Learning Equations using a Computerised Balance Model: A Popperian Approach to Learning Symbolic Algebra

James Aczel

Thesis submitted to the University of Oxford for the degree of DPhil

Keble College

Michaelmas Term 1998

Learning Equations using a Computerised Balance Model: A Popperian Approach to Learning Symbolic Algebra

James Aczel

Thesis submitted to the University of Oxford for the degree of DPhil

Keble College *Michaelmas 1998*

Abstract

This study investigates, using a perspective based on the work of Karl Popper, how students aged 10-15 can learn about simple linear equations, with particular reference to the use of a computerised balance model of an equation.

Popperian epistemology implies a conjectural view of knowledge, in which rigour is dependent on the potential for intersubjective criticism. A Popperian approach to psychology is advocated, in which "understanding" is viewed as *problem-solving* rather than sense-making, imagining or re-enactment; and learning occurs through trial-and-improvement of *strategic theories* in response to *concerns*, rather than through the development of context-free modes of thought. From this perspective, explanatory constructs from research into learning algebra such as "letter interpretations" and "equation metaphors" are seen as recontextualised meta-algebraic theories rather than as slowly maturing "underlying" algebraic cognitive structures.

A Popperian reinterpretation of the research literature into the problem of learning algebra enables the development of an instrument to detect learning in a range of principal algebraic concerns - representation, interpretation, transformation and utilisation. A computer program called EQUATION is also constructed, which acts as a research tool to explore the educational limitations of the balance model of an equation.

Fieldwork is carried out to test conjectures relating to the program, involving around 100 students. Analysis involves reconciliation of classwork learning and pre-post testing. It is argued that a concern for symbolic algebra can be initiated firstly by using the balance model to promote formal operations on equations and secondly by encouraging the formulation of equations to find an unknown number in a word problem. In addition, by providing progressive challenge and feedback on the effects of operations, it is possible for students to create, test and improve strategic theories for a number of transformation and representation problems.

Acknowledgements

To Soft Warehouse for allowing the pilot study to research the use of Derive.

To Scott Forbes for carefully nudging me in the direction of games, and for backing up the idea with a coherent rationale for holistic education.

To Professor Richard Pring and Dr. Lyn English for their helpful suggestions.

To the teachers and students who took part in the empirical studies, whose questions, ideas, patience, enthusiasm and hard work helped me to see algebra from their point of view. Thank you also to John Pawsey, Mike Sarson, Phil Corbett and Mike Pepper at KHS, and to Mukesh Mistry, all of whose dedication to students and openness to new ideas set me on this path.

To John Wilson for the many valuable discussions and the always engaging criticism. Thank you also for offering me a lifeline that kept me going in a time of crisis.

To Dominic Tilley, for his inspiration in so many ways. Thank you in particular for the log file idea, and for reminding me to be Popperian in the face of unreality.

To Professor Joan Solomon, whose advice, support, criticism and insight were invaluable. Thank you for your help.

To Donald Cudmore (even though he thinks Popper is dead!) for his example and camaraderie.

To my parents and grandparents, without whom my studies would not have been possible.

To baby Francesca for putting life into perspective; and to my wife Bee, who endured my worries and attic-dwelling above and beyond the call of duty. Thank you for everything.

Learning Equations using a Computerised Balance Model: A Popperian Approach to Learning Symbolic Algebra

Contents Page

Chapter 2: Identifying Algebraic Theories and Concerns..38

List of Abbreviations

7CON	Year 7 control class - not using EQUATION
7EQ	Year 7 class using EQUATION
10CON	Year 10 control class - not using EQUATION
10EQ	Year 10 class using EQUATION
AI	Artificial Intelligence
BVSR	Blind Variation and Selective Retention; see Campbell (1960)
CAS	Computer Algebra System
CIA	Computer Intensive Algebra; see Heid (1990), Heid & Zbiek (1995)
CMF	Children's Mathematical Frameworks; see Dickson (1989)
CSMS	Concepts in Secondary Mathematics and Science; see Hart (1981)
GCSE	General Certificate of Secondary Education
IEA	International Association for the Evaluation of Educational Achievement; see Robitaille (1989)
ILS	Integrated Learning System; see NCET (1994b)
IQ	Intelligence Quotient
IT	Information Technology
JMC	Joint Mathematical Council
LHS	Left-Hand Side
NCET	National Council for Educational Technology
NFER	National Foundation for Educational Research
OFSTED	Office for Standards in Education
RHS	Right-Hand Side
SCAA	School Curriculum and Assessment Authority
SESM	Strategies and Errors in School Mathematics; see Booth (1984)
SMP	School Mathematics Project
TIMSS	Third International Mathematics and Science Study; see Keys, Harris & Fernandes (1996)
TOAN	Think Of A Number
VBA	Visual Basic for Applications

Introduction

The latest international studies and numerous cognitive research studies suggest that there is room for improvement in the algebraic knowledge of a large proportion of students. Particular difficulty is found with handling expressions and equations; and many students never seem to appreciate that algebra is a tool rather than some arcane ritual. How can more students be helped to appreciate the power and challenge of symbolic algebra? This research aims to explore using a Popperian perspective how students aged 10-15 can learn about simple linear equations, with particular reference to the use of a computerised balance model of an equation.

Chapter 1 defines the research problem, taking account of associated background theories. In particular, "symbolic algebra" is taken in this research to refer to the use of letters to stand for numbers, and Popperian approaches to both epistemology and psychology are advocated. Popperian epistemology implies a conjectural view of knowledge, in which rigour is dependent on the potential for intersubjective criticism. Empirical validity is therefore concerned with the extent to which conjectures are tested by data, rather than with the origins of assertions. Popper also distinguishes World 3 - which contains published theories, problems and arguments - from the purely subjective World 2. Meanwhile, a distinctive psychological line is developed: "understanding" is viewed as *problem-solving* rather than sense-making, imagining or re-enactment; and learning occurs through trial-and-improvement of *strategic theories* in response to *concerns*, rather than through the development of context-free modes of thought. From this perspective, explanatory constructs from research into learning algebra such as "letter interpretations" and "equation metaphors" are seen as recontextualised meta-algebraic theories rather than as slowly maturing "underlying" algebraic cognitive structures.

Chapter 2 examines various studies into the learning of algebra, especially CSMS (Hart, 1981) and the student-professor problem, in order to identify target algebraic theories and concerns. It is stressed that in explaining incorrect responses to a problem, a lack of understanding of the problem is a possible explanation. It is also argued that identifying *strategies* may be a valuable way of analysing learning. This reinterpretation of the literature using Popperian psychology enables the development of questions intended to detect learning in a range of principal algebraic concerns - representation, interpretation, transformation and utilisation.

Chapter 3 looks at a range of algebraic learning activities, including the use of syncopated language, computer algebra systems, spreadsheets, the mathematics machine of SESM (Booth, 1984), "Think of a Number" games and the arithmetic identities of Herscovics and Kieran (1980). Arising out of this review is an instructional proposal that focuses on the limitations of the balance model for learning formal methods for solving equations and the formulation of equations to represent situations. However, examples of apparent "transfer" of strategic theories between concerns are also noteworthy - not because transfer indicates underlying cognitive structures or overarching conceptions, but because it suggests that theories have partial autonomy from the concerns that generated them. In particular, the thesis is developed that promoting the simplification of equations and easing the formulation of equations to represent word problems can provide a purpose for algebraic symbolism that can assist in other transformation and representation problems.

Chapter 4 draws together threads from the previous chapters by arguing for the development of a computer program called EQUATION. This program is intended not so much as some sort of ideal learning environment, but as a research tool to illustrate theoretical arguments and to provide prima facie tests of certain conjectures. The program promotes a simplification strategy through a game-like balance model; it introduces algebraic notation as a convenient abbreviation, which enables negative signs and negative answers to break with the model; and it then promotes algebra as a tool for solving word problems. It also logs what students see on the screen, and what they click and enter. The development of EQUATION should be considered an integral part of the research; the reader is therefore advised to try out the program once the first four chapters have been read.

Chapter 5 describes the development and execution of fieldwork to test conjectures relating to the balance model. It also discusses practical research design decisions, including the development and piloting of EQUATION and the pre-post test instruments.

Chapter 6 analyses data from the fieldwork in an attempt to identify improvements in students' equation theories and concerns as a result of using EQUATION, and to provide some clues as to when and why such learning might have occurred. Research data includes interview recordings, responses to written tests and questionnaires, recordings of conversations of students working at the computer, logs of students' interactions with the program and word problems posed by the students themselves. Around 100 students in Years 6, 8 and 10 were involved.

Finally, chapter 7 attempts to use the analysis and the Popperian psychological perspective to reconcile the classwork and pre-post testing, to explore limitations of the balance model and to compare EQUATION with other initiatives.

Chapter 1
Defining the Research Problem

1.1 Introduction

This chapter initially sets out the broad problem of learning symbolic algebra by means of a brief historical outline. The subsequent section focuses on what many see as an indicator that the current generation of algebra students is on the cusp of a new era. An attempt to define the research problem follows, but certain aspects then require elaboration in later sections, in particular the research's position on epistemology and psychology. The final section then redefines the research problem, but more specific research questions are developed as the argument of the thesis progresses in succeeding chapters.

1.2 The Problem of Learning Symbolic Algebra

A large-scale and influential UK study (CSMS: Hart, 1981) found wide variations in algebraic achievement and few signs of substantial student progress over two years. Sleeman (1986) asserts that "The difficulties of learning algebra have been greatly underestimated." (p. 52). Herscovics (1989) affirms that UK and US assessment studies indicate that "algebra is a major stumbling block for many students in secondary school. Only a minority of pupils completing an introductory course achieve a reasonable grasp of the course content. Even fewer manage to build up enough courage for a second course." (p. 60). More recent international studies such as TIMSS (reported in Keys, Harris & Fernandes, 1996) seem to corroborate this. Although performance in algebra for both England and Scotland was below the international mean - and well below Singapore, Japan and Hungary in particular - this could result from curriculum differences. The recent Royal Society / JMC report (1997) on teaching and learning algebra suggests that expectations in England and Wales of what students can achieve in algebra are generally lower than expectations in, for example, France and Germany. Even so, there is little difference in TIMSS algebra scores between the US, England and Germany. On the other hand, Robitaille (1989), reporting on the second IEA survey, notes that for *all* countries "performance levels on the Algebra subtest as a whole are a cause for concern internationally. Students' achievement, even on what appear to be quite straightforward items, was frequently very poor." (p. 114).

More in-depth research confirms that this is a problem not just with standardised written assessments and curriculum emphasis. English & Sharry (1996) summarise much of the research using classroom-observation and interviews when they describe the "major difficulties students experience in algebraic learning" as "well documented" (p. 135). In particular, Bednarz *et al.* (1992) note that "students have difficulty acquiring and developing algebraic procedures in problem-solving." (p. 65); while Geary (1994) concludes, "Most people find the solving of algebraic word problems a cumbersome task." (p. 127). The obstacles are generally considered two-fold. Firstly "students encounter major difficulties in representing word problems by equations" (Herscovics, 1989, p. 63). Secondly, many studies report that solving equations is found difficult (for example Carry, Lewis & Bernard, 1980), and this is corroborated by international assessment. When 15-year-olds were asked to solve $5x + 4 = 4x - 31$ (an equation on the curriculum of all the countries involved), the success rate for countries ranged from 9% to 58% (Foxman, 1992). There is also evidence suggesting that a quarter of UK schoolchildren *never* end up able to solve equations as simple as $2x + 4 = 10$ (Foxman *et al.*, 1991). Is it therefore surprising that numerical examples of a rule carry more conviction than algebraic proof (Lee, 1987)?

Meanwhile anecdotal evidence suggests that the formalism of algebra may be a contributory factor in declining mathematical confidence and enthusiasm between primary and secondary school. Certainly SCAA (1996) suspects that lack of algebraic manipulation skills at GCSE may hold back students from taking mathematics beyond the compulsory years. Moreover, despite the fact that many more students are taking A-level, the number taking mathematics A-level has declined by over a quarter since 1985.

Attempts to address these difficulties have had minimal success. English & Halford (1995) note, for example, that "a large proportion of students do not acquire the expected levels of proficiency, even after several years of study" (p. 219). Moreover, Arzarello (1991) writes that "an effective construction of algebraic knowledge as a net of operative ideas and algorithms is difficult" (p. 80); while Arcavi (1994) points out:

> "many high school students make little sense of literal symbols, even after years of algebra instruction. Even those students who manage to handle the algebraic techniques successfully, often fail to see algebra as a tool for understanding, expressing and communicating generalisations, for revealing structure, and for establishing connections and formulating mathematical arguments." (p. 24)

Has learning algebra always been a problem?

Sutherland (1990) writes, "Since the beginning of the century the teaching of algebra has been a central concern of mathematics educators, who have always been aware of the inherent difficulties in the subject." (p. 155). English & Halford argue that fashionable ideas in psychology are often used to justify changes in curriculum and teaching methods. For example, Thorndike's "law of exercise" was used to justify drill. Although "connectionism", "associationism" and "S-R bond theory" are no longer part of the vocabulary of mathematical education, the idea that mathematics is a set of behavioural habits which can be linked with specific stimuli using repetitive conditioning perhaps still influences practice. "Number bonds", for example, are still in fashion in primary schools, and the notion of a bond's "readiness to act" was one of Thorndike's themes. He could be said to be responsible for the original argument that a teaching programme in mathematics should consist of fragmenting the knowledge into tiny nuggets of facts and skills, sequencing them in the right order, and then ensuring sufficient practice with each bond. Kieran & Wagner (1989) conclude from reviewing research into the learning of algebra between 1900 and 1930, "All that seems clear from these studies is that the larger the number of steps required to solve an equation, the less often students were able to solve it correctly." (p. 2).

Sutherland notes a reaction against this early 20th century emphasis on definition, notation and exercises in substitution and simplification. Lack of enthusiasm or progress was sometimes attributed to students finding such a diet "uninteresting", "meaningless" and "artificial" (see, for example, Mathematical Association, 1934). As a result, symbolic algebra started to be introduced in the context of practical problems, especially those in geometry and mechanics. Not only would such "applied" mathematics provide reasons for learning algebra, but it would provide a less abstract framework through which to understand the principles. Yet there was a basic quandary in this pedagogical movement. Because such practical problems are notoriously messy and complex, either they were artificially simplified - which meant that they could then probably be solved without the use of algebra anyway - or else the students were restricted merely to describing a situation using algebra and leaving solution aside. Both these scenarios tended to defeat the reasons for learning algebra in the first place, and made the whole activity appear rather unreal. Moreover, "Algebra word problems have been a source of consternation to generations of students." (Berger & Wilde, 1987, p. 123). One reason, argues Sutherland, is that "what gives algebra its power is the potential to work with symbols without reference to 'external' meaning." (p. 162), but without such reference students can find algebra literally "meaningless". It is a classic Catch-22.

English & Halford cite Wheeler (1935) as a proponent of more "meaningful learning" through experience rather than systematic instruction; and Brownell (1935, 1945) as a proponent of aiming for an understanding of mathematical relations and structure, before any drill and practice takes place. For example, "a child who promptly gave the answer '12' to the fact '7+5' would not be regarded as having demonstrated a knowledge of the combination unless she understood why 7 plus 5 equals 12 and could convince others of its correctness" (English & Halford, 1995, p. 4). However Kieran & Wagner can find little research from that period into how this might work for algebra. The influence of the Gestaltists, English & Halford suggest, was felt more in the discrediting of connectionist and structuralist models of learning mathematics than in any positive theory of behaviour, cognition or effective teaching. But perhaps Pólya's convincing discussion of "heuristics" is strongly redolent of the "saturation - incubation - inspiration - verification" model that some Gestaltists professed. English & Halford go on, however, to point out that Wertheimer's (1959) distinction between productive thinking (grasping structural relations and combining them into a dynamic whole) and reproductive thinking (repeating learned responses to individual subparts) may have encouraged the theory that the giving of ready-made steps can in fact be harmful for learning.

The movement kick-started by the 1959 Woods Hole conference emphasised the laying of "foundations" for a proper grasp of modern mathematics; this included deductive reasoning, precision of language and unifying structures like sets, variables and functions. Widespread anecdotal evidence would indicate the unpopularity among students of what the Americans called "New Math". Freudenthal was critical of this approach, especially of the use of letters to denote both sets and members of sets; and according to Sutherland, Chevellard (1984) criticised the loss of the dialectic between arithmetic and algebra. Whatever the reasons, Sutherland points out that "almost all traces" of New Math seem to have been removed from the National Curriculum for England and Wales. Kieran (1992), writing about the US experience agrees: "… those changes that have persisted into today's algebra curricula have been more cosmetic than substantial." (p. 392). Perhaps this trend stems from Bruner's (1960) work into children's developmental levels of representation (enactive, iconic, symbolic) and Dienes' (1960, 1963) investigation of structured learning experiences, in particular their stress that algebraic formalism should come *after* more direct experience of objects and relations, and that a spiral of extending and elaborating concepts was preferable to a linear exposition of "foundations". Clearly, too, English & Halford's view of psychology's impact on learning theory is supported by the authority given to Piaget's work on developmental thinking, even though one aspect of his work (the invariance of stages of sensorimotor, pre-operational, concrete operational and formal operational thinking) has come in for particular criticism. Nevertheless, his view that "there is a parallelism between the progress made in the logical and rational organization of knowledge and the corresponding formative psychological processes" (Piaget, 1971, p. 13) has certainly sent many educators scurrying to the history books for ideas. Misinterpretations of Lakatos (1963) may hinge on this point. Kieran & Wagner note a dramatic increase in research from the late 1960s onwards, focusing more on mathematics education, and less on mathematics as a convenient domain for studying general questions of memory and skill development.

English & Halford describe the "constructivist paradigm" as emphasising "children's active involvement in their own learning" (p. 1). The introduction of "mathematical investigations" in Britain may be a part of this emphasis. Educators talk about a change in views of mathematics from a formal, external body of knowledge that has to be transmitted to learners, to a dynamic field in which each student can explore the processes through which mathematics develops by constructing knowledge on the basis of his or her own experience. Consequently children should be provided with activities that provide rich experiences, and opportunities for reflection, questioning and discussion. However, if active participation within a meaningful context does not guarantee "strong acts of construction" (Noddings, 1990, p. 14) and conversely "traditional instruction does not necessarily imply 'weak constructions' on the part of students." (English & Halford, p. 12) then is it the nature of the activities that makes the teaching constructivist, or the teacher's intentions, or the "strength" or "desirability" of constructions?

Although there have been many research studies into algebra, none has perhaps been more influential in the UK than CSMS. Sutherland, for example, argues that it is the most obvious influence on the algebra components of the National Curriculum, particularly in that "many pupils in Britain are… introduced to algebraic ideas at a later stage of their secondary schooling and with more caution than they were some years ago. This seems to be a direct consequence of i) the finding that pupils have difficulty with algebra and ii) a belief in Piaget's theory concerning formal operations and the related idea that pupils will not be able to cope with algebra until they reach the stage of formal operations." (p. 159). SCAA (1996) notes that a deliberate broadening of the pre-16 curriculum over the past twenty years has corresponded to a decrease in "manipulative algebra" (p. 25) but an increase in graphical representation of functions. Sutherland does not entirely attribute to CSMS the loss from the National Curriculum of the "close interrelationship between algebra and arithmetic" (p. 159). The "hierarchy of concepts" used for assessment is another matter: O'Reilly (1990) suggests that there are many possible "pathways" through mathematics, and yet the levels of CSMS have been largely enshrined as the hierarchy of the National Curriculum. Sutherland argues that the "suppression of symbolism" and the greater emphasis on "number patterns arising from a whole range of situations, as opposed to number patterns arising from the rules of arithmetic" (p. 159) probably owe more to a reaction against the early 20th century than to specific findings about children's difficulties with algebra. For example, the highest marks in GCSE practical investigations are expected to be given to work containing algebra, sometimes even if the context does not lend itself to algebra.

The consequence of this suppression of symbolism is that during the 1980s "pupils' first introduction to algebra was more likely to be in the context of expressing generality" (Sutherland, 1990, p. 160) rather than manipulating symbols. For example, Mason *et al.* (1985) emphasise expression of generality over the learning of isolated, apparently purposeless skills:

> "Algebra is firstly a language - a way of saying and communicating… algebraic language is a powerful means of communicating abstract and complex ideas. It is especially suited to expressing generalities. A second important feature of the language of algebra is that it contains its own manipulative rules which need to be learnt and practised. But the central of feature of algebra is that it is an ideal medium through which one can see and express general statements." (p. 1)

Scientific laws and relationships can be described concisely; patterns in numbers can be demonstrating using a "general" case. By widening the definition of algebra to encompass awareness of pattern and generality (which even babies can have to some extent), Mason *et al.* argue that "lack of facility in expressing generality renders formal algebra totally meaningless." (p. 63).

> There once was a 'gebra called Al,
> Who said 'aren't these numbers banal:
> To you the specific
> May seem just terrific
> But my statements are très général'. (Mason *et al.*, 1985, p.1)

However when MacGregor & Stacey (1993a) state that the "ability to perceive a relationship and then formulate it algebraically is fundamental to being able to use algebra." (p. 181), it is not clear whether the *perception in itself* is algebraic. The Royal Society / JMC report judges not (p. 21).

Moreover, while this increased emphasis on variables through the encouragement of generalisation may be welcome, Sutherland points out:

> "There are many who believe that operating on the unknown is the most crucial and difficult aspect of algebra (Filloy & Rojano, 1989). In other countries pupils are faced with this idea when they solve algebraic equations. Algebraic equation solving has been de-emphasised within the National Curriculum for mathematics and we need to ask whether or not equation solving had a crucial role in helping pupils accept the important algebraic idea of operating on the unknown." (p. 163)

Because of OFSTED's judgment that at GCSE "even the highest attainers produced poor answers to questions involving algebra" (SCAA, 1996, p. 9), GCSE syllabuses have now been forced to give greater emphasis to manipulation (p. 28-9).

According to English & Halford, cognitive science "provides the most scientific method yet devised for analyzing the real psychological processes that underlie mathematics. By detailing the way "concepts are represented" it offers great promise for increased efficiency in mathematics education" (p. 14). Nevertheless, Chaiklin (1989) admits out that within a cognitive approach "As a rule, the instructional experiments conducted to date have not been particularly successful in helping students to develop word-problem-solving abilities. However, they illuminate the performance model and highlight issues that need to be addressed in developing effective instruction." (p. 103). A related approach he discusses is that of artificial intelligence in education (see especially Wenger, E., 1987), which aims to apply programming techniques of knowledge representation and transformation to model teacher interventions and student learning.

Yet despite all this curriculum variation, pedagogical innovation and research investigation, the problem of learning algebra remains. The Royal Society / JMC report summarises it thus: "The issue for mathematics education is how to re-emphasise the role of symbols without precipitating a return to the traditional and often ineffective means of teaching algebra which were prevalent 20-30 years ago. These methods were ineffective because they only worked for a very small proportion of the school population and, as Cockcroft [1982] pointed out, actually alienated many pupils from mathematics." (p. 5).

Given this forbidding weight of contemporary and historical evidence that students find learning difficult, the question has to be asked: *is algebra worth learning?*

The Royal Society / JMC report notes concerns from a number of bodies about the decline in numbers of A-level mathematics students and about "a serious lack of essential technical facility" in algebra among incoming undergraduates. These bodies include the London Mathematical Society; the Institute of Mathematics and its Applications; the Royal Statistical Society; and the respective institutions of chemical, civil, electrical and mechanical engineers. It would appear that it is not just the supply of future mathematicians that is threatened - the scientific subjects for which A-level mathematics is a "service course" may no longer be able to assume that students have adequate mathematical knowledge. If a large proportion of the population is not to be excluded at an early age from scientific occupations, "algebra for all" would appear to be an appropriate goal. On the other hand, there are also some who argue (see later) that new technology means there is less need to learn the technical skills that computers and calculators can carry out much more quickly and accurately. Moreover, Lewis (1981) points out that perhaps not even those professional mathematicians using elementary algebra every day become flawless equation-solvers (p. 107). The Mathematical Association (1934) also notes the importance of mathematics for science and technology occupations, but doubts that using symbolic algebra is a vital part of skills for other occupations and for day-to-day living.

Nevertheless, the Royal Society / JMC report suggests, with respect to the "life-skills" required by students on vocational courses, that "it is important for all students to become confident and competent with certain areas of pre-algebra and algebra in order to function effectively as citizens within a society which is increasingly shaped by mathematics and where problems are increasingly converted to forms for which there is a calculable solution." (p. 24). However, it could be argued that none of the life-skills examples given in the report (setting up spreadsheet formulae for administration, numerically estimating timber production using a known relationship, having a "feel for number", understanding percentages and interpreting visual presentation of data) involve symbolic algebra, which perhaps explains the report's conclusion that "algebra for citizenship" should not focus on "formal" algebra.

Moreover, even though the Mathematical Association similarly asserts that an educated person should learn "something of the part that mathematics has played and continues to play in the development of the modern world." (p. 8), the fact that a society is "increasingly shaped by mathematics" does not imply that such sociological inquiry has to involve much algebraic activity in school. It may very well be desirable for citizens to understand the forces that shape their society, but there are many such forces: why single out algebra as having a favoured position in the general curriculum? The Association also notes that custom plays a large role in keeping topics in the curriculum. But just because a good pass in a school-leaving mathematics certificate (with its attendant algebra requirements) is currently a prerequisite for entry to a very wide range of courses and jobs, this does not mean that the curriculum *should* continue to require algebra.

Are there any other possible arguments for keeping symbolic algebra in the curriculum? Beyond a probably small group of professional and recreational mathematicians, algebra does not appear to be valued universally as a beloved, aesthetically-pleasing, cultural artefact. However, if it is decided that there are reasons that mathematics is worthy of study other than the needs of everyday life or jobs - for example because it also provides training in analytical techniques - then there is a strong argument that one cannot appreciate mathematics properly without appreciating algebra. The Royal Society / JMC report asserts, for example, "algebra and the algebraic language are central to mathematics and if we do not teach algebra then we are not teaching mathematics." (p. 5).

From a pragmatic perspective, algebra is in virtually all secondary school curricula the world over, and that in itself is a good enough reason for researching ways of helping students (whether or not they intend to pursue scientific occupations) to appreciate the power and challenge of algebra. That is therefore the research focus for this thesis.

1.3 The Potential of Information Technology

1.3.1 Some Questions

Fey (1989b) writes that new technological developments “challenge every traditional assumption about what we should teach, how we should teach, and what students can learn.” (p. 237, emphasis removed). For school algebra, these technological developments include spreadsheets, graphics calculators, computer algebra systems, tutoring programs and dynamic geometry packages. Spreadsheets such as Microsoft Excel can ease repetitive numerical calculations, statistical analysis and sophisticated simulations; graphics calculators such as the Casio fx-7400G can provide instant views of functions and tables of values; computer algebra systems such as Derive by Soft Warehouse can solve many types of equations and carry out all the algorithms a mathematics student could imagine; tutoring programs can make mundane exercises more interactive; and geometry packages like Geometer’s Sketchpad or Cabri Géomètre make exploring geometry more fun. The TI-92 calculator from Texas Instruments, released in 1995, combines all the above facilities in a hand-held machine. Soon very few schools will be without access to the Internet.

How does such information technology (IT) affect what should be taught? English & Halford suggest that improvements in technology and changed demands in society place more of an emphasis on analysing and interpreting data for making decisions than on routine algorithms. Tall (1989), for example, writes “the manipulation of algebra to solve equations will be less important for that class of problems for which a numerical solution is appropriate and for which simple numerical algorithms on the computer will suffice” (p. 91). Will there be a need to learn the product rule for differentiation, or the Newton-Raphson method for solving equations numerically, or the pencil-and-paper long-division algorithm, or even how to sketch a graph, if these can all be done faster and more accurately using technology? “Much of what goes on in higher secondary algebra is learning skills. Are these needed when supercalculators and computer algebra systems can do the work for us? Even more to the point, will the relief from concentrating on routine skills free students and teachers to study concepts in more depth?” (Monaghan, 1994, p. 201). Should the curriculum and assessment adapt to reflect new technological capabilities? For Fey (1989a), “The unanswered question standing in the way of reducing the manipulative skills agenda of secondary school algebra is whether students can learn to plan and interpret manipulations of symbolic forms without being themselves proficient in the execution of those transformations.” (p. 206-7). Are there learning processes or outcomes that IT can render safely obsolete? Are there learning processes or outcomes that IT *cannot* render obsolete?

Aside from the issue of what should be taught, how does IT affect teaching strategies? The number of questions rapidly proliferates. How is IT being used in the classroom? Can any general conclusions be drawn about what works and what doesn't? Are there important differences in IT effects for distinct social groups? Is the apparent success of computers symptomatic of a fascination for novelty, or of something deeper? Do the students' predominant modes of learning affect the success of IT-based activities? Teachers are rapidly adapting new technology and writing their own custom software to assist their students' learning. How can these innovations be evaluated and shared? To what extent does IT replace, support or undermine the role of the teacher? Could a problem-solving approach to introducing algebra be made more realistic by computers handling the complexity?

How does IT affect what students can learn? How is "understanding" affected by technology? Is it true that IT improves investigative skills and understanding at the expense of competence by reducing the opportunity for practice? Programming languages, such as Logo and BASIC, have been around in schools for over two decades; they are becoming higher level, easier to learn, more visual and more powerful. What impact have they had on students' learning? Papert (1980) writes that computers can be used to "challenge current beliefs about who can understand what and at what age." because history is no longer a good guide to potential intellectual development - we have a technologically richer environment than ever before. Can technology really refute the attribution of understanding to cognitive development (Sutherland, 1991)?

Papert also suggests that computers could "contribute to mental processes not only instrumentally but in more essential, conceptual ways, influencing how people think even when they are far removed from physical contact with a computer" (p. 4). How does the nature of the tool influence the nature of responses to the problem? How is the notion of a "skill" determined by the available technology? If "square root" is now considered an elementary operation rather than a complicated algorithm because of calculators, how soon before "solve" moves that way? Why would "model" or "prove" not become operations? Monaghan (1994) suggest that it is not just the content of school mathematics that will change, but that "Computer-based technology is changing the character of mathematics... Computers not only introduce new areas of mathematics but bring with them new ways of thinking about mathematics." (p. 193). In Papert's words, computers can be "carriers of powerful ideas and of the seeds of cultural change" (p. 4). Does technology encourage a particular view of the nature of mathematics?

Of course the availability of this IT raises many further important questions. For example: What are the major obstacles to IT usage? Access, tradition or ignorance? How should technological change be implemented? In which areas should money be spent on research and development, in equipment purchase, and in training? Should a handheld computer be the right of every child? Ernest (1991) describes a number of views of the mathematics curriculum - attributed to

"Industrial Trainers", "Technological Pragmatists", "Humanists", "Progressive Educators" and "Public Educators". To what extent does IT policy serve the interests of these viewpoints? More generally, how are different sociological groups affected by new technology? Is inequality of technological expertise and access class-based? Is there a gender differential?

1.3.2 Some Answers?

Given the large number of questions raised by the apparent potential of new technology to challenge assumptions about the learning of mathematics in general - and algebra in particular - it is perhaps unsurprising that there has been an enormous flurry of research activity in this area over the past few years (see, for example, Bednarz, Kieran & Lee, 1996). Yet Kaput & Thompson (1994) roundly condemn the mathematics education research community's "lack of technological engagement" (p. 680), concluding that the likelihood of mathematics educators shaping the roles of educational technology (let alone its development and research) is small:

> "To use technology in mathematics education research is intellectually demanding - one must continually rethink pedagogical and curricular motives and contexts. To exploit the real power of the technology is to transgress most of the boundaries of school mathematics practice. And normally, a powerful technology quickly outruns the activity-boundaries of its initial design - students and teachers, as well as its designers, generate activities that were not conceived in the design process. This renders classroom-based research, especially research that extends beyond brief interventions, difficult - and makes direct comparison and tightly controlled experimental studies inappropriate." (p. 681)

Oldknow (1995) charts a recent history of IT for mathematics in the UK. Special national and local government funding for technology and training, the annual BETT conference, the teacher journal *MicroMath* and initiatives by the National Council for Educational Technology are all signs of "this unexpectedly rapid improvement in the IT provision for mathematics", and evidence perhaps of a desire to adapt tools originating outside education for educational purposes.

He makes a case for:

1. An evaluation of the ways that IT is currently being used to support mathematics education.
2. A research programme into the quality of learning with IT, and into attitudes towards IT.
3. A curriculum review, taking account of the changing uses of IT in industry, commerce and research, and the power of IT to aid understanding and promote challenges.
4. A forum for the international exchange of ideas.
5. Teacher support in the learning of mathematics using IT.

Oldknow warns that it is "all too easy to be carried away on a wave of enthusiasm", which is perhaps a reminder that research studies involving enthusiasts, considerable external support and time for planning do not represent typical experience of IT. However, although this thesis is clearly attempting to contributing to the second proposal, the idea that educational technology innovations can be research-driven also seems important. But Fey (1989b) is pessimistic:

commenting on the myriads of innovations reported in the previous five years, he writes, "it is very difficult to determine the real impact of those ideas and development projects in the daily life of mathematics classrooms, and there is very little solid research evidence validating the nearly boundless optimism of technophiles in our field." (p. 237).

Hammond (1994) redresses some of the enthusiasm by noting:

> "… two major field studies, the *Impact* report (Watson, 1993) and the *Plait* report (Gardner *et al.*, 1992) have… thrown some doubt on the value of using IT. The *Impact* study was a major longitudinal study of 2300 children in primary and secondary schools across 19 local authorities in England and Wales. The major empirical finding was that 'IT did make a contribution to learning, but the contribution was not consistent across subjects or age groups.'.
>
> "… In the *Plait* study 235 pupils in nine schools were given their own portable computers to use for one year… The study found that 'The impact of personal access to laptop computers on pupils' performance was not significant or at best marginal over one school year.'." (p. 252)

Hammond makes the point that although there is already a large body of research into IT use in education - part of the broader agenda of research into school innovation in general (Fullan, 1982) - such research cannot hope to offer a productive evaluation of any "contribution to learning" if it does not pay attention to the design of the particular software, the topic-specific aims of the teachers, the way the software is used in class and the distinctive learning outcomes for individual children. Pea (1987), on the other hand, believes that it is possible to identify "transcendent functions" of cognitive technologies - "tools of the intellect provided by the culture" (p. 91).

Berry, Graham & Watkins (1994) identify five ways in which Derive - and most other CASs - can be used in a mathematics course: as a mathematical tool, a problem solving assistant, an investigative environment, a demonstration aid, and an interactive tutor. Perhaps these categories apply to all IT that can be used to assist in the learning of algebra.

Mathematical tool. Berry *et al.* argue that as a mathematical tool for backing up pencil and paper skills, a computer algebra system can save time, but may lead to students relying on it too heavily, and consequently failing to master the pencil and paper skills they are expected to learn. Nevertheless, Love (1995) suggests that "as more and more mathematical techniques are able to be carried out by software there is less and less need for anyone to learn to be able to carry them out, even with the aid of software." (p. 116) and therefore the future curriculum requires radical reconceptualisation.

Problem solving assistant. In courses where there is a greater emphasis on problem formulation and interpretation than on solution, a computer algebra system can lessen the time needed for solution, extend the range of problems, permit more complicated models, increase the accuracy of solutions and allow a focus on the validity and reliability of solutions (Heid, 1990). Kaput (1992) indicates that he, at least, is convinced that studies such as Hembree & Dessart (1986) demonstrate that "heavy use of calculators in the early grades… does not harm computational ability and frequently enhances problems-solving skill and concept development." (p. 534). Ruthven (1995) offers a critical review of evidence from research over the past twenty years concerned with the effects of calculator use. This evidence would appear to indicate, as Monaghan (1994) says, that "calculators are an aid in the process of solving problems" because they ease time-consuming calculations, thus allowing a focus on the type of mathematics required. Similarly, Sutherland (1995) concludes that the spreadsheet "frees pupils from the process activity of evaluating an expression, thus enabling them to focus more on the structural aspects of a situation." (p. 285). The formalism of programming languages can also provide a new entry point to the formalism of algebra (Sutherland, 1990; NCET, 1994a).

