Rodgers, C., and Tolmie, A. (1990) Physics in the primary school; peer interaction and the understanding of floating and sinking. European Journal of Psychology of Education, 459-75.

Hoz, R., & Harel, G. (1989). The facilitating role of table forms in solving algebra speed problems: Real or imaginary. In Proceedings of the 13th International Conference for the Psychology of Mathematics Education, vol. 2 (pp. 123-130). Paris.

Hughes, M. (1986). Children and number. Oxford: Blackwell.

Janvier, C. (1987). Translation processes in mathematics education. In C. Janvier (Ed.), Problems of representation in the teaching and learning of mathematics (pp. 27-32). Hillsdale, NJ: LEA.

Kaput, J. J. (1987). Towards a theory of symbol use in mathematics. In C. Janvier (Ed.), Problems of representation in the teaching and learning of mathematics (pp. 159-196) Hillsdale, NJ: LEA.

Kaput, J. J. (1989). Linking representations in the symbol systems of algebra. In S. Wagner & C. Kieran (Eds.), Research issues in the learning and teaching of Algebra (pp. 167-194).Hillsdale, NJ: LEA.

Kaput, J. J. (1992). Technology and mathematics education. In D. A. Grouws (Ed.), Handbook of teaching and learning mathematics (pp. 515-556) New York: Macmillan Publishing Company.

Kaput, J. (1994). Democratizing access to calculus: New routes using old roots. In A. Schoenfeld (Eds.), Mathematical thinking and problem solving (pp. 77-156). Hillsdale, NJ: LEA.

Kaput, J., & Maxwell-West, M. (1994). Missing-value proportional reasoning problems: Factors affecting informal reasoning patterns. In G. Harel & J. Confrey (Eds.), The development of multiplicative reasoning in the learning of mathematics (pp. 237-292). Albany, New York: State University of New York Press.

Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in Geometry. Cognitive Science, 14, 511-550.

Laborde, C. (1996). Towards a new role of diagrams in dynamic geometry? In Proceedings of the European Conference of Artificial Intelligence in Education, (pp. 350-356). Lisbon.

Lampert, M. (1986a). Knowing, doing, and teaching multiplication. Cognition and Instruction, 3(4), 305-342.

Lampert, M. (1986b). Teaching multiplication. Journal of Mathematical Behaviour, 5(3), 241-280.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. American Educational Research Journal, 27(1), 29-63.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11, 65-99.

LeFevre, J., Greenham, S. L., & Waheed, N. (1993). The development of procedural and conceptual knowledge in computational estimation. Cognition and Instruction, 11(2), 95-132.

Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. J. Derry & S. P. Lajoi (Eds.), Computers as cognitive tools (pp. 75-105). Hillsdale, NJ: LEA.

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 (pp 33-40). Hillsdale, NJ: LEA.

Levine, D. R. (1982). Strategy use and estimation ability of college students. Journal for Research in Mathematics Education, 13, 350-359.

Lohse, G. L., Biolsi, K., Walker, N., & Rueler, H. (1994). A classification of visual representations. Communications of the A.C.M., 37(12), 36-49.

Luchins, A. S., & Luchins, E. H. (1950). New experiemental attempts at preventing mechanization in problem solving. Journal of General Psychology, 42, 279-297.

Markovits, Z. (1989). Reactions to the number sense conference. In J. T. Sowder & B. P. Schappelle (Eds.), Establishing foundations for research on number sense and related topics: report of a conference, (pp. 78-81). San Diego: San Diego State University Center for Research in Mathematics and Science Education.

Markovits, Z., & Sowder, J. T. (1988). Mental computation and number sense. PME-NA: Proceedings of the Tenth Annual Meeting, (pp. 58-64). Northern Illinois University: DeKalb.

Markovits, Z., & Sowder, J. T. (1994). Developing number sense: An intervention study in grade 7. Journal for Research in Mathematics Education, 25(1), 4-29.

Milheim, W. D., & Martin, B. L. (1991). Theoretical bases for the use of learner control: Three different perspectives. Journal of Computer-Based Instruction, 18(3), 99-105.

Millwood, R. (1996). Issues in multimedia educational software design. Paper presented at the British HCI Group Usability and Educational Software Design London.

Morgan (1990). Factors affecting children's estimation strategies and success in estimation. In Proceedings of the 14th International Conference for the Psychology of Mathematics Education, vol. 3 (pp. 265-272). Mexico.