Investigative environment. But Sutherland also stresses that computer environments can support pupils' development of an algebraic approach to problems. NCET stresses that the fast, reliable, non-judgmental and impartial feedback that computers can provide can encourage conjecturing and testing. Multiple representations can encourage the seeking of connections. Berry *et al.* suggest that as an investigative environment, a computer algebra system can help students to gain an "intuitive feel" for the ideas before they are formally introduced. π, i and e can arise naturally, for example. Similarly, diSessa (1995) describes how Boxer - a piece of software that allows visual programming - is successful as a collaborative medium. LaTorre (1995) describes how programmable graphics calculators "carry the computational burden" to allow undergraduates to experience linear algebra in an "active, constructive environment" and to "achieve understanding of certain concepts better". Laborde (1995) says "The graphical and computing possibilities of some software now allow a reification of abstract objects and in particular of mathematical objects as well as numerous possible operations on these objects and various feedback." (p. 36). But she also argues that these objects, although "embodied in a material environment", are not necessarily accessible to the learner.

Demonstration aid. Computers can create images to help visualise ideas. For algebra this can mean dynamically linking "multiple representations" - numbers, letters, graphs, icons and natural language, for example (see Kaput, 1989 & 1992; Fey, 1989b). However, Hunter *et al.* (1995) found that, for graphical work, there seemed to be little effective difference between use of a computer algebra system and the traditional methods of sketching and drawing. Mention should also be made of word-processors, reference programs, multimedia authoring software and the Internet - all of which involve communicating ideas in some way. Simulations enable complex models to be developed and monitored.

Interactive tutor. Finally, as an interactive tutor (Wenger, E., 1987), computers can present problems to students, assess the knowledge of a student based on inputs, and manage the future instruction, activities and problems accordingly. They make mundane exercises more interactive, and can provide contextual help suited to the student. NCET (1994b) found that children working on a particular ILS "performed significantly better than children working in control groups, making gains of twenty months over the six-month period." (p. 6). However, although one student said "It doesn't go off and help someone else." (p. 16) and the privacy was appreciated ("No one knows when you make a mistake", p. 18), perceptions of progress by students were more negative than the tests would indicate. The research did not focus on the algebra components. Various software under the labels of "Computer Assisted Learning", "Computer-Based learning", "Expert Systems", "Intelligent Computer-Assisted Instruction" and "Integrated Learning Systems" are in development, given new lease thanks to more graphically-oriented operating systems and programming languages. Following Olds *et al.* (1980), Kaput (1992) emphasises educational computer games as teaching tools:

> "The role of motivation [in educational games] has been extensively studied by the social psychologist Lepper and colleagues at Stanford University (Lepper, 1985; Malone & Lepper, 1987), although without much attention to the curricular value of the educational objects of the games involved." (Kaput, 1992, p. 519)

Love (1995) makes the point that the design intentions of software do not confine its category of use. For example, "Logo is valued not only for its geometric aspects, but as an opportunity to learn about the use of variables, or the modular construction of solutions to a problem." (p. 111). He warns that, to some extent, the mathematics that the teacher sees in software is not *inherent* in software, and perhaps not even in the eye of the beholder. It would be hard to describe Visual Basic, say, as "containing" a finite number of mathematics concepts; but perhaps it would be slightly easier to delineate the potential of a given program.

1.4 The Research Problem - A First Attempt

The fact that one can imagine all five of these roles for technology within one program suggests that attempting to constrain the research by "type" of technology might be rather unavailing. But with regard to a mathematical focus, since the motivation for this research derives from wanting to help more students appreciate the power and challenge of algebra, one key element could be the equation. Equations could perhaps demonstrate the utility of algebra, because they are often used in tackling word problems. As we have seen, representing word problems using equations and then solving the equations are found to be very difficult.

So an initial (and tentative) formulation of the research problem would be "What is the impact of information technology on students' concept of equation?". The substance of this thesis, therefore, is likely to relate to the psychology of learning about equations in the context of IT and the methodology for obtaining evidence relating to such learning.

However, as soon as a focus arises, there is immediately a serious problem to be addressed: How is it possible to ignore the other issues of technology and algebra raised earlier? How is it possible to research the impact of technology on students' thinking, without considering its impact on teaching, resource policy, the curriculum and assessment? Can one really focus on equations without considering functions and graphs? For example, the demand to have "functions replace equations as the fundamental objects of algebra" (Chazan, 1993, p. 22) seems to be gathering momentum for many reasons (Pimm, 1995, p. 104-6). Yet if I pretend (for the sake of having a focus) that these issues need not concern me in this particular quest, is it not the case that the research is thus dependent on my unconscious prejudices? Surely all these subjective preconceptions must be made explicit? This issue is a very serious point to address, and cannot be properly debated until the nature of knowledge has been considered.

1.5 Popperian Epistemology

This research takes a robustly realist stance - there *is* a world of stones, sounds, trees, children and so on; and therefore the central questions that concern epistemology are: How is it possible to obtain knowledge about the world? How can we separate knowledge from opinion? A theory of knowledge is required as a rationale for two categories of knowledge here: the knowledge that the research itself finds or aims to find; and the algebraic knowledge that children are supposed to find.

1.5.1 Obtaining Knowledge

How is it possible to obtain knowledge about the world? There seem to be diverse views in the mathematics education research community on epistemological foundations. Hammersley (1995) discusses the relative merits of (1) *positivism*, including behaviourism; (2) *interpretive research*, including radical constructivism, hermeneutics, and phenomenology; and (3) openly *ideological research*, including Marxist, social constructionist, post-structuralist, feminist and anti-racist approaches. The major sources of authority or "foundations" upon which each paradigm is based appear to be (1) experiment; (2) personal experience; and (3) collective experience, respectively.

A major difficulty that all three of these positions attempt to address can be illustrated by such claims as "The child thought that 3*x* meant thirty-*x*". If, it is argued, such statements are not to be accepted *dogmatically*, we must be able to *justify* them; for as Hume said "If I ask why you believe any particular matter of fact… you must tell me some reason; and this reason will be some other fact, connected with it. But as you cannot proceed after this manner, *in infinitum*, you must at last terminate in some fact, which is present to your memory or senses; or must allow that your belief is entirely without foundation.". This *trilemma* is elaborated by Popper (1934):

> "If we demand justification by reasoned argument, in the logical sense, then we are committed to the view that *statements can be justified only by statements*. The demand that *all* statements are to be logically justified (described by Fries as a 'predilection for proofs') is therefore bound to lead to an *infinite regress*. Now, if we wish to avoid the danger of dogmatism as well as an infinite regress, then it seems as if we could only have recourse to *psychologism*, i.e. the doctrine that statements can be justified not only by statements but also by perceptual experience. Faced with this *trilemma* - dogmatism vs. infinite regress vs. psychologism - Fries, and with him almost all epistemologists who wished to account for our empirical knowledge, opted for psychologism. In sense experience, he taught, we have 'immediate knowledge': by this immediate knowledge, we may justify our 'mediate knowledge' - knowledge expressed in the symbolism of some language." (p. 93-4)

Hence we have the view that knowledge has to be related to sense perceptions and thus to our personal experiences, in particular our observations of the physical situation. Sfard (1994b), for example, appears to assent to the assertion that "our bodily experience is the only source of understanding." (p. 46). So a researcher would be able to say, perhaps "It seemed to me that the child thought that 3*x* meant thirty-*x*." or, better maybe, "I heard the child say 'If *x* is 7 then 3*x* is 37'" but not "The child thought that 3*x* meant thirty-*x*", because we do not have direct access to the child's thoughts.

To avoid subjectivity, argue the positivists, these observations need to be *experimental*. For example, Carr & Kemmis (1986) describe this view as asserting that "science provides the methods of enquiry that educational research should seek to emulate and that scientific theories provide the logical criteria to which educational theories should aspire to conform." (p. 51). However, the interpretivists point out the crucial role of subjective and social factors in the production of knowledge, with Kuhn (1962) being a key reference. In particular, observation is *theory-laden* - there can be no neutral description of facts. The interpretive approach therefore aims to unpack the personal preconceptions that cloud and shape observations, and indeed the whole conduct of the research; while the ideological approach might stress the normative, political aspects, particularly with respect to how the research might be used. The Hegelian solution to this difficulty is to unveil hidden ideologies; whereas those who are more pessimistic about the success of such "sociotherapy" sometimes use "oppression", "democracy", "equality" and so on as touchstones to help locate "underlying" assumptions

Whether research is positivist, interpretive or ideological, it is possible to pursue a naturalistic line with respect to *data collection* - a paradigm that points out that the artificiality of research conditions can work against observations being valuable at all, and so conducts research in a situation free as far as possible from constraints imposed by the researcher (Lincoln & Guba, 1985). On the other hand, if one wants to explore *particular* aspects of a situation such intervention may sometimes be necessary (Hammersley, 1995).

It is also possible within all three positions to pursue a naturalistic line with respect to *data analysis* - a paradigm that refuses to prejudge a situation, and so seeks to use theoretical constructs that unfold as the research progresses rather than imposing *a priori* constructs (Lincoln & Guba, 1985). On the other hand, if one wants to explore empirically the current theoretical situation with respect to a particular research problem (and it would appear unwise to ignore previous studies completely), then such constructs may sometimes be necessary (Hammersley, 1995).

Popper discusses in great detail the view that knowledge has to be related to sense perceptions, and there is not space here to look at it in the depth required to compare it with other attempts to solve this problem of knowledge. However, one important point (against Hume) is that personal experiences are not necessarily any more reliable than other forms of data. In fact, as numerous court cases seem to suggest, eyewitnesses can make mistakes. Moreover:

> "Every witness must always make ample use, in his report, of his knowledge of persons, places, things, linguistic usages, social conventions, and so on. He cannot rely merely upon his eyes or ears, especially if his report is to be of use in justifying any assertion worth justifying. But this fact must of course always raise new questions as to the sources of those elements of his knowledge which are not immediately observational." (Popper, 1963, p. 22-3)

Every empirical claim goes far beyond "what can be known with certainty 'on the basis of immediate experience'". For example: how do we (as readers of this particular claim) know that it was the child who said something? How do we know that it seemed to the researcher that the child said "If x is 7 then $3x$ is 37"? How do we know that what was claimed was said was said? How do we know it was English being spoken? How do we know that the child was not describing someone else's view, or giving a *reductio ad absurdum*? How do we know that the child was not making up conventions to solve a particular problem? How do we know that what the researcher intended by the claim is what we understand by it?

Popper makes the distinction between questions of origin ("How do you know?", "What is the basis of the claim?", "What is the source of your assertion?") and questions of validity ("How can we test this claim?", "What error-elimination has taken place?", "What evidence is available to exclude possibilities?"). Questions of origin are "entirely misconceived: they are questions that beg for an authoritarian answer" (Popper, 1963, p. 25), and Popper says that we should be critical of *all* sources of authority because they are theory-laden and fallible. What has to be done in analysing the validity of claims, he argues, is to *test* claims rather than seek their sources. In doing so, of course, various sources *will* be cited, and their value has of course to be questioned; and moreover what constitutes evidence must be constantly re-assessed. But the point is that we can never know we are right; and so it is not the *making* of claims that depends on the reliability of the sources, but the *testing*, because we *can* sometimes know we are wrong; and hence we can learn. The possibility of mistakes prevents the slide into relativism, and thus perhaps we may get closer to the truth. So knowledge is possible despite Fries' trilemma if it is conjectural rather than based on some sort of inductive logic from foundational statements. Although this formulation sounds simplistic, it seems difficult to grasp, perhaps because it relies on the rejection of the "intuitive" Wittgensteinian idea that a term (such as "truth", "reality", "value-neutrality") is meaningless without a criterion for application. For example, Popper points out that the following assertions appear to be self-contradictory from an epistemic point of view, but clearly true from a conjectural point of view:

> "a theory may be true even though nobody believes it, and even though we have no reason for accepting it, or for believing that it is true; and another theory may be false, although we have comparatively good reasons for accepting it." (Popper, 1963, p. 225)

Furthermore, with respect to the ideological approach, Popper (1945) notes that the "sociology of knowledge hopes to reform the social sciences by making the social scientists aware of the social forces and ideologies which unconsciously beset them. But the main trouble about prejudices is that there is no such direct way of getting rid of them. For how shall we ever know that we have made any progress in our attempt to rid ourselves from prejudice? Is it not a common experience that those who are most convinced of having got rid of their prejudices are the most prejudiced?" (p. 210-1). There is no substitute for a critical examination of the empirical and theoretical claims.

However, despite Popper's trenchant criticisms of positivism, subjectivism and the sociology of knowledge, his robust defence of empirical inquiry does not appear to have stemmed the post-modern loss of confidence in the search for truth (as evidenced, for example, by the popularity of Rorty, 1980). Sfard (1994a) describes how the "message of relativity, sometimes interpreted as an eulogy of irrationality" was brought by such writers as Kuhn, Feyerabend and Foucault. The promise of progress was replaced by the hope for endless recapitulation.

For some, instrumentalism replaced truth:

> "… the criterion of truth and validity was replaced by the ideas of solidarity and of usefulness. People should no longer ask whether anything is objectively true; rather, they are expected to judge knowledge according to whether it can bring them together and whether it can do anything for them." (Sfard, 1994a, p. 252)

The prevalence of relativistic ideas in research in the last few decades has, according to Hammersley (1995), been encouraged by the later views of Wittgenstein that "science is just one form of 'language game' among others, having its own distinctive rules but not being in any general sense superior." (p. 13). The work of Husserl in phenomenology was influential in promoting the view that "our understanding of the world is constructed on the basis of assumptions, rather than being a reflection of how the world actually is" and hence "different cultures produce different realities." (p. 14). Similarly, Gadamer's hermeneutics argued that understanding is based on interactions between culturally and historically based assumptions and phenomena. "The implication of this is that there is no method by which universally valid knowledge can be produced because all knowledge reflects the socio-historical context of its production." (p. 14). The zeitgeist seems to be "the idea that knowledge is paradigm-dependent, in other words that its validity is relative to a set of assumptions about what the world is like and how it can be understood, those assumptions being beyond rational justification." (p. 16).

Nevertheless, given the Popperian view that knowledge is conjectural, it is not clear that a lack of a criterion for truth should necessitate the abandonment of the possibility of knowledge about the world. Following Tarski, truth can still be seen in Popperian epistemology as "correspondence to the facts" rather than based on strength of belief or utility. Although there cannot be guarantees of correspondence, truth is still (as Russell stressed) above human authority.

1.5.2 Rigour and Objectivity

Lincoln & Guba (1985) ask "How can an inquirer persuade his or her audiences (including self) that the findings of an inquiry are worth paying attention to...?"; "How can one establish the degree to which the findings... are determined by the subjects... and conditions of inquiry and not by the biases, motivations, interests, or perspectives of the inquirer?" (p. 290). Hammersley (1995) notes a trend in social science away from the view that it is vital to guarantee that other researchers obtain the same results (for example through controlled experiments), or even that other researchers would have obtained the same interpretation when faced with the same data. This is sometimes characterised as a move from "validity" (in the senses of Campbell & Stanley, 1963) to "credibility": it is still expected that the reader is made fully aware of the *basis* upon which interpretations are made, but it is also important to convince the reader that the interpretations are reasonable by describing as much of the experiences that contributed towards the making of those interpretations as possible. However, if, as Popper says, questions of origin are "entirely misconceived" is the search for objectivity misguided? Moreover, is a solitary individual in a position to separate "data" from "interpretation"? Even a thorough research biography has to be selective about the assumptions it details - the researcher's aims and preconceptions, the variety of analyses that might be carried out, the phenomenological issues, the theoretical models that might pertain, and so on. Finally, convincing the reader of one's conclusions is not in itself desirable, because this can be done by rhetoric rather than by reasoning about evidence.

The vision of rigour that comes from Popper (1934) is clear (see also Phillips, 1989), but it is eminently distinguishable from the three paradigms described above. Given that observation is theory-laden; that accounts are selective; that knowledge is produced in a socio-historical context; that situations have multiple perspectives; that knowledge about learning is neither handed down from on high by scientists in white coats, nor by mystics in coloured robes, nor by politicians in primary-coloured ties; Popperians conclude that it is not in general methods of production that warrant knowledge, nor even the authenticity or authority of the knowers, but the extent to which knowledge is tested critically - theoretically or empirically.

However, while a study in which explicit conjectures are tested to the limit might be desirable, hypotheses about World 2 are often difficult to make and test; so whether a researcher is rigorous, in a Popperian view, depends on the extent to which steps are taken to try to eliminate error. This means that the limitations of testing (with particular regard to major alternative conjectures that might answer the research questions or fit the data) have to be apparent. Often the best way to test something is to put it into a form that can be checked by others. "It seemed to me that the child thought that $3x$ meant thirty-x." is not easy to check, because it is a statement about my perceptions. However, "The child thinks that $3x$ means thirty-x" is more easily checked. It will probably turn out to be false, because by putting it in the present tense, the description has been removed from the particular situation of the child that prompted the claim to be made. However, if it is false then that may be revealing in itself: it would be very interesting to try to explore the situations in which the *child does or doesn't think* that $3x$ means thirty-x, rather than those situations in which the *researcher thinks* that the child does or doesn't think that $3x$ means thirty-x. But error elimination can take many forms: contrast, for example, the ethnographic analysis techniques described by Eisenhart (1988), the multi-level statistical analysis techniques used by Sammons (1995) and the theoretical analysis of Ernest (1991).

Popperians do not view the products of research (observations, theories, characterisations, interpretations, and so on) as inherent in the data. They do not arise *naturally* from the data. They are not implied or given by the data in some inductive sense. They are *invented* by researchers. However, because knowledge does not progress by *establishing* or "grounding" theories (cp. Glaser & Strauss, 1967), we need tests that our theories may fail (so we are challenged to find better theories) or pass (so we are challenged to find better tests or alternative theories). There is no absolute requirement, therefore, for every piece of research to include what Ball (1990) calls a "reflexive research biography", unless, for example, accounting for the researcher's emergent thoughts and feelings through introspection is one of the explicit objects of the research. Moreover, for some studies, it may be the case, as Phillips (1989) suggests, that "processes involved in, and even central to, the *making* of discoveries during the pursuit of a research program may not be involved - and might be counterproductive if allowed to intrude - when the discoveries are *checked* and *tested* and *critically evaluated*.". One objection to this view is that without such a process of grounding, another researcher might reach different conclusions. However, if it is the case that another researcher might reach different conclusions, no amount of "grounding" will help because *interpretation* of evidence is probably also in dispute. What will help is the devising of crucial tests to decide between alternative interpretations. However, the nature of a "crucial test" for a descriptive theory of mental processes is highly problematic.

Every statement is, in a sense, a claim. Since not all claims can be tested, some knowledge has to be taken for granted. In practice, one only attempts to produce a justification or evidence for claims that are in some way controversial. The choice of what claims are considered necessary for testing for the purposes of this research (and by implication, that which is considered "known") is subjective in the sense that they are the decisions of the researcher. However, these decisions are objective in the sense that they are, in principle, open to intersubjective criticism. Because decisive refutation is rare, sometimes the best test available is (theoretical) intersubjective criticism. Therefore, it is not the making of unjustified claims that is wrong, it is the idea that a claim has been sufficiently tested. Popper argues that our presuppositions can be changed by decision and by experience; not all at once, of course, but *piecemeal*. This process of breaking down individual assumptions is endless - "Any assumption can, in principle, be criticised. And that anybody may criticise constitutes scientific objectivity." (Popper, 1945, p. 209). The objectivity of a theory rests upon its criticizability. A theory that has survived tests it could have failed is to be preferred to an irrefutable theory, or one defended by *ad hoc* explanations. The search for "objectivity" - in the sense of the removal of the personal, subjective or partial from theories and research - is an endless quest for truth by a *community* committed to open criticism. Objectivity is not a product of the individual researcher's impartiality, because it is "closely bound up with the social aspect of scientific method" - it is dependent on the intersubjective criticism provided by the research community. By this means, a study that might apparently only relate to a particular context might be germane to other contexts. All means of testing are considered - the alleged opposition between qualitative and quantitative techniques dissolves.

While this *rationale* of research methodology is very different from the positivist paradigm (because the search for certainty is abandoned, and multiple perspectives are actively sought), from the interpretive paradigm (because the researcher's experience is not seen as central, and the social aspect of scientific method is emphasised) and from the ideological paradigm (because priority is given to seeking empirical tests), the *practice* may be difficult to distinguish sometimes. For example, the theoretical "justification" of a claim can be seen, perhaps, as an argument that the claim can be derived from less hotly debated conjectures. The empirical "basis" for a claim could be that the claim succeeded where competing conjectures failed. On the other hand, "validity" unquestionably concerns the relationship between conjectures and data that might test those conjectures rather than the relationship between propositions and their origins, their adherents or their political role. So it is vital in a research report to be fair in claiming whether the data collected constitutes a critical test that decides between conjectures.

1.5.3 The Role of Background Theory

Rather than providing an "underlying framework" or the foundations for empirical research, Hammersley (1995) argues that philosophy functions as background theory which, once explicit, is open to "correction and modification in the light of what we learn in practice". Some of these theories are empirically testable (for example that children grow into adults, or that the curriculum includes equations); others are not (for example that other people exist, or that $2 + 2 = 4$). Lincoln & Guba (1985) distinguish three means by which a hypothesis may be tested: empirical (to see if it is "consistent with nature"); logical (to see if it is consistent with other knowledge); and methodological (to see if the person who asserts the hypothesis is conforming to certain ethical or professional standards of conduct). Moreover, there are certain "metaphysical beliefs" that cannot be tested by any of these means.

Furthermore, Lincoln & Guba assert that these metaphysical beliefs "must be accepted at face value" (p. 14) - they are axiomatic "basic" beliefs - and so they "represent the ultimate benchmarks against which *everything else* is tested" (p. 15). However, Popper (1963) points out that the idea that the truth of a theory can be inferred from its irrefutability is an obvious mistake - there may be two incompatible theories that are equally irrefutable. He cites determinism and indeterminism; but there are many others: for example one could also have the axiom of choice and its negation; or "There is a simple incantation that makes everyone who hears it know all algebra instantly." versus the view that there is no such incantation.

So is it rational to believe that there is no incantation if we cannot prove it? Popper points out that we can try to refute an empirical statement or its negation, but there are still critical questions we can ask even if refutation is too tall an order. "A theory is comprehensible and reasonable only in its relation to a given *problem situation*, and it can be rationally discussed only by discussing this relation.". So we can ask things like: "Does it solve the problem? Does it solve it better than other theories? Has it perhaps merely shifted the problem? Is the solution simple? Is it fruitful? Does it perhaps contradict other… theories needed for solving other problems?". Irrefutability does not mean there are no good critical arguments.

Although there will be ideas taken for granted in any research, the next section makes explicit some important background theory.

1.6 Some Background Theories

1.6.1 Autonomous Knowledge

Popper calls the physical world "World 1"; while "World 2" is the subjective world of our conscious experiences. But according to Sfard (1994a), the unattainability of truth challenges "the assumption that science aims at a discovery of mind-independent reality.". Similarly, Lincoln & Guba criticise the assumption that "there is a tangible reality" by conflating realism with the view that "experience with [reality] can result in knowing it fully" (p. 82). This conflation entails not the orthodox idea that each of us creates theories about the world rather than having direct access to it (our knowledge of World 1 is a World 2 construction); it entails the more radical view that there is no world except that which is constructed by humans (World 1 is part of World 2). The difference is crucial for research. In the second case, one could only ever claim to be musing on one's own dreams - it would not be legitimate to talk of pursuing knowledge about some sort of free-standing educational setting. The ideas that "participants" distinguishable from oneself could have "rights" over one's data, or that there is some sort of "responsibility" to a wider "community of inquiry" would be unnecessary elements in the one-player game. Nevertheless, given a conjectural epistemology, it is not clear why the reality of World 1 has to be denied.

Popper (1972) argues that if it is possible for theories about World 1 to clash with World 1 itself, then it must be possible to talk about their *logical content*. Theories that are accessible to other people (rather than just being inside a person's head) are members of "World 3". Examples of knowledge in World 3 include "theories published in journals and books and stored in libraries; discussions of such theories; difficulties or problems pointed out in connection with such theories; and so on." (p. 73). Arithmetic, methods for solving equations, gravity, astrology, flat-earth theory and fractals belong to World 3. World 3 contains false theories as well as true, and problems as well as arguments. It is not unchanging, because knowledge grows.

Popper argues that World 3 is "a natural product of the human animal, comparable to a spider's web" (Popper, 1972, p. 112), and is largely autonomous (which is why he refers to "objective" knowledge). For example, although humans created natural numbers, whether there is an infinite number of prime numbers is not something we make up - this problem and its resolution are unintended consequences of the original creation. Another argument for autonomy consists of a thought experiment in which all works of art and technology - and the knowledge in people's heads about them - are destroyed: does the survival of *libraries* affect the re-emergence of civilization? Even von Glasersfeld (1995), who starts "from the assumption that knowledge, no matter how it be defined, is in the heads of persons" and therefore denies the existence of World 3, also wants to argue that ideas "should never be personal property".

But this autonomy is not complete - new problems and creative thinking may lead us to new theories and thereby add to World 3, creating "new unintended facts; new unexpected problems; and often also new refutations" (Popper, 1972). This process (albeit rather simplified) can be described using the schema:

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

P_1 is the problem from which we start; TT is a tentative theory - an imaginative conjectural solution; EE is error-elimination, involving critical discussion or experimental tests; and P_2 is the resulting problem situation, perhaps containing new problems. We can attempt to gauge progress by comparing P_2 with P_1. Note also that a problem situation incorporates background theories (some inherent in the structure of language). This is important for understanding Lakatos' notions of progressive and degenerating problem shifts.

This view of knowledge as autonomous is undoubtedly out of kilter with the prevailing philosophy of the mathematics education research community (Ernest, 1991) because of its perceived associations with four positions:

1. The belief that mathematical language aims to capture unchanging "essences" in response to mathematicians' direct intuitions of a Platonic realm of Forms (thus forgetting the human role in creating mathematical objects, rules and practices).
2. The view that words and symbols carry fixed meanings (described by Sfard, 1994b).
3. The arrogance that assumes that mistakes are never made in mathematics.
4. The naïve version of positivism that ignores Fries' trilemma and the subtlety of human thought and interaction by maintaining that controlled experiments are the only way to obtain knowledge about teaching and learning mathematics.

The Popperian formulation outlined earlier would suggest that this association is unnecessary. Only if knowledge were foundational rather than conjectural would the existence of World 3 lend any support to these positions. Whatever the philosophical status of World 3, the unfortunate conflation of autonomous knowledge with these rather unsophisticated views (Edwards & Núñez, 1995) means that the researcher can feel uncomfortable talking about mathematical ideas unless they are portrayed as being in the head of a particular participant in the research (a hopeless situation if one wants to *improve* students' existing knowledge). This conflation can also entail a significant failure to understand students' thinking by ignoring their very real struggle in maintaining a consistent, memorable and workable set of mathematical ideas. At a minimum, World 3 should be seen as a construct for respecting the "weekday Platonism" (Davis & Hersh, 1980) of those doing mathematics.

1.6.2 What is Algebra?

Mathematics in the Popperian view is an archetypal member of World 3. Although it is a human creation, we can discover things about it. This fits very well with Davis & Hersh (1980), but does not answer the question "What is mathematics?", let alone "What is algebra?".

Popper (1945) offers very striking arguments against the Aristotelian view of definition as the description of an essence. In this view, "A puppy is a young dog" answers the questions "What is a puppy?" or "What does 'puppy' mean?" or "What do we mean by 'puppy'?". Popper argues for using in philosophy the scientific method of *labelling*. For example, "A puppy is a young dog" would answer the question "What shall we call a young dog?". A summary of Popper's nominalist position would be that arguments about terminology are philosophically barren, and that we should ensure that nothing important depends on the meaning of terms. This is clearly directly opposed to Wittgenstein's view of philosophy as the dissolving of linguistic confusions through the clarification of terms. Whether too much or too little has been loaded into terms depends entirely on the task to which they are put.

Nevertheless, it should be made clear to avoid confusion that in this thesis, "algebra" is usually a shorthand for "symbolic algebra", which is taken here (following Kieran, 1989a) to refer to the use of letters to stand for numbers (for example when students solve equations, represent word problems or express laws of arithmetic). Of course what needs justification in this thesis is not that the word "algebra" *essentially means* or *should mean* "the use of letters to stand for numbers" (because such discussions are ultimately barren unless one is interested in the history of words), but why this research focuses on the use of letters to stand for numbers, particularly in equations. This decision is because of the earlier formulation of the problem of learning algebra as being centred on difficulties with the symbolism, rather than because the use of letters is "the most obvious feature of algebra" (Booth, 1989a, p. 57). There are reasons (as will be seen) why the definition assumed here does not make use of a distinction between algebra as syntactical form and the semantic processes - mathematical or psychological - that "lie behind" the syntax (Hewitt, 1985, Davis, 1986b); in particular why the definition is not in terms of a mode of thinking, such as "awareness of generality", "handling the as-yet-unknown" or "appreciation of mathematical structure". However, it is to be expected that each person has his or her own preferred definition, and to avoid futile arguments about terminology one would not want to defend particular definitions too strongly. For example, Radford (1995) shows clearly that the 9th century work of al-Kwharizmi, which predates the use of letters standing for numbers, involved procedures - such as treating the unknown as if it were known, and using arguments for the equivalence of two ways of calculating whatever the particular numbers selected - that are recognisably part of what students do in the classroom under the label "algebra". Moreover, modern algebra need not even refer to arithmetic in that it can be generalised to the study of objects and the rules between them. Algebras such as Boolean or geometric transformational are not considered here.

So there are good reasons for alternative formulations. For example, the Royal Society / JMC report takes the view that "the mere use of algebraic symbols does not imply algebraic activity." (p. 29), and consequently excludes trial-and-improvement solution of equations as an algebraic activity because it involves "moving forwards from a 'known' starting number to the 'unknown' number, whereas algebraic methods involve working backwards from an 'unknown' number to a known number." (p. 8).

As for identifying "algebraic problems and ideas", it seems reasonable to claim that the problems described in the literature on research in the learning of algebra may be intended to involve some engagement with these problems and ideas. Of course the originators of the activities were making certain assumptions about what common algebraic goals are appropriate for students to attain, and these assumptions depend on theories about what problems of algebra are accessible, valuable and interesting for the particular students, and on the perceived constraints on the learning environment. So although one can start by looking for activities that are explicitly labelled "algebraic", decisions can be made to exclude or include based on further criteria. In particular, the focus is largely determined by various conjectures that arise out of the reinterpretation of the research literature into the identification and improvement of students' algebraic knowledge (chapters 2 and 3). But functions and graphs - which would probably form part of many people's algebra curriculum - do not feature much in this research for reasons of time and thesis length.

1.6.3 The Issue of Focus

We have seen in section 1.5 how "that anybody may criticise" constitutes scientific objectivity. For example, the failure to address the questions in section 1.3 may fuel a criticism of a particular claim, method or bias in this research. If the criticism is a good one, this is to be encouraged. What constitutes a "good" criticism is of course debatable, but the point is that this research should not try to immunise itself from criticism. To take a particular case, a partial resolution of the question "To what extent is research into IT across the curriculum applicable to mathematics?" is to *assume* that, although IT is used across the curriculum for measurement and data-handling, there must be some research done which is specific to mathematics. This research is essential, it could be argued, if one is to adequately illuminate the relationship between IT and the specially mathematical processes of investigating pattern in number, shape and function, using the language of algebra and studying geometry. Now this argument is debatable; but it is not debated here. Rather, it is part of the background theory that is open to criticism. So why offer the argument? Simply because by exposing more theory to scrutiny, it makes it more likely that an error will be found.

1.6.4 Ethics

What rights do relevant parties have over the data? The research is based on *co-operation* with students, teachers and schools; therefore each participant should be made aware of the nature of the research and given the option to exercise the rights of anonymity and confidentiality. This is all very well, but what about validity problems caused by data omissions? Such omissions would place limitations on the extent to which it would be possible to test theories, but this in itself would not necessarily completely negate the value of the research. These principles may sound like drastic limitations on the "right to know", but while an ethical framework cannot cope with *every* moral dilemma, it is important to know the rules so that breaking them is then a conscious decision. Although it is potentially arrogant to fit others into our own theoretical frameworks, it is even more arrogant to suppose that this can be avoided. However, arrogance all round can be minimised by an epistemological approach that suggests there are always alternative theories.

Another issue is whether it is fair to potentially advantage one group of students by providing them with technology that might (if a control group design were adopted) be prohibited from another group. Since this research is not longitudinal (see later) it will be possible to allow any students to eventually have access to any benefits. Conversely, the collaborative nature of the research means that the teachers involved could at any time terminate any activity that appeared to be unproductive.

1.7 Popperian Psychology

"I suggest that one day we will have to revolutionize psychology by looking at the human mind as an organ for interacting with the objects of the third world; for understanding them, contributing to them, participating in them; and for bringing them to bear on the first world." (Popper, 1972, p. 156)

"How do we think?" can be addressed on a number of levels. It is important to point out that we do not require here a description of neuronal brain activity, or solutions to the problems of consciousness, intelligence, personality, social behaviour, trauma, etc. Rather, because the research is directed towards the knowledge about mathematics inside students' heads, we require a *rationale* for the relationships between World 2 knowledge, experience, learning and understanding. Note that "World 2 knowledge" here includes the products of understanding; subjective knowledge of "facts" (that is: "what is the case", as opposed to propositions); and "skills" (the capability to solve problems). Note too that without something like the distinction between World 2 and World 3, any difference between epistemology and psychology would have to be denied.

1.7.1 Learning

The commonsense theory of commonsense knowledge is called by Popper the "bucket theory of the mind": we open our eyes, prick up our ears, and information streams into the mind, accumulating and then being digested as knowledge. He argues that this theory is completely mistaken, and yet exerts a powerful influence on some theories of teaching, particularly the behaviourist notion of conditioning. Knowledge is treated as consisting of "thing-like" elements (ideas, impressions, sense data) *in* us, that we passively receive (unless we actively create error by interfering with or "going beyond" these given elements); higher level knowledge establishes itself by the repetitive association of these elements. Opposed to this, "As children we learn to decode the chaotic messages which meet us from our environment. We learn to sift them, to ignore the majority of them, and to single out those which are of biological importance for us either at once, or in a future for which we are being prepared by a process of maturation." (Popper, 1972, p. 63). This learning "consists of the modification (possibly the rejection) of some form of knowledge, or disposition, which was there previously; and in the last instance, of inborn dispositions." (*ibid.*, p. 71). Moreover, modification occurs by a process of trial-and-error-elimination. That is, we somehow *jump* to a World 2 theory and then test it in the hope of getting nearer to the truth.

So if learning *starts* from existing knowledge, but is *discontinuous* with existing knowledge, can there be any notion of "natural" (or even "proximal") algebraic development? If increases in ability arise only through learning activities rather than by any age-related process of maturation, can commonalities between students be explained by common experiences? Sutherland (1991) suggests that IT has the potential to challenge the attribution of students' misconceptions to cognitive development. If learning is creative and conjectural, can there be *cognitive* gaps between theories? (cp. Herscovics & Linchevski, 1994).

1.7.2 Understanding

Three popular (overlapping and partial) theories of understanding are:

Sense-making: an act or process of constructing meanings, connections and rationale; a reflection on these; or a holistic perception (perspective, view, interpretation, "seeing as", construal, reading, noticing, jigsaw) of these. See Sierpinska (1994) for some discussion of this theory. Kaput (1989), for example, writes, "meaning is the foundation of mathematics learning" (p. 168).

Imagining: a construction or possession of an image (mental representation, model, metaphor, intuition, embodied schema, first approximation mental picture, result of reflection); or an examination of the properties and relations of an image. The Pirie-Kieren (1989) model is an explicit theory of mathematical understanding along these lines. English & Halford (1995) develop a sophisticated theory of cognitive models. Larkin (1989), for example, suggests that a large part of "understanding an equation" is being able to "construct a good internal representation for it" (p. 130); while Sfard (1994b) refers to non-propositional "image schemata" (following Lakoff, 1987 & Johnson, 1987) as a primary means for organising experience.

"*Re-enactment*": an act or process either of re-enacting through empathy another's coherent experience, or of social construction. This is perhaps best exemplified by the work of Lave in the case of situated cognition, and Collingwood in the case of historical research.