Mulholland, P., & Domaigne, J. (in press). Teaching programming at a distance: The internet software visualisation laboratory. Journal of Interactive Media

Mulligan, J. T. (1992). Children's solutions to multiplication and division word problems: A longitudinal study. In Proceedings of the 16th International Conference for the Psychology of Mathematics Education, vol. 2 (pp. 144-151). New Hampshire.

National Council of Teachers of Mathematics, (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: NCTM.

National Research Council (1989). Everybody counts: A report to the nation on the future of mathematics education. Washington, DC: National Academy Press

Nesher, P. (1987). Towards an instructional theory: The role of students' misconceptions. For the Learning of Mathematics, 7(3), 33-40.

Nunes, T. (1992). Ethnomathematics and everyday cognition. In D. Grouws (Eds.), Handbook of teaching and learning mathematics (pp 557-574) New York: Macmillan.

Nunes, T., Schliemann, A.-L., & Caraher, D. (1993). Street mathematics and school mathematics. New York: Cambridge University Press.

Nunes, T., & Bryant, P. (1996). Children doing mathematics. Oxford: Blackwell.

Oberlander, J., Cox, R., Monaghan, P., Stenning, K., & Tobin, R. (1996). Individual differences in proof structures following multimodal logic teaching. In G. W. Cottrell (Ed.), Proceeding of the Eighteenth Annual Conference of the Cognitive Science Society, (pp 201-206). Hillsdale, NJ: LEA.

Oliver, M., & O'Shea, T. (1996). Using the model-view-controller mechanism to combine representations of possible worlds for learners of modal logic. In Proceedings of the European Conference of Artificial Intelligence in Education, (pp. 357-363). Lisbon.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), Cognition and categorization (pp 259-303). Hillsdale, NJ: LEA.

Plötzner, R. (1995). The construction and coordination of complementary problem representations in physics. Journal of Artificial Intelligence in Education, 6(2/3), 203-238.

Petre (1993). Using graphical representations requires skill, and graphical readership is an acquired skill. In R. Cox, M. Petre, J. Lee, & P. Brna (Eds.), Proceedings of AI-ED93 workshop Graphical Representations, Reasoning and Communication, (pp. 55-58). Edinburgh.

Petre, M., & Green, T. R. G. (1993). Learning to read graphics: Some evidence that 'seeing' an information display is an acquired skill. Journal of Visual Languages and Computing, 4, 55-70.

Philipp, R. A., Flores, A., Sowder, J. T., & Schappelle, B. P. (1994). Conceptions and practices of extraordinary mathematics. Journal of Mathematical Behaviour, 13(2), 155-180.

Pimm, D. (1995). Symbols and meanings in school mathematics. London: Routledge

Price, M., & Foreman, J. (1989). A prIME experience. Mathematics in School (January), 2-5.

Resnick, L. B. (1983). A developmental theory of number understanding. In H. P. Ginsburg (Eds.), The development of mathematical thinking (pp. 109-151). Orlando:FL: Academic Press.

Resnick, L. B. (1989). Defining, assessing and teaching number sense. In J. T. Sowder & B. P. Schappelle (Ed.), Establishing foundations for research on number sense and related topics: report of a conference, (pp. 35-39). San Diego: San Diego State University Center for Research in Mathematics and Science Education.

Resnick, L. (1992). From protoquantities to operators: Building mathematical competence on a foundation of everyday knowledge. In G. Leinhardt, R. Putnam, & R. A. Hattrup (Eds.), Analysis of arithmetic for mathematics education (pp. 375-429). Hillsdale, NJ: LEA.

Resnick, L., & Omanson, S. (1987). Learning to understand arithmetic. In R. Glaser (Ed.), Advances in instructional psychology volume 3. (pp. 41-95). Hillsdale, NJ: LEA.

Reys, R. (1989). Some personal reflections on the conference. In J. T. Sowder & B. P. Schappelle (Ed.), Establishing foundations for research on number sense and related topics: report of a conference, (pp. 65-66). San Diego: San Diego State University Center for Research in Mathematics and Science Education.

Reys, R. E., Rybolt, J. F., Bestgen, B. J., & Wyatt, J. W. (1982). Processes used by good computational estimators. Journal for Research in Mathematics Education, 13, 183-201.

Reys, R. E., Reys, B. J., Nohda, N., Ishida, J., Yoshikawa, S., & Shimizu, K. (1991). Computational estimation performance and strategies used by fifth and eight grade Japanese students. Journal for Research in Mathematics Education, 22(1), 39-58.