The relationships between these theories in research studies are complex (see Davis, 1992, for example), especially with regard to language - as a factor in structuring thought; as an insight into thinking; as a means for understanding; as sharing of meaning; etc. There are also difficult issues: Are knowing *how* or *when* aspects of understanding? What is the role of understanding in developing and executing strategies? For example, one of the teachers involved in the research argued that "problems with manipulation all come from problems of not understanding… if you understand it then you can follow the rules". Is to understand something to have the theories that enable one to be able to use it appropriately, or is it to have a theory relating it to simpler "knowns"? The sense-making theory is associated by Popper (1963) with Wittgensteinian notions of meaning and warrantability; but it captures in informal language the sensations of gradual piecing together and sudden insight. The imagining theory has much to recommend it, particularly its emphasis on models and their limitations, and the potential of similes such as "an equation is like a balance" to reduce the "cognitive load" demanded by memorisation or execution of algorithms. But can it be assumed that one can characterise failure to "perform" as the grasp of an inadequate image (or the inadequate grasp of an appropriate image)? The re-enactment theory helpfully emphasises situations and their influences.

One difficulty is that if computers cannot have understanding (and therefore knowledge), the operational definition of understanding has to tighten to compensate for computers' capability to achieve tasks that previously were thought to require understanding. On the other hand, behaviour is not always an accurate guide to understanding. It is possible to devise a strategy, understand it in any of the senses above, even write a program to carry it out, but then find it difficult to execute, because of (say) the number of variables - some practical knowledge is needed (such as an image or pencil-and-paper notation) to keep track of the algorithm. Understanding a problem, understanding how to solve it and being able to solve it are quite distinct: researchers at NFER, analysing the mathematics scripts of KS1 students concluded that many of the errors were not caused by failure to read the problem, to understand the problem, or to represent the problem, but ultimately by transformation or process errors. In other words, the children either used an inappropriate strategy or executed a good strategy badly.

As a general heuristic, Popper suggests that, contrary to expectations, if we want to know about how and why something is produced we can often learn more by studying the products themselves than by studying production behaviour. We can regard the *outcome* of the subjective process of understanding as a theory - a World 3 object; and the *process* of understanding as a sequence of theories. A psychological process can (and should, perhaps?) be analysed in terms of the World 3 objects in which it is anchored. Popper's view is that understanding a theory is, at heart, *understanding the problems it intends to solve and understanding why other solutions fail.* This applies equally to language, science, religious belief, or any history of human action. His general problem-solving schema: then applies: $P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$. In connecting understanding and problem-solving, this view is perhaps closest to that of cognitive science (Chaiklin, 1989, p. 96).

If understanding a theory is a recursive *metaproblem* of understanding an underlying problem situation, what are the consequences for the desire for pedagogical knowledge, and the notions of a linear, spiral, modular or hierarchical curriculum? Under what circumstances could we talk about "degrees of understanding", or "levels" or "types"? Rather than seeking students' "grasp" of meanings and images, we should look at their *theories*.

1.7.3 BVSR

Campbell (1960) offers a mechanism for increases in knowledge that can be seen as an extension of Popperian ideas to creative thought. Blind-Variation-and-Selective-Retention (BVSR) is "fundamental to… all genuine increases in knowledge, to all increases in fit of system to environment." (p. 380). Even mechanisms that shortcut BVSR have themselves been originally achieved by BVSR. Campbell also argues that such shortcut mechanisms still involve BVSR at some level, although this will not be assumed here.

By analogy with evolution by natural selection, BVSR in understanding would require "a mechanism for introducing variation", "a consistent selection process", and "a mechanism for preserving and reproducing the selected variations" (p. 381).

In line with Popperian psychology, the variations can be characterised as theories. The "selection process" is presumably (following Locke) provided by Worlds 1 and 3, and can be characterised (following the model of understanding outlined above) as a problem of special interest to the individual - a "concern". Concerns would include desires, motivations and fears - in short: anything that exerts a selection pressure on the formation of theories. A concern is a World 2 construction like a theory, and can be considered as incorporating what English & Halford (1995) refer to as the "student's problem-situation model".

Campbell notes that the BVSR model of thought "joins the Gestaltists in protest against the picture of the learning organism as a passive induction machine accumulating contingencies. Instead, an active generation and checking of thought-trials… is envisaged. … Poincaré's (1913) aesthetic criteria and the Gestalt qualities of wholeness, symmetry, organised structure, and the like can be regarded as built-in selective criteria completely compatible with the model." (p. 389). However, "While 'insight' is accepted as a phenomenal counterpart of the successful completion of a perhaps unconscious blind-variation cycle, its status as an explanatory concept is rejected, especially as it connotes 'direct' ways of knowing." (p. 390). Theories do not spring fully formed into being.

Within such a model, learning need not necessarily be entirely conscious, in contrast to the view that "only things that students are fully aware of can form the basis of their learning and it only these awarenesses that can be used by the teacher to develop their learning." (Hewitt, 1985, p. 15).

While decrying the deification of "the creative genius to whom we impute a capacity for direct insight instead of mental flounderings and blind-alley entrances of the kind we are aware typify our own thought processes" (p. 391), Campbell lists four ways in which thinkers may be expected to differ, according to the BVSR model:

1. The accuracy of representations of World 1.
2. The number and range of variations in thought trials produced.
3. The accuracy and number of selective criteria.
4. The ability to retain solutions.

Of course there are difficulties with the model: Campbell himself notes it is difficult to test (giving a number of reasons for this) and it leaves variation itself unexplained. Nor do we have any idea to what extent shortcut mechanisms might be more influential in practice than simple BVSR. However, it might be reasonable to suggest (following Kant) that human physiological development has a very strong early influence on the propensity to construct certain theories - for example for space, time, quantity, quality and relation - that might then act as templates or constraints on future constructions.

One shortcut mechanism for transfer between problems could be analogical reasoning (Rumelhart & Norman, 1981; English & Sharry, 1996), in which "The ability to form an effective problem-situation model can… assist students in recognising similarity in problem structure between a known (base) problem and a new (target) problem. This can facilitate transfer of a known solution model. Such analogical transfer involves constructing a mapping between elements in the base and target problems, and adapting the solution model from the base problem to meet the requirements of the target problem" (English & Halford, 1995, p. 244). Moreover, "Successful transfer to the target problem can lead to the induction of more general models encompassing the source and target problems; these models can facilitate solution of subsequent analogous problems" (*ibid.*).

1.7.4 Some Elaboration

Popperian psychology has several strong competitors, and there is much theoretical work to be done in this area. It is beyond the scope of this thesis to discuss in any detail its relationships to other perspectives, or to offer anything approaching a critique - this research can at most act as a preliminary illustration of some of its themes. However, it may be helpful to compare it briefly with a recent theoretical framework that shares some affinities, in order to bring out some of its features, in particular its congruence with Popperian epistemology.

Drouhard & Sackur (1997) describe a developing framework for understanding the learning of algebra, influenced by Piaget and Vygotsky. They propose a "Triple Approach" that involves a subject, a social group and "reality" (either material or conceptual). Piaget's idea of knowledge resulting from interactions between psychology and "reality" was a starting point. However:

> "it seemed necessary to us to introduce a socially-related dimension, related to the society of past and present mathematicians and teachers amongst others. This dimension relies on the idea that mathematics is a social construction, and that if even some basic (e.g. logical) knowledge may be constructed just by the interaction of the child with his/her environment, it is highly improbable that s/he could build up by him/herself advanced mathematical ideas (those which have been built through a long and uncertain historical path) just by interacting with his/her environment.
>
> Obviously such ideas are related to those developed by Vygotsky [1962]. One may take care however that we are not focusing here on the *social* construction of the mathematical knowledge, but rather on the construction *by an individual* of the *socially* - and historically - *already constructed* (advanced) mathematical knowledge, that is not the same thing (even if related)." (p. 226)

This latter distinction is essentially that between Popperian epistemology and Popperian psychology. Moreover, the two "realities" that Drouhard & Sackur mention are clearly Worlds 1 and 3. But the Popperian approach would count in World 3 not just the "already constructed" mathematics but also the classroom's locally produced knowledge (written, spoken or implied). So in focusing on the individual's learning, the "social" - in either the local or the more extensive sense - is very far from being ignored. Although the teacher is usually the one responsible for setting the local rules of the game of mathematics, he or she will often have regard to the more widespread rules of the game that have developed over the centuries.

Drouhard & Sackur have a rather different view of knowledge to Popper's. For them, knowledge is all *local* - which seems to contradict their view of public mathematics - and *true* inside given limits. In the psychological area, "true" means *coherent* inside the domain where the subject may use it. In the social area, "true" means *valid*, in the sense of having been validated by a social group. In the "area of reality", "true" means *efficient.*

As previously discussed, knowledge in the World 3 sense is conjectural and "true" simply if it corresponds to what is the case. On the other hand, it can also be relatively adequate in solving a problem, and so in that sense might be considered "efficient". However, false but efficient conjectures would hardly be acceptable as "true" in either science or mathematics. Newtonian mechanics, for example, is very useful for snooker table calculations, but with the advent of quantum mechanics would not be described as "true". Similarly, the false claim that $x^2 + 1 = 0$ has no solutions is very "efficient" in early algebra classes. Drouhard & Sackur would therefore count "snooker tables" and "early algebra classes" as outside the respective "truth domains". A logical argument or an empirical conjecture could also be considered "valid" vis-à-vis its relationship with the rules of the game; but although the rules can be created and checked by social groups, there is no sense in which either a valid argument from false premises or a valid conjecture from erroneous data is a desirable truth. Finally in this analysis of "truth", it is possible to argue that knowledge in the World 2 sense is *also* "true" if it corresponds to the facts. A false idea which perfectly coheres with all one's other ideas is better than a completely contradictory false idea, but not usually as good as a true idea. A false idea may equally well be "authentic", "understandable" and "useful".

It may also be useful to point out another respect in which Popperian psychology as a tool for analysis is harmonious with Popperian epistemology as a model for research. English & Halford (1995) describe their approach as being…

> "more specifically related to cognitive science than to any school or ideology. We have done this because… we believe cognitive science provides the most accurate account so far of the actual processes that people use in mathematics and offers the best potential for genuine increases in efficiency. However, it also entails an attitude of scientific inquiry. All the tenets we propose in this book are subject to verification in further research, both in the laboratory and in the classroom. … Some of the chief tenets of the constructivist position are… incorporated, but in a form that makes their underlying processes explicit and open to scrutiny." (p. 306).

This "attitude of scientific inquiry" is at the heart of Popperian epistemology: it entails trying to put our theories into a criticizable form.

1.8 Preliminary Implications of Popperian Psychology

1.8.1 Concepts

Talk of "concept" is confusing, because (as we have seen) it can refer to very different theories of understanding; to Aristotelian essentialism; or merely to the realm of discourse. Having made a decision to follow Popper's advice to consider the act of understanding as being essentially the creation of theories in response to a problem, treating "equation" as a precise psychological construct is not useful. Nor are students' *post hoc* rationalisations about equations central. Perhaps referring to students' theories rather than their "concepts" would be more appropriate. This has the advantage that it emphasises that *what students perceive* is a personal construction (rather than a direct record of reality), but one which is potentially at odds with what is there. Therefore some theories can be better than other theories. Although it is likely for there to be commonalities between different people's theories because they have tackled similar problems, a World 2 theory is *implicit* because the knower is perhaps rarely aware of more than a small part of it at any one time and it cannot be transferred directly from one person to another. It is also *imprecise* because it is not concrete or formulated in logical terms, and its vagueness can help us to explore new situations and subtleties, by selecting perhaps mutually incompatible theories in different contexts. However, there are assumptions built into knowledge, and we can discover new things in existing knowledge that we did not realize by means of reflection. Although remembering what has been learned in lessons can seem difficult if it does not play a major role in everyday living, theories can linger.

1.8.2 The Nature of the Research

Rather than Collingwood's "empathy" with the (World 2) emotions and desires of participants, it is necessary to attempt to reconstruct participants' (World 3) problem situations: What problems were they trying to solve? What was the situation as they saw it? What *was* the situation? What theoretical solutions were proposed? Did they work? In the case of research into individuals' learning mechanisms, this does not mean restricting analysis to the individual's immediate physical and social environment, because it is the *intellectual* environment that is of most interest and this is more than just the "taken-as-shared" meanings and practices of a particular classroom community (cp. Cobb, Yackel & Wood, 1992). For the learner, such local contributions to World 3 are often indistinguishable from more widely known contributions, yet the researcher who is interested in the improvement of theories rather than purely their variety has to be concerned with *both* sets of World 3 contributions. Methods such as introspection, observation of behaviour, interviewing and studies of cultural artefacts are welcomed as sources of insight into mental processes, but they are not error-free.

A second implication of a Popperian psychology for research is that there can be no law-like cause-and-effect relationship between IT and cognition. If one is interested in the learning mechanisms by which a student's equation theories and concerns change, such broad slogans as "IT improves problem-solving at the expense of algebraic competence" have to be abandoned in favour of more subtle illuminations of the problem situation of the student. When students grapple with learning activities, therefore, the technology is portrayed as an inherent part of the problem situation. Treating "technology" or "algebra" as monolithic entities is simplistic. As Johnson, Cox & Watson (1994) discovered firsthand (also citing Niemiec & Walburg, 1992), large-scale studies involving thousands of students, several school subjects, numerous topics and a wide variety of software can sometimes tell us less about these learning mechanisms than in-depth case studies.

Davis (1992) describes how modern trends in mathematics education research are away from studying students solving routine tasks with which they are already familiar, towards "'evaluation' studies that map the student's mental representations, ability to use these representations, and ability to build up new representations in order to deal with novel tasks." (p. 239). In other words, from an emphasis on ranking students to an emphasis on understanding them:

> "Where possible, the new approach prefers not to try to infer mental processes from looking only at what a student has written on paper, but prefers to use videotaped records of task-based interviews, or of small group work, where what the student does, says, questions, revises, decides (and on what basis) can be observed more directly." (p. 239)

The research involves "cognition" is the sense that it tries to understand mental operations and their products in specific algebra tasks, but it is not directed towards finding "basic cognitive processes" that might underlie thought in general (cp. Chaiklin, 1989). This research does not focus primarily on how cognitive changes are dependent on characteristics of the student (such as social background, sex, or typical modes of learning), the teacher (for example, teaching strategies, experience, classroom organisation, or confidence) or the environment (such as setting, resources, class size), but on how changes occur as a potential result of particular *activities*. This does therefore not imply a focus on the "typical" experience of students in an algebra classroom; or on the sociological role of words that students or teachers use to talk about what they are doing (cp. Pimm, 1995); or on the metaphysical role of technology in human interactions (cp. Bolter, 1984); or on the attitudes and perceptions that might make one a better algebra teacher (cp. Mason *et al.*, 1985).

1.8.3 Algebraic Thinking

Love (1986) describes certain "modes of thought" as "essentially algebraic": for example, "handling the as-yet-unknown, inverting and reversing operations, seeing the general in the particular." (p. 164). Kieran (1989b) adds "knowledge of structures, use of variables, understanding of functions, symbol facility / flexibility,… ability to formalise arithmetic patterns" (p. 163) as further potential dimensions of algebraic thinking. Lins (1992) characterises "algebraic thinking" (distinct from "algebra") as "thinking arithmetically, thinking internally, and thinking analytically". He then uses this characterisation to explain the tensions involved in the historical production of algebraic knowledge. He also tries to show that the characterisation is an adequate framework for distinguishing students' solutions and for identifying their sources of errors. Love goes further and claims that there is "no longer a series of topics called algebra, but rather a range of ways of thinking that need to be spotted, developed and encouraged." (p. 50). It is interesting that both he and Davis (1985) contrast this purely World 2 view of algebra with "learning to manipulate meaningless symbols by following rules that you learned by rote" (Davis, p. 203) - the idea of representing situations using symbolic algebra does not seem to feature.

However, the Popperian view would suggest that psychological processes relating to algebra have to be seen as an interaction with World 3 objects, and that the context-specific nature of these processes negates the possibility of a "mode" of thought that is somehow independent of the problem situation. (cp. Linchevski & Herscovics, 1996; Arzarello, 1991, Bednarz *et al.*, 1992). Similarly, from a Popperian perspective, can there be "didactic cuts" between different *types* of thinking, such as that between "arithmetic thinking" (situation-specific strategies and values) and "algebraic thinking" (general objects, relationships and methods) as described by Filloy & Rojano (1989) and others? Talk of an "algebraic approach" may, however, be appropriate, if it describes heuristics to be contrasted with (say) a "whole-part" approach, a trial-and-error approach, or a spreadsheet approach to particular problems (Sutherland & Rojano, 1993). On the other hand, seeing operational or structural conceptions (Sfard & Linchevski, 1994) as explanatory devices could carry risks if one thought these were generalised context-free modes of thought that students can either exhibit or not exhibit. Trying to remove the context from students' thinking might obscure the possibility that it is the nature of *problems* experienced rather than the "type of conception" that is crucial for appreciating students' theory-construction.

It is clear, then, that this research is *not* primarily about how new technology may influence ways, habits, paths, styles, modes or forms of thinking. It is rather less general than that, being more concerned with the learning of particular ideas in specific contexts.

1.8.4 Behavioural Conditioning

Nevertheless, at the other extreme, modelling algebraic knowledge as a set of behaviours, with or without the use of propositional facts also seems unsatisfactory (cp. Nicaud, 1992). Firstly, as has been seen earlier, such a model does not seem to reflect hypothesised ways of thinking, such as the use of schema, imprecisely formulated ideas, analogies, images or metaphors. Secondly, it offers no plausible mechanisms for learning, beyond passive reception and repetitive association.

How procedures and propositions might be related to the active generation and checking of thought-trials, as characterised by BVSR, remains to be elaborated; and will be considered further in the next chapter. Also, it cannot be denied that repetitive exercises can sometimes serve to impose a problem as a concern. Nevertheless, it is clear that from the Popperian perspective, portraying learning as the absorption of knowledge states is unhelpful.

1.8.5 Symbol Meanings, Images, Interpretations & Metaphors

The words that students use to talk about symbols or "the meanings that students ascribe to letters" are not, in the Popperian view, fundamental insights into students' cognitive processes, but rather the theoretical consequences of past problems (especially ones solved using arithmetic or proportion). Sfard & Linchevski (1994) write: "Algebraic symbols do not speak for themselves. What one actually sees in them depends on the requirements of the specific problem to which they are applied." (p. 192). Theories from past problems may be brought to bear on new algebraic problems, or on the problem of how to respond to an interviewer asking what x means. Such recontextualised "meta-algebraic theories" should not be seen as "causes" of difficulties or as "cognitive obstacles" but as genuine (albeit sometimes inadequate) attempts to solve problems. They are, however, valuable in that they may help teachers and researchers to conjecture students' reasoning in a given problem situation.

Similarly, metaphors such as the balance model do not reflect "understanding" but derive from the problem of providing a ready-made meta-algebraic theory to rationalise certain practices. Popperian psychology would suggest that Sfard & Linchevki's expression interpretations are, in fact, theories that have been artificially separated from the problems they are intended to solve. For example, the theory that $3(x + 5)$ is a *computational process* would be relevant to a problem such as: "I think of a number, add 5 and then multiply by 3. If the answer is 99, what was my number?". The theory that $3(x + 5)$ is a *number* might result from the problem: "Predict the x^{th} number in this sequence: 18, 21, 24, 27, 30, …". The theory that $3(x + 5)$ is a *function* might be relevant to the problem "If $f(x) = x + 5$ and $g(x) = 3x$, what is $g(f(x))$?". Another example of recontextualisation is the claim of Behr *et al.* (1976) that some children "see" 2 + 4 as an instruction to add; others see it as a number. But when one is faced with a page of sums, the theory that 2 + 4 is an instruction is a pretty good one; whereas when one is solving the problem "What kind of entity results from adding two numbers.", the theory that 2 + 4 *is* a number (rather than *produces* a number) seems not to be more relevant.

Again, from a Popperian perspective, what Pirie & Kieren (1989) are describing is not therefore *per se* an image-based "growth of understanding" but a recontextualisation - the transformation of a strategy that solves a problem into a theory whose properties and relationships can be examined. This characterisation of interpretations, images and metaphors as meta-algebraic theories brought about by recontextualisation could be contrasted with the "procept" analysis of Gray & Tall (1993).

1.9 The Research Problem - A Second Attempt

There are three elements to the research problem:

1. identifying (for individual students) improvements in equation-related theories and concerns as a result of engaging in learning activities;
2. relating the improvements to the problems found in the activities; and
3. exploring the implications of Popperian psychology for the learning of algebra.

In essence, the question is: "How can students' equation theories and concerns improve?"

This concludes the definition of the research problem and associated background theories. The next chapter examines various studies into students' algebra, in an attempt to identify the algebraic theories of concern to teachers and students. The third chapter then looks at initiatives to improve students' theories and concerns.

Chapter 2
Identifying Algebraic Theories and Concerns

2.1 Introduction

Because of the importance of CSMS, this chapter starts by examining that study from the perspective of the problem of identifying students' theories and concerns. After all, a snapshot (albeit a generation ago) of students' abilities in solving algebra problems must give some indication of the theories they used, perhaps even the concerns which generated them (even though the study did not consider how students were taught).

Further studies in the 1980s on students' difficulties with letters in algebra, especially in the US, focus on the use of variables to represent situations. Although allegedly revealing deep insights into the conceptual structures used in solving algebraic problems, this chapter will argue that these studies may reveal rather more about the expectations of students as to the purposes to which symbolism is put (and thus both students' theories and concerns). Such a result might apply to many studies which have researched the nature of algebraic thinking, and this is not to denigrate them; far from it - knowledge of such expectations may help teachers to hone them; and hence the section is titled "Recontextualised Theories and their Value as Insights into Concerns". But the Popperian approach to psychology would suggest that undue authority should not be given to claims about having unearthed the fundamentals of thought. Incidentally, such a suggestion must also apply to Popperian psychology itself.

Another strand of research relates to the *theories used as strategies* in solving problems, particularly solving equations, finding patterns in number, and tackling so-called "algebraic word problems". This forms section 2.5. Finally, the implications of these studies for this research are discussed.

2.2 CSMS

The research programme "Concepts in Secondary Mathematics and Science" (CSMS) was based at Chelsea College, University of London, 1974-9. It considered the problem of the difficulty of various concepts for children at secondary school in the UK. The aim was to delineate a hierarchy of "levels of understanding" of concepts. The study was described in Hart (1981):

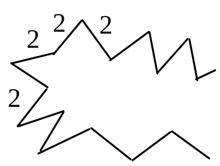
> "In many cases the order in which mathematics is presented to children is dictated by the needs of mathematics e.g. division before trigonometry or linear equations before quadratics. In other cases there is no clear order of presentation apparent in the mathematics and decisions have to be made by the teacher, based on experience and the dictates of the school syllabus. The research … took individual topics and attempted to form a hierarchy in each based on what the children tested appeared to understand."

In mathematics, this was tackled by administering pencil-and-paper tests in a variety of topics - including algebra. The tests were developed on the basis of "concepts" which the researchers considered important, an analysis of textbook treatments, discussions with teachers and interviews with children. The tests were intended to probe understanding rather than recall of standard methods, and were therefore free as far as possible of technical words. Trial items were provided to avoid the need for specialised knowledge (for example, to demonstrate the use of the arrow in $x \rightarrow 3x$ and the use of a in $a + 4$ or $4a$).

In algebra, 27 children aged 13-15 were interviewed to trial the tests and to indicate some of the methods being used to tackle the items; and 3550 children from around 25 schools took the final test.

Here are some examples of items, with the facilities for Year 8 and Year 10 given in brackets:

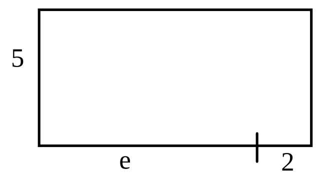
- What can you say about r if $r = s + t$ and $r + s + t = 30$ (30%, 39%)
- Part of this figure is not drawn:



There are n sides altogether all of length 2.

What is the perimeter of the figure? (24%, 41%)

- Add 4 onto $3n$ (22%, 41%)
- If cakes cost c pence each and buns b pence each, and 4 cakes are bought and 3 buns are bought, what does $4c + 3b$ stand for? (14%, 30%)
- L + M + N = L + P + N is true:
  Always Sometimes (when: _____) Never (Circle one) (11%, 27%)
- Multiply $n + 5$ by 4 (8%, 25%)
- What is the area of this rectangle? (7%, 16%)



- If $(x + 1)^3 + x = 349$ when $x = 6$, what value of x makes $(5x + 1)^3 + 5x = 349$ true? (4%, 16%)
- Blue pencils cost 5p each and red pencils cost 6p each. I buy some blue and some red pencils and altogether it costs me 90p. If b is the number of blue pencils bought and if r is the number of red pencils bought, what can you write down about b and r? (2%, 13%)

- Which is larger, $2n$ or $n + 2$? Explain. (4%, 10%)

The levels were found by a procedure which grouped items together by similarity of content and association of facility. While the levels might be interpreted as indicative of discrete stages of cognitive development, it is difficult to believe that one's taught experiences are independent of the abilities to understand and solve problems. Is it really possible to distinguish clearly "understanding of algebra" from "acquisition of methods" as is claimed? How is fluency in English controlled as a variable? O'Reilly (1990) makes a good case that "while the CSMS study contains valuable information concerning the errors and strategies which children make and adopt in learning mathematics, its 'hierarchies of understanding', rather than being universal in application are *at best* the results of particular teaching methods and conditions in England in the 1970s." (p. 77-78). He cites attempts in Germany and Taiwan to replicate the results as evidence, and describes studies demonstrating that the facility of an item varies according to presentation, context and language.

Nevertheless, in the light of Popperian psychology, the CSMS algebra results are still interesting, as they can be treated as indicating that students have tackled a range of different problems, with wide variation in degrees of success between students compared with other topics. This raises the question of how the variation between students can be explained.

2.2.1 Explaining Variation in Facility

One explanation of variation might be experience - the older the child, the more school mathematics they have experienced (facts, techniques, notation, and so on), and so the higher the level of attainment. There is a large gap between the performance of 13 year olds and 14 year olds in the study, while the 14 and 15 year-olds are rather closer together (p. 180), so surely this is a result of school experiences? Algebra being a more "abstract" topic - than, say, fractions or ratio - implies that children are less likely to have acquired the requisite knowledge for themselves without explicit teaching.

However, a longitudinal study of 105 students carried out as part of CSMS suggested that degree of success tends to change little as the child moves from Year 8 to Year 10. Only 10% moved up more than a level (although the results have to be treated with caution as there was a 50% loss of respondents). Experience does not seem to count for much.

A possible clue as to an explanation of variation was that certain errors were made very often. For example, a quarter of students in Year 8 gave some number between 32 and 42 as an answer for the perimeter of the *n*-sided figure. Almost half gave an answer of $3n4$ or $7n$ for "Add 4 onto $3n$", and over half stated that $L + M + N = L + P + N$ is *never* true.

Although several thousand children were tested, many of them would have been in the same classes, with the same teachers, in the same schools, or in similar LEAs. However, in "interviews it was found that children from very different educational backgrounds attending very different types of schools made the same type of error and often used the same methods." (p. 6). Could these errors explain children's lack of progress in attaining the higher levels?

Küchemann's analysis of the errors, influenced by the work of Collis (1975, for example) on Piagetian notions of concrete and formal thinking, produced six categories of letter usage:

Letter evaluated: the letter is assigned a numerical value from the start.

Letter not used: the letter is ignored (or acknowledged, but without meaning).

Letter as object: the letter is treated as a shorthand for an object or as an object in its own right.

Letter as specific unknown: the letter is regarded as a specific but unknown number.

Letter as generalized number: the letter is seen as able to take several values, not just one.

Letter as variable: the letter represents a range of unspecified values, with a systematic relationship existing between two such sets of values.

These categories were used to explain the popularity of certain errors. For example, the tendency to give a numerical answer as the perimeter of the n-sided figure might be indicative of treating letters as having a value which must be pre-determined, or of the ignoring of letters. The answers $3n4$ or $7n$ to "Add 4 onto $3n$" could be interpreted as the child treating letters as objects which can be collected up or as entities without meaning - the child has difficulty in accepting an "unclosed answer" like $3n + 4$ *as* an answer. The reason why $L + M + N = L + P + N$ is never true could be because L is always 12, M is always 13, and so on (letter evaluated); or because there is an M on one side but a P on the other (letter as object). Even success can be explained, to some extent, in these terms: "Using a letter as an object, which amounts to reducing the letter's meaning from something quite abstract to something far more concrete and 'real', allowed many children to answer certain items successfully which they would not have coped with if they had had to use the intended meaning of the letter." (Küchemann, p. 107).

One view is that these categories indicate the presence of deep-rooted conceptions for children at different stages of cognitive development, and these conceptions dictate the way the children then respond to items. An alternative view would be that while Küchemann's letter interpretation categories are a useful way of cataloguing the variety of strategies that students may deploy, their ubiquity could be explained simply by commonality of prior experience. Meanwhile, Olivier (1988), who found that ¾ of 13-year-olds appeared to have the misconception that $L + M + N = L + P + N$ is never true, portrays the misconception as a bad inference from "the same letter stands for the same number" (so the L on the left must represent the same number as the L on the right) to the converse "the same number stands for the same letter" (so the number that M represents could not also be represented by P) or even the inverse "'Not the same letter' stands for 'not the same number'" (p. 513).

Although the interviews were not used to probe understanding to any great extent but to trial the tests, evidence from them (and from a wide range of other research studies) indicates that children tend to use their own "informal" strategies to tackle mathematics questions rather than the standard "formal" methods taught in the classroom. These strategies work on easy items, but fail in fairly predictable ways on harder items. Students therefore do not apparently realize the power of the formal methods they have been taught. Since the focus in the study was on "generalized arithmetic" - that is, the use of letters for numbers and the writing of general statements representing arithmetical rules and operations - failure might be related to the procedures used in *arithmetic*.

Hart suggests that "achievement is closely linked to the IQ score of the child" (p. 210). In algebra:

> "There was a distinct difference in attainment (by the age of 15) between those with an IQ score below 100 and those with IQ score above 100. There were no children of IQ score greater than 100 who were still at Level 1 whereas 40% and 17% of those in the IQ ranges IQ ≤ 89 and 90 ≤ IQ ≤99 respectively were still only achieving this level." (p. 185)

If algebra makes more demand on intelligence than other topics (perhaps because of its abstract nature), and the interpretation of letters and usage of informal methods are responsible for particularly common errors, perhaps these facts are sufficient to explain variation in facility. What would follow from this for teaching?

2.2.2 Suggested Implications for Teaching: The Notion of "Readiness"

1. Hart suggests that teaching must be individualised:

> "The type of mathematics given to the children must be tailored to their capabilities. It is *impossible* to present abstract mathematics to all types of children and expect them to get something out of it. It is much more likely that half the class will ignore what is being said because the base on which the abstraction can be built does not exist. The mathematics must be matched to each individual and teaching a mixed ability class as an entity is therefore unprofitable." (p. 210)

2. She argues that algorithms should be introduced only when the child appreciates the need for them, and that they must build on informal methods:

> "We appear to teach algorithms too soon, illustrate their use with simple examples (which the child knows he can do another way) and assume once taught they are remembered. We have ample proof that they are not remembered or sometimes remembered in a form that was never taught… The teaching of algorithms when the child does not understand may be positively harmful in that what the child sees the teacher doing is 'magic' and entirely divorced from problem solving. … In order to avoid teaching rules which the child cannot apply we must first discover what the question we are asking him to answer means to *him*. If he does not see it as 'multiply' then giving him an algorithm for multiplication is not apposite. Secondly we must find out what method he normally uses to solve the problems of this type and build on that." (p. 212)

3. She suggests that teachers presenting a topic must be aware of the likely errors and take appropriate action:

> "They may even be able to build into their presentation examples which show the illogical outcome of the incorrect method. Instead of always asking a child to do a series of very nearly identical problems a useful exercise is to present a problem done in different (and erroneous) ways and ask the child to state which are wrong answers obtained from which wrong methods. To correct or discuss homework or set exercises simply by repeating the 'teacher method' seems to be of limited value. If that is all that is needed then the child would have been able to take in the method on its first

> presentation. … Perhaps we should get away from 'I'll show you' and into 'let us discuss what this means'" (p. 214)

4. Topics must be taught in an "order of difficulty":

> "It is hoped that the presentation of a hierarchy in each topic in this book will give teachers some guidance on the sequencing of the topic they teach" (p. 216)

5. More time could be spent consolidating foundational work, such as whole number arithmetic:

> "Is there really any point in teaching something we know most children will not understand? One reason given for doing this is that the child will become familiar with the idea and understand it later. We have no proof of this, in fact our results show that the understanding does not 'come'. Surely all that happens is that the child becomes familiar with a lack of success and that mathematics is something you do but it makes no sense. Is the possible answer to state that fractions and decimals are topics fit only for secondary school children and so encourage the primary school to limit its number work to whole numbers? … This does not mean that the brighter child is penalised. In order to stretch him we do not need to present him with a new set of numbers but with new situations in which he can use the numbers he knows how to handle." (p. 217)

Küchemann concurs, suggesting that the majority of 13, 14 and 15 year olds were, in Piagetian terms, at the stage of concrete operations, "which means that for most children, the teaching should be firmly rooted in this level whether the aim is to consolidate their understanding or to ease the transition to formal operational thought." (p. 118).

6. Finally, we have in summary the key message of the research:

> "All children make some progress but it is very *slow*. The less able pupil may be slowly moving through the type of mathematics which occurs in stages 1 and 2 and if he is given the type of mathematics which is essentially abstract, he is unlikely to assimilate any of it until he is 'ready'. The problem is always to find what his 'knowledge base' is and build on that. As teachers we have expectations of what a child 'should' know, very often based on intuition and usually very different from the actuality. (Hart, p. 217)

However, there is a tension in these recommendations, at which Küchemann hints:

> "It … seems sensible to base the teaching given to children at Levels 1 and 2 on the meanings for the letters that these children readily understand. On closer examination this is by no means a straightforward task, for example, the use of letters as objects totally conflicts with the eventual aim of using letters to represent numbers of objects. However, it may well turn out to be the case (see Inhelder *et al.*, 1974) that it is precisely through being made aware of such conflicts that children see the need to reorganise their thinking and thereby move towards a higher level." (p. 119)

In what sense are students not "ready"? Is it in the sense that a new-born baby is not yet able to walk - some biologically-based maturation needs to take place? Or is it in the sense that a 1-year-old is not yet able to talk - more experimenting is needed before sufficient behaviours are available? Or are students not "ready" in the sense that a child who does not know either side at the Battle of Hastings would not know who won - some prerequisite facts are needed? Or is it in the sense that a colour-blind person might not be able to tell which cloth is red and which green - certain innate faculties are missing? Or are there other possibilities?

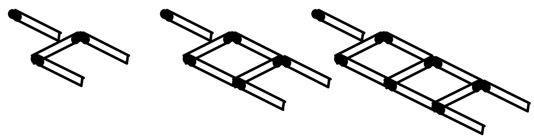
The Popperian formulation "Understanding a theory is understanding the problems it is intended to solve or which it solves, and understanding why other potential theories are not solutions." might look like a sleight of hand. After all, if we want to know the pre-requisites for a theory, we then have to know the pre-requisites for the problem - which presumably are theories in their turn - and so on. It does not seem to get us any further forward. However, if one looks at the examples given in CSMS, it could be argued that the reason that 92% of Year 8 students could not multiply $n + 5$ by 4 was that they did not understand the problem.

Following this line of argument, virtually *all* of the items in CSMS contain algebra in their formulation and require familiarity with mathematical conventions for their understanding. For example, the problem “multiply $n + 5$ by 4” is a demand (by convention) to find an equivalent but simpler expression. A student who gives the answer “$n + 20$” is apparently either not aware that equivalent expressions can be checked by substituting particular numbers, or is not concerned to check (for whatever reason). In other words, the success criteria for this problem are not apparently accessible to students.

On the other hand, a problem such as…

Here are some piles of matches:

How many matches are in the 100th pile?
What is the number of the pile with 568 matches?



Pile 1 Pile 2 Pile 3

… does not require the use of algebra in its formulation; but it is easily grasped, and those who have grasped algebraic techniques can demonstrate that they can use them. The objection that this problem can be solved by arithmetic means is not helpful; the formula governing the matches can be made sufficiently complicated so that algebra is an advisable shortcut (after all, why bother with algebra if arithmetic will always do?), or students can be asked to justify their solutions and compare them with other students’ reasoning. The use of this sort of problem for learning is praised by Küchemann: “… the task of finding the relationship, of representing it economically and unambiguously, and of comparing the equivalent representations provides children with a worthwhile challenge.”. Hart notes that in the interviews, replies to the question “why?” were “often tautologous” or “an expression of feelings” (p. 217). Such an activity could help students to clarify what counts as a valid explanation.

The above discussion is intended to show that evidence from the CSMS algebra test does not allow us to decide between the hypothesis that students understand the problems they are being asked to solve but lack the intellectual resources to find an answer, and the hypothesis that students do not understand the problems.

Sutherland makes a different, although related criticism:

> “I suggest that the difficulties which pupils develop with algebra will at least be partly related to the ways in which algebra has been introduced to them in school. It seems surprising therefore that a major study, the CSMS algebra study, did not in any way attempt to describe or link the algebraic practices from which these difficulties stemmed, especially since at the time of the study it is likely that some schools were using the ‘new mathematics’ approach and others a more traditional approach.” (Sutherland, 1990, p. 159)

O’Reilly (1990) was “struck by the lack of controls in a major research project of this nature” (p. 84), particularly with respect to teaching schemes and styles. The schools in the sample were self-volunteering, and the only variable attended to was IQ: comprehensive schools were mostly used, but grammar schools used when the "IQ distribution departed from the normal curve and extra children with high IQ were needed" (Hart, 1981, p. 5).

2.2.3 Further Questions and Issues

There are several questions arising from CSMS for which further evidence would be very useful. How can variation in facility between students be explained? Why are certain errors common? Is algebra especially difficult? But this research will not be focused on these issues. It concerns itself rather with the question: given a lack of progress in algebra, "what experiences, if any, might bring about a more substantial change?"(Küchemann, p. 118).

Of special interest to this research is the experiential pathway of equation solving described earlier by Filloy, Rojano and Sutherland; a pathway distinct from the pathways of simplification using the general rules of arithmetic (as in the early 20th Century), representing situations symbolically (as in CSMS), and exploring patterns using letters (as in the 1980s). The issue of how children's difficulties with symbolism should be managed is of crucial importance in comparing such experiences. Before examining this issue in detail in chapter 3, further important research delineating students' theories is considered, especially those used as strategies and - in the next section - those that appear to give insight into cognitive processes.

2.3 Recontextualised Theories and their Value as Insights into Concerns

Pereira-Mendoza (1987) asserts that "Comprehending the solution of equations, factorisation, polynomials, etc., depends on students' comprehension of algebraic symbolisation. Without this comprehension, algebra will be internalised as a set of disjoint and meaningless rules." (p. 331). Popperian psychology requires a reinterpretation of this position. Interpretations of expressions and letters have already been described, in sections 1.8 and 2.2 respectively. This section considers the interpretations of letters in expressions, interpretations of the equals sign, and interpretations of equations.

2.3.1 Interpretations of Letters in Expressions

Küchemann's categories of letter interpretations are frequently used in research as an explicative device. Indeed, in a sequel to CSMS, called SESM, Booth (1984) found in interviews that "children may handle letters as 'objects', especially in more abstract examples" and that they "may interpret letters as specific unknown numbers in situations which require consideration of them as generalized number." (p. 38). However, for problems such as "Add 3 to 5*y*":

> "It would… seem that the level of letter interpretation may bear little relation to success… in that children interpreting letters as objects may get the item right as well as wrong, while children who recognize letters are representing number may still produce the erroneous answer. It appears, therefore, that the meaning of letters may not be something which children take into account in answering such items; and in fact during the interviews each of 16 children who gave the erroneous answer '$8y$' did so by applying the 'rule' of adding the numbers and writing down the letter, regardless of the meaning which they then went on to ascribe to the letter itself." (p. 28)

This raises the rather important question: what *is* the interpretation - that which is evidenced by the answer or that which is brought out in response to a direct request for an interpretation? Ursini (1990) found, in a test of 65 11-14-year-olds' abilities in symbolising situations and interpreting symbolisations, that these skills were not obviously interdependent. Bell (1988) provides evidence with a class that "the distinction between x as a specific unknown and x as a generalised number did not appear to relate to conceptual difficulty." (p. 152). Moreover, there are a larger number of ideas that students may have than are covered by the categories. Booth describes, for example, interpretations such as $4y$ representing not just $4 \times y$ but also 4 y's, forty y or $4 + y$, which might explain substitutions for $y = 3$ of 12, 7 or 43.

Kutscher & Linchevski (1997), meanwhile, observed that their students found it easier to treat multiplicative algebraic expressions (such as $4x$) as a single number-like entity than additive algebraic expressions (such as $x + 6$). Some students found brackets useful in this regard. For example, the sum of x and "x•4" (meaning $x \times 4$) could be represented by these students as $x + 4x$, but the sum of x and $x - 31$ proved more problematic until the teacher demonstrated $x + (x - 31)$. Kutscher & Linchevski suggest one possible explanation: that the "manipulation of x•6 or 6•x to $6x$ assists in the evolving of the structural aspect of the expression." (p. 171).

However, characterising the "interpretation" or "meaning" as a *theory* which can be used or not used, be made explicit or remain unconscious in various situations gives an added versatility to Küchemann's categories. This perspective acknowledges that to some extent students' thinking is being "recontextualised" by such analysis (recall Sfard & Linchevski on expressions in chapter 1), and that it is difficult to tell whether this recontextualisation is giving us a deeper insight into cognition at a basic level, or a broader and richer, but shallower, panorama. For example, early empirical work for this research indicated some extra interpretations to be considered at A-Level: "unknowns" are constants to be found; "constrained variables" vary according to external rules; "unconstrained variables" can be varied by the mathematician; and "parameters" specify particular constraints. Do these belong to the "deep" or "panoramic" variety of recontextualisation? Perhaps one should be wary of letting anything that is formally taught into the "deep" category.

On the other hand, Booth suggests that misconceptions can be attributed not only to "inadequacies in the teaching-learning situation" (p. 87), but "some of the difficulty which children have appears to be related more to a 'cognitive readiness' factor. … the view that there is a deeper basis to this conception appears to be supported by the data which indicated a strong resistance, on the part of children in the present study, to the assimilation of the idea of letter as generalized number even within the context of a teaching programme specifically designed to address this aspect of algebra." (p. 87). The dangers of such an inference were made clear in the section on CSMS. Booth wants to link this finding to Collis' suggestion that appreciating the generalized nature of letters is characteristic only of those at the stage of formal operational thinking. Nevertheless, although propounding this classically Piagetian view, she goes on to say about the acceptance of unclosed expressions that "the apparent effectiveness of the teaching programme in restructuring children's thinking in this regard would suggest that the notion was not beyond the conceptual grasp of these children.". However, this is indicative, for her, that "the acceptance of lack of closure, and the view of letters as generalized rather than particular number, may relate to different levels of conceptual difficulty, rather than be manifestations of a single cognitive structure as suggested by the Collis-Piaget formulation." (p. 91). The way to challenge this formulation, therefore, would be to succeed where SESM failed. SESM is discussed in more detail in chapter 3.

Nevertheless, whatever the merit of particular categories of letter interpretations, or the truth of developmental theories, it should be apparent that students' concern to use literal symbols for algebra is a desirable, but often elusive goal. For example, one of the teachers associated with this research commented about GCSE students' difficulties with the idea of x standing for numbers that "They can't visualise it, they can't do anything with it. They can do lots of things with a 2, but what can you do with an x? I was amazed that it took so long for them to realize that if there is a $2x$ somewhere, it actually means 2 times x.". Similar difficulties are reported in Gallardo & Rojano (1987). The students' question "What can you do with an x?" can be explicated, in the Popperian view, as "What problems can you solve with an x?". The raising of the question would suggest that the students lack *concerns* for letters.

Kieran (1989a) writes that "High school algebra usually starts with instruction in the concept of variable." (p. 40) and suggests this may be a remnant of the New Math view of it as a unifying idea. Pereira-Mendoza (1987) writes for example that "The basis of algebra is the concept of a variable and its associated notations." (p. 331). Meanwhile, according to Sakonidis & Bliss (1990), "The concept of a variable is one of the keystones of the discipline of mathematics" (p. 133). Schoenfeld & Arcavi (1988) see "The concept of variable" as "central to mathematics teaching and learning in junior and senior high school." (p. 420). Mason *et al.* (1985) seem to see unknowns as special cases of variables. But in unifying unknowns, constrained variables, unconstrained variables and parameters, the distinctions may have been blurred (Matz, 1982). Variables are often seen as more cognitively sophisticated entities than unknowns (English & Halford, 1995, p. 220), yet they address different concerns.

2.3.2 Interpretations of the Equals Sign

How do young children interpret the equals sign? Behr *et al.* (1976) carried out non-structured individual interviews with children from 6 to 12 years old. They found that the equals sign is seen as a request for an answer, a "do-something signal", rather than a relational symbol comparing two expressions.

Some considered the sign acceptable only when one or more operation signs precede it (so $2 + 4 = 6$ is acceptable, $6 = 6$ is not). Meanwhile, when faced with non-action sentences such as

$3 = 3$ or $3 = 5$, children wanted to make an action out of it, for example by re-writing as $3 + 3 = \square$ or $5 - 3 = \square$. Many young children thought the answer should always be on the right, so $\square = 2 + 4$ is not acceptable because it appears to be written backwards; and some children even rewrote it as $2 + 4 = \square$ or $\square + 2 = 4$. "2 and 4 make 6" would be a typical reading.

Identities such as $2 + 3 = 3 + 2$ were rejected by 6-year-olds, and rewritten as separate sums ($2 + 3 = 5$ and $3 + 2 = 5$) or as one big sum ($2 + 3 + 3 + 2 = 10$).

> "[They] do not view sentences like $3 + 2 = 2 + 3$ as being sentences about number relationships. They do not see such sentences as indicating the sameness of two sets of objects. Indeed, it appears that the children considered these as "do something" sentences. In most cases the presence of a plus sign along with two numerals, suggests that another number, an answer, is to be found." (p. 15)

To a few, $2 + 3 = 3 + 2$ could be acceptable, because both sides contain the same numerals; but then $2 + 3 = 4 + 1$ would not be acceptable.

Kieran (1981) writes about the notions of the equals sign as a "do something signal" or "makes":

> "It can be argued that these notions reflect the kind of instruction that these children have received. One might then assume that later exposure to equality sentences involving the commutative and associative properties might broaden the elementary school child's notion of the equal sign. However this does not appear to be the case."

To back up this claim she cites the conclusion of Behr *et al.* (1976) that "there was no evidence to suggest that children changed in their thinking about equality as they progressed to upper grades; in fact, even sixth graders seemed to view the equal sign as a 'do something signal'." (p. 319). Kieran also refers to Denmark *et al.* (1976), who "designed a teaching experiment to teach the concept of equality as an equivalence relation to a group of first grade students" (p. 319) using a balance and written equations, and they understood $3 = 3$, $3 + 2 = 4 + 1$, $5 = 4 + 1$, etc. But they still saw the equals sign as an operator primarily, not as indicating a "relation between two names for the same number". More recently, Dickson (1989) asked 11-13 year-olds the question "What does '=' equals mean?" to test for the relational notion of equality, as opposed to an operational "gives the answer". The operational interpretation was found to be common among and persistent throughout a period of instruction in solving equations by formal methods.

Herscovics & Kieran (1980) found similar results:

> "It is telling that when we asked seventh graders to provide an example in which they used the equal sign, they limited themselves to an operation with two numbers on the left side and the result on the right." (p. 573)

But Herscovics & Kieran managed to find an approach that successfully broadened the notion of the equals sign within arithmetic using identities (see section 3.4.7). This suggests that experience *does* seem to be the crucial factor in expanding notions. Moreover, how would the specifically 20th century use of the equals sign in programming ($n = n + 1$ is an operation to increment n, not a relation governing n) fit into some "natural" conception of the equals sign? In any case, in many contexts, the equals sign *is* a "do something signal" - it means give a simpler answer; in others it means "makes the same as". Admittedly, $3 = 3$ can never be a "do something signal", but what sort of exposure have children had to such cases? How is such a statement useful to them? What concerns would such identities address?

Kieran describes high schoolers as being in transition "between requiring the answer after the equal sign and accepting the equal sign as a symbol for equivalence.", and this is corroborated by Cooper *et al.* (1997).

Kieran also considers the use of the sign in such cases as $1063 + 217 = 1280 - 425 = \ldots$ may conceal a "tenuous grasp of the underlying relationship" (p. 317). "Even college students taking calculus, when asked to find the derivative of a function, frequently seem to be using the equal sign merely as a link between steps." (p. 324). But this could alternatively demonstrate that the concern to "show working" in as concise a way as possible outweighs any concern to demonstrate equivalence statements - perhaps because students have not appreciated the importance of this algebraic convention.

Cortés, Vergnaud & Kavafian (1990) give four meanings of the equals sign: introducing a result, equivalence, identity and definition; while Kieran (1992) distinguishes sameness", "equality" ("comparative" or "operator"), "equivalence" and "identity". Early empirical work for this research found uses of the equals sign among A-Level students including both as a relation (identity of references, similarity of references, equality of values) and as a process ("Makes", "Leaves", "Is made by", "Is worked out using the formula", "Gives the answer"). But does such a variety of uses give us any insight into students' "concepts"?

Kieran is in no doubt: "The importance of the equal sign in high school algebra cannot be overestimated. Most children identify algebra with equations. In fact, the absence of the equal sign seems to create huge conceptual problems.". For example, "*a* means nothing... there is no equal sign with a number after it." (p. 324). "Perhaps this explains why students have difficulty in dealing with polynomials later in high school when they are introduced as indeterminate forms." (p. 324). This is corroborated by Chalouh & Herscovics (1988) who documented students believing that expressions, representing areas (say), were somehow incomplete unless they included an equals sign: Area $= 4x + 4y$ and $4x + 4y = 10$ were legitimate; but the expression $4x + 4y$ by itself was not.

2.3.3 Interpretations of Equations

Little research has been carried out into interpretation of equations, as opposed to interpretations of symbols within an equation. The view that interpretations must be considered the "root" of students' difficulties with equations would make such research a priority. Popperian psychology dissents.

However, there are some interesting meta-algebraic theories about equations. SMP (1978), for example, introduces equations initially as a description of a flow diagram turning the input into the output; but this is generalised to a relation containing an equals sign and "at least one letter (usually x) representing some number whose value is not known to begin with" and "correct for only a few values of x". (p. 2). A formula, meanwhile, is "a relation between two or more variables associated with a particular object or situation" (p. 29). On the other hand, SMP (1983) introduces equations in the context of balances (to ease the rule "What you do to one side, you do to the other."), and the unknown weights are eventually represented by letters. But in addition to equations being portrayed as a flow of operations, a relation and a balance, it is also possible to emphasise how one might use an equation. For example, Cortés, Vergnaud & Kavafian (1990) quote 7th and 8th grade students in France describing an equation as "it translates a text", "it simplifies like a shorthand", "the unknown can be put on one side and it is possible to calculate it more quickly" (p. 30). Additionally, there were early indications in this research that students would associate equations with particular solution methods, such as "balance", "formula" and "function". Such methods will be considered in section 2.5.

There are also some serious misconceptions reported. For example, Stacey & MacGregor (1997) and Dickson (1989) found students believing that an x in one part of the equation might have a different value to an x in another part of that equation. In Booth (1984), in response to the question "Can $x + y + z$ ever equal $x + p + z$?", several 15-year-olds justified their answer "Never" by something akin to "But if y and p were the same, you'd have thought they would have put x, y and z; instead of x, p and z.". Moreover, Dickson (1989) records some students thinking that when $x + y = 10$, x and y cannot *both* be 5; they also fail to consider that x could be zero or negative. Wagner (1977) asked the question "Would different solutions be obtained from $7w + 22 = 109$ and $7n + 22 = 109$?" of a variety of children. Some said that the solution of the first one would always be greater; and others that they couldn't tell until they had solved the equation. Those who had no doubt that the same solution would be obtained from both equations (and thus realised that a change of letter does not affect the solution) were called "conservers". About half of 12-year-olds and a fifth of 14 and 17-year-olds failed to "conserve equation". Dickson (1989) notes that some children could quite happily see 109 as the answer.

Kieran & Wagner (1989) point out Thorndike's (1923) preference for keeping separate from the start the two aspects of an equation: "a thing to be solved" and "an expression of a general relation among variables". But the blurring of all usages of letters as aspects of variables (described above) may have been unfortunate for the understanding of equations. For instance, Clement (1982) suggests that "understanding an equation in two variables appears to require an understanding of the concept of variable at a deeper level than that required for one variable equations." (p. 22). This suggestion arises from a particular set of predominantly US studies.

2.4 The Student-Professor Problem

Various studies examined the difficulties that students have with translating word problems into algebra. One classic problem originates with Kaput & Clement (1979):

Write an equation, using the variables S and P to represent the following statement: 'At this university there are six times as many students as professors.' Use S for the number of students and P for the number of professors.

Rosnick (1981) gave the problem to around 150 entrant engineering students at the University of Massachusetts. Over a third could not write the correct equation, S = 6P. An even more incredible ¾ of the students failed when the ratio was 4:5 instead of 6:1. Social scientists did worse still. The most common error was the "reversed equation", 6S = P.

Rosnick used evidence from interviews to argue that S was being used to stand for "students" rather than "number of students". In other words the letter was being as a label for objects rather than numbers. To test this, students were given the following information in a questionnaire "At this university, there are six times as many students as professors. This fact is represented by the equation S = 6P" and asked what the S stood for using a list of options (professor, student, students, number of students, none of the above, more than one of the above, don't know). Almost half failed to select "number of students", including 22% who thought it stood for "professor"…!

Rosnick suggests that his results show that "the misconceptions that students have surrounding the use of letters in equations contribute significantly" (p. 418) to translation difficulties from English sentences into algebraic expressions. If college students struggle to distinguish letter as variable from letter as object label, no wonder they struggle to understand when letters are labels, variables, constants, parameters, etc.

> "The curriculum of mathematics... follows a path of increasing abstraction. As the curriculum becomes more abstract, the symbols used become more obscure. For many students, as was true for me, unfamiliarity with mathematical symbols and the abstract concepts to which they refer breeds contempt for mathematics." (p. 418)

Gibbs & Orton (1994) point out that "research in the past has been more adept at identifying the problem issues than at providing guidance for effective intervention in the classroom." (p. 102). Rosnick bemoans the apparent resistance of this error to whatever teaching strategies were tried. Perhaps its persistency and prevalence are indicative of some deep-rooted psychological structure or obstacle that must be altered or overcome before true algebraic enlightenment is attained. Philipp (1992) is in no doubt, for example, that his study aimed not just to find theories used to solve this sort of problem, but to "gain further understanding of students' conceptions of algebraic variables".

Wagner (1983) analysed, in this vein, the similarities and differences between letters and numerals, and letters and words which might hinder the learning of algebra. This sobering exchange is given:

Teacher: Suppose we use *x* to represent an unknown integer. How can we write the next consecutive integer after *x*? That is, how can we represent the number we get when we add 1 to *x*?

Student: (*without hesitation*) *y*

Wagner argued that "Literal symbols [are] easy to use but hard to understand." (p. 474), and cites many potentially confusing characteristics of letters. For example, a letter can be used as an abbreviation for a word, as a name for an object, or as a representative of a number subject to numerical operations and relations. Letters are like numerals in many ways; however "numerals represent a single number but letters can represent, simultaneously yet individually, many different numbers, as in $0 < n < 20$ or $y = 3x + 2$" (p. 475). This "property of simultaneous representation" helps to produce general, but concise and unambiguous statements. Different letters can represent the same number but (in a given context) one letter should stand for the same number or numbers each time. -1 is negative, but $-x$ may be positive *or* negative. There is no connection between alphabetical order and numerical order. $3mn$ means $3 \times m \times n$, whereas 347 means 300 + 40 + 7, and this relies on the fact that, because letters can represent multidigit numerals, there is no ambiguity. Letters are like pronouns in that "both can act as placeholders in certain expressions" (p. 476). For example, the "He" in "He is a mathematics teacher" can be replaced by the names of different men to make true or false statements; just as the x in $3x^2 = 10$ can be changed. Yet if we changed the "He" to a "She" this would, in itself, make a difference to the reference, whereas changing the x to a y would not. Hence perhaps Wagner's (1981) results on "conservation of equation" referred to above.

Wagner concludes:

> "The more characteristics [of literal symbols] we can identify, the better able we are to devise teaching strategies to help students understand and appreciate these things we call variables."
> (p. 474-5)

Other studies (such as Rosnick & Clement, 1980; Clement, 1982; Kaput & Sims-Knight, 1983; Wollman, 1983; Cooper, 1984; Cooper, 1986; Lochhead & Mestre, 1988; Fisher, 1988) found similar results to that of Rosnick; with secondary school and university students and teachers; in the US, Fiji and Israel; with situations of varying familiarity; with and without pictures or tables; with different variable names, and so on.

2.4.1 Examining Strategies

However, myriads of explanations have been proposed that do not see the cause of the student-professor error as misconceptions deriving from abstract characteristics of letter usage. For example, Gibbs & Orton (1994) go along with the argument that "It is the very structure and grammar of the English sentence that leads to mistranslation to the algebraic equivalent." (p. 102). The "natural order of the sentence is followed in translating to the algebraic equivalent". Thus one thinks: "There are 6 times as many students (that is, 6S) as professors (that is, = P)". It is only "thinking about the meaning of the sentence or substituting real numbers" which reveals the error.

Clement (1982) described four strategies used by students to construct the equation:

- **word order matching** - This is also called the "direct-translation approach" or "syntactic translation" (Herscovics, 1989) and has been identified by a number of studies (see Kieran, 1992). It "involves a phrase-by-phrase translation of the word problem… Some semantic knowledge is often required… but solvers typically use nothing more than syntactic rules" (p. 403). In the student-professor problem, it appears to be the formulation "There are…

times as many… as…" that ensures this strategy fails. If the sentence were "The number of students is 6 times the number of professors." then word order matching would succeed.

- **static comparison** - the number is placed next to the letter representing the largest group: "There are 6 students (that is, 6S) for every professor (that is, = P)"; or perhaps "One professor (that is, P) is worth 6 students (that is, = 6S)". One could argue that this strategy also assumes an interpretation of the equals sign as "is worth" or "for every" rather than the usual "is the same as" (Wollman, 1983), but interviewees might equally say, "6 students equals one professor"; so even the word "equals represents correspondence rather than equivalence. The number acts as an adjective.
- **operative approach** - a operation is sought that turns one number into another: "multiply the number of professors by 6 in order to get the number of students (that is, 6 × P = S)". Clement says that the equation in this case "does *not* describe the situation at hand in a literal or direct manner; it describes an equivalence relation that would occur if one were to perform a particular *hypothetical operation*, namely, making the group of professors six times larger than it really is." (p. 21). Other operative approach solutions include S/6 = P, S/P = 6 and P/S = 1/6. This produces correct responses because, unlike word order matching and static comparison, it corresponds to mathematical usage. The 6 is a multiplier on a variable quantity P to produce another quantity S.
- **substitution** - the student seeks likely equations (such as 6S = P and 6P = S), chooses some numbers which fit the situation described (such as P = 10, S = 60), substitutes the numbers into the equation, and chooses the one which works. The use of the word "times" next to the S in the problem, always suggestive of multiplication, may be used by some to choose an equation rather than trying a substitution.

Word-order matching could be eliminated as a strategy, perhaps, by using pictures or well-known relationships that do not have to be stated explicitly. Pictorial representations actually seemed to increase the difficulty (Mestre & Lochhead, 1983; Sims-Knight & Kaput, 1983). Sims-Knight & Kaput (1984) found that problems using familiar quantitative relationships (5 fingers on a hand, for example) were more difficult than unfamiliar relationships (such as 6 students to 1 professor); however, they did not take steps to avoid explicitly stating the relationship. MacGregor & Stacey (1993b) produce further evidence that word-order matching is not, after all, a common strategy. "In test items designed so that syntactic translation would produce a correct equation, most students did not translate words to symbols sequentially from left to right" (p. 217). For example: "'The number *y* is eight times the number *z*.' Write this information in mathematical symbols." (p. 222) was answered correctly by only 37% of 281 Australian 14-year-olds drawn from 68 classes and 21 schools. 52% produced some sort of reversal, although this includes expressions and inequalities in which the 8 was associated with the *y* rather than the *z*. Better performance was seen on an item asking "'*z* is equal to the sum of 3 and *y*.' Write this information in mathematical symbols.", and MacGregor & Stacey conjecture that "is equal to" is a better prompt for *word-order matching* than "is", and a clear test of this would be useful. The facility of the following problem was half that of the previous one: "*s* and *t* are numbers. *s* is eight more than *t*. Write an equation showing the relation between *s* and *t*." Could this be because "more than" is harder to translate than "is equal to"? Only a third of around a thousand 13-15 year-olds could answer the following question correctly: "I have \$$x$ and you have \$$y$. I have \$6 more than you. Which of the following equations must be true? $x = 6y$, $6x = y$, $x = 6 + y$, $6 + x = y$, $x = 6 - y$" .

MacGregor & Stacey also take care in their items to avoid the potential use of letters to refer to objects rather than numbers, but they have not entirely eliminated the use of a strategy based on static comparison of objects, as they imply. For example: "The number y is eight times the number z." could be translated as "Whatever y counts, there are 8 of them for every thing that z counts. So 8 y's for every z. So $8y = z$". Even so, if a static-comparison strategy were to be caused by a misinterpretation of letters as standing for names rather than numbers, one would expect that warnings to this effect or the use of neutral letters would diminish the number of reversals. Most studies show this not to be the case. Fisher (1988), for example, used N_s and N_p as variables and the results were even worse. So how can the popularity of static comparison be explained.

2.4.2 Another Perspective: Problem Expectations

Now the important point made about the CSMS problems - that they predominantly require familiarity with mathematical conventions for their understanding, and hence their success criteria may not be easily accessible to students - also applies here. The student-professor problem is a professor's problem, not a student's. However the answer is found, the only way it can be checked is by reference to the ways in which one expects it to be used.

How will it be used? Presumably to answer the question "How can we find S given P, or P given S?". If the question were to predict how many professors there would be if there were 120 students, no doubt anyone with the equation 6S = P would be able to think "Right, 20 times as many students, so 20 times as many professors.". If the question were to predict how many students there would be if there were 10 professors, no doubt anyone with the equation 6S = P would be able to think "Right, 6 students for every professor, and there are 10 professors so 6 × 10 students.". In other words, so long as the interpretation of 6S = P remains "There are 6 times as many students as professors." or "There are 6 students for every professor.", it will also possible to use ratio arguments - or imagining one group of people being multiplied or divided to get another - to solve problems requiring a numerical answer. That the resulting *form* is "wrong" stems more from convention than "understanding". For those who got it wrong, 6S (in the sense of "6 students") *is* a number, but S *isn't* (it is "students"). Wollman (1983) suggests that computation is straightforward, even for those who reversed equations; while Seeger (1990) appears to show, for around 550 students between the ages of 13 and 24, that those students giving the reversed equation were mostly able to produce a correct computed solution from an equation.

However, if the expected usage is to solve a problem demanding algebraic (as opposed to ratio) arguments, conflict and possibly errors will occur. For example, one might expect to be asked to substitute the value S = 72 *into the equation* (as opposed to finding the number of professors if there are 72 students); or to make S the subject of the formula; or to compare this university with another for which (say) S = 4P + 9; or to communicate the fact to another mathematician who knows the convention that equations should be amenable to algebraic operations such as substitution or re-arrangement and may therefore interpret 6S = P as "6 times the number of students is the number of professors". Wollman (1983) suggests the question "Can you tell from the sentence which number is greater?" can enable many students to correct reversals, a result which at first glance conflicts with this analysis. However, almost 80% of Wollman's students were using word order matching rather than static comparison - a much higher percentage than most other studies, especially MacGregor & Stacey's. Indeed, Wollman writes that static comparison reversals "were less readily corrected during interviews. Moreover, the students who made these errors seemed less confident of their corrected answers." (p. 76).

MacGregor & Stacey "explain the forms of students' responses by a theory of cognitive models" (p. 221). They are unambiguous about the process of formulating equations from situations: "Students generally try to make sense of the text of problems, and in doing so they intuitively construct mental images or models. What they then write on the page is an attempt to reproduce the content and form of the model." (p. 230). However, if one has shown that there is no mathematical or pedagogical justification for an apparent and persistent distinction in student behaviour, the distinction does not have to be explained by postulating an "underlying cognitive model". The incorrect responses could be equally well interpreted as indicative of a lack of concern for algebraic convention. By seeing students' mathematical knowledge in terms of strategic theories and concerns, such reasoning is unnecessary. If one wants to view these theories as symptomatic of anything, it should be of *an attempt to solve a problem*, rather than of "non-linguistic conceptual structures based on fundamental relations such as grouping, comparison and contrast." (p. 228). A confident assertion that the response tells us something about pre-existing causative psychological entities has to be tempered by the reminder that we may not even know what problems students are trying to solve.

To summarise: getting the reversed equation perhaps tells us more about students' expectations of the concerns that they think are relevant than about cognitive structures or obstacles. These concerns will be central to this research. The strategies they use - such as *word-order matching* or *static comparison* - are symptomatic of such concerns.

2.4.3 A Sample Analysis - Philipp (1992)

One of the most recent investigations of the student-professor problem was carried out by Philipp (1992) with 295 high school students. One variation of the problem replaced obvious letters like S and P with "neutral" letters like X and W. Another variation involved implicit relationships:

Pennies-Dime Problem

You have a pile of pennies and another pile of dimes. The value of the pile of pennies is the same as the value of the pile of dimes. Write this as an equation, using P for the number of pennies and D for the number of dimes.

The students were from two schools, and involved 13 classes (4 first-year algebra, 5 geometry, and 4 second-year algebra). One of four questions was given to each student, presumably at random; so although there is no way of comparing outcome in one problem with outcome in another, we can compare proportions succeeding with the different problems. Philipp concluded that problems involving implicit relationships are significantly harder than those requiring an explicit statement of the relationships; while the choice of letters made little difference to the proportion of correct results in either problem. But the data he gives can be analysed to bring out more subtle conclusions.

For the student-professor problem, 33% were correct, 41% reversed and 25% made some other error. Using neutral letters in the algebra classes seemed to increase the reversals, but decrease other errors. Why? Unfortunately, no classification of other errors is given, but could it be that it is easier to check an equation is correct by substitution or the operative approach if it is obvious which letter refers to which group? Does the need to move one's head between the statement of the relation to the equation to the definition of the variables, while perhaps keeping an image in mind, increase the chance of forgetting the need for such comparison with algebraic convention? But none of this explains why the proportion of other errors decreases. The only way to make progress with this is to find a way of determining which strategy students are using, which Philipp indeed attempts (see below).

Meanwhile, compare these results for high-school students with Kaput & Clement (1979), in which 63% of first-year college engineering majors were correct and 43% of social science students were correct. It may be that those who can succeed in such problems go on to do engineering, or that those who want to do engineering tend to improve their performance dramatically from Algebra 2 (40% correct). It may be that those who end up doing social science do not improve from Algebra 2, or that there is a wide variation from class to class.

Turning to the penny-dime problem, 13% were correct, 38% reversed and 49% made some other error. Again, using neutral letters in the Algebra 2 classes seemed to increase the reversals, and decrease other errors. Why is the penny-dime problem so much harder than the student-professor problem? One would perhaps expect that a problem involving both frequency and value would be harder than one only involving frequency. Seeking an operation is fraught with extra difficulties, because imagining a pile of dimes being multiplied up by 10 seems to be harder than multiplying a group of professors - a dime is inherently worth more than a penny; whereas a professor has no numerical value in terms of students. Are any of the students writing "P = D" to indicate the piles have equal *value* rather than frequency? What about "a dime (that is, D) is worth 10 pennies (that is, = 10P)"? Wouldn't we therefore expect more reversals?

2.4.4 Results from Interviews

There is a serious difficulty with the quantitative approach to researching conceptions of variables. Varying word order, the letters used, or the presentation of the problem neglects to take account of the fact that many students will no doubt, in translating the information given into a form of words familiar to them, end up writing an equation based on a very different form from the one given.

Even showing the information in the form of a table (right) would suffer from the same problem. Many students would no doubt want to codify a proposition expressed in natural language, and hence we are back to "translation" issues.

P	S
1	6
2	12
3	18
4	24

Even asking students how they know their answer is correct (hoping they will say something like "I substituted the numbers P = 10…" or "I imagined a small group of professors being multiplied…") is not as straightforward as one supposes, as Philipp's second study demonstrates.

Half-hour interviews were carried out with 7 adults, aged 25-35, of diverse mathematical backgrounds. One wonders why only adults were interviewed. The obvious thing would have been to interview some of the students, to gain an insight into their strategies. Are adults expected to be more eloquent? There are good reasons for expecting different results to the quantitative study. Surely adults will make less use of automatic (and potentially meaningless) syntactic processes such as those practised in high-school algebra classes? And if they do not use mathematics regularly in their work, surely they will be less attuned to the mathematical conventions?

Liz describes how she (correctly) answered the question: "There are more pennies. Ten times as many pennies as dimes. So we must multiply the number of dimes by ten to get the number of pennies.". Moving from "Multiply the number of dimes by 10 to get the number of pennies" to P = 10D (i.e. using the operative approach) appears to be less fraught with dangers than moving directly from "There are ten times as many pennies as dimes.", which could go either way. Pat ("computer scientist with strong mathematics background"), for example, starts off: "There are six times as many students as professors. So P is (*writes* P = 6S) equal to six… (*pauses*) … whoops. Let's see, there are six times as many students as professors. So six times P (*writes* 6P = S) is equal to S.". Perhaps he started off seeking a literal representation of the ratio relation, and then corrected himself by imagining multiplying a group of students by six. While being questioned as to his thought processes, he was suddenly gripped by the thought that he was right the first time, and then had to substitute numbers to reassure himself. Does the realization that ratio relations do not mix with algebraic relations avoid confusion? When he then attempted the pennies-dime problem with X for the number of pennies and W for the number of dimes, he said "there are 10 times as many pennies as there are dimes" and immediately wrote the reversed equation 10X = W. Philipp points out that, after a long struggle with trying to make sense of this reversed equation in terms of both numbers of coins and of values of piles, Pat managed (at different times) to describe X as representing the number of pennies (as stated in the question), the value of the pennies (which is the case), and the value of a single penny.

> "Pat demonstrated something common in all five people who wrote incorrect equations for the penny-dime problem. They all tried to justify the meaning of their equation by changing the meaning of the variables. The definition of the variable, though explicitly stated in the problem, was often changed to fit the equation written." (p. 172)

The interviewer constantly asked the subjects about the meaning of letters or expressions such as 10P. This in general seemed to confuse people. Pat got very annoyed at his confusion because "these are very simple", but it could be considered that it is not at all simple to have to handle relations and operations, *and* values and frequencies in the same problem. Pat thinks his confusion is just between value and number, but perhaps the confusion would not have arisen had he sought an *operation* rather than a *relation* (as Clement, 1982, conjectures). Finding an operation for the *value* does not make sense because the values of the piles are equal. So one seeks an operation for the numbers; the values are equal but dimes are worth more than pennies, so there must be more pennies; there must be 10 times as many pennies as dimes; 10 times the number of dimes gives the number of pennies; 10W = X. Seeking operations easily extends to more complicated equations such as $2x + 3y = x^2 + z$, because "Represent the relationship between x, y and z using an equation" is but another way of asking "What sequence of operations connects x, y and z?".

In each transcript given, it is noticeable that it is only when the subjects considered how they can obtain one number from another (i.e. looking for an operation) that they made progress. Considering meanings actually got in the way. Although Philipp acknowledges the divide between algebra and ratio equations - described by Davis (1986a) as the difference between "numerical variables equation frame and the labelling frame" - and emphasises the number of meaningful referents for variables as a possible reason for the increased difficulty of the penny-dime problem, his conclusions seems to miss the crucial point that mixing up meanings may be a *symptom* of having different problem expectations to the interview and not a *cause* of errors.

2.4.5 Future Research on Variables

Philipp's questions for future research are:

> "… how do students actively change the meaning of variables within problems? Second, how is problem difficulty affected by explicit versus implicit information in the problems? Finally, what effect do the variable representations have on students' performance? Answers to some of these questions will help us to understand the many problems students have with algebra word problems in particular and algebra in general." (p. 175)

Against this agenda, I would suggest that students' changing of meaning is an *inherently* contextual phenomenon, dependent on their particular expectation of the problem. However, research could profitably ask students how they might *use* the equations they produce. This would both demonstrate students' expectations of equation usage and test the view that the reversed equation was only "wrong" from the point of view of the conventions of mathematics professors rather than of mathematics students. Nevertheless, it could still be argued that a particular student was using the original situation rather than the equation to solve the problem he or she posed. This scenario would perhaps indicate that the student's algebraic concerns are better tackled through non-algebraic means, and would raise the question "How can we better make *algebraic* problems concerns?", rather than some demand for reflection on meaning.