Reeves, T. C. (1993). Pseudoscience in computer-based instruction: The case of learner control literature. Journal of Computer-Based Instruction, 20(2), 39-46.

Roberts, M. J., Wood, D. J., & Gilmore, D. J. (1994). The sentence-picture verification task: Methodological and theoretical difficulties. British Journal of Psychology, 85, 413-432.

Rubenstein, R. N. (1985). Computational estimation and related mathematical skills. Journal for Research in Mathematics Education, 16(2), 106-119.

Santos, M. S. (1994). Students approaches to solve three problems that involve various methods of solution. In Proceedings of the 18th International Conference for the Psychology of Mathematics Education, vol. 4 (pp. 193-200). Lisbon.

Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work? International Journal of Human-Computer Studies, 45, 185-213.

Schoen, H. L., Frisen, C. D., Jarrett, J. A., & Urbatsh, T. D. (1981). Instruction in estimating solutions of whole number computations. Journal for Research in Mathematics Education, 12, 165-178.

Schoenfeld, A. (1986). On having and using geometric knowledge. In J. Hiebert (Eds.), Conceptual and procedural knowledge: The case of mathematics (pp. 225-265). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A. H. (1987). What's all this fuss about metacognition. In A. H. Schoenfeld (Eds.), Cognitive Science and Mathematics Education (pp 189-215). Hillsdale, NJ: LEA.

Schoenfeld, A. H. (1988). Problem solving in context(s). In R. I. Charles & E. A. Silver (Eds.), The teaching and assessing of mathematical problem solving (pp. 82-92). Hillsdale, NJ: LEA.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem Solving, metacognition and sense making in mathematics. In D. Grouws (Eds.), Handbook of teaching and learning mathematics (pp. 334-270). New York: Macmillan.

Schoenfeld, A. H., Smith, J. P., & Arcavi, A. (1993). Learning: The microgenetic analysis of one student's evolving understanding of a complex subject matter domain. In R. Glaser (Ed.), Advances in instructional psychology volume 4 (pp 55-175) Hillsdale, NJ: LEA.

Schwartz, B., & Dreyfus, T. (1993). Measuring integration of information in multi-representational software. Interactive Learning Environments, 3(3), 177-198.

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. The Journal of the Learning Sciences, 4(3), 321-354.

Self, J. A. (1990). Bypassing the intractable problem of student modelling. In C. Frasson & G. Gauthier (Eds.), Intelligent tutoring systems (pp. 107-123). Norwood, New Jersey: Ablex Publishing Company.

Snow, R. E., & Yalow, E. (1982). Education and Intelligence. In R. J. Sternberg (Eds.), A handbook of human intelligence (pp 493-586). Cambridge: Cambridge University Press.

Sowder, J. T. (1989). Discussion notes. In J. T. Sowder & B. P. Schappelle (Ed.), Establishing foundations for research on number sense and related topics: report of a conference, (pp. 7-24). San Diego: San Diego State University Center for Research in Mathematics and Science Education.

Sowder, J. (1992). Estimation and number sense. In D. A. Grouws (Eds.), Handbook of research on teaching and learning mathematics (pp 371-389) Macmillan Publishing Company.

Sowder, J. T. (1992). Making sense of numbers in school mathematics. In G. Leinhardt, R. Putnam, & R. A. Hattrup (Eds.), Analysis of arithmetic for mathematics education (pp. 1-51). Hillsdale, NJ: LEA.

Sowder, J. T. (1995). Instructing for rational number sense. In J. T. Sowder & B. P. Schappelle (Eds.), Providing a foundation for teaching mathematics in the middle grades (pp 15-30). Albany: State University of New York Press.

Sowder, J. T., & Wheeler, M. M. (1989). The development of concepts and strategies used in computational estimation. Journal for Research in Mathematics Education, 20(2), 130-146.

Stacey, K., & MacGregor, M. (1995). The influence of problem representation on algebraic equation writing and solution strategies. In Proceedings of the 19th International Conference for the Psychology of Mathematics Education, 2 (pp. 90-97). Brazil.

Steinberg, E. (1989). Cognition and learner control. A literature review. Journal of Computer Based Instruction, 16(4), 117-121.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: logic and implementation. Cognitive Science, 97-140.

Svendsen (1991). The influence of interface style on problem solving. Internation Journal of Man-Machine Studies, 35, 379-397.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12, 257-285.