Philipp asks:

> "(1) Why do students encounter difficulty with the student-professor problem?"
> "(2) Will the knowledge gained in pursuing the answer to question (1) help mathematics educators develop a better understanding of the difficulties students have with the concept of variable?" (p. 165)

The discussion above indicates that the answer to question (2) is *No*, because their difficulties in the student-professor problem give us insights not into their "concept of variable", but into their concerns and the strategies they use. As for question (1), students find the student-professor problem difficult because it is not clear to them what the equation is *for*. This is partly corroborated by Wollman (1983) in that those students who used static comparison corrected their reversals once they were told that the equation should represent the mathematical operations for obtaining the number of students from the number of professors. Moreover, they "still felt that their original 'equation' was an equally valid rendition of the meaning of the sentence" (p. 176). Likewise, Pawley & Cooper (1997) found that giving the students worked examples (presumably of operational calculations) similarly reduced the reversals; and also found that it did not make much difference if students were asked in addition to check that the larger quantity in the sentence is also the larger quantity in the equation, or if they were asked in addition to check by using trial numbers. In other words, this corroborates the claim that seeking operations rather than relations may help to clarify the purpose of an equation.

2.5 Theories used as Strategies

Clement's strategies for the student-professor problem can be seen as theories about what mechanisms can generate a legal and valid equation from a description of a situation. Küchemann's letter interpretations are theories about what symbols represent in algebra; but they can also be seen as elementary strategies for dealing with letters in certain types of problems.

For a given problem, there are myriads of theories that may be used as distinct strategies or approaches for solving the problem; and so no particular classification can be considered definitive as either complete or unique. Moreover, it is inevitable that it is easier to consider publicly-known strategies which have proved valuable over time than the implicit, individual, contextual strategies that students actually use in a particular situation; because the latter are often implicit, highly individualistic and contextual. A thorough analysis can therefore only take place when a small number of problems has been selected. Nevertheless, these "public strategies", highlighted by research, can serve as a starting-point for analysing students' knowledge. So without attempting to be encyclopaedic, this section considers six further important sets of strategic theories: for simplifying expressions, solving equations, representing situations using simple linear equations, representing situations using equations with two variables, solving word problems and proving using symbols. These theories, while focusing on linear equations, relate to a large part of what Kieran (1992) describes the "core" of school algebra:

> "Many first-year algebra courses begin with literal terms and their relation to numerical referents within the context of, first, algebraic expressions and, then, equations. After a brief period involving numerical substitution in both expressions and equations, the course generally continues with the properties of the different number systems, the simplification of expressions, and the solving of equations by formal methods. The manipulation and factoring of polynomial and rational expressions

> of varying degrees of complexity soon become a regular feature. Interspersed among the various chapters are word problems, thinly disguised as 'real world' applications of whatever algebraic technique has just been learned. Students eventually encounter functions and their algebraic, tabular, and graphical representations." (p. 395)

The aim here is not to summarise all the research that has ever been done relating to these problems; but to identify strategies that may help us to understand students' algebraic knowledge, and if possible to find questions that help to identify these strategies.

2.5.1 Simplifying Expressions

CSMS considered the problem of simplifying expressions, and found strategies of evaluating letters from the start, ignoring letters and treating letters as objects. Only a brief outline of additions to this research is given here.

A number of studies (including Wenger, R H. 1987; Becker, 1988; and Lewis, 1981) have looked at students' simplifying of expressions and have found many errors (Kieran, 1992, also cites Davis, Jockusch & McKnight, 1978; Greeno, 1982; and Carry, Lewis & Bernard, 1980). Some of the more common strategies include:

- Assume implicit addition. For example, $39x$ means $39 + x$ rather than $39 \times x$, so $39x - 4$ becomes $35x$; similarly, $2yz - 2y = z$; and if $x = -3$ and $y = -5$ then $xy = -8$
- Treat letters as digits. For example, if $x = 3$ and $y = 2$ then $xy = 32$
- Read expressions from left to right. For example, $a + a + a \times 2$ becomes $3a \times 2$; $a + a \times 2 + a$ becomes $5a$
- Give higher priority to simplifying letters than numbers. For example, $a + a + a \times 2$ becomes $3a \times 2$; $a + a \times 2 + a$ becomes $6a$
- Treat brackets as optional. For example, $4(6x - 3y) + 5x$ becomes $4(6x - 3y + 5x)$, $4 + 5x$, $24x - 3y + 5x$ or $24x - 12y + 12x$
- Associate a term with the sign that appear *after* it. For example, $2x + 9 - x$ becomes $3x - 9$
- Ignore signs. For example, $2x + 9 - x$ becomes $3x + 9$
- Assume that the coefficient of x is 0. For example, $3x \div 3 = 0$
- Collect like terms.
- Multiply out brackets before collecting terms.
- Multiply out brackets and collect terms in one pass.

These (and many other) strategies can coincide in the same problem with varying consistency of application, making a definitive identification difficult.

Robitaille (1989) states that "On the whole, the range of achievement scores [in algebra] is very similar to that on the Arithmetic subtest." (p. 111) which raises the issue of the extent to which algebraic proficiency is dependent on arithmetic strategies (Lee & Wheeler, 1989). Pereira-

Mendoza (1987) describes a complex set of relationships between algebraic and arithmetic "space" that can be "faulty" or misperceived. On the other hand, Kirshner (1987a, 1989) sees the problem as one of using the visual cues of ordinary notation (as opposed to rules about the order of operations) to make syntactic decisions. Kieran (1988a) argues that it is a failure to recognise and use "surface structure" (the arrangement or disposition of the terms and operations) that lies at the heart of students' errors. Sleeman (1984) distinguishes "manipulative mal-rules" (where steps in sound algorithms are omitted or varied) and parsing errors (where aspects of notation are misinterpreted and algorithms are therefore misapplied). Matz (1982) suggests that "errors are the result of reasonable, although unsuccessful, attempts to adapt previously acquired knowledge to a new situation." (p. 25), and that errors must be relied upon as the central evidence for a model of algebraic competence" (p. 26).

Moreover, these inappropriate strategic theories for simplifying equations may indicate that while students may realize that "something simpler" is required, they perhaps do not understand that such simplifications must be true for *all* values of the unknowns, and that therefore trial substitutions can reveal incorrect simplifications (Davis *et al.*, 1978; and Lee & Wheeler, 1989). One has to say "perhaps" because it is also possible that many students are unwilling to check using substitution because it would reveal an error that they are hoping does not exist, and for which they would not want to spend any more time worrying about; or it is also possible that they do not know how to substitute; or it is also possible that they are not sufficiently confident of their ability to substitute that substitution would act as an adequate check; or it is also possible that they have substituted incorrectly, because their simplification heuristics are based on flawed substitution rules. This range of possibilities suggests that using errors in simplification to draw conclusions about inappropriate strategic theories is problematic.

As for *successful* strategic theories, the fact that simplifications are usually one-step acts, without intervening working or accompanying explanations, means that it is very difficult to conjecture *how* students might be arriving at their simplifications. Presumably, though, they are often making implicit use of fairly basic identities (Davis, 1986b). One might conjecture a set of heuristics something along the lines of the following: "Identify terms. Those that are numbers can be added or subtracted. Those that are x's can be added or subtracted. Those that are y's can be added or subtracted… Those that are x^2's can be added or subtracted… If there are brackets it may be possible to multiply out; if not, perhaps factorisation may be a good idea…". Even so, this still requires further strategies for combining x's, expanding brackets and so on. Alternatively, one might conjecture something like: "3 times x plus 5 times x. Now what is that the same as? Well if we have 3 of something and 5 of something, then another way of saying this is that we have 8 of something…". Another approach might be: "Multiplication is distributive over addition. So $3x + 5x \equiv x(3 + 5)$. The bracket simplifies to 8, and now the brackets can be removed, making $8x$.".

On another level, Sproule (1988) provides evidence that both the speed and accessibility of such strategies are student concerns. "Some students were inclined to persist with more secure, if more time consuming strategies. Alternatively, some students, upon reflection, recognised the time consuming or inefficient nature of their strategy and proceeded to search for a more efficient approach." (p. 309). Lewis (1981) suggests that mathematicians routinely "set their own trade-offs among speed, effort and accuracy" (p. 107).

In conclusion: there is a very large number of possible strategic theories for simplifying expressions, and it is difficult to find questions that help us to distinguish these strategies. Hence one must be cautious when interpreting student responses to problems of this type. It may sometimes be only possible to gather whether a student has a successful strategy or not.

2.5.2 Solving Equations

Radford (1995) identifies an aesthetic component to mediaeval mathematics; and the "resolution of problems and difficult riddles… constituted an *ad hoc* instrument of social recognition for the master." (p. 31). The exploration of problems that bore only a cursory relation to *practical* problems became highly refined. This refinement brought about the idea of an equation: a description of the problem situation that was sufficient to contain the information required to solve the problem, that used standardised notation, and that could therefore be tackled using standard techniques. Bednarz *et al.* (1996) state that current teaching often reduces algebra to transformation rules, the historical role of problem-solving degraded to tacked on "applications". However Chaiklin (1989) points out that solving equations in isolation may not be the same as solving equations that the student has generated from a situation (p. 102).

What strategies do students use to solve equations? CSMS and SESM steered clear of equation-solving on the whole, but many of the transformational errors recur in equation-solving; for example, $\frac{3x}{2x}$ simplified to $2x$. Bell, Costello & Küchemann (1983) describe a number of research studies comparing equation-solving methods. However, they conclude "It is difficult to point to any generally useful conclusions from such studies: they are neither extensive nor convincing." (p. 142).

However, these studies do, at least, tend to identify commonly occurring errors. Matz (1982), Kieran (1989b, 1992), Payne & Squibb (1990), Moncur (1994) and Becker (1988) provide further examples of such errors. Bell, Costello & Küchemann assert "perhaps the most obvious conclusion is that many errors are caused by children who have developed skill in manipulating meaningless symbols being disinclined to think in terms of meaning or to consider that the symbols represent numbers." (p. 144)

Of course equations can be often solved numerically. Kieran (1992) reports on studies that have shown students using recall of number facts (for example, solving $5 + n = 8$ by remembering the fact that 5 plus 3 is 8), counting techniques (for example, solving $5 + n = 8$ by counting on "6… 7… 8…" and noting that 3 numbers were involved) or a "cover-up" method (for example, solving $2x + 20 = 6x$ by reasoning "Since $2x + 20$ totals $6x$, the 20 must be worth $4x$, so x is 5."). Dickson (1989) notes in relation to numerical skills that "Some pupils, particularly lower attainers, depended on non-generalisable strategies to solve problems." (p. 188). This may limit the range of cases that can be considered.

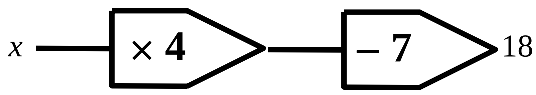
Forms of trial-and-improvement using substitution can be time-consuming and prone to error unless a means is found to record guesses systematically. A recent variation on trial-and-improvement is to use a spreadsheet or calculator to produce tables of values for both sides of an equation, and vary the value of the unknown until equality is reached.

Graphs, too, can be used effectively, but they are outside the remit of this research.

When there are multiple unknowns or non-numerical coefficients all the above strategies are trickier to implement (see for example Carraher & Schliemann, 1987); and in many cases solutions may be missed, or it may be impossible to demonstrate beyond reasonable doubt that all the solutions have been found.

There is also the method of "undoing" or "inversing" which can solve many linear equations (for example $4x - 7 = 18$). One version is to interpret the equation as a "think of a number" problem (see for example Dickson, 1989): I think of a number, multiply it by 4 and then subtract 7. If the answer is 18, what was my number? x can be found by the following (analytic) reasoning: the number from which was subtracted 7 to make 18 is $18 + 7 = 25$. The number which was multiplied by 4 to make 25 is $25 \div 4 = 6\frac{1}{4}$, and so we have x. Alternatively, this reasoning can be automated by writing down the expression $4x - 7$ and working out the operations that reduce it to x: i.e. $+ 7$ then $\div 4$. These operations are then applied to 18 to give the answer.

Another version of the method is demonstrated by SMP (1981) and Austin & Vollrath (1989), and involves showing the equation as a "function machine" or "flow diagram":



The idea is to run the flowchart backwards with *inverse operations*: -7 becomes $+7$; and $\times 4$ becomes $\div 4$. So $x = (18 + 7) \div 4$. This strategy is synthetic because it does not require one to manipulate an unknown. Both the analytic style of reasoning and the synthetic use of inverses fail to solve equations with unknowns appearing on both sides - those equations that Filloy & Rojano (1989) call "non-arithmetical" and therefore beyond the "didactic cut" between arithmetic and algebra. But even multiple unknowns on just one side can cause problems: Kieran (1988b) found that some students solved $3x + 4 - 2x = 8$, for example, by taking 8, dividing by 2, adding 4 and then subtracting 3. In other words, using inverse operations blindly. Bernard & Cohen (1989) point out that a simple equation like $14 - x = 8$ is not directly solvable by undoing. Herscovics & Linchevski (1991), for such reasons, prefer to characterise the cut as "the student's inability to operate with or on the unknown" (p. 180) rather than in terms of mathematical form.

Sometimes students are found using strategies that were intended by teachers as an accessible introduction to the *idea* of solving equations rather than as a general method. For example, the problem $5x + 20 = 7x$ can be modelled by a balance:

where the white blocks represent unknown weights and the black blocks are (say) kilograms. Then, pairs of blocks of unknown weight can be removed (one from each side), which maintains equilibrium, until none remain in the left hand pan. We then know that 2 of these weights weigh 20kg, and so we have the unknown. Unfortunately this can be time-consuming to draw, and there is no easy way for equations involving negative signs or negative solutions to be rationalised in this model, although attempts have been made.

However, the usual focus of attempts to teach algebra are the *formal* methods that allow not only the solution of a wide range of simple equations, but also easily extend to "changing the subject of the formula" or "re-arranging the equation". They involve treating unknowns as if they were known, refer to no particular problem situation and make use of the rules governing arithmetic.

Radford shows that mediaeval mathematicians simplified unknown terms on the same side of the equation (for example, $12t + 9t$ becomes $21t$). Thus the simplification of expressions, as described previously was devised to make the solution of equations easier, by condensing the number of cases to be considered. They also "restored" equations which were "broken" by the presence of a subtracted term. For example $3 - 2x = 10x$ could be restored to $3 = 12x$. There are two standard formulations of a general rule for this. One is "operating on both sides of the equation with additive and multiplicative inverses", which Filloy & Rojano (1989, p. 20) attribute to Euler. Schliemann *et al.* (1992) provide a quotation from Leibniz that seems to capture very well the spirit of this method of formal operations:

> "if from two equal things the same quantity be taken away, the things will remain equal; likewise… if in a balance everything is equal on the one side and on the other, neither will incline, a thing which we foresee without ever having experienced it" (p. 298)

Radford also suggests that, as a late historical development, mediaeval mathematicians *transposed* terms from one side to the other. This would allow all the knowns to end up on one side and the unknowns on the other. Filloy & Rojano point out that in "pre-symbolic algebra textbooks" of the 13th to 15th centuries, most solution strategies for $x^2 + c = 2bx$ and $x^2 = 2bx + c$ are completely different; and "This difference would not exist if the authors had had recourse to the rule of transposing terms from one side of an equation to the other" (p. 19). However, an arithmetic without negative numbers would see such a difference as natural. Filloy & Rojano attribute the transposition strategy to Viète.

Kieran (1981) makes the point that: "not only does equation solving involve a grasp of the notion that right and left sides of the equation are equivalent expressions, but also that each equation can be replaced by an equivalent equation (i.e. one having the same solution set). Unfortunately, very little research has addressed itself to the question of how this concept is acquired by high schoolers." (p. 323). See also Kieran (1984) and Steinberg, Sleeman & Ktorza (1991). Within both the formal methods, strategies such as "get the unknown by itself on one side of the equation" and "collect like terms" would make sense.

The transposition strategy of moving terms from one side of the equation to the other is mathematically equivalent for most purposes to the Leibniz strategy of operating on both sides of an equation. But a question inspired by Lins (1992) occurs: are there important psychological consequences that flow from the choice of formal strategy? Take for example the transformation of the equation $120 - 2x = 315$ into the equivalent equation $120 = 315 + 2x$. Is there a major difference between "adding $2x$ to both sides" (i.e. *thinking up an operation* to use on an equation) and "move the $-2x$ over to the other side and it becomes $+2x$" (i.e. *deciding on an object* in the equation to move)? There seems little detailed research on this. Davis (1986b, 1989a) associates the transposition strategy with an emphasis on notational rules divorced from understanding. But of course this need not be the case: if A is 6 more than B ($A = B + 6$) then B is clearly 6 less than A (so $A - 6 = B$). This justification for transposition seems as "reasonable" as: if A is the same as B plus 6 ($A = B + 6$) then 6 can be subtracted off both A and $B + 6$ (so $A - 6 = B$). It seems just as likely that the justification for "doing the same thing to both sides" is forgotten (algebraic operations on equal things do not change the equality) as that the justification for "change the side; change the sign" is forgotten (adding onto one side is the same as subtracting off the other; multiplying one side is the same as dividing the other). However, English & Halford (1995) consider that transposing "does not emphasise the symmetry of an equation." (p. 227).

In any case, many studies report that the problem of finding a solution to an equation by formal methods is a major source of student difficulties:

> "... for a large number of high-school students, there are many cognitive obstacles involved in perceiving an equation as a mathematical object on which they can perform operations. It is equally difficult for them to grasp the idea of equivalent equations and construct a meaning for the formal solution procedures" (Linchevski & Herscovics, 1996)

These two "formal" methods - the Leibniz and transposition strategies - can also be contrasted with the "informal" methods outlined earlier, such as using balances, flowcharts and numerical trial-and-improvement. The 1997 Royal Society / JMC report points out that "trial and improvement seems to be becoming the preferred and probably only method which the majority of pupils are confident with in pre-16 education." (p. 7). Some students have consequently been noted using trial-and-improvement on linear equations, and failing to realise that there are two solutions to quadratics. The report also suggests that "When pupils have become proficient with trial and improvement methods for solving equations they are unlikely to want to learn algebraic [formal] methods.". This is not to say that trial-and-improvement should not be taught - the report describes it as a "valuable technique" that allows a wider class of equations to be tackled and, moreover, Davis (1986b) notes that it has troubleshooting potential when learning formal methods - but that it may have edged out the formal methods. SCAA (1996), for example, noted an increased emphasis on such iteration in O-level and GCSE since 1975.

Kieran (1992) reports studies showing that although substitution is generally dropped when formal methods are eventually learned, students "also seem to drop it as a device for verifying the correctness of their solution" (p. 400). Kieran (1988a) points out that, just as for simplifying expressions, students often "do not know how to show that an incorrect solution is wrong, except to re-solve the given equation. They do not seem to be aware that an incorrect solution, when substituted into an incorrectly transformed equation will yield different values for the left and right sides of the equation. Nor do they realize that is it only the correct solution which will yield equivalent values for the two expressions in any equation of the equation-solving chain." (p. 437). That is: they do not know that an incorrect solution can be tested, and they do not know that a "value which works" must be a solution. Kieran (1992) cites Lewis (1980) on this.

Interestingly, Mayer (1985) found that students given equations to solve such as $(8 + 3x) \div 2 = 3x - 11$ tended to use a formal strategy to isolate x; whereas students given syncopated equations such as "Find a number such that if 8 more than 3 times the number is divided by 2, the result is the same as 11 less than 3 times the number." attempted to "reduce" the problem to something simpler systematically - for example a first step might be to consider the problem "Find a number such that if 8 more than it is divided by 2, the result is the same as 11 less than it.".

The problem for research of identifying the individual steps in strategies is often profound. Studies undertaking this can be put into three broad categories according to research method.

The first method is to use a "clinical interview" (as in Herscovics & Linchevski, 1991). Clarifications, justifications, extensions and so on can be requested. Models are then formulated to explain each student's responses, and then generalisations attempted across students. Herscovics & Linchevski were able to identify strategies of inverse operations, grouping terms, guessing an operation, checking, trial-and-error, and so on, for a range of types of simple linear equations.

The second method involves examining student responses to written items, and formulating models to explain the responses. Sometimes individuals' responses are compared qualitatively; sometimes facilities for different items are compared quantitatively.

However, with both these methods, one soon runs into the same difficulty as for simplifying expressions - the number of potential explanations for errors on an item depends exponentially on the number of postulated "attributes" (procedures, skills, facts, etc.) required to answer the item; and the data at our disposal is limited. The mind seems to work in such a way that it often does not seem to be able to examine *how* a step was taken; let alone articulate the method. So even asking students how they did it does not usually help. Sometimes a *post hoc* rationalisation is the best that can be obtained (noted by Drouhard & Sackur, 1997; see also Dennett, 1993 and Popper & Eccles, 1977).

The third method is to describe a range of problems in terms of "attributes" (such as associativity, cancellation of elements, collection of like terms) and use *all* the problems to model students' skill. Birenbaum *et al.* (1993), for example, used a sophisticated statistical refinement ("rule space"), to handle the complexity. For example, the study cited found that, for 231 Israeli 14-15 year-olds taking a 32-item diagnostic test in solving linear equations, a set of attributes that took into account that students might try to minimise operating with negative numbers if legitimately possible was better able to explain the variance of responses than a set of attributes that ignored this strategy.

The big disadvantage of this "attributes approach" is that it depends very much on having an accurate set of the required attributes for each item. However, it has the advantage that errors that are either "wild" or result from "slips" do not have as great a potential to lead researchers astray. It also provides a means of testing models of students' theories that is more open to scrutiny and criticism than the traditional "item-by-item" approaches of interviews and written tests: one can compare how well different models appear to account for the range of responses that a student or a group of students makes. On the other hand, even when we have identified the attributes that we are interested in, an item-by-item approach is perhaps less likely to fall into the trap of assuming that the same set of attributes are appropriate for all problems. An attributes approach would be useful when we have a large range of very similar problems which differ in subtle but identifiable and known ways, and students can solve some but not others. An item-by-item approach would be useful when we have a small number of problems for which we want to examine the range of possible strategies.

The choice of research methods for this research depends largely on the level of detail that is required to fill the gap in research that I hope to have identified by the end of chapter 3, given the constraints of research resources. However, an item-by-item approach is likely to be preferable in a study with a large number of *sufficient* (as opposed to *necessary*) strategic theories.

Related to the attributes approach is the AI diagnostic approach. Sleeman (1986), for example, catalogued a very wide variety of rewrite "mal-rules" in the context of solving equations, using a diagnostic computer program. But he found that the program was unable to determine the mal-rules deployed merely on the basis of the answer. The sheer number of possible rules at each step of a solution process made precise determination difficult. Moreover, several inferred mal-rules were wrongly identified as the result of the student making wild guesses. On the other hand, when a human took over the role of identifying the problematic mal-rules (interviews were find to be the most effective way), two tests of a group of students four months apart appeared to show students' consistency. Sleeman concludes from a variety of studies (including Paige & Simon, 1966; Lewis, 1980; Davis, Jockusch & McKnight, 1978; Matz, 1982 & Sleeman, 1984) that "Pupils have a great facility for inferring their own rules, or sometimes higher level schema, and then using them consistently and often in *inappropriate* situations." (p. 52).

On the other hand, Payne & Squibb (1990) found that there are many *infrequent* mal-rules, and few frequent ones; and that it is rare for students to use a mal-rule in as many of half of its applicable conditions (p. 478). They also discovered that identifying mal-rule is easier with students who make fewer errors.

Sleeman used his results to question the assumption of Brown & Burton (1978) that all errors are the result of perturbations to correct rules (p. 45). On the other hand, Payne & Squibb demonstrated firstly that the 99 separate mal-rules that they identified could be grouped into more abstract principles; and secondly that the ten most frequent mal-rules in schools diagnosed 67% of the errors in that school, but fewer than 10% in another. Payne & Squibb concluded that "specifics of educational experience seem to have a heavy influence on acquired mal-rules." (p. 455).

Given a set of "Domain Rules" (such as $3x = 6 \Rightarrow x = 6/3$, $2 + 3x = 6 \Rightarrow 3x = 6 - 2$ and $2x + 3x = 15 \Rightarrow 5x = 15$) and a set of mal-rules (such as $7x = 4 \Rightarrow x = 7/4$, $2 + 3x = 6 \Rightarrow 5x = 6$ and $4x = 2x + 3 \Rightarrow 4x + 2x = 3$), AI techniques can also allow software to model student reasoning. However, a big difference between an AI approach and an educational approach is that for educational purposes, the details by which students develop and implement strategic theories are of interest only insofar as there appear to be difficulties in addressing target concerns.

AI models of cognition can be criticised for artificiality (Chaiklin, 1989), but is characterising students' thinking as something like a number of modular, interacting programs any more artificial than the abstract metaphors of "awareness", "concept", "relation", "experience", "process", "idea", "expression", "skill", "meaning", "error", "behaviour" and so on that permeate current discussions of learning? Popperian psychology shares with the "competing rules" model of Payne & Squibb a critique of the simplistic account of errors that student performance can be partitioned into "Executing the right rule correctly", "Executing the right rule incorrectly" and "Executing the wrong rule". Indeed, BVSR suggests that strategic theories are "in competition" in a given problem situation. However, the student's perception of the problem itself must also be taken into account, and AI studies to date do not appear to have modelled students' problem representation in equation solving contexts.

2.5.3 Representing Situations Using Simple Linear Equations

Wollman (1983) notes that "The ability to translate sentences into algebraic relationships figures heavily among the problem-solving skills required in quantitative science courses and mathematics courses in secondary school and college." (p. 169). Yet…

> "Generating equations to represent the relationships found in typical word problems is well known to be one of the major areas of difficulty for high school algebra students." (Kieran, 1992, p. 403)

Similarly, Lochhead & Mestre (1988) write:

> "It is well known that word problems have traditionally been the nemesis of most mathematics students. The translation process from words to algebra is perhaps the most difficult step in solving word problems." (p. 134)

According to Chaiklin (1989), there are two main approaches to representing relations given in sentences as equations: "direct-translation" and "principle-driven". Phrase-by-phrase translation is rather limited in the range of pertinent problems. Paige & Simon (1966) also found that students using this approach failed to detect the contradictions in certain problems, whereas those using the principle-driven approach could. This latter approach involves using cues to recognise and organise the relations according to various schemata or templates, such as whole-part, rate, age and multiplicative (Mayer, 1981). These templates can be vital in students developing future problem-situation models (English & Halford, 1995, p. 240-245).

Mayer (1982a) distinguished the propositions in word problems as: assignments, relations and questions. He found (Mayer, 1982b) that, for college freshmen, relations were harder to remember than assignments; and when they were asked to construct their own word problems, the students rarely made use of relations. Yet Chaiklin observes that "to solve algebra word problems successfully, students must be able to interpret and understand the mathematical relations in these problems." (p. 93).

Berger & Wilde (1987) distinguish three tasks in representing a situation:

- Value assignment: setting a noun phrase equal to a numerical value or to a symbol.
- Value derivation: operating on assigned values to produce new values, using relationships and formulae that are either given in the question or expected to be known.
- Equation construction: creating a computational representation of the structural relationship between variables.

Although it may seem as though this order is the order in practice, "if one can determine the form of the final equation, the range of possibly appropriate value assignments and derivations may be reduced." (p. 127). This is borne out by a study of high school students, which also suggested that value assignment was found to be easier than value derivation, which was easier than equation construction. In addition, making use of formulae that are expected to be known appeared to be harder than using a given relationship.

Cortés (1988) notes that situations to be represented generally make use of objects, lengths or prices "whose additive and multiplicative properties students, in general, know." (p. 209). She suggests that the concepts of equivalence and "numerical function" (which can be interpreted here as "expression") and the principle of "homogeneity of terms" are implicitly required cognitive demands in such representation. She identifies three strategies that partition the problems found in 8th-grade French textbooks:

> "I - substituting unknowns with given numbers and units into a given formula or into a constructed function.
>
> II - substituting unknowns with linear functions into a two (or more) unknown equation.
>
> III - equating two linear functions." (p. 210).

2.5.4 Representing Situations Using Equations with Two Variables

"It is tempting to believe," writes Herscovics (1989), "that after extensive work with equations in one unknown, students will have an easier time constructing expressions and equations in two variables. But, the presence of more than one variable appears to compound their difficulties." (p. 63).

The earlier discussion about the student-professor problem has clearly introduced some of the relevant strategies for this class of problems, but there are further strategies to be considered. Reed, Dempster & Ettinger (1985), for example, found that "Students had considerable difficulties in specifying the relations among variables." (p. 123).

As has already been pointed out, in the UK such representation is more likely to be from number patterns found by the student in objects, pictures or tables than from sentences. The term "situation" here covers all four of these contexts, but note that this section does not consider the generalising of arithmetic rules such as $(x + y)^2 \equiv x^2 + 2xy + y^2$.

MacGregor & Stacey (1993a) gave problems like the following to 143 Australian students aged about 14, in a pencil-and-paper test:

(A) Look at the numbers in this table and answer the questions:

x	y
1	5
2	6
3	7
4	8
5	9
6	..
7	11
8	..
..	..

(i) When x is 2, what is y?
(ii) When x is 8, what is y?
(iii) When x is 800, what is y?
(iv) Describe in words how you would find y if you were told what x is.
(v) Use algebra to write a rule connecting x and y.

(B) The results of an experiment that measured two quantities L and Q were:

L	Q
3	9
5	15
9	27
21	63

(i) What would you expect Q to be when L is 30?
(ii) What would L be when Q is 99?
(iii) Describe in words how you would find Q if you were told what L is.
(iv) Use algebra to write a rule connecting L and Q

(C) Look at the numbers in this table and answer the questions:

x	0	1	2	3	4	5	6
y	2	5	8	11	14	17	..

(i) When x is 6, what is y?
(ii) What x is 10, what is y?
(iii) When x is 100, what is y?
(iv) Describe in words how you would find y if you were told what x is.
(v) Use algebra to write a rule connecting x and y.

They found that "more students could find and use a relationship for calculating than could describe it verbally or algebraically." (p. 181). This seems to confirm Herscovics (1989). MacGregor & Stacey suggest that this is because the students were primarily "focusing on recurrence patterns in one variable rather than on relationships linking two variables" - the facility of C (iii), for example, was much lower than that of C (ii). It is also pointed out, nevertheless, that many students who *could* use a functional relationship for calculating were unable to describe the rule in either words or symbols. So whereas the facility for A (iii) was 61%, that for A (iv) was 49% and that for A (v) was 35%. Moreover, in interviews with 15 students, the researchers were struck by the "variety of patterns perceived and the large proportion of generalisations expressed verbally that cannot be expressed in the elementary algebra that students are learning." (p. 184). It is interesting that although half of the students could describe a correct rule verbally for problems A and B, only a third wrote down a correct algebraic formula. For problem C, 17% could describe a rule verbally, and 12% wrote down a correct formula. Students found it more difficult to find a rule for calculating in B, perhaps because of the absence of constant differences.

From the report of this research, a number of strategies for solving pattern-seeking problems can be identified:

- find the numerical value of one variable from another by using proportion (so for A (iii), to find y when $x = 800$, multiply 800 by 5 because this works for the first row of the table; or multiply by 100 the value of y when $x = 8$, because this direct proportionality has worked in previous problems);
- find the next number in a sequence by using a recurrence rule;
- find a recurrence rule by looking at differences, ratios, the number of integers in the gap, and arithmetic operations (this is not termed an *algebraic* problem);
- express a recurrence rule algebraically;
- find a functional rule by looking at differences, ratios and so on (again, not an algebraic problem);
- express a functional rule algebraically;
- find a functional rule by finding the "inverse" of a known rule between the variables.

2.5.5 Solving Word Problems

Radford (1995) describes how mediaeval mathematicians used algebra to solve problems arising primarily from commerce in which an unknown quantity was to be found. The main "goal of the works or chapters dedicated to algebra is not to explain the geometrical algorithms nor to learn how to carry out calculations on binomials but to show how to use the *techniques* of algebra to solve *word problems*." (p. 30). This suggests that "algebra was intended to be, above all, a problem-solving tool" and one which is more "fertile" than those based on numerical or geometrical approaches. For example:

Denarii Problem

Two men have a certain amount of money. The first says to the second: if you give me 5 denarii, I will have 7 times what you have left. The second says to the first: if you give me 7 denarii, I will have 5 times what you have left. How much money do they have?

One of the teachers taking part in this research suggested that it is more sophisticated to see algebra as a tool than as someone else's game. Word problems may provide a test of this sophistication, because unlike the student-professor problem, the use of algebra is not a prerequisite. Hence a "symbolic algebraic approach" (defining letters to represent unknown quantities, creating equations to represent the relations between these quantities, solving the equations and then interpreting the answers in the context of the problem) can be contrasted with (say) a numerical "trial-and-error" approach or a contextual "whole-part" approach.

However, Reed (1984) showed that averages or intuition are often preferred strategies to algebra, perhaps even in circumstances in which the student suspects that the intuition may not be sound. Lee (1987) describes research indicating that "Students who can competently handle the forms and procedures of algebra rarely turn spontaneously to algebra to solve a problem even when other methods are more lengthy and less sure." (p. 317).

The Driving Problem

Mr. Sweetmann and his family have to drive 261 miles to get from London to Leeds. At a certain point they decided to stop for lunch. After lunch they still had to drive four times as much as they had already driven. How much did they drive before lunch? And after lunch?

Lins (1992) used word problems like this to test out the extent to which his characterisation of "algebraic thinking" could be useful to "distinguish and understand, *on the fly*, the thinking and learning processes which are developing on the part of the learner." (p. 21). The essential characteristics of algebraic thinking, he argues, are that it is "arithmetic", "internal" and "analytical". Taking the driving problem as an example; thinking "The distance plus 4 multiplied by that distance makes 261" rather than "That bit of road combined with 4 similar bits of road has a length of 261" is illustrative of *arithmetic* rather than *whole-part* thinking. Suppose that a student decided to use the equation $x + 4x = 261$. If, in solving this equation, no further reference to the situation is made, this is a hint of *internal* thinking - the "meaningfulness of each expression produced is related only to the perceived correctness of the process that produced it" (p. 202) rather than to the relationships described in the situation. So "In most of the solutions using equations we could reasonably establish that the reference to the problems' context was abandoned, in particular through the generation of expressions where the minus sign could not be given an immediate *non-algebraic* interpretation". *Analyticity*, meanwhile is manipulating the unknown ("$x + 4x$ is $5x$… Divide $5x$ by 5 and 261 by 5…").

Lins' extensive analysis of word problems results in a description of an algebraic approach:

1. forming an equation by analytically converting a series of arithmetical calculations "*analogically* associated with the problem's 'story' or context" into expressions;
2. linking these expression using equalities, again "analogically associated with the context";
3. treating each equation internally (all p. 227).

Now clearly, from a Popperian psychological perspective, this formulation is not acceptable because it makes use of "modes of thinking". However, it is perfectly possible to distinguish, as Lins does very effectively, different *strategies* for solving problems. Non-algebraic approaches include a whole-part approach ("1 lot of x plus 4 lots of x is equal to 261") and a trial-and-error approach ("10 is too small, 100 is too large, …"). Some scripts also showed that algebraic and non-algebraic approaches can be used at "different stages of the same solution process, highlighting the possibility of usefully combining algebraic and non-algebraic models, and at the same time emphasising the dissimilarities between them." (p. 303). For example, a student could find an algebraic representation of the situation by representing a whole-part relationship, but then solve it by non-algebraic means such as trial-and-error. One can view such strategies as theories which aim to solve the problem; and one could even use Lins' characteristics of algebraic *thinking* as indicators of an *algebraic strategy* that is distinct from a *symbolic* algebraic strategy (although this is not done in this thesis).

Lins shows not only that the facility for word problems depends crucially on the numbers involved (2/3 of students got the driving problem correct, but only 16% got the problem right when the integer 4 is replaced by the decimal 2.7) but also that the choice of *strategy* is similarly dependent (because adding 1 part and 2.7 parts and then dividing 216 by 3.7 parts is not as familiar a strategy as adding 1 part and 4 parts and then dividing 216 by 5 parts).

His method was to give two question papers to each of 10 classes of Year 9 and Year 10 students from England and Brazil (228 students in total). There were three varieties of these papers, so varying the numbers used and the "situational context" could be studied. Each paper contained 6 questions, was attempted in a 50 minute session, and did not mention the word "algebra".

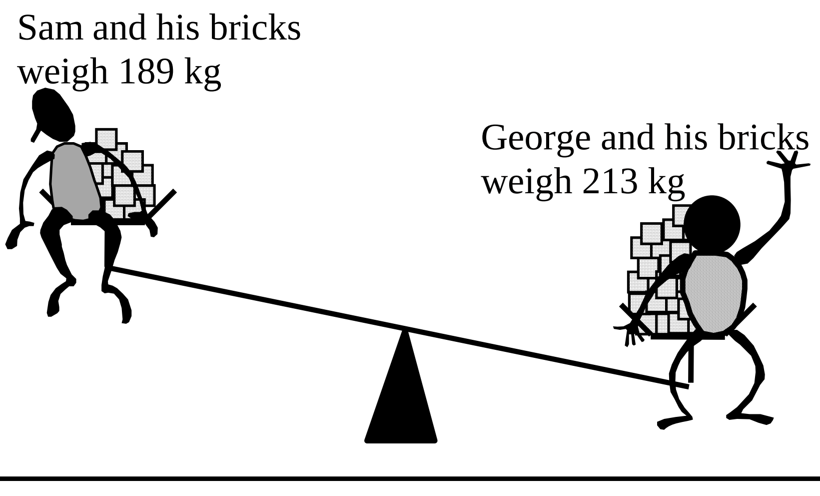
Lins' results show that facility also depends crucially on the *context* of the problem. For example, the tickets problem was consistently easier than the driving problem in whatever form:

Tickets Problem

Sam and George bought tickets to a concert. Because Sam wanted a better seat, his ticket cost four times as much as George's ticket. Altogether they spent 74 pounds on the tickets. What was the cost of each ticket?

Lins replicated these results with problems based on different "underlying equations":

The Seesaw Problem



Sam throws away 5 bricks and George throws away 11 bricks. Now they are balanced. What is the weight of one brick?

[NB This problem has been slightly adapted]

Sale Problem

Maggie and Sandra went to a records sale. Maggie took 67 pounds with her, and Sandra took 85 pounds with her (a lot of money!!). Sandra bought 11 LP's and Maggie bought 5 LP's. As a result, when they left the shop both of them had the same amount of money. What is the price of an LP?

Now, from the Popperian epistemological perspective, one would not want to talk about equations "underlying" a situation any more than one would want to claim that theories are "inherent" in a situation. One would say, rather, that these problems *can* be solved by means of a common equation; and the relationship between problem and equation is no stronger than that. Mayer (1981) termed such problems as belonging to the same "family". So the results of Reed, Dempster & Ettinger (1985) and Reed (1987) would be interpreted as potentially indicating a lack of transfer of understanding between families. Moreover, Lins seems to show (if I re-interpret his results) that the facility of a problem will depend not only on whether each student has an adequate strategy to solve the problem, but also on the student's understanding of the problem, and on the student's expectation of the ease of execution of each strategy that the student expects might solve the problem.

This result presents a seemingly insurmountable obstacle for those wanting to graduate word problems in order of difficulty. However, in practice it often turns out to be possible (see for example, Robertson, 1994; Hinsley, Hayes & Simon, 1977) to conjecture, for a particular group of students and a given problem, the understanding that students in that group might have of the problem, the strategies that they might have available, and their expectations about the ease and adequacy of those strategies in solving the problem. For without such conjectures, teachers would be unable to make judgments about questions that would be helpful for students to consider. Moreover, Hinsley, Hayes & Simon provide compelling evidence that students not only use a line-by-line procedure in attempting to comprehend word problems, but they also often try to classify the problem as soon as possible in order to make use of strategies associated with standard problems. Word problems that appear so similar to the student that none of these expectations are substantially different could be said to belong to the same "class". In this way, a single problem can be used as a research item to test for ability to solve a class of problems.

Returning to Lins' work to find some suitable problems, the main distinction he makes about the Tickets and Driving problems (that for the whole-part approach, the visualisation of the expensive ticket as four tickets is easier than the visualisation of the longer distance as four short distances) is not significant for this research because the change of facility by context is taken for granted. The 2.7 version of the Driving problem can be used as a potential indicator of an ability to use equations to solve problems, because it is the hardest one of the four relating to equations like $x + 4x = 74$.

The Seesaw problem is superficially a balance problem of the type often used to introduce equations, but in fact is not of the standard variety; because even though what is modelled is still a balance, it does not use the weights and bricks as they appear on the scale, but an act of balancing itself. It brings a superb opportunity for negative signs to be brought into use, and is not easily soluble by trial-and-error. Those with a superficial grasp of the potential of algebra may very well represent the problem by an equation which includes the weights of the people. On the other hand, it may also be possible to find a useful equation such as $189 - 11b = 213 - 5b$ (where b is the weight of a brick) while still interpreting the $11b$ as "11 bricks". This would not affect the result so long as the operations on the equation were carried out. There is also a similar problem in which we are told not about the number of bricks thrown off, but only that George throws away four times as much weight as Sam does and we are asked for how many kilograms George throws away. There seems (at first glance) to be not enough information here. We don't know the weights of either person, of a single brick, of all the bricks on either side, or of the bricks removed. Yet we can still magically solve the problem by using the equation $189 - x = 273 - 4x$. Comparing the two Seesaw problems, the two Sale problems and a Secret Number Problem ("I am thinking of a secret number. I will only tell you that 181 − (12 × secret no.) = 128 − (7 × secret no.) What is my secret number?"), the 4× Seesaw problem was the hardest (facility 22%) and the 11-5 Sale problem the easiest (facility 39%) - a difference which Lins says may be explained perhaps by the potential number of referents for letters. However he also points out that those with the most experience of solving equations - the Brazilian 8th grade students - obtained a facility for the secret number problem of 88%. Only 4% of the English Year 9 students could tackle it effectively. This may indicate that "the development of *algebraic thinking* is a process much more akin to *cultural processes* than to age-related stages of intellectual development." (p. 228-9). This is reinforced by the English students' immense comparative success with a pattern-seeking type problem in which a formula was given, with which they, but not the Brazilians, had had much experience. Many of the successful English students justified their solution by "reversing the formula" (which is, incidentally synthetic, not analytic) whereas most of the successful Brazilian solutions used equations.

Lins concludes that solving the "secret number" problem depends heavily on the use of equations - only 5 out of 146 students apparently managed to solve it by using a non-algebraic model. For the other problems, most of the incorrect solutions by the Brazilian 7th graders do not attempt to use an equation; whereas most of the incorrect solutions by the Brazilian 8th graders *do* represent a mistaken use of equations. "This suggests that for the Brazilian 7th graders the 'default' approach is non-algebraic, and for the 8th graders it is an algebraic one, namely the use of equations." (p. 199) The younger Brazilians used non-algebraic approaches when they could, and only switched to algebraic approaches when they couldn't solve the problem any other way. They did this even when they couldn't solve the equations produced. The problems in this group do not seem to tell us substantially different things about students' abilities or strategies. So can a single, sufficiently difficult, problem act as a test for an algebraic strategy? For example, as far as it was possible to tell from the scripts, 6% correctly used an equation to solve the 11-5 Seesaw problem; 26% got it right without an equation (for example, by analysing the situation qualitatively, or by transforming the situation - "it is as if George threw away 6 and Sam didn't throw away any, so the weight of 6 bricks is the difference in weights"); and 68% could not solve the problem. If a set of activities could be found that enabled more to get it right, the change in percentage using equations could be used as a measure of the extent to which the activities were making an algebraic approach more accessible.

Another problem inspired by Lins is the following: "A limousine and two minis end-to-end are 9.8m long. The limousine is 1.7m longer than three minis. How long is the limousine?" which seems like a situation one could envisage being of interest to students, in contrast to the traditional application of algebra to problems such as "Kate thinks of a number, and Liz thinks of a number. It turns out that the numbers add up to 162. Also, three times Kate's number minus twice Liz's number is 16. What was Liz's number?". Nevertheless, Lins' results seem to suggest that the former problem may be more accessible to non-algebraic strategies than the latter because car lengths may have more "meaning" to students than imagined numbers. The evidence also suggests that $\{x + 2y = 9.8, x - 3y = 1.7\}$ would be even less accessible, because students would find it difficult to "model back" the problem into a context in which whole-part operations would be possible.

This section can be summarised by noting that a number of problems have been identified that might indicate whether students can use equations *as a strategy*.

2.5.6 Proving using Symbols

Harper (1987) gave to 144 secondary school students from Years 1 to 6 problems such as:

Sum and Difference

If you are given the sum and difference of any two numbers, show that you can always find out what the numbers are.

Harper identifies three strategies for solving this problem:

- *rhetorical*: specification of a general procedure. For example: "You divide the sum by 2 then divide the difference by 2; then to get the first number add the sum divided by 2 to the difference divided by 2; to get the second number take the difference divide by 2 away from the sum divided by 2." (Harper, 1981, p. 81)
- *Diophantine*: letters represent unknown quantities. For example: "If the sum is 8 and the difference is 2, then $x + y = 8$ and $x - y = 2$; then solve for x and y. This works for any numbers."
- *Vietan*: letters represent unknowns and *knowns*. For example: "Let the numbers be x and y, the sum be m and the difference be n. Then $x + y = m$ and $x - y = n$. Solve these to get $x = (m + n)/2$ and $y = (m - n)/2$."

> "In Year 1, all of the correct solutions were rhetorical. In years 2 and 3, the rhetorical solutions continued to outnumber the other two types. It was only from Year 4 onward that the balance shifted in favour of, first Diophantine and, then, Vietan solutions." (Kieran, 1992, p. 407)

Kieran (1992) suggests that "the use of the letter as a Diophantine 'unknown' is more cognitively accessible than is the use of the letter as a 'given'" (p. 407); and that "procedural conceptions of literal terms precede structural ones" (p. 407); but Harper's study could alternatively be interpreted as demonstrating students' growing confidence in algebraic manipulation to prove results. This problem, then, could be a useful test of concern to use algebra.

Another problem given could be similarly useful:

Consecutive Numbers

Take three consecutive numbers. Now calculate the square of the middle one; subtract from it the product of the other two… Now do it with another three consecutive numbers… Can you explain it with numbers?… Can you use algebra to explain it?

Kieran quotes Lee & Wheeler's (1987) finding that out of 118 grade 10 students given the following problem, only 9 provided a satisfactory answer:

Think of a Number Proof

A girl multiplies a number by 5 and then adds 12. She then subtracts the original number. She notices that the answer she gets is 3 more than the number she start with. She says "I think that would happen, whatever number I started with.". Using algebra, show that she is right.

The overwhelming strategy, even amongst these 9, was to use numerical evidence rather than algebra. Lee (1987) provides corroborating evidence that "For most students, numerical instances of generalisation carry more conviction than an algebraic demonstration of the generalisation."; and that "Many students do not appreciate that a single numerical counter-example is sufficient to disprove a hypothesised generalisation." (p. 316). Thus we have another test to distinguish between strategies.

2.6 Discussion of Research Implications

An action, strategy, plan, heuristic, procedure or process can be considered as theory. Being able to use such a theory might be called a "skill" or an "ability". Such a theory can be used without being consciously devised, willed or recognised. Concern to use it can therefore vary. Being concerned about a problem may lead to a concern with strategic theories that are perceived as potential solutions. It has been concluded that students have myriads of complicated, implicit and context-related theories (perhaps not all strategic), produced by a wide variety of mathematical experiences and concerns; and therefore generalised analysis of these theories and experiences is difficult. However, this research is considering "algebraic concerns" - that is, those concerns which are related to algebra and its use in some way. An example is a concern for algebraic convention: the need for mathematical expressions to be unambiguous, for example, constitutes a problem to be solved when using algebra (albeit an elementary one for those with experience).

Another is the concern to work out the number of students given the number of professors and the student-to-professor ratio. Another is to avoid looking stupid in front of one's peers. Now it might be objected from this that "algebraic" does not really limit the concerns one might be interested in improving, because *any* concern that might happen to be important to a student at any time when using algebra could be relevant. This is a good point; but while there are clearly many concerns that might be relevant to an analysis of the use of symbolic algebra in a classroom situation, the "target" algebraic concerns are those problems of elementary algebra itself. For example, to find an unknown quantity; to demonstrate a conclusion beyond reasonable doubt; to discover rules governing arithmetic; and so on. A number of these concerns have been distinguished in this chapter. These, and some others, can be put into a table:

Types of algebraic problems

Represent situations using equations and expressions. Graph data and functions. Symbolise arithmetic identities.	Solve and rearrange equations. Simplify, factorise and expand expressions. Substitute.	Find unknowns in situations. Find patterns & relationships. Justify, prove, predict, explain and pose problems.	Interpret equations and expressions within particular situations. Interpret graphs.

The first column can be labelled as "representation" (Kieran, 1996, describes them as "generational activities"); the second as "transformation"; the third as "utilisation" in that they are considered here to be "algebraic" only if symbolic algebra is used as a tool (Kieran calls them as "global, meta-level activities"); and the fourth as "interpretation". Representation problems tend to start with a situation external to algebra, and end with something more formal, conventional or symbolic; for interpretation problems it is usually the other way round; and in transformation problems there is usually no reference to any situation external to algebra. There are also the "meta-algebraic" problems of considering meanings, interpreting symbols, discussing metaphors, creating rationales, and so on. These types are not intended to be a complete characterisation of algebraic activity, nor is the typology unique, nor are the types mutually exclusive - tackling one problem will bring in others. Note that these problems are, of course, subject to quality criteria such as: complexity (for example, more complicated equations, variables or solution steps), applicability, discrimination (for example, more relevant application), facility (less effort), speed, accuracy, precision, reliability and completeness. The various studies into children's learning of algebra also place differing emphases on these problems.

Turning to the assessment of students' knowledge, understanding is not viewed as sense-making, imagining or re-enactment; it is problem-solving. It has therefore been argued that symbol interpretations, images, meanings and metaphors are not necessarily deep insights into cognitive processes and structures: but are theoretical consequences of past problems. These "meta-algebraic" theories can allow students to examine properties and relationships of processes; and they can give us an insight into students' experiences, rationalisation of experiences, and concerns. Identifying meta-algebraic theories may thus help us to conjecture students' thinking in a given problem situation; but they might not be a good guide to future mathematical behaviour.

In analysing incorrect student responses to mathematical problems, it is important to attempt to distinguish those that are based on a misunderstanding of the problem, from those that are caused by strategic theories (they don't work; they haven't been followed through correctly; they involve an absent-minded slip; or they are an incoherent response stemming from an inability to think of a strategy). Donaldson's (1978) comment about one of Piaget's experiments seems pertinent: "the questions the children were answering were frequently not the questions the experimenter had asked." (p. 49). A student's concern includes a view of the problem situation that may not coincide with that of the person who posed the problem. But the concern provides the basis on which the student tackles the situation. In the language of English & Halford (1995), this signifies the crucial importance of the student's "problem-situation model" (especially p. 241-2).

It ought to be clear that merely solving (or failing to solve) a problem cannot constitute sufficient evidence for a concern (or a lack of one). A guiding principle is that there must be some sort of struggle, an element of trial-and-improvement. What else can help identify concerns? Much work has been done within and across school subject boundaries about motivation, rewards, emotion, attitudes, affective considerations, etc.; and it is true that "identifying concerns" falls within such research. Yet it seems that there exists no consensus that one may draw on to help here. There is only one guideline that is regularly offered. Papert (1980) puts it: "Anything is easy if you can assimilate it to your collection of [cognitive] models. If you can't, anything can be painfully difficult.". The mechanism of what Piagetians call "assimilation" is, in Popperian psychology, that of trial-and-improvement *in response to a concern*. In other words, *learning* a theory is easy if it addresses a concern. So not only must children's strategies be taken seriously, but their concerns and errors must be seen as the fuel to power learning.

Teaching a theory is a kind of "trick" to make the problems into concerns, however briefly. By attempting to address one's concerns, theories are produced; these may contain further problems which may become concerns; and the original concerns do not necessarily disappear. Given the conjectural nature of learning, the attribution of students' misconceptions to "natural" cognitive development is considered a claim requiring the highest standards of testing. The hypothesis that some target theories are beyond the "conceptual grasp" of certain students can be challenged by finding activities that not only give students on opportunity to grasp and demonstrate their grasp of these theories; but that are also effective in improving students' theories and concerns as evidenced (through interview or written test) by the items identified within this chapter. What constitutes an "activity"? Because it is *engagement with algebraic problems* that is crucial for making target public theories more graspable, an activity is an experience that promotes such engagement.

CSMS appears to have been influential in shaping school algebra. The Royal Society / JMC report suggests, for example, that activities directed towards representation problems are prioritised over transformation and utilisation problems, and "fluency with symbols has become confused with rote manipulation" (p. 5). However, many of the suggested "implications for teaching" of CSMS would be challenged by a study that could find activities making a dramatic impact on students' theories - activities that do not attend explicitly to the apparent psychological or mathematical "prerequisite" foundations for these theories, but instead engender genuine concerns to use equations. This is the starting point for the next chapter.

Chapter 3
Improving Students' Equation Theories and Concerns

3.1 Introduction

Answers are sought to the question "What activities can improve students' equation theories?". Algebraic theories have been categorised in the previous chapter by the problems they are intended to address: utilisation, representation, interpretation and transformation. Cortés, Vergnaud & Kavafian (1990) describe the learning of algebra as being an "epistemological jump" from elementary arithmetic to algebra; and ask the questions "How to start teaching algebra? With which types of problems?" (p. 27). Given that the theories-concerns instrument has identified a particular student with a wide range of difficulties with *all* problem types, is it possible to predict a sequence of activities that could best improve the equation theories of this student? This is asked independently of any particular teaching or learning *styles* or *strategies* as such - "reflection on experience", "whole-class teaching", "small groupwork", "individualised learning", "investigational approaches", "waiting for cognitive readiness" - although, of course, it is extremely difficult to separate these from learning activities in practice.

The baseline sequence is to practise tackling each of the problem types individually. For example, to promote utilisation, students are given larger numbers of examples and exercises, progressively growing in complexity and abstractness, in which equation is used to represent situations in which unknowns must be found, to prove theorems, to explore patterns, etc. In the sections that follow, deviations from this baseline are considered. Why, after all, should success or experience in one problem type transfer at all into others? Given that this research starts from the presumption of conjectural solutions to problems, every prima facie case of transferability is most interesting, as it may indicate commonalities between theories used for different problems or even identical theories. Geary (1994) concludes from various studies in cognitive science: "it is clear that metacognition in mathematical problem solving requires extensive experience in solving similar types of problems" (p. 126). The levels in CSMS could be interpreted, for example, as suggesting that simplification, the finding of unknowns and representation illustrate Geary's different "types" of problem. Studies which demonstrate an activity that improves students' ability to solve more effectively problems *other* than those inherent in the activity chip away at Geary's baseline assumption. On the other hand, studies that demonstrate an activity easing just one or two of these problem types reinforce the idea of channelling instruction via separate types.

There is an attempt below to organise broad instructional schemes by means of the most relevant problem type, although this is somewhat arbitrary in a few cases. It ought to be made clear that even though the studies described here may not have explicitly set out to address the question of improving theories, they may perhaps contribute something to answering it. This may be no more than a suggested "implication for teaching" arising out of a priori analysis, but it could be as much as a controlled comparison of activities. The discussion is constrained by the potentiality of the theories-concerns instrument, developed in the previous section, to gauge improvement.

The first section discusses activities that aim to promote meta-algebraic theories. The second section focuses on activities that aim to promote equations as a mathematical tool, in particular by easing transformation. The third section considers how representation and interpretation theories may be improved (these two problem types are considered together because of their close relationship). The final section includes attempts at promoting equation solution strategies.

There are a number of reasons for gloom about the notion of comparing activities:-

- *Different things work for different people*

People have their own ways of learning. No single activity will be appropriate for everybody. It will always be possible to point out people for whom a particular activity is inappropriate.

- *Contexts can limit*

Knowledge does not automatically transfer between contexts. It will always be possible for students to gain from an activity theories that are appropriate to that particular experience and little else.

- *Understanding is always incomplete*

Knowledge is infinite. No single activity could hope to trigger all the learning one would want. It will always be possible to point to a deeper level of understanding that students have not acquired.

- *Learning cannot be timetabled*

No-one can tell in advance how long an activity should take, or after how long a particular theory will have been grasped.

- *An activity does not depend on its rationale*

The success of activities does not depend on the claims used to support their proposal. Other claims can support a proposal, and the rejection of a reason does not necessitate a rejection of the activity.

- *Benefit is intangible*

Benefit from an activity is difficult to measure. There will always be theories that are not learned, or theories that are learned and then forgotten, theories that were wrongly identified as learned, and theories learned that were never identified.

Whether there is any hope left after such gloom remains to be seen.

3.2 Meta-Algebraic Theories for Equations

In the previous chapter a number of suggestions were quoted from various researchers suggesting that students' meaning for algebraic symbolisation might be the "root" of students' difficulties with algebra. However, there is a large ambiguity in such common phrases as "the meaning that students have for letters", "the meaning of the equals sign", "meaning for formal procedures", "meaningless rules", "meaningless symbols", "meaningful referents", "meaningful learning", "external meaning", "sharing of meaning", "negotiating meaning", and so on. It is rarely made clear whether the author means by "meaning" something like "intention", "denotation", "entailment", "purpose", "importance", "definition", "idea", "interpretation", "rationale", "understanding", "gist", and so on; or sometimes even whether the word is a noun or verb. It is even rarer to find criteria for detecting such meaning in practice.

Even when authors have taken great trouble in their papers to spell out clearly what they denote by meaning, quotations are easily ripped from their context. This is rather important when claims are made such as: "[The students] seem unable to encode meaning from natural language… And they seem not to be able to recognize meaning in an algebraic sentence either." (Burton, 1988, p. 4); "[It] seems sensible to base the teaching… on the meanings for the letters that these children readily understand." (Küchemann, 1981, p. 119); "acts of practice and exercise… are invigorating and stimulating only if they have meaning and purpose" (Griffin & Hirst, 1989, p. 20); "Algebra takes on a much clearer meaning in the solution of problems which are insoluble or difficult to solve through arithmetic." (Cortés, Vergnaud & Kavafian, 1990, p. 27); "an understanding of generalised arithmetic is a necessary prerequisite for meaningfully operating a computer algebra system" (Hunter *et al.* 1995, p. 318); "Children need to be encouraged to reflect upon the meaning of the mathematical expressions" (Booth, 1994, p. 93); "Word problems are essential to create *relevance* for algebra; however, they may fail to develop *meaning* for equations." (Herscovics & Kieran, 1980, p. 572); "[The] issue around which all the others can be organised, is that of *meaning*." (Lins, 1992, p. 276); "meaning is the foundation of mathematics learning" (Kaput, 1989, p. 168).

Among all the conflicting meanings for meaning, it is necessary here to draw a distinction between "meaning for letters" as referring to *what sort of thing* students perceive the letters to be (how students interpret them, how students describe what they are, students' notion or concept or idea or understanding of letters, what metaphors or images are invoked by students in connection with letters, and so on); and as referring to what *purpose* students perceive the letters to have (how important students think the letters are, what significance students attach to them, what students think letters can be used for, and so on). The former category is considered to refer to meta-algebraic theories about letters; the latter refers to concerns about letters.

So what is being discussed in this section are activities presupposing that students' meta-algebraic theories (or lack of them) cause students' difficulties with algebra. Such theories would include awareness of imagery, beliefs about the role of language, characterisations of misconceptions and perceptions of the nature of algebra.

> "The teaching of high school algebra usually begins with the topics: variables, simplification of algebraic expressions, equations in one unknown, and equation solving. Students' difficulties with these topics have been found to centre on (a) the meaning of letters, (b) the shift to a set of conventions different from those used in arithmetic, and (c) the recognition and use of structure." (Kieran, 1998a, p. 433)

This is suggesting, reinterpreting this using Popperian psychology, that students' difficulties centre on inadequate meta-algebraic theories for algebraic symbols, particularly letters and expressions.

Griffin & Hirst (1989) is a "professional development resource pack" for helping teachers to "come to an understanding of the notion of an equation" and to "feel confident about what is needed" in preparing to teach equations (p. 4). It deliberately eschews the provision of either "ready-made classroom activities" in favour of promoting the teacher's meta-algebraic theories.

The book begins with the following questions:

> "Is 7 + 3 = 10 an equation?
>
> What is the difference between an equation and a formula?
>
> Does every equation with an unknown have a solution?
>
> Have you ever noticed that pupils who are able to follow the necessary rules for manipulating an equation and who are often able to obtain the correct answer are unaware of the meaning of the solution?
>
> What is an unknown? To whom is it an unknown?" (p. 4)

The algebraic problems that were identified in chapter 2 are present, it seems, only to the extent that they "help teachers in assisting their pupils to develop a better understanding of the nature of an equation" (p. 5). The justification for this emphasis is apparently the following:

> "Quite a large part of the secondary mathematics syllabus is concerned in some way or another with equations and their solutions, but the idea of an equation involving an unknown quantity, which can be manipulated according to certain rules, is a source of great mystery to many pupils. Even though many of the standard methods and techniques of solving equations can be automated and performed with apparent success, many errors and misconceptions still remain." (p. 4)

Note that it is the *idea* of an equation that is causing difficulties and must therefore be promoted, rather than the "skills of manipulation" which "most teachers would like their pupils to achieve" (p. 20). The main method for developing such understanding is by reflecting on the language, experiences, questions, feelings and imagery that are evoked by tackling the problems. Such reflection might "help assess the meaning of what you are doing" (p. 6). Implicit, then, are the ideas that such reflection is difficult, that it is not always undertaken, and that it can be encouraged by forthright demands to do so.

However, it should be asked to what extent inadequate meta-algebraic theories inhibit further progress with algebra more than difficulties with transformation; and also whether the best way to learn the conventions of algebraic notation that are required for representation is to develop meta-algebra theories first. In early empirical work for this research, one student said vehemently, "I think you have to definitely have the basic understanding. You don't really *need* to know, but I think it's so much better that you do. Anybody can say that 'that equals that'. But ask them what it means, and they're stumped… Anybody can plug a few numbers into an equation and get an answer. But if you ask them 'What does that answer mean?' then you haven't got a hope in hell, if they don't understand what they're doing. It just appears a jumble of letters. Well, what the hell does that mean?".

This is perhaps more precisely put by Davis (1986b) when he distinguishes an emphasis on notation (for example, the use of letters, knowledge of convention, skills in symbolic manipulation and simplification), from an emphasis on "the *ideas* which the symbols are supposed to represent" (p. 21) - what Booth (1989a) calls the "semantic" aspects of algebra. English & Halford (1995) suggest that although the syntactic components are a vital part of learning algebra, they are insufficient:

> "What is also needed and is frequently not acquired by students is the semantic component. An understanding of just what algebraic statements represent, and of why we can make certain transformations on these statements, is essential. … When students lack this semantic understanding, they simply manipulate symbols with little sense of purpose or meaning." (p. 220).

From a Popperian epistemological perspective, this distinction between the form and content of language is clearly very attractive. But Popper does not make the distinction in order to argue that the form is somehow just an adjunct to content - this is completely incorrect. It is only through objectifying ideas by means of public discourse that contents can be critically examined. Popper's nominalist position as regards language would suggest that the implications, constraints and roles of specific language forms do *not* have to be fully appreciated before they are used, because it is *through* usage that such meta-linguistic theories are best developed.

This research is concerned with letters standing for numbers because, as has been suggested by numerous authors in the preceding chapters, many children's difficulties centre precisely on creating and dealing with such representations. As was noted earlier, the Royal Society / JMC report follows Lins (1990, 1992) in distinguishing symbolic algebra from algebraic thinking; and so it suggests that one reason "why symbolic aspects of algebra have been under-emphasised is that it is clear that the mere use of algebraic literal symbols does not imply that pupils are acting and thinking algebraically." (p. 4). However, a Popperian reinterpretation of this suggestion is that the mere use of symbolic algebra does not imply that students have algebraic concerns (i.e. they do not understand the *purposes* of algebraic symbols) or algebraic theories for why symbolic transformations work, even when they are creating and transforming representations apparently successfully.

Such concerns and theories are clearly very important factors in future learning, such as representing new situations and solving new equations; and so these *algebraic* concerns and *algebraic* theories are perhaps better candidates for "causes" of algebraic difficulties than are the *meta-algebraic* theories described in section 2.3 and above. But even if meta-algebraic theories were *per se* responsible, Popperian psychology would suggest that they are better developed by using symbolic algebra in the tackling of algebraic problems than by explicit demands for reflection.

Other proposals for improving meta-algebraic theories will therefore be discussed in the appropriate sections below.

3.3 Equation Utilisation

Burton (1988) describes student responses to problems such as:

Combines Problem

> The Cornhusker Combine Company has manufacturing facilities in Omaha and Irkutsk. Management finds that for every 5 combines made at the Omaha plant, 3 are produced at the Irkutsk plant. Each year 75 combines from Omaha, and 325 combines from Irkutsk, are found to be defective as they come off the assembly line. Management wishes to market, in all, at least 8000 combines annually. How many should be produced in Omaha?

Typical responses avoid algebra, and use proportional reasoning to produce a 5000:3000 split or (to compensate for rejects) a 5075:3025 split. Even where students use algebra, they tend to be unsuccessful.

Why do students find such word problems difficult? How could students be helped to choose an equation as a tool? Kieran (1992) remarks, “cognitive studies in algebra problem solving have, up to now, been unable to explain why certain methods of instruction in the learning of schematic relations for solving word problems are more effective with certain students than with others.” (p. 403).

The activities for promoting utilisation theories (or, to put it another way, to improve concern to use equations) can be divided into two broad categories. The first assumes that provided the skills of transformation are available, students will naturally see the value of using symbolic algebra. Transformation provides the *purpose* for representation, so transformation theories must be promoted. Activities for doing this are considered in section 3.5 below.

However, Burton argues that “a major component of student difficulty with algebra is the inability to make sense of the algebraic symbol system as a language, and accordingly that remedies should be sought by considering algebra in a linguistic sense.”. (p. 2). More precisely, “Whatever their syntactic facility in manipulating the expressions of the algebraic language, many students cannot attach meaning to an algebraic expression.” (p. 2). Both syntax and semantics must become available: “If the young student receives only a quick and abstract initial encounter with variables followed by practice in the essentially syntactic skills of manipulating algebraic expressions, the semantic component of the new language may never be realized.” (p. 6).

So the second category of activities assumes that “The primary source of difficulty in solving algebraic word problems is translating the story into appropriate algebraic expressions.” (Geary, 1994, p. 127). Cortés, Vergnaud & Kavafian (1990) state that the "first step in solving a problem algebraically is to express it as an equation." (p. 27-8); and it is the failure of translation to and from algebraic symbolic language that Burton argues is responsible for the retreat from semantics: students' "remaining option is to use algebraic language simply to carry out formal manipulations on patterns of symbols. Their algebraic language is empty, having only syntax." (p. 4). It is often suggested that "for many pupils, a greater problem [than solving equations] is to formulate the equation which needs to be solved." (Mathematical Association, 1992, p. 65). So in order to promote utilisation, representation theories must be promoted. Activities for doing this are considered in section 3.4 below.

3.4 Equation Representation and Interpretation

Griffin & Hirst (1989) advocate asking students for a "story" to go with an equation, which may encourage more fluidity between context and symbols. In this way, activities that promote representation can also promote interpretation, and vice-versa; so the two are considered together here.

3.4.1 Syncopation

Burton suggests that to help students to create equations from natural language problem descriptions, there are two approaches:

1. define unknowns and then seek to arrange the information given in such a way that it forms an equation;
2. assemble the whole sentence in English first, and gradually transform the sentence via (what others have called) "syncopated" sentences into algebra.

Take, for example, the following problem:

Wallet Problem

A wallet contains \$460 in \$5, \$10 and \$20 bills. The number of \$5 bills exceeds twice the number of \$10 bills by 4, while the number of \$20 bills is 6 fewer than the number of \$10 bills. How many bills of each type are there?

1. The traditional strategy is to let x = number of \$5 bills, and so on; write down the side conditions ($x = 2y + 4$, $x = y - 6$) and the values ($5x$, $10y$, $20z$); and finally construct the central equation $5x + 10y + 20z = 460$
2. Burton's proposed approach is to reason as follows:
   TOTAL MONEY is \$460

so VALUE of FIVES + VALUE of TENS + VALUE of TWENTIES = $460
so $5(FIVES) + $10(TENS) + $20(TWENTIES) = $460
so 5[2(TENS + 4)] + 10[TENS] + 20[TENS – 6] = 460
and now notice that we have an equation in one unknown that can be solved.

The advantage of the second method, is it argued, is that the algebra seems to "flow" from expression in natural language. Nevertheless, it hinges on starting with the total value; whereas some students may start with, say "NUMBER OF FIVES is twice NUMBER OF TENS plus 4". So the argument rather begs the question: how do students know what constitutes a useful syncopated form? Lins (1992), for example, warns that if letters are used as abbreviations for words, different quantities can sometimes end up being represented by the same letter; and that the order of verbal syncopation does not necessarily match the conventions of arithmetic. Similarly, MacGregor & Stacey (1993b) caution "Students should be made aware that some relationships (such as "eight more than") are easy to express in natural language and easy to comprehend but must be paraphrased, reorganized, or reinterpreted before they can be expressed mathematically." (p. 229).

3.4.2 Recording Trial-and-Error

Rubio (1990) activity followed this sequence:

1. use of numerical trial-and-error;
2. representation of the trial-and-error process using letters to stand for unknowns;
3. algebraic and/or numerical resolution of the equation.

The selection of problems presented took account of the number of times unknowns appeared in "the equation associated with word problem" (p. 128); whether the unknown might appear on both sides; the difficulty of expressing an unknown as a function of the others; equations in which the unknown is a divisor; and the need for brackets. However, for a trial-and-error approach, these attributes were difficult to characterise, particularly as the accessibility of the context seemed to have a greater impact on difficulty. Rubio suggests that this use of a numerical approach "generates consciousness of the equivalence between two expressions" (p. 131). However, students sometimes failed to formalise operations carried out mentally. Moreover, those students who had not already grasped the transposition method from previous lessons had great difficulty in attempting to "understand the equation as an equivalence relationship and therefore to establish and assign meaning to the equations derived from the word problem." (*ibid.*).

3.4.3 Tables

One way of encouraging representation is through the use of tables of values, as described by (for example) Demana & Leitzel (1988) and Kutscher & Linchevski (1997). The latter study asked students to fill in a table of values based on a word problem, and encouraged them to find general terms in the last row:

Basketballs and T-Shirts

Danny bought 3 basketballs and 12 T-shirts for his basketball team and paid a total of \$243. A ball costs \$6 more than a T-shirt. What is the price of a ball and of a T-shirt

Price of a ball	Price of a shirt	Cost of 3 balls	Cost of 12 T-shirts	Total Paid
10	10 – 6	3•10	12•(10 – 6)	3•10 + 12•(10 – 6)
14	14 – 6	3•14	12•(14 – 6)	3•14 + 12•(14 – 6)
…	…	…	…	…
…	…	…	…	…
x	*x* – 6	3•*x*	12•(*x* – 6)	3•*x* + 12•(*x* – 6)

The researchers noted that students were "not solely generalising vertically, but using a spontaneous combination of both vertical and "horizontal" generalisation." (p. 170). That is: they could use previous columns to find an expression as well as previous rows.

Syncopation, the recording of trial-and-error and this usage of tables are special cases of the formalisation of method.

3.4.4 Formalising Method for a Variable

In the project SESM, Booth (1984) describes the difficulties of representation as being partly caused by children's use of implicit, "informal", context-dependent arithmetic methods that are troublesome to symbolise. However, she also recognises that children may not see that formalising is "an appropriate thing to do." (p. 38). A teaching experiment designed to alleviate these difficulties used the metaphor of a "mathematics machine" that can accept instructions to solve problems. This focused students' attention on the "need to make explicit the procedure by which a problem is to be solved, and on the need for precision in representing that procedure" (p. 41), and provides a rationale for the use of expressions. So, an example problem would be "I want the machine to add 5 to any number I give it. How will I write the instructions?", expecting the answer $n + 5$. Another example is "Find the area of any square.", expecting something like $a \times a$. So the programme essentially consists of tackling representational problems, although some comparison of equivalent expressions (and hence transformation) was initiated by the teachers.

The experiment was conducted over five or six 40-minute lessons with seven classes. Variations on the CSMS algebra test were used as tests - about half were representational and most of the rest transformational.

In one school, a comparison (of "parallel" classes) between the "mathematics machine" approach and the usual algebra programme for that school suggested that, although both programmes improved performance on the test, the "gain for those children who followed the experimental teaching programme was significantly greater than that of those who participated in the control programme." (p. 73). Meanwhile, a comparison in another school demonstrated that parallel groups that received no instruction made little progress from pre-test to a delayed post-test some three months later. However, one class that was taught using the usual approach achieved comparable results to a parallel class taught using the experimental approach.