Tabachneck, H. J. M., Koedinger, K. R., & Nathan, M. J. (1994). Towards a theoretical account of strategy use and sense making in mathematical problem solving. In A. Ram & K. Eiselt (Eds.), 16th Annual Conference of the Cognitive Science Society, (pp 836-841). Hillsdale, NJ: LEA

Tabachneck, H. J. M., Leonardo, A. M., & Simon, H. A. (1994). How does an expert use a graph? A model of visual & verbal inferencing in economics. In A. Ram & K. Eiselt (Ed.), 16th Annual Conference of the Cognitive Science Society, (pp 842-847). Hillsdale, NJ: LEA

Thompson, P. W. (1992). Notations, conventions and constraints: Contributions to effective uses of concrete materials in elementary mathematics. Journal for Research in Mathematics Education, 23(2), 123-147.

Threadgill-Sowder, J. (1984). Computational estimation procedures of school children. Journal of Education Research, 77(6), 332-336.

Trafton, P. R. (1986). Teaching computational estimation: Establishing an estimation mind-set. In H. L. Schoen & M. Zweng (Eds.), Estimation and mental computation (pp. 16-30). Reston Virginia: National Council of Teachers of Mathematics.

Underwood, J., & Underwood, G. (1987). Data organisation and retrieval by children. British Journal of Educational Psychology, 57, 313-329.

Vygotsky, L.S. (1978) Mind in society: The development of higher psychological processes. (M. Cole, V. John-Steiner, S. Scribner, E. Souberman, Eds. and Trans.). Cambridge, MA: Harvard University Press.

Watson, J. M., Campbell, K. J., & Collis, K. F. (1993). Multimodal functioning in understanding fractions. Journal of Mathematical Behaviour, 12, 45-62.

White, B. (1993). ThinkerTools: Causal models, conceptual change, and science education. Cognition and Instruction, 10(1), 1-100.

Williams, D., Duncomb, I., & Alty, J. L. (1996). Matching media to goals: an approach based on expressiveness. In M. A. Sasse, R. J. Cunningham, & R. L. Winder (Ed.), People and Computers XI Proceedings of HCI' 96, (pp. 332-347). London.: Springer.

Winn, B. (1987). Charts, graphs and diagrams in educational materials. In D. M. Willows & H. A. Houghton (Eds.), The psychology of illustration (pp. 152-198). New York: Springer-Verlag.

Xploratorium (1991). Broken Calculator. Anglia Polytechnic University: Xploratorium.

Y1 (1979). SevenOaks: Hodder and Stoughton.

Yerushalmy, M. (1989). The use of graphs as visual interactive feedback while carrying out algebraic transformations. In Proceedings of the 13th International Conference for the Psychology of Mathematics Education, vol. 3 (pp. 252-260). Paris.

Yerushalmy, M. (1991). Student perceptions of aspects of algebraic function using multiple representation software. Journal of Computer Assisted Learning, 7, 42-57.

Young, R., & O'Shea, T. (1981). Errors in children's subtraction. Cognitive Science, 153-177.

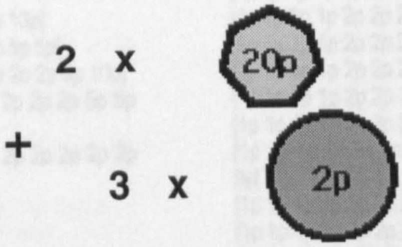
Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. Cognitive Science, 18, 87-122.

APPENDIX ONE

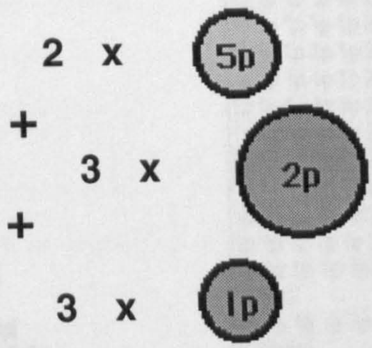
A COPPERS Pen and Paper Test

How many ways can you make this much money?

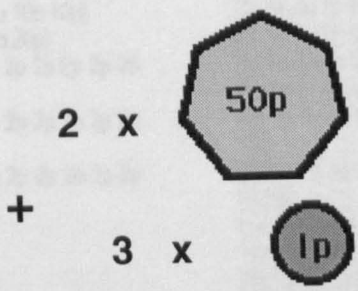
1)



2)



3)



Here are the 350 possible correct solutions for question one.