Booth concludes that the programme was "effective in improving children's general level of understanding in elementary algebra, as measured by a sustained improvement in performance on test items in this topic. At the same time, gains observed were not great, being in the order of an average gain of three to seven items correct out of a total of 21." (p. 84).

Booth's argument that "some of the difficulty which children have appears to be related… to a 'cognitive readiness' factor." (p. 87) because "the observed similarities in the nature of the informal methods used by different children… suggest some generality in cognition which requires explanation." (p. 95) would if valid, have important implications for teaching. It is clearly a bold argument, since as Olivier (1988) notes (about "L + M + N never equals L + P + N") that "certain experiences (instructional interventions) may well address this misconception successfully, disproving the developmental hypothesis" (p. 512). SESM itself provides an example of how swift improvement with respect to a particular strategic theory can challenge the hypothesis that grasp of the theory is dependent on maturation factors: Booth "noted that the idea of an unclosed, non-numerical answer was initially not accepted by children in the age range investigated here, namely 12 to 15 years. However, the apparent effectiveness of the teaching programme in restructuring children's thinking in this regard would suggest that the notion was not beyond the conceptual grasp of these children." (p. 91). Nevertheless, there is a way out for the neo-Piagetians, and it is to conjecture (as Booth does) that "acceptance of lack of closure" and "seeing letters as representing generalised numbers" are on different "cognitive levels". Nevertheless, Booth hints strongly that it may be simpler to admit that traditional Piagetian cognitive development does not ensure the growth of algebraic understanding, and so provides a number of recommendations:

> "1. Since children appear to be predisposed to the idea of letters as specific unknown values, it may be useful to adopt the generalized number interpretation of letters from the time that letters are first introduced." (Booth, 1984, p. 92).

Although this initially appears to be a quite dramatic implication, Booth subsequently seems to imply that, aside from the use of a "mathematics machine" model, this suggestion deviates from current practice solely in seeing that the x in $x + 5 = 8$ (say) can take a whole range of values, only one of which makes the equation true. It could be argued that the distinction is merely technical, because even if one views x as representing just *one* specific (but unknown) value, it is a necessary consequence that one can imagine many other *proposed* values for x. It is then a simple step when one gets to quadratics or equations in more than one variable to discover that there are multiple correct values for these equations.

On the other hand, Kieran (1992) mentions the Soviet approach described by Davydov (1962), based on extensive teaching with "part-whole relations and the use of problems involving only literal data" (Kieran, 1992, p. 404). Pimm (1995) describes an activity called "Rulers" (a version of which was later computerised as "Grid Algebra") for first recording and then generating moves around a grid in which horizontal moves involve adding or subtracting something to the current cell and vertical moves involve multiplying or dividing. By starting at a cell with y in it, for example, one can move two cells left to $y - 2$ and one cell down to $2(y - 2)$. Expressions name specific squares, but also tell you what to do to get from one privileged location (x or y, or whatever) to the one named by the expression." (p. 94). So one could also start from y and move one cell down to $2y$ and two cells left to $2y - 4$. This is the same cell as previously, so $2(y - 2)$ must equal $2y - 4$. In introducing the grid, the teacher used certain noises and gestures to focus attention and aid recall.

Booth notes that a "de-emphasis on the requirement for a final numerical answer has been suggested to be a prerequisite for children's acceptance of the unclosed expression as a legitimate 'answer'." (p. 94). An activity that blatantly disregarded this suggestion and yet was successful in improving students' tolerance of expressions would constitute a useful test of the advice.

> "2. Since many children confuse the arithmetic and algebraic usage of literal expressions, it may be useful to make this distinction explicit by discussing, for example, the alternative meanings of terms such as $3m$ ('3 times m' or '3 metres'). This discussion may usefully include consideration of the meaning of such symbols as the equals sign." (p. 92)

On the other hand, we have seen how attempted instruction based on such consideration of the meaning of symbols was of limited success in connection with the student-professor problem (Rosnick, 1981; Philipp, 1992).

> "3. Several work schemes use algebraic initialisation, such as 'V = the number of vertices' as a mnemonic device. However, since children do not always appear to make a distinction between '*v* for vertices' and '*v* for the number of vertices', such usage may be best avoided, at least initially."

Evidence from the student-professor research would appear to be equivocal on this point. So only if instruction that persisted in using initial letters as variables were able to produce significant improvements in such items, would the advice be compromised.

> "4. Children need to be encouraged to reflect upon the meaning of the mathematical expressions they meet. This is essential to an appreciation of the need for rigor in symbolising different mathematical operations (such as division and subtraction expressions) and so must form a part of any attempts to help children's understanding of the formalisation of mathematical procedures… It is also useful to the child's handling of algebraic simplification exercises of the kind '$2a + 5b + a$'. The consistently successful handling of examples of this type can only proceed from an awareness of the letters as representing possibly different numerical values." (p. 93)

This is the clearest possible demand for students to *consider meaning* (that is: active development of meta-algebraic theories as opposed to algebraic concerns). "Repeated interpretation of the symbol's meaning" (p. 94) is advocated. But representations only have power when they are *seen to represent*, so if a student cannot relate the symbols to a concern, then he or she will not understand those symbols.

3.4.5 Formalising Method for an Unknown

The "Think of a Number" (TOAN) activity is another way of concentrating students' minds on representing method, but by aiming to find an unknown number (usually by the reverse-flowchart approach described in section 2.5.2) rather than by aiming to produce a representation of a method applicable to "any number". Morelli (1992), for example, shows how concrete and iconic approaches to TOAN could lead to exploring the distributive law, the collecting of like terms and eventually symbolic algebra. Pimm (1995) describes how a teacher used TOAN to generate a notation; he then aimed to refine that "functioning and consistent notation into the one that is conventionally employed, rather than striving for the fully-fledged, conventional one from the outset." (p. 93).

The software-based "Marble Bag Icon Laboratory" of Feurzeig (1986) and Roberts *et al.* (1989) is similar, in that students are shown how to create and solve stories about bags containing an unknown number of marbles. An iconic representation of the story problems can be converted into natural language; but students are also introduced to algebraic notation as a shorthand way of writing the stories. For example (see Thompson, 1989), the student could be asked to construct a story ending in $2(4x) + 2$. They could also be asked to find the number of marbles in a bag, given the total at the end. The inversing method for solving equations is used; although Kieran (1992) comments that in doing so "the student is able to operate exclusively with numbers" and thus avoid operating on equations.

Another example is the proposal offered by Meira (1990), who suggests "the value of dynamic physical systems as powerful aids in promoting students' understanding of symbol systems and concepts. (p. 107). But it is arguable from the case study that while the students' understanding of the system grows and the students' representation of the system (a winch pulling a block, which follows the rule Final Position = Number of Spool Turns × 4 + Initial Position) gets closer to conventional notation, the representation does not appear to play much of a role in that understanding. Nevertheless, it is suggested that the creation and solution by students of their own marble bag stories, TOAN puzzles or dynamical systems could assist the introduction of algebraic notation; but representation of equations with unknowns on *both* sides should not be ignored.

3.4.6 Belief-Revision Software

Aziz (1996) applied AI techniques in a program (TRAPS) that models student's thinking in representing situations using equations with two variables (like the student-professor problem), and in interpreting such equations. The program either confronts students with contradictions between their representations and interpretations; or, if there is no contradiction, provides "canned text" explaining, for example, that "S stands for 'number of students', not 'students'.".

However, Aziz points out that such an approach may help students get the symbols the right way round in the short-term (although the wordiness of the explanation may be counter-productive), but this could easily be because they deliberately put the equation in the form they know the computer thinks is correct rather than the form that they themselves think is correct. In particular, one must suspect (based on the arguments in section 2.4) that those who are consistently using static comparison would not value the argument provided by the canned text.

Although the program's student model was effectively pre-programmed, Aziz gives examples of other programs that can act as a tutor to help students do the translation - such as Hawkes *et al.* (1993) and Singley & Anderson (1989). However, software that mimics what teachers do is of less interest in this chapter than software that can improve students' equation representation theories in ways that teachers cannot. A more exciting use of AI would be the student *teaching the computer* to solve new problems (Papert, 1980), perhaps by building on programs that can translate only a small set of word problems into equations at first (such as Bobrow, 1968).

3.4.7 Step-by-Step Extension of Existing Representations

Herscovics & Kieran (1980) suggest that traditional approaches to introducing equations can fail to "establish a meaning for equations.". Hence children are lost when recall of arithmetic facts is unable to solve "open sentence" problems like $3.1 \times \square = 296$. Moreover, the "Think of a Number" type approach is beyond "those who cannot accept the representation of a number by a letter." (p. 572); and translating word problems into equations is akin to translating into an unknown language. "Word problems are essential to create *relevance* for algebra; however, they may fail to develop *meaning* for equations." (p. 572). Here, an "expansion of meaning" for the equals sign and a "construction of meaning" for algebraic equations are examined; a "construction of meaning" for operating on equations is discussed in section 3.5 (below).

First, students are asked to give an example of an equation. Should it be of the typical form "two numbers on the left side and the result on the right", they are asked "Can you use the equal sign with an operation on both sides?". Most students responded with an example involving commutativity, for example $5 \times 4 = 4 \times 5$. The next question, then, is "Can you give me an example with a different operation on each side?". So answers are obtained such as $6 + 3 = 3 \times 3$. Then, "Can you give me an example in which you have more than one operation on each side?"; for example $2 + 2 + 2 = 2 \times 3$. Some of the equations did not follow the conventions of operation order; for example $2 + 1 \times 5 = 3 \times 4 + 3$, so students were also asked for an example involving brackets.

By creating these "arithmetic identities ", students had in effect "expanded the meaning" of the equals sign from being a "do-something signal" to indicating that the operations on each side yield identical values:

> "If this expansion were not done first, the student would be bringing with him into the study of algebraic equations the idea that the result is always on the right side of the equal sign. Thus, equations such as $3x + 5 = 26$ might fit in with his existing notions, but $3x + 5 = 2x + 12$ would not. Not only would the presence of this multiple operation on the right side be foreign to him, but also seeing it for the first time within the context of an algebraic equation would add to the cognitive strain." (Kieran, 1981, p. 321)

English & Halford (1995) note that "The important point that emerges from analogy theory is that learning algebra will depend crucially on how well arithmetic relations are learned, because arithmetic relations are the source for starting to understand algebraic relations" (p. 72). This is supported by Booth (1989b) who found that *arithmetic* notions of inverse operations, commutativity, and associativity by beginning algebra students were poorly understood. She therefore suggests that "students' difficulties in algebra are in part due to their lack of understanding of various structural notions in arithmetic" (p. 141).

Some might see this approach as a pre-algebraic attempt to avoid a "cognitive obstacle". However, it can be re-interpreted in Popperian terms: although it does not initially appear to involve algebraic problems, it challenges the theory that equations can only solve problems like "What number is 6 × 8?" and opens up the possibility that equations can solve problems like "Is 6 × 8 the same number as 12 × 4?"; and this has clear ramifications for the range of situations that can be symbolised algebraically.

Turning to the "construction of meaning" for algebraic equations:

> "It is now possible to define the concept of an equation by expressing the mathematical idea involved without resorting to unnecessary formalism.
>
> "We take an arithmetic identity and cover up one of its numbers with a finger. Thus we define an equation as *an arithmetic identity with a hidden number*." (p. 575)

Then the empty box notation is introduced; and finally the box is replaced with a letter of the alphabet. Many letters are used. Children soon learn that a given arithmetic identity can lead to many different equations. They are also taught that they can put two letters in the equation. If the same letter is used more than once, it has to hide the same number each time; but there is, of course, nothing to stop us hiding the same number in different locations with different letters (another "cognitive obstacle" avoided).

The advantage of this approach is that students' knowledge of arithmetic can be "transformed gradually so that they can build for themselves the notion of an algebraic equation." (p. 573), although there is an initial lack of generality (all equations have a solution). But:

> "… that such equations develop an existence of their own is evidenced by the fact that our students quickly went on to constructing equations without first writing arithmetic identities." (p. 576)

For example $2x + 3 = 2x + 4$ can be invented, which has no solution. So the definition of equation moves from "an arithmetic identity with a hidden number" to "any algebraic expression of equality containing a letter (or letters)" (p. 577).

So by an "expanded meaning for the equals sign" and a "construction of meaning for algebraic equations" one understands the *step-by-step extension of existing representation theories to a wider range of relevant problems*. What is lacking from this improvement is any sense of *why* a student would be at all interested in such problems, unless they are being posed by a researcher.

3.4.8 Automatic Equation Solvers

Lins (1992) concludes: "we have shown that non-algebraic models used as primary ways of dealing with problems involving the determination of a number or numbers do constitute an obstacle to the development of an algebraic mode of thinking" (p. 328). If one considers learning algebra to be about acquiring a set of potentially useful strategic theories (rather than "developing a mode of thinking") then this conclusion suggests that the availability of *non-algebraic* means for solving problems may inhibit the perceived need for algebraic representation much as a pen might be an "obstacle" to proficiency with the word-processor. One might agree, in that case, with Cortés, Vergnaud & Kavafian (1990) that the value of algebra becomes more apparent when tackling problems that are insoluble or difficult to solve through arithmetic; and that therefore students learning utilisation should not be given problems for which non-algebraic methods are available. However, this has the twin difficulties that such problems may be difficult for students even to start thinking about, because arithmetic strategies are inapplicable; and that the transformations required to solve the program may be rather difficult. A way out of this hole may be to let technology deal with the transformation, and so make an algebraic approach more attractive than, say, trial-and-improvement or a whole-part approach.

Word problems offer one method of providing a purpose for symbolic algebra, yet "Describing the problem situation in an algebraic form may be high on the teacher's agenda, but not on that of the pupils." (Ainley, 1995b, p. 26). Can technology provide an incentive for students?

Heid (1990) and Heid & Zbiek (1995) describe "Computer-Intensive Algebra" (CIA) in which transformation (including graph plotting) is entirely relegated to computers so that realistic problems can be explored, with the aim of developing students' concepts of function and variable. Heid & Zbiek cite research results suggesting that "with access to tools, CIA students can solve typical algebra word problems and perceive similar structures in word problems as well as or better than traditional algebra students." (p. 655). It seems from the examples given that *equations* tend to arise only as meeting points of functions, and in situations where numerical solution would be sufficient; but this may not be typical of the actual activities undertaken. On the other hand, the use of automatic equation solvers to solve numerically can enable students "to tackle a very much wider range of equations than has been possible with the purely algebraic techniques which have hitherto been customary in most mathematics classrooms, and which have usually limited problems to those which lead to linear or quadratic equations." (Mathematical Association, 1992, p. 65). The possibility of more realistic problems may promote the value of symbolic algebra.

This focus on tackling interesting problems leading to mathematical ideas rather than learning mathematical ideas leading to subsequently interesting problems clearly fits very well with the Popperian psychology emphasis on concerns as the driving force for learning. Davis (1992) expresses this very clearly, when he rejects "application" as just some sort of afterthought to mathematical theory:

> "Instead of starting with 'mathematical' ideas and then 'applying' them, we would start with *problems* or *tasks*, and as a result of working on these problems the children would be left with a residue of 'mathematics' - we would argue that mathematics is what you have left over *after* you have worked on problems. We reject the notion of 'applying' mathematics, because of the suggestion that you *start* with mathematics and then look around for ways to use it. Fourier did not develop Fourier series, and then decide to 'apply' them to the study of heat flow; he set out to study heat problems, and when he had worked on this for a while he had Fourier series as one of the by-products. Presumably, in the remote past, someone set out to keep track of something - sheep, perhaps? - and when they had invented a way to do this they had begun to develop the idea of 'counting'." (p. 237)

He gives a classroom example in which children developed the idea of mathematical proof as a by-product of their work on a task. "They could talk about mathematical proofs in an intelligent way *because they were speaking from experience*." (p. 238). That is: experience with *problems*.

However, when Hunter *et al.* (1995) used Derive with 14-15 year old students, it was concluded that the computer algebra system failed to develop students' understanding of generalised arithmetic: "A [computer algebra system] can be of benefit for the students' learning of algebra as long as the students are mathematically ready to use it. However a more traditional approach with its attention to constructive detail is more appropriate up to that stage." (p. 322). On the other hand, it is noted that the "investigative intention of the experimental groups' work was not fully realised in practice. This is in contrast to the results of the pilot studies reported in Hunter & Monaghan (1993), where students regularly created and tested their own rules." (p. 321). This is attributed to better motivation in the pilot groups and to "the smaller amount of student-student debate when they used palmtops as opposed to laptops." (p. 322). Interviews with the students revealed that "Opinions as to the usefulness of Derive were mixed. Some students felt that it had assisted their progress in the topic whilst others felt that they had spent their time simply pushing buttons. Several students expressed their frustration with the technicalities of operating the machine." (p. 317).

Does experience of automatic equation solvers follow through to environments in which the technology is not available? The evidence is still ambiguous at this stage. Since there are now calculators available that incorporate such facilities, further research in this area seems important.

3.4.9 Word Problem Representation Software

Garançon, Kieran & Boileau (1990) believe that the use of computable algorithms to represent word problems "constitutes an intermediate step in the development of standard algebraic representations." (p. 51). To explain a rule to a computer requires a formality that is not necessary to explain a rule to humans, and this can be an opening to symbolic algebra. Kieran, Boileau & Garançon (1989) used the software tool CARAPACE to allow students to express relationships in a form closer to syncopated natural language than equations. Students then explored expressions using substitution. Kieran (1992) suggests that "students were moved toward thinking in terms of forward operations rather than inverse operation" (p. 405). English & Halford (1995) identify thinking in terms of forward operations as "a major cognitive shift for the beginning algebraic student" (p. 240). Kieran continues: "The teaching approach used in this study was effective in helping students develop a problem-solving method that they could formalise with apparent ease." (p. 405). Problems that, if represented as an equation would have two occurrences of a variable "appeared to be no more difficult to represent and solve than traditional algebra word problems involving only one occurrence of the variable." (*ibid.*). This, Kieran states, is in contrast to the findings of Filloy & Rojano and others "which showed students experienced considerable difficulty not only in setting up a single equation involving two occurrences of the variable but also in solving it by formal methods." (*ibid.*).

Similarly, Thompson (1989) found substantial progress in quantitative reasoning when middle school students entered relationships between quantities into the "Word Problem Assistant"; while, the "Algebraic Proposer" (Schwartz, 1987) is another software tool that can assist in the representation of problem situations; but in this case only if the student already knows algebraic conventions. Abidin (1997) found gains in problem solving performance when students used the "FunctionLab" to represent word problems using a schematic language.

3.4.10 Programming

Research has been conducted into ways of using the formal (but not quite algebraic) notation of computer programs as a stepping-stone to symbolic algebra:

> "Like algebra, Logo is a formal system with precise syntax and rules and pupils must perceive it as such before they can use it in any meaningful way." (Sutherland, 1990, p. 163)

Environments and languages such as Boxer (diSessa, 1995), Visual Basic, calculator programs and spreadsheet macros are similarly valuable tools. Sutherland suggests:

> "One of the difficulties with 'traditional' algebra is that it is not easy to find introductory problems which need the idea of variable as a problem solving tool. Many introductory algebra problems can be solved without using algebra. This is not the case in the Logo programming context. Logo is a language for expressing generalities and in order to express the generality it is essential to name and operate on a symbol as representing a variable." (p. 164)

Nevertheless, the Mathematical Association (1992) makes the point:

> "There are, of course, differences between the ways in which variables are used in programming and in symbolic algebra. In programming, for example, variables are used as placeholders for numbers which will be known at the time at which the program is operating. There is no need for algebraic manipulation and so, when writing programs, pupils may make use of algebraic notation without necessarily understanding that $2(a + b)$ is the same as $2a + 2b$ or that $x - (y - z)$ is the same as $x - y + z$." (p. 67)

Of course programmers are often striving for ever faster, neater and more economical ways of achieving the same end; and this striving *can* lead to them learning about simplification of expressions: the lack of memory on my first computer meant that $x - y + z$ would have been evaluated quicker and used up 2 fewer bytes than $x - (y - z)$; which is a significant consideration when the entire program has to fit into less than about 900 bytes *and* land the spaceship on the moon in real-time (or whatever). However, to what extent programming assists the fostering of theories for representing problems using *equations* is questionable. Moreover, Soloway *et al.* (1982) found that students tackling the student-professor problem who were encouraged to write a program to model the relationship were more successful than those writing an equation. (On the other hand, this finding supports the earlier suggestion that seeking an *operation* rather than a *relation* might be preferable when representing a situation).

In Logo, meanwhile, the variable in the procedure title line can be subjected to lines of programming without the need for evaluation until runtime. Sutherland (1987) found some evidence of students accepting algebraic expressions as useful objects after using Logo, and of an understanding that a variable name represents a range of numbers. But there seems to be little evidence that theories for representing situations using equations are improved by using Logo (Sutherland, 1988 and 1989a; Hoyles & Sutherland, 1989).

3.4.11 Spreadsheets

Sutherland (1995) reports on a study into how students' experiences with and spreadsheets influence their learning of algebraic ideas (see also Sutherland, 1993; Rojano & Sutherland, 1991; Sutherland & Rojano, 1993).

> "At the beginning of the study most of the pupils said that they could not answer the algebra questions because they had never seen anything like them before. Many of them interpreted a letter as representing the position in the alphabet. In subsequent interviews some pupils began to refer spontaneously to their Logo or their spreadsheet work when presented with the algebra questions." (p. 279)

For example:

Marbles Problem

If John had J marbles and Peter had P marbles what could you write down for the number of marbles they have altogether?

Rachel moved from adding the alphabet positions to saying "Well J marbles could be anything and P marbles could be anything… say J could be 10 and P could be 12… so the answer could be any number.". Sutherland suggests that "this represents an important move in the development of algebraic thinking and that this development has been mediated by the algebra-like symbols of the computer environment." (p. 279).

Perimeter of Field Problem

The perimeter of a field measures 102 metres. The length of the field is twice as much as the width of the field. How much does the length of the field measure? How much does the width of the field measure?

None of the 14-15 year olds were able to solve the field problem at the beginning, even using methods such as a whole-part strategy (i.e. 102 ÷ 6) or trial-and-improvement; but all could solve it using a spreadsheet by the end and many could solve it without a computer. The spreadsheet method is something akin to: put a trial width in one cell (A2, for example); a rule for the length of the field in another cell (for example the formula =A2*2 could go in B2); a rule for the perimeter in a third cell (for example A2 + B2 + A2 + B2 could go in C2); and then vary the trial width (A2) until the perimeter (C2) is 102. The spreadsheet method can be portrayed as similar to the algebraic approach (Let width of field = x. Then length of field = $2x$. Then perimeter = $x + 2x + 2x + x$. So perimeter is $6x$. $6x = 102$, so $x = 17$); except that the last step is trial-and-error rather than algebraic manipulation.

A previously unseen problem was used in the post-interviews:

Chocolates Problem

100 chocolates were distributed between three groups of children. The second group received 4 times the chocolates given to the first group. The third group received 10 chocolates more than the second group. How many chocolates did the first, the second and the third group receive?

Jo outlined on paper a spreadsheet-type solution:

A	B	C	D
first group	second group	third group	Total 100
=B1-4=	=A1×4 Return	=B1+10	=B1+A1+C1

Similarly, Ainley (1995b) used spreadsheets and graphs to explore ways of introducing children to "the power of generalising through formal algebraic notation" (p. 27). Two 11-year-olds who did not know algebraic notation and had never devised a spreadsheet formula were shown how to enter and replicate a given spreadsheet formula; with little further help they were eventually able to come up with the expression = 30 – B11 * 2 to help them solve the sheep-pen problem. The cell reference was initially "little more than an alternative *name* for the value of the width. Later Jordan at least used it as a placeholder for a *potential number* soon to be realised. … Finally he seemed to be using the cell reference as a placeholder for a *range of numbers*" (p. 32).

How easy would it be to move from spreadsheet representations to symbolic algebra? When Sutherland asked Jo "If we call cell A2 X what could you write down for the number of chocolates in the other groups?", Jo wrote down: =X, =X×4, =X×4+10.

> "Jo, who had always been unsuccessful with school mathematics, had successfully carried out what is considered to be the most difficult part of solving an algebra story problem, that is representing the problem in algebraic code." (p. 285)

Expressions are suddenly involved in Jo's strategies for word problems: "Results so far suggest that it is not as problematic as we might have supposed to transfer from, for example, a spreadsheet expression (3A5 + 7) to an algebraic one ($3x + 7$). This seems to be because both the spreadsheet symbol and the algebra symbol come to represent "any number" for the pupils." (p. 285). However, Dettori *et al.* (1995) point out that the equals sign in *spreadsheets* assigns a value to a cell, but in *algebra* it represents a relation: "The inability to write relations in a spreadsheet implies that it is not possible to use it to completely represent algebraic models." (p. 265). Moreover, they argue, as with the Mathematical Association earlier about programming, that the spreadsheet is ultimately limited by its inability to handle unevaluated relations, and hence equations. On the other hand, the links with expressions are strong (Rojano & Sutherland, 1997).

3.4.12 Graphical Representations

Although graphs lie outside the scope of this thesis, The potential for graphing software to make the relationships between symbolic and graphical representations more transparent has been recognised for some time:

> "… the dynamic nature of the medium supports dynamic changes in variable values that renders the underlying ideas of variable and function more learnable, which should make them accessible to a younger population, and which in turn makes possible a much more gradual and extended algebra curriculum, beginning in the early grades." (Kaput, 1989, p. 192)

Much work has been done in this regard (see Kaput, 1989 and 1992, for a survey of the literature).

Fey (1989b) points out the dynamism offered by computer representations in carefully constrained exploratory environments; and he also notes that "while some multiple embodiment computer programs might be viewed as poor simulations of more appropriate tactile activity, it has been suggested that this electronic representation plays a role in helping move students from concrete thinking about an idea or procedure to an ultimately more powerful abstract symbolic form." (p. 255). One interesting innovation is the software environment of Nathan, Kintsch & Young (1992), which is able to animate aspects of equations when instructed by the student. Such an approach would allow students to learn about representation by examining the effect on animations.

3.4.13 Expressing Generality

Mason *et al.* (1985) present a book full of ideas for encouraging the "expression of generality" (see also Mason & Pimm, 1984). They argue about algebra that "… like all language learning, it is best learned by constant use. That means using it to express the otherwise inexpressible." (p. 53).

As Sutherland (1990) has already noted: "During the 1980s and influenced by the CSMS research findings… the introduction of symbol manipulation was delayed and pupils' first introduction to algebra was more likely to be in the context of expressing generality" (p. 160). Very often this was in the context of seeking patterns, as in the matches example in chapter 2. MacGregor & Stacey (1993a), for example, cite official recommendations for the investigation of patterns as a route to algebra. Typically, the student would be asked to predict the hundredth number, find a general method and then express this method algebraically. Some argue forcefully that the activity of expressing relations algebraically is necessary to avoid over-stressing the importance of symbolic manipulation. Booker (1987) states that "Rather than focus on [the] procedural side of algebra from the outset, it would be more appropriate to build up an awareness of the need for a concise representation of relationships and, indeed, to focus on the determination of these general relationships." (p. 278). Davis (1984) suggests that students could be introduced to algebra through expressing concisely such relationships as: $1 + 0 = 1$, $2 + 0 = 2$, $3 + 0 = 3$, …, $1066 + 0 = 1066$, … Lamon (1998) describes a systematic method for representing a situation: identifying the variable quantities in a situation, making explicit necessary assumptions, describing the relationships between quantities verbally, representing the relationships using arrows between tabulated quantities and finally classifying structurally similar situations using algebraic notation.

However, Lee (1987) concludes from one research project that "A majority of students do not appreciate the implicit generality of algebraic statements involving variables." (p. 316). Moreover, "From the child's point of view, it is difficult to see any purpose in formalising the pattern in algebraic terms: a verbal description of the pattern, or a generic method for calculating values, may seem just as efficient for giving the solutions required." (Ainley, 1995b, p. 27). See also Ursini (1991).

But it was suggested in the discussion of Popperian psychology in chapter 2, and in section 3.2, that concern for symbolic algebra crucially involves having a genuine purpose for it. It is potentially disastrous if "once generalised statements are produced most students do not invest them with any meaning or see any use for them other than as a condensation of the problem statement." (Lee, 1987, p. 316). Cortés, Vergnaud & Kavafian (1990) support Lee's claim: "Most pupils are not familiar with the concept of the equation. For them an equation is an abbreviated way of writing the terms of the problem: a summary. The purpose of the equation largely escapes them." (p. 28). Ainley (1996) writes that "the lack of any sense of *purpose* for the use of formal algebraic notation in traditional approaches to beginning school algebra may contribute to children's difficulties in accepting formal notation." (p. 405). Sutherland (1991) suggests that, from a Vygotskian perspective, "If algebra is a language which can structure thinking then we might predict that methods which present the algebraic language as a final translation of an already understood process will restrict pupils in their development." (p. 44).

Moreover, as Popper (1972) points out, while "it is always possible to 'explain' every linguistic phenomenon… as an 'expression' or a 'communication'", human interactions also depend on the 'descriptive' and 'argumentative' functions of language. In the case of algebra, it is not sufficient for appreciating the value of algebra to practise the *expressive* function, or even the *signalling* function. Even if it is granted that representations can be *true* or *false* (for example 2, 4, 8, 16, 31 being described by 2^n), the *argumentative* function of language (even at the elementary level of seeking to omit the false) is required to analyse the validity of representations.

Take, for example, the programming of a computer. The significance of learning a formal language is not just that students identify relationships within a problem and then attempt to *express* them by *signalling* to the computer that it should execute various instructions. It is also that in executing the program or pressing the return key students are also *running their argument*, subjecting their theories to the higher authority of the problem situation. In the Logo and spreadsheet studies, it was not just the fact that students *used* variables in their symbolic code that helped them to greater success in some of the CSMS tasks; it was also that in attempting to find a way to represent their strategic theories for solving the problems, they had to put their representations to the test. Considered linguistically, the external protocols of formal languages such as symbolic algebra, Logo and spreadsheets are not just expressions, instructions or assertions - they are also tools for argumentation.

If this emphasis on the argumentative function seems abstruse, one practical question that it informs is whether it is easier to start learning symbolic algebra by:

1. thinking of x as a *particular (but unknown) number*; and then move on to the possibility of x representing several values simultaneously; and then to the idea of x representing "any" number;

or by

2. thinking of x as a *variable* used to express generality, and exploring what that might mean in a variety of contexts; one of which would be when a simple linear equation acts as a constraint on x and the "process of solving an equation can be seen as seeking a succinct explicit description of its scope." (Mason *et al.*, p. 60).

However one seeks to characterise the choice ("specific unknown versus generalised number", "procedural versus structural", "seeing particularity versus expressing generality") it is clear from the literature that the whole thrust of 20th century school algebra (curriculum and pedagogy) has been away from the first approach, towards the second. It is time to question this move.

This is not to suggest that the second approach is not valuable, but that the first approach should not be rejected for poor reasons. The line of reasoning above would suggest that the crucial factor in learning symbolic algebra is not whether the letter represents an unknown or a variable (for clearly different problems call for a wide variety of different roles for letters), but that learning takes place through tackling concerns in which the use of symbols is a genuinely convenient reasoning tool. The obvious inference is that a context must be found in which the equation is a genuine problem-solving tool rather than part of the required background theory of the problem situation.

This hypothesis can be illustrated by means of the compelling examples that have been seen in this chapter that the *mediating role* of symbols - in helping students to treat the unknown as if known, to extend in a step-by-step fashion the range of problems amenable to algebraic strategies, to explore relationships between variables and to explore arithmetic identities - is a role that cannot just be "grafted on" as a final stage in the algebraic process (Sutherland, 1991).

Moreover, if one goes along with Hewitt (1985) in claiming that "Algebraic statements are only the proof that some algebra has already taken place." (p. 15), or with Lins (1990) in claiming that "the 'symbolic calculus' of algebra was but a consequence of the development of a body of knowledge that already embodied the calculus" (p. 95-6), then there may be a temptation to view symbolic algebra as merely grafted onto students' pre-symbolic algebraic understanding. The symbolic language would play no role in the development of that understanding.

3.4.14 Word Problems

It is considered that the case that representation difficulties can be addressed by experience of word problems has much merit, because such problems can provide a concern for algebraic symbolism. In simple terms: algebra can be useful. Kieran (1992) reports that Bell, Malone & Taylor (1987), for example, encouraged students to construct equations for problems such as:

Piles of Rocks

There are 3 piles of rocks. The second has two more than the first; and the third has 4 times as many as the first. There are 14 rocks in total. Find the number of rocks in each pile.

All students initially started with the first pile as x; but they were then encouraged to solve the problem using the second pile as x. When they wrote $x - 2$ and $4x - 2$ for the other two piles and found that the resulting equation produced a different result, a discussion ensued about the need for brackets. When they used the third pile as x, most students wrote $x \div 4 + 2 + x \div 4 + x = 14$, collected the $3x$ together and wondered what to do with the numbers. Thus concerns for both representation and transformation arise naturally from the situation, rather than artificially from the teacher.

Problems that are of clear practical, everyday relevance will often involve accessible success criteria, familiar relations and objects with easily visualised properties. Such characteristics improve the chances that a problem can be converted into a concern. It seems to be this "graspability" - not the realism of the context - that is crucial. The reality of a problem may entail messy issues that can actually detract from the target theories; and while experience of such messiness is sometimes helpful, Lins points out that it may not be helpful for promoting an algebraic approach. Conversely, we have seen how programming (an accessible, but hardly "realistic" activity) can be used to encourage the use of variables. Therefore, English & Halford (1995) suggest that "In solving algebraic word problems, students require guidance in forming an appropriate problem-situation model that informs and constrains the formal expressions required for solution." (p. 303). This suggestion would fit with the strategies suggested by Mayer (1981), Berger & Wilde (1987) and Chaiklin (1989), as described in section 2.5.3. Berger & Wilde note that "Novices are much more likely to stop after they have generated a list of value assignments, unable to see relationships inherent in the structure of the problem. A key advantage that experts hold is that they are familiar with a large number of problem forms." (p. 135). On the other hand, it is not at all clear, they argue, how such expertise is developed - it could be that symbols play a crucial mnemonic role in standard problem-situation models.

What formalisation and expressing generality have in common is that "pupils are usually asked to express a mathematical relationship in natural language *before* they are asked to express it in algebraic language." (Sutherland, 1995, p. 276). Moreover, apart from the authority of the teacher or textbook, "Beginning algebra students have no resources for interrogating the appropriateness of their algebraic construction." (Sutherland, 1993, p. 43). But there is "no evidence which suggests that expressing a mathematical relationship in natural language necessarily comes before being able to express it in symbols." (Sutherland, 1990, p. 170). As has been seen here from discussions of the use of automatic equation solvers, programming languages, spreadsheets and word problem representation software, "Work with computers is provoking a re-questioning of this dominant anti-symbol ideology. Algebra-like computer languages support pupils in their problem solving constructions." (Sutherland, 1995).

The reinterpretation of the activities described so far using Popperian psychology would suggest the following conclusions. There is a set of software tools that have a number of desirable characteristics. Each tool allows students to enter representations; it checks the syntactical legality of those representations according to some predefined conventions; it provides feedback that enables students to improve their strategic theories for the syntax of representations; it enables transformation of those representations to be made easier in some way (such as by automatically solving equations or supporting numerical trial-and-improvement), which enables students to put their representations of the problem situation to the test and thereby to improve their strategic theories for representing that problem situation.

In short: such "representation software" can allow the student to learn representation theories as a *by-product* of assisting the student in tackling word problems, by enabling the representation to function as a tool that addresses concerns rather than as an end in itself.

However, there are three additional characteristics, one or more of which each software tool lacks:

1. the use of standard algebraic notation and explicit equations in the representations allowed;
2. accessibility to students who have no prior knowledge of how to represent the situation using the predefined conventions;
3. the promotion of formal equation-solving.

The search for representation software that also incorporates these additional characteristics - a search that is the subject of the next section - can be designated by the following question:

> "Can we develop a school algebra culture in which pupils find a need for algebraic symbolism to express and explore their mathematical ideas?" (Sutherland, 1991, p. 46)

3.5 Equation Solving

Researchers refer to equation solution as being a "source of great mystery" (Griffin & Hirst, 1989, p. 4), to it requiring "extensive practice" (Geary, 1994, p. 125), to it being "not easily acquired" (Kieran, 1992, p. 402) and to it being not easily accessible by "spontaneous developments" from students' "initial grasp of operational algebraic behaviour" (Filloy & Rojano, 1989, p. 24).

Moreover, Cortés, Vergnaud & Kavafian (1990) note the lack of concern: "the algebraic treatment of equations is initially a response to the teacher's request. Pupils learn, for sure, but the introductory process is slow and rests entirely on the pupils' acceptance of the didactical contract." (p. 27).

3.5.1 Informal Methods

One way to start equation-solving is with students' own informal methods. Kieran (1992) notes Petitto's (1979) point about the lack of generality - but greater practical success - of informal methods. Meanwhile, Geary (1994) describes the work of Sweller *et al.* (1983):

> "asking students to simply solve for X, as is done in most classrooms, did not greatly influence their problem-solving approaches. Even after extensive practice, they still used problem-solving approaches that are commonly used by novices. However, more general goals, such as asking students

> to find different ways to solve the same problem, did lead to the use of problem-solving approaches typically used by experts." (Geary, 1994, p. 126)

For example, Demana & Leitzel (1988) continued the sequence of instruction based on tables (described in section 3.4.3) with questions such as: "Find w if $w^2 + 4w = 45$", where 45 would be in the table. This was followed by asking similar questions, but in which the numerical values would *not* be in the table. The students at this point tend to develop their own solving methods.

Booth (1984) similarly suggests that teachers, when introducing equations, should define the activity as that of "developing general equation-solving procedures which can be used to find the unknown value in a whole range of problems" (p. 94) rather than as that of 'finding x'. On the other hand, "Only when children become aware of the limitations of their own methods… will they be prepared to contemplate the value of the more formal methods which the teacher is attempting to teach." (p. 93).

Dickson (1989) identifies a number of informal ideas and methods that might stand in the way of learning a formal method, including the "letter as object" interpretation; an operational interpretation of the equals sign; the use of "counting on" rather than subtraction; the theory that different letters necessarily represent different values; and the theory that letters have to represent integers.

For example, Olivier (1988) discusses a student who did not permit equal values as a solution to the problem "If $a + b = 4$ what values of a and b will make the sentence true?". He was encouraged to construct the expression $4a + 3b$ to represent the total points scored by a team in a rugby match, when they scored a tries and b penalties. When he successfully used the expression to find the total points when a team scores 3 tries and 3 penalties, the inconsistency with his earlier response was pointed out. The result of this attempt to induce cognitive conflict resulted in confusion. Of 22 students with the misconception, cognitive conflict resulted in an even split between persistence in the misconception (for example insisting on excluding equal values in the rugby match), total confusion and successful remediation (for example realising that a and b could *both* equal 2).

3.5.2 Expressions

One of the difficulties with analysing proposals for teaching equation solving is the wide range of approaches and contexts. It does, however, seem common that (with the exception of simple open sentences for which the recall of number facts or the use of counting techniques are appropriate) simplification of expressions and substitution into expressions tend to precede formal solution of equations. Note this order in Kieran's account of introductory algebra:

> "Many first-year algebra courses begin with literal terms and their relation to numerical referents within the context of, first, algebraic expressions and, then, equations. After a brief period involving numerical substitution in both expressions and equations, the course generally continues with the properties of the different number systems, the simplification of expressions, and the solving of equations by formal methods. The manipulation and factoring of polynomial and rational expressions of varying degrees of complexity soon become a regular feature." (Kieran, 1992, p. 395)

And in SMP (1981):

1. Substitution of numbers into expressions.
2. Simplification of expressions.
3. Solving equations
   (i) using flow diagrams and trial-and-error;
   (ii) using flow diagrams and inversing;
   (iii) by operating on both sides of an equation using inverse operations.
4. Multiplying out brackets.
5. Solving inequalities using inverse operations.
6. Solving simultaneous equations.
7. Solving equations using graphs.
8. Solving inequalities using graphs.
9. Construction of a formula from (i) data; (ii) description of relations
10. Using formulae.
11. Changing the subject of a formula.
12. Finding gradients and equations of straight lines.

Does there lurk in this order of expressions and equations a belief that knowledge of an entity depends on knowledge of its components' behaviour? Do students therefore need to learn about variables before expressions before equations? Leitzel (1989) certainly asserts that "before students see algebraic equations, they need to have considerable experience with mathematical expressions arising from concrete problem situations." (p. 30).

Linchevski & Sfard (1991) contrast:

- thinking of an equation as a *propositional formula*, and of solving an equation as finding the truth set that makes the equation a true proposition, by applying elementary operations to the equation;

with

- thinking of an equation as *two computational procedures* on numbers, and of solving an equation as finding the inputs for which the procedures give the same results.

The first approach, they argue, can lead to a "pseudostructural" conception, with which students of an equation would be unable to relate elementary operations on equations to arithmetic operations. They would have failed to reify the "primary processes" and "Once the developmental chain has been broken… the process of learning is doomed to collapse" (p. 318). Without the abstract objects of expressions, students' understanding is instrumental, in Skemp's terminology. An illustration of this is given in an empirical study investigating whether students aged 15-17 used transformability of equations as a criterion of equivalence (rather than, presumably, "having the same solution set"):

> "The findings seem to reinforce the impression that for many respondents, an equation or inequality was nothing more than a string of symbols which can be manipulated according to certain arbitrary rules." (p. 323).

The framework for this research has questioned the value of generalising about "conceptions" in this way. Moreover, the study's definitions of "transformable" and "equivalent" are debatable.

Expressions can be made easier to understand: Thompson & Thompson (1987) used expression trees displayed on a computer screen with eight 7th-grade students (12-year-olds). Expressions and equations could be operated upon by a variety of algebraic identities. The students' success in transforming expressions is attributed by Kieran to an improved "ability to recognize the form of surface structure of an algebraic equation" (p. 436). Thompson & Thompson, on the other hand, point out that the students were not assessed outside the computer environment and so "It is quite conceivable that had these students been left to their own devices, they would have committed errors on paper and pencil that they learned not to make while using the computer." (p. 253). Even so, it does appear from the study that the availability of automatic transformations and a visual display of the structure of an equation assisted the exploratory learning of where these transformations might be appropriate. For example, one theory that was challenged was $(a + b) \times c \Rightarrow a + (b \times c)$. The students also seemed to prefer the tree displays to sentential strings of characters (Thompson, 1989). Moreover, Thompson suggests that students were not "bothered by the introduction of letters in expressions" because they only had to manipulate expressions, rather than evaluate them (p. 153).

Nevertheless, although it is usual to define equations in terms of expressions, it is certainly possible to introduce them via a context that does not feature algebra in its composition but can make use of algebra (i.e. a word problem). There is, moreover, a good reason for introducing solution before simplification: when equations and analytic operations are used to solve the problem "What numbers satisfy these conditions?", potential solutions can be checked against the original problem; whereas when expressions and a simplification calculus are used to solve the problem "What is the standard way of representing this arithmetic procedure?", potential solutions can only be checked for a sample of values, and are dependent on mathematicians' conventions for their accuracy.

So why does simplification precede methods of solution? Perhaps it is argued that the use of equations to find an unknown in a situation is dependent on having a standard method of solution; or that formal methods *require* a simplification calculus. Perhaps it is simply because of the contrast between the diversity of solution methods and the apparent simplicity of combining "like terms".

3.5.3 Substitution

Whatever the reason for teaching a simplification calculus before solution methods, when students are eventually taught how to solve equations, it is apparently common for trial-and-error substitution to precede the transposition and Leibniz methods (Kieran, 1992; Bernard & Cohen 1988; Zehavi, 1988).

Numerical computation was seen as providing insight into the order of operations and into simplifying expressions. Kieran (1992) suggests that "students who use substitution as an early equation-solving device… possess a more developed notion of the balance between left and right sides of an equation and of the equivalence role of the equals sign than do students who never use substitution as an equation-solving method… this awareness is helpful in successfully making the transition to the formal method of equation solving" (Kieran, 1992, p. 400). She proposes the use of computer-generated tables for solving equations (as produced by spreadsheets and some calculators) as one way of easing the memory demands of trial-and-error substitution. Thomas & Tall (1986, 1988 & 1989), Tall & Thomas (1991) and Graham & Thomas (1997) found evidence that exploring equivalent expressions using substitution (via BASIC, specialised software or calculators, for example) could assist in CSMS questions.

Dettori *et al.* (1995) similarly argue that spreadsheets can be used "to understand what it means to solve an equation, even before knowing what an equation is" (p. 272). "Tables are potentially dynamic, that is, columns and rows can be added" (p. 267) and so spreadsheets can be used to "make and to test conjectures about a function's trends". And the "spreadsheet can be useful to introduce the concept of generalisation of a problem and to learn to distinguish between variables and parameters." (p. 268).

However:

> "… the resolution approaches of algebra and spreadsheets are strongly different: in algebra the solution of a problem is found by formal manipulation of equations describing it, while with a spreadsheet successive numerical approximations must be performed until a numerical solution is reached. This basic discrepancy can even lead students to misunderstand what is algebra if they are told that, using a spreadsheet, they are learning algebra." (p. 265)

This is perhaps overstating the case, unless one takes a very restricted view of algebraic activity as *requiring* formal manipulations of equations. But Dettori *et al.* also make the point that the opportunity for numerical solution might actually discourage students from making an effort to manipulate equations. It ought to be pointed out that formal manipulation is an important skill, because problems involving multiple variables are not in general easily solved by trial-and-error: for example, $\{x + y = 100, 8x + 6y = 650\}$ is difficult to solve unless one manipulates the constraints so that one of the variables is calculated rather than varied. Even if we succeed in solving the equations without manipulation, we cannot guarantee with a spreadsheet that the solution obtained is the only one. Another example is $4x + 3y = 100$, for which the solution is an infinite set of pairs of values.

Nevertheless, Booth (1984) suggests that difficulties with algebra are actually founded on a lack of experience with explicitly considering method in arithmetic and with dealing with general objects. It could be that just investigating the effect of varying a cell value on the result of a spreadsheet formula could be enough to raise sufficient awareness of method and generality to improve students' theories for formal methods. Additionally, if the computer can free students from having to evaluate expressions, perhaps they will be more able to consider structural aspects of the situation. Even so, there is no evidence yet that representing situations using equations has been helped in this way.

Computer programs and substitution software such as CARAPACE can be used to find numerical answers, in a similar way as for spreadsheets; but again no use is made of formal methods. Sutherland (1988) explicitly avoids suggesting that programming in Logo will help pupils to solve algebraic equations. Moreover, Blume & Schoen (1988) concluded that programmers and non-programmers did not differ, when solving word problems in the "frequency or effectiveness of their use of variables and equations, or in the number of correct answers to problems." (p. 154). However, "Programmers also checked for and corrected more errors in their potential solution." (p. 153), which one might have expected to provide an advantage. But the susceptibility of the problems to algebra may be a factor here, because Blume & Schoen also noted that "programmers used systematic trial more frequently than non-programmers did." (*ibid.*).

3.5.4 Flowcharts

One approach studied by Dickson (1989) followed this sequence:

1. Representation of arithmetic procedures as flow diagrams, such as:

   $$2 \xrightarrow{\times 3} 6 \xrightarrow{+ 4} 10$$

2. Representation of those flow diagrams as expressions (for example $3x + 4$) or functions ($x \to 3x + 4$).
3. Representing "Think of a Number" type problems by finding the input of a given flow diagram from the output (using inverse operations).
4. Finding the operations missing from a flow diagram starting at x:

   $$x \xrightarrow{?} x + 4 \xrightarrow{?} 3(x + 4)$$

5. Finding the operations missing from a flow diagram ending in x:

   $$2(x + 3) \xrightarrow{?} x + 3 \xrightarrow{?} x$$

6. Solving equations using the "undoing" method mentioned in chapter 2. For example, solving $2(x + 3) = 18$ by finding the operations that reduce the expression $2(x + 3)$ to x (using a flow diagram) and then applying these operations to 18. The checking of an answer by substitution in $2(x + 3)$ or in its flow diagram can be emphasised. An alternative version of this step, as used in SMP (1981) is to write the flow diagram for $2(x + 3)$ and then reverse it to find the inverse operations.
7. Solving equations by finding the required inverse operations in the same way, but this time abandoning the flow diagrams for applying the operations to the isolated constant term and using standard algebraic notation. For example, for example $2(x + 3) = 18$; so $x + 3 = 18 \div 2$; so $x + 3 = 9$; so $x = 9 - 3$; so $x = 6$.

8. Solving equations, but omitting the flow diagrams when the students felt they did not need them.

Dickson found that brackets could cause a few difficulties, and that students tended to be unconfident in the transition from the context of functions to the idea of an equation and its solution. The students were not able to explicate the connections between the written symbols and the flow diagrams all that well. Five of the six students interviewed who had previously solved equations using a version of transposition had difficulty interpreting transposition in terms of flowcharts. Although they were able to use both transposition and flowchart methods at the end of the series of lessons, only one of the five could use flowcharts three months later. The sixth pupil interviewed used a slightly modified form of the flowchart method successfully. Dickson claims that the flowchart approach is "more mathematically sound" (p. 188) than transposition, but it is not clear why.

One advantage of this sort of approach is that students could, in principle, solve a very wide range of equations without further assistance; for example $(x + 11)/6 = 1$ or $2(x/5 + 6) = 14$. On the other hand, it relies on there being an isolated number on one side of the equation.

Kieran quotes research by Whitman (1976) with six US 7^{th}-grade classes comparing the cover-up and formal methods, which found that "students who learned to solve equations by means of only the cover-up method performed better than those who learned both ways in close proximity, whereas students who learned to solve equations only formally performed worse than those who learned both techniques." (Kieran, 1992, p. 400). Kieran infers from this that "the students who had been taught to solve equations by the formal method alone were not conceptually prepared to operate on equations as mathematical objects with formal, structural operations" (*ibid.*). On the other hand, the result surely suggests the possibility that they had not learned the formal method properly. A proposed activity to promote a formal method would enable a re-test of Whitman's result.

3.5.5 Arithmetic Identities

The method of Herscovics & Kieran (1980) for using arithmetic identities to motivate the idea of "solution" may be useful in introducing operations on equations in a "natural" way. For example: suppose we start with $2 \times 5 = 10$. Students are asked "What happens if 7 is added on the right hand side?", which hopefully elicits a reply along the lines of "It's no longer an identity". They are then asked to make it an identity again by using only addition, and hence we have the idea of operating on an equation. However, one has to be careful with the order of operations: if we start with $4 + 5 = 9$ and multiply the right hand side by 3, students sometimes conclude $4 + 5 \times 3 = 9 \times 3$ instead of $(4 + 5) \times 3 = 9 \times 3$.

When numbers are hidden, concern to solve arises naturally: "The process of hiding a number seemed to carry inherently the reverse process of uncovering." (p. 576) and the idea of multiple solutions emerged easily. For example, hiding the 2s in $3 + 4 = 10 \div 2 + 2$ might lead one to realise that 5 could also work. The idea of an infinite number of solutions comes fairly easily: hide both 3s in $2 + 3 = 3 + 2$.

To solve $7x + 48 = 139$, Herscovics & Kieran propose writing the identity $7 \times \boxed{13} + 48 = 139$ next to the equation, and then asking "What must be done to the 139 to get back to 13?" (p. 578). This then focuses on the method, rather than the answer.

So for example:

$$\begin{array}{rcl} 7 \times \boxed{13} + 48 = & 139 \\ -48 = & -48 \\ 7 \times \boxed{13} = & 91 \end{array} \qquad \begin{array}{rcl} 7 \times x + 48 = & 139 \\ -48 = & -48 \\ 7 \times x = & 91 \end{array}$$

Students' attention is drawn to the fact that you have to subtract 48 on *both* sides, because otherwise it would not remain an arithmetic identity. The – 48 = – 48 notation saves re-writing the equation and reminds one to operate on both sides.

$$\frac{7 \times \boxed{13}}{7} = \frac{91}{7} \qquad \frac{7 \times x}{7} = \frac{91}{7}$$

$$\boxed{13} = 13 \qquad x = 13$$

This approach was tried in individual interviews with 6 students of varying abilities in grades 7 and 8 in different schools. Post-tests 6 weeks later showed that students had retained a clear understanding of arithmetic identities, equations and the justification of the algebraic rules.

However, Kieran (1988a) reports a study in which she used this activity with "six average ability 12-year-olds who had not had any previous algebra instruction" (p. 438). She found that "those students who had initially preferred inversing… were in general unable to make sense of the solving procedure being taught, that is, performing the same operation on both sides of an algebraic equation." (p. 438) and, in fact, this undoing method "seemed to work against them" (Kieran, 1992, p. 401) when an unknown appeared more than once. Those who had initially preferred trial-and-error substitution did not have the same problem. She describes the relationship between left-and right-hand expressions of equations as a "cornerstone of much of the algebra instruction taking place." (Kieran, 1988a, p. 439) and yet "it has been found that for some students, teaching methods based on this aspect of the structure of equations often do not succeed. For these students, who tend to view the right side of an equation as the answer and who prefer to solve equations by transposing, the equation is simply not seen as a balance between right and left sides, nor as a structure that is operated on symmetrically." (p. 439). As a result of these findings, Kieran (1992) proposes substitution as an "intuitive basis" for structural solving methods as opposed to inversing. Although inversing "seems much closer to the problem-solving methods used in arithmetic" (p. 401), it "appears to encourage the learner to bypass the algebraic symbolism rather than deal directly with the equation as a structural object." (p. 401).

Moreover, the above method does not offer a mechanism for operating on unknowns; only on numbers. The method runs therefore into difficulties when there are unknowns on both sides, just as for the flowchart method. Filloy & Rojano (1989) have investigated such equations, and found that students tend to resort to trial-and-error methods rather than attempting to operate on the unknowns. "Suitable interventions from a teacher at the point of transition may be crucial for students learning algebra for the first time." (p. 19) Is this always true? Could there be an activity that inspires students to operate on the unknown, rather than to use inverse operations or trial-and-error?

Herscovics & Linchevski (1992), reasoning that operations on the equation as an object cause difficulty, attempted instruction in which a term is decomposed into a sum or difference with the aim of cancelling identical terms on both sides of the equation. So, for example, $5n + 41 = 8n + 5$ becomes $5n + 41 = 5n + 3n + 5$. Unfortunately, students became confused when decomposing a term into a difference.

3.5.6 Standard Forms

Dickson (1989) describes a sequence of activities used by one teacher, which has striking similarities to the typologies of equation that Radford (1995) describes mediaeval mathematicians as using:

1. The class was shown that equations like $3x = 20$ and $25x = 12$ could be solved by dividing the constant term by the coefficient of the unknown. Students practised solving many of these types of equations, using a variety of numbers and letters. This made the equation type $ax = n$ familiar.
2. A justification for (1) along the lines of "What you do to one side you have to do to the other" was provided. For example, to justify the move from $3x = 20$ to $x = 20/3$, it might be said: "You divided $3x$ by 3 to get x. So you had to divide the 20 by 3. This gives you x.". However, based on the classroom observation, Dickson would dispute the existence of this step.
3. The teacher now told the class that equations of the type $ax + b = n$ could be solved by reducing them to the familiar form $ax = n$. "So, for example, he described the equation $3p - 2 = 7$ as having something 'funny', 'wrong', 'different about it', and continued 'I'm going to get shot of this minus two. I'm going to lose it… How can I get rid of that minus two?'." (p. 175). The idea then presented was of adding 2 to each side. Students practised such problems.
4. A justification for both (1) *and* (3) along the lines of "If you want to get rid of something from one side, you can put it on the other side so long as you change the sign." was now provided.
5. Finally, the students practised representing "Think of a Number" situations using equations to find the number. For example: "When a certain number is multiplied by 6, and 9 is subtracted, the result is the same as when it is multiplied by 4 and 11 added. Find the number."

This method is basically one of training the students to recognise standard forms, and supplying them with appropriate procedures to apply to those forms. Justifications are used to tie the forms and procedures into existing ideas to make recall easier.

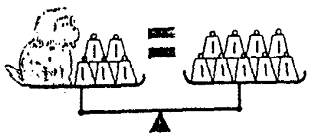
Dickson highlights, however, a number of difficulties that students had with representation and interpretation after this sequence. There is little evidence, nevertheless, that these difficulties were *caused* by the sequence. Moreover, nearly all the students (11-12 year-olds in a top set of three) were successfully able to solve equations of the forms they had practised. There were difficulties among some students, however, with equations containing unknowns on both sides.

Interestingly, although the meta-algebraic justifications were used by the teacher to introduce and reinforce new techniques and forms of equations, the students seemed not to be able to repeat these justifications later to the researchers, even when the techniques were successful. Note also that, although the teacher referred to a seesaw and being "out of balance" in his justification, there are two ideas here that could potentially become confused. There is the idea of "doing the same thing to both sides" which is the point of using the balance models discussed in chapter 2 (and below). But the teacher also seemed to refer to the equation $3p = 7$ as being in the balanced form, and the equation $3p - 2 = 7$ as being "out of balance": the minus two is "throwing it all out of balance". In early empirical work for this research, some students referred to an equation "as a balance" because "what you do to one side, you've got to do to the other"; others referred to an equation "as a balance" because "you've got to balance it". Those students that referred to both types of balance ideas did not seem to be confused by it in interview; and it is possible that they would not be confused by it when solving equations, because the role of these justifications or stories is largely in binding together apparently different techniques *after* they have been thoroughly automated. This binding is a preamble to new forms or techniques, but it does not *necessarily* interfere with those already learned.

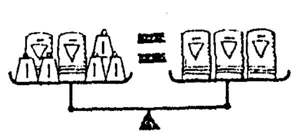
It is a moot point whether, in operating on equations, including a step with the operation unevaluated is an improvement. On the one hand, the additional transcription errors introduced by this method have been cited many times as being a major drawback; on the other hand, Cortés, Vergnaud & Kavafian (1990), for example, suggest that it "allows the pupils more easily to check their work; it permits the construction of a script-algorithm which is used to provide guidance in the very beginning" (p. 33).

3.5.7 Concrete Models

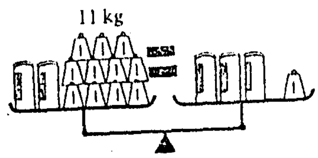
Soh (1995) studied using the Pirie-Kieren model (Pirie & Kieren, 1994) two 11-year-olds tackling the balance puzzles that SMP (1983) used to introduce equations:



Stage 1: Single object on only one side of the scales



Stage 2: Weights on only one side of the scales



Stage 3: Weights and objects on both sides

The textbook intended to teach pupils to solve a balance puzzle by removing identical objects or known weights from both sides of the balance. One of the students - called Dale - grasped the weights strategy for "Stage 1" puzzles (above), and the objects strategy for the Stage 2 puzzles, but was unable to combine them for the Stage 3 puzzles. Instead, she took off weights and objects from either side separately so as to be *left* with the same objects and weights on each side. The intended purpose of the method - simplifying the puzzle without changing the equality - was lost. It seems possible that she was using a matching strategy which became unwieldy, and would also be far more difficult to translate into operations on equations than a subtraction strategy would be. But when the textbook insisted on cancellation, she showed great anxiety about the untidiness that resulted from the repeated drawing of the balance puzzles after each step.

Unlike Dale, the other student Ken could not work out what to remove in any of the Stages, nor what to do with the objects and weights left on the balance. On video, the tension in these children is noticeable. Ken found that the hedgehog in one puzzle apparently weighed more than the dog in a previous puzzle; and thereafter in the work he seemed to feel that his common-sense was out of place. He would take off just two objects from both sides because that is what the example showed, or because that is what worked last time, even when more could have been taken off. At Stage 3, he developed the strategy of dividing the total weight in the picture by the total number of objects in the picture. If this did not produce an integer, he re-read the textbook example (if this happened to Dale, she tended to discard the fractional part of the answer). Soh writes, "Although this strategy was successful in only one out of ten questions, he was not deterred from using it in every question that came up.". Again, the purpose of the taught method was unclear to him. This child - who if faced with a real balance scale would probably clear all the objects off, put the unknown weight on one side and start putting known weights on the other until equilibrium - seemed to see little connection between the exercises he had to do and problems he might care about. The answer was right or wrong, but that had nothing to do with the relationship presented on the scales, and everything to do with accurate adherence to the unwieldy rigmarole (as far as he was concerned) that the textbook wanted him to perform.

Schliemann *et al.* (1992) investigated Brazilian children's ability to find unknown weights on marketplace balance scales. 75 children were involved, between the ages of 5 and 12. When unknowns appeared on just one side of the scale, only 5 and 6 year olds had difficulties. When unknowns were on both sides, the 5 and 6 year olds could not solve the puzzles at all; older children, meanwhile, tended to use trial-and-error. When the puzzles got harder, the children were shown the cancelling strategy. The 11 and 12 year olds were more inclined than the younger subjects to use it. With a subsequent series of questions aimed at discovering whether they had grasped the Leibniz theory, it was the 11 and 12 year olds who tended to give justifications along the lines of "They will still balance, because you are taking the same amount from each side.", whereas the younger children attempted to compute the weights left on the scales. This study is not given here as evidence for a developmental stage theory (it is not clear whether those who gave computational justifications failed to give an explanation involving the Leibniz theory because they had not grasped it, or because they were deliberately seeking to illustrate what they considered obvious to an interviewer who apparently did not find it obvious); rather it is given to suggest that spontaneous cancellation is rare for children at least up to the age of 12, and that even after prompting the strategy is not necessarily readily adopted. Of course, it is likely that children from different backgrounds, with different experiences, and under different circumstances may contradict this suggestion. The point is that one should not *expect* the cancellation strategy to be part of every child's theoretical arsenal.

Austin & Vollrath (1989) describe how balances, washers and containers (which weigh, when empty, the same as a washer) can be used to represent equations. By removing paired washers, the number of washers in each container can be found, and the solution steps are then represented using symbols. Equations such as $2(x + 1) = 8$ can be represented by using two piles on the left-hand side - each containing one washer and one container. Two washers are removed from each side; and it can then be checked that each container contains 3 washers. A balance set up to represent $2(x + 1) = 2x + 2$ confused many students, and only after putting different numbers of washers in the containers did students realise that some equations have many solutions. Austin & Vollrath also describe two-balance problems using different coloured containers. "Whereas students readily accept that the colour of the containers does not affect the solution, some students think that the letter used to represent a variable affects its value" (p. 609) - recall Wagner (1981). They suggest that using physical objects in this way can make it easier for some students "to understand equations and why each transformation is used in solutions." (p. 611).

However, Lins (1992) uses his characterisation of algebraic thinking to conclude that the use of scale-balances in learning algebra contribute "to the constitution of obstacles to the development of an algebraic mode of thinking"; this is "not only for very quickly becoming a complex net of what are in effect different models, but also for not fostering a frame of mind adequate for the development of an *algebraic mode of thinking*." (p. 209). It would be useful to try to test this claim by finding activities that use the balance model extensively and exploring the extent to which students were able to adopt an algebraic approach.

Moreover, Mordant (1993) suggests that the balance model is a "very poor substitute" (p. 22) for starting with expressions, because quadratics cannot be solve by physical operations with a scale pan and because "a scale-pan presents the student with an exceptionally poor notion of the algebraic expression". A related point that should be made here is that the model does not in itself promote *conventional* algebraic notation rather than idiosyncratic or syncopated representations. After all, representing a balance as "4a5:3a10" might be rather confusing for others, but it could be good enough for one's own solution method. How do students know what constitutes an "appropriate" form of algebraic representation? "Appropriate" for what?

A teacher in Dickson (1989) used a balance model followed by practice in using the Leibniz method. When later interviewed by the researcher, the students did not seem have a relational view of the equals sign, a recognition that one does not need to work with specific values, improved representational or substitution abilities, a numerical interpretation of letters or a concern to use symbols. Moreover, a concern for the Leibniz method appeared to have been grasped by only a few students, and the ability to carry them out by even fewer. Dickson concludes that the "Think of a Number" type questions would, in relation to five students under scrutiny "perhaps have been a better starting point for their introduction to algebra in general and equations in particular. The container mode of representation could perhaps be retained indefinitely to ensure the appropriate development of the notion of a variable." (p. 166). However, recall that the "Think of a Number" type problem introduces inverse operations on numbers but not operations on unknowns or equations; secondly note that the students were only *shown* the balance model - they did not actually practise solving balance problems; and thirdly it is important to realise that the balance model used in the study involved balancing the values of symbols (for example $b + 2$ on one side and 5 on the other) rather than the more familiar idea of weights.

Another teacher balanced numbers of objects (loose apples, and apples in apparently weightless boxes), but again the link between practice in erasing extraneous objects and subtracting terms from algebraic representations does not seem to have been sustained, as evidenced by interviews some months later. In the delayed interviews, it was noticed that "On the whole the pupils did not employ the formal written approach for solving equations" (p. 171), their approaches seeming to "vary according to the type of equation under consideration and the mode of presentation. The balance diagrams usually elicited a 'crossing out' or 'grouping / matching' strategy and the algebraic symbolisation usually invited a mental version of a 'formal' approach or a trial and error strategy." (*ibid*.).

Booth (1987) concludes from Dickson's study that "'concrete or ideographic approaches, though designed to help children gain in understanding of the formal procedures, may be unsuccessful in doing so if children never see the connection between the two.". She also notes that it was only the students "who had perhaps least need of the ideographic approach in the first place" who appeared to see and make use of this connection. Kaput (1987) interprets the study as emphasising that concrete models are weak because "of the inherent particularity of such models a particularity which runs entirely opposite to the inherent generality and abstractness of algebraic statements." (p. 352).

There are other concrete models that could be described here, each of which have their various advantages and limitations. For example, Bruce *et al.* (1993) describe an activity using cups and beans; which enables representation of equations such as $4x - 3 = 2$ and $3(5 + x) = 15 + 3x$; but does not easily extend to powers. Vollrath & Austin (1989) describe the use of line segments.

F. M. Thompson in Coxford (1988) shows how a pictorial representation of an equation using coloured stickers can assist solving.

But Booker (1987) has further reservations:

> "While the use of materials to represent [unknown] values appears attractive by analogy with the number situation, in reality the material does not serve as a forerunner to the use of letters; rather letters label the material which is manipulated. It also leaves the question of why these letters should themselves be the object of mathematical manipulation unanswered." (p. 278).

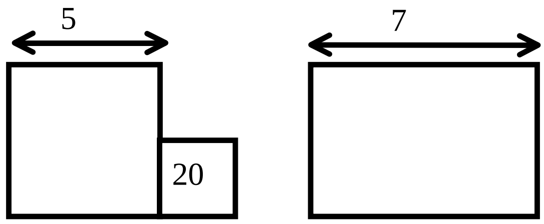
> "Only when the development of a generalised arithmetic has established the need for and power of algebraic symbols can algebra be extended to a topic in its own right and meaningful procedures for manipulating the symbols be considered." (p. 279)

This lack of generalised arithmetic is crucial:

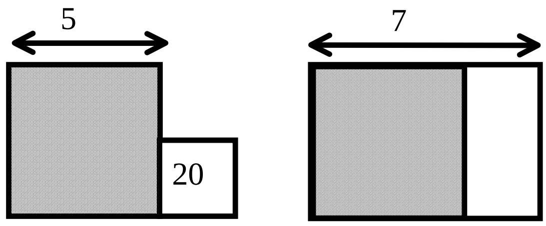
> "Algebraic symbolism should be introduced from the very beginning in situations in which students can appreciate how empowering symbols can be in expressing generalisations and justifications of arithmetical phenomena... By displaying structure, algebraic symbols are not introduced as formal and meaningless entities with which to juggle, but as powerful ways to solve and understand problems, and to communicate about them." (Arcavi, 1994, p. 33)

Nevertheless, according to Nolder (1991), although the mathematics can be taught without metaphors, imagery is "something concrete and familiar to help them to understand an unfamiliar, abstract idea... linking a new concept to the learner's past experience. ... the teacher aims to make that piece of mathematics more secure and hopes that the metaphor by its novelty may enhance memorability." (p. 108). Similarly, in early empirical work for this research, A-Level students claimed about the equation as a balance something along the lines of "It helps you to make sense of 'What you do to one side, you have to do to the other'.". This is a compelling reason for at least attempting to find an appropriate metaphor. On the other hand, other usages were also met: "The left-hand side is equal in value to the right-hand-side."; "If you take the same thing from each side, the sides are still equal."; "If you take the same thing from each side, the answer is still the same."; and "Balancing is taking equal things from each side.". Also recall the teacher in Dickson (1989) talked about the -2 in $3x-2=7$ "throwing it all out of balance". Maybe the differences between these usages are not large, but they do mean that there is at least the potential for confusion when invoking the metaphor of a balance.

Moreover, Filloy & Rojano (1989) state their belief that "there are theoretical reasons for believing that a semantic approach to learning algebra is more likely to lead to good algebraic performance in later years than a purely syntactic one." (p.20). In other words, that modelling analytic operations, in some concrete, familiar context - in which "the first elements of an algebraic syntax are constructed on the basis of the behaviour of the model" (p. 20) is more effective in the long run than learning the syntactic rules by rote. They presented students with concrete models involving the use of either balances or areas. For example, the following picture models the equation $5x + 20 = 7x$:



Solution follows by comparing areas:



So we have a new equation $2x = 20$ which can be solved by inspection - simple reasoning or recall of multiplication facts.

Filloy & Rojano noticed a "temporary loss of previous abilities, coupled with behaviours fixated on the models." (p. 21). But an alternative interpretation could be that the students expected that they had to *use the model* to solve for example $3x + 30 = 9$, even if they already knew other methods. The researchers also noted both successful and unsuccessful attempts to transfer operations within the model to operations on the equation. For example, for $15x + 13 = 16x$, a student says "$16x$ minus $15x$ equals 1. So 1 and 13 equals 14.". The extent to which students seemed inclined to break away from the model varied from student to student, with preferences for solutions methods "ranging from the most operative and algorithmic to the most semantic and analytic." (p. 23).

They also point out some limitations of the models, especially the "lack of meaning" for negative solutions; and hinted that using the geometric model might lead to the obstructive expectation that integer answers should be obtainable. In comparing the models, they explained that "With the balance model, the step from solving the type $Ax + B = Cx$ [where the capital letters are given] to solving the type $Ax + B = Cx + D$ is a small one since 'iterated cancellation' will reduce both types to arithmetic equations." (p. 24); this may not be true for the geometric model. On the other hand, "… the transition to types $Ax - B = Cx$ and $Ax - B = Cx + D$, while impossible for the balance model can be accomplished in the geometric model with the introduction of the operation of removing areas when negative terms are involved. This extension does not do violence to the semantics of the model." (p. 24). For a concrete model to be successful, and enable students to abstract a rationale for the operations without undue reliance on the particularities of the model, there must be a drive to detach aspects of the semantics of the model. Unfortunately, mastery of translation of an equation into a concrete model can inhibit this detachment of the semantics, and hence "delay the construction of an algebraic syntax" (p. 25). And in detaching the semantics, essential elements of the solution process may be lost.

In a similar vein, Carraher & Schliemann (1987) suggest with regard to 25 market vendors using weighing scales that "transference from the practical setting to a hypothetical one with unknowns on only one side of the scale was observed in all cases. … transference to situations with two unknowns is observed less frequently" (p. 294). Whether this is indeed attributable to "transference" is questionable, without controls.

Boulton-Lewis *et al.* (1997) argue that use of different, unfamiliar models that are not explicitly related to the target theories are confusing. They found that none of a class of students who had been taught a containers and objects approach to solving linear equations spontaneously used the method - the majority used inversing. Only four out of 21 were able to reproduce the concrete method. The researchers suggest that this supports the view that equations must be understood in terms of sequences of operations rather than a "structural" concrete approach, because the latter entail a "heavy cognitive load" (p. 190).

Margolinas (1991), who comes close to the Popperian view of mathematical objects being dependent on problems, as opposed to existing in distinctive complementary dualistic forms (cp. Gray & Tall, 1993), describes, with respect to the solving of equations, some of the subtlety in the relationship between the result of the solution and the answer to the problem. She shows how such subtlety can be lost in textbook or classroom discussions; and explains the consequential difficulty in conveying the notion of algebra as a tool. However, pessimism with respect to concrete situations in this regard could neglect their potential analogical role. On the other hand, English & Halford (1995) appear to share this pessimism: "concrete analogs do not appear as effective and as versatile in teaching algebraic concepts as they are in promoting arithmetical understanding." (p. 240).

Meanwhile, Kieran (1988a) describes a study carried out by O'Brien (1980) with two groups of 3rd