[1p 2p 2p 2p 2p 2p 5p 10p 20p]
[1p 2p 2p 2p 2p 2p 5p 10p 10p 10p]
[1p 2p 2p 2p 2p 2p 5p 5p 5p 10p 10p]
[1p 2p 2p 2p 2p 2p 5p 5p 5p 20p]
[1p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p 10p]
[1p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p 5p 5p]
[1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p 10p]
[1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 20p]
[1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 10p]
[1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p]
[1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p]
[1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 5p 10p 10p 20p]
[1p 5p 10p 10p 10p 10p]
[1p 5p 20p 20p]
[1p 5p 5p 5p 10p 20p]
[1p 5p 5p 5p 10p 10p 10p]
[1p 5p 5p 5p 5p 5p 10p 10p]
[1p 5p 5p 5p 5p 5p 20p]
[1p 5p 5p 5p 5p 5p 5p 5p 10p]
[1p 5p 5p 5p 5p 5p 5p 5p 5p 5p]
[1p 1p 2p 2p 5p 5p 10p 20p]
[1p 1p 2p 2p 5p 5p 10p 10p 10p]
[1p 1p 2p 2p 5p 5p 5p 5p 10p 10p]
[1p 1p 2p 2p 5p 5p 5p 5p 20p]
[1p 1p 2p 2p 5p 5p 5p 5p 5p 5p 10p]
[1p 1p 2p 2p 5p 5p 5p 5p 5p 5p 5p 5p]
[1p 1p 2p 2p 10p 10p 20p]
[1p 1p 2p 2p 10p 10p 10p 10p]
[1p 1p 2p 2p 20p 20p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p 10p 10p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p 20p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 10p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p 5p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 10p 20p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 10p 10p 10p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 10p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p 10p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 20p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 2p 2p 2p 2p 5p 10p 20p]
[1p 1p 1p 2p 2p 2p 2p 5p 10p 10p 10p]
[1p 1p 1p 2p 2p 2p 2p 5p 5p 5p 10p 10p]
[1p 1p 1p 2p 2p 2p 2p 5p 5p 5p 20p]
[1p 1p 1p 2p 2p 2p 2p 5p 5p 5p 5p 5p 10p]
[1p 1p 1p 2p 2p 2p 2p 5p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p 10p]
[1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 20p]
[1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 10p]
[1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p]
[1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 2p 5p 5p 10p 20p]
[1p 1p 1p 1p 2p 5p 5p 10p 10p 10p]
[1p 1p 1p 1p 2p 5p 5p 5p 5p 10p 10p]
[1p 1p 1p 1p 2p 5p 5p 5p 5p 20p]
[1p 1p 1p 1p 2p 5p 5p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 2p 5p 5p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 2p 10p 10p 20p]
[1p 1p 1p 1p 2p 10p 10p 10p 10p]
[1p 1p 1p 1p 2p 20p 20p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p 10p 10p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p 20p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 10p 20p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 10p 10p 10p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p 10p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 20p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 2p 2p 2p 5p 10p 20p]
[1p 1p 1p 1p 1p 2p 2p 2p 5p 10p 10p 10p]
[1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p 20p]
[1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p 10p]
[1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p 20p]
[1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p 20p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 10p 20p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p 10p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 20p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 5p 5p 10p 20p]
[1p 1p 1p 1p 1p 1p 5p 5p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 5p 5p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 5p 5p 5p 5p 20p]
[1p 1p 1p 1p 1p 1p 5p 5p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 5p 5p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 10p 10p 20p]

[1p 1p 1p 1p 1p 1p 10p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 20p 20p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 10p 20p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 10p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 10p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 10p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 10p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 10p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 20p]

[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 10p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 10p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 10p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 10p 10p]

[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 2p 2p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 2p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 2p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 2p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 2p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 10p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 20p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p 2p
2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p 2p 2p
2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p
5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 5p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p 2p
2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p 2p
5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 5p
5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p 2p
2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p 2p
2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 2p
2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 5p
5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
10p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
2p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 2p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 2p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 2p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 5p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 2p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 2p]
[1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p 1p
1p 1p 1p 1p 1p 1p 1p 1p 1p]
[2p 2p 2p 5p 5p 10p 20p]
[2p 2p 2p 5p 5p 10p 10p 10p]
[2p 2p 2p 5p 5p 5p 5p 10p 10p]

[2p 2p 2p 5p 5p 5p 5p 20p]
[2p 2p 2p 5p 5p 5p 5p 5p 5p 10p]
[2p 2p 2p 5p 5p 5p 5p 5p 5p 5p 5p]
[2p 2p 2p 10p 10p 20p]
[2p 2p 2p 10p 10p 10p 10p]
[2p 2p 2p 20p 20p]
[2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 10p 10p]
[2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 20p]
[2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 10p]
[2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p 5p 5p]
[2p 2p 2p 2p 2p 2p 2p 2p 10p 20p]
[2p 2p 2p 2p 2p 2p 2p 2p 10p 10p 10p]
[2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 10p]
[2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p 5p 5p]
[2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p 10p]
[2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 20p]
[2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 5p 5p]
[2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 10p]
[2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p 2p]

PAGES MISSING IN ORIGINAL

APPENDIX TWO

A CENTS Pen and Paper Test

Estimation

Name ______________________ Class _______________ Date of Birth_____________

This booklet has 20 estimation questions.

Can you tell me what estimation means ?

____

____

____

____

I'm going to give you questions like this:

Estimate 13 x 28

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

1
Estimate: 17 x 34

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

2
Estimate: 285 x 687

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

3
Estimate: 525 x 386

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

4
Estimate: 59 x 69

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

5
Estimate: 17 x 33

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

6
Estimate: 613 x 521

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

7
Estimate: 335 x 562

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

8
Estimate: 74 x 24

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

9
Estimate: 27 x 46

my estimate is [ ]

I think that my estimate is ....

very much less	much less	less	just less	exactly the same	just more	more	much more	very much more
30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above

than the exact answer

10
Estimate: 16 x 33

my estimate is

I think that my estimate is	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	than the exact answer
	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	

11
Estimate: 82 x 48

my estimate is

I think that my estimate is	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	than the exact answer
	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	

12
Estimate: 867 x 356

my estimate is

I think that my estimate is	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	than the exact answer
	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	

13
Estimate: 31 x 88

my estimate is

I think that my estimate is	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	than the exact answer
	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	

14
Estimate: 745 x 234

my estimate is

I think that my estimate is	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	than the exact answer
	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	

15
Estimate: 16 x 34

my estimate is

I think that my estimate is	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	than the exact answer
	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	

16
Estimate: 22 x 62

my estimate is

	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	
I think that my estimate is	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	**than the exact answer**

17
Estimate: 491 x 209

my estimate is

	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	
I think that my estimate is	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	**than the exact answer**

18
Estimate: 338 x 164

my estimate is

	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	
I think that my estimate is	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	**than the exact answer**

19
Estimate: 23 x 66

my estimate is

	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	
I think that my estimate is	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	**than the exact answer**

20
Estimate: 18 x 34

my estimate is

	very much less	much less	less	just less	exactly the same	just more	more	much more	very much more	
I think that my estimate is	30% or below	30% to 20% less	20% to 10% less	10% to 0% less	0 %	0% to 10% more	10% to 20% more	20% to 30% more	30% or above	**than the exact answer**

APPENDIX THREE

Publications Based on this Thesis

Technical Reports

Ainsworth, S. E. (1994a). COPPERS: A mathematical learning environment (technical report number 8). ESRC Centre for Research In Development, Instruction and Training, University of Nottingham.

Ainsworth, S. E. (1994b). An experimental evaluation of COPPERS: A mathematical learning environment (technical report number 10). ESRC Centre for Research In Development, Instruction and Training, University of Nottingham.

Ainsworth, S. E. (1995). CENTS: the design and implementation of a learning environment to teach computational estimation (technical report number 29). ESRC Centre for Research in Development, Instruction and Training, University of Nottingham.

Ainsworth, S. (1995). The role of multiple representations in promoting understanding of estimation: an empirical investigation (technical report number 30). ESRC Centre for Research in Development, Instruction and Training, University of Nottingham.

Conference Papers

Ainsworth, S. E., Bibby, P. A., & Wood, D. J. (1996). Combining and translating between representations. In I.E.E Colloquium on Thinking with Diagrams Digest no 96/010, London.

Ainsworth, S. E., Wood, D. J., & Bibby, P. A. (1996). Co-ordinating Multiple Representations in Computer Based Learning Environments. In Proceedings of the European Conference of Artificial Intelligence in Education, (pp. 336-342). Lisbon.

Journal Articles

Ainsworth, S. E., Bibby, P. A., & Wood, D. J. (in press). Information Technology and Multiple Representation: New Opportunities - New Problems. Journal of Information Technology for Teacher Education, 6(1).

Ainsworth, S. E., Wood, D. J., & O'Malley, C. (in press). There's more than one way to solve a problem: Evaluating a learning environment to support the development of children's multiplication skills. Learning and Instruction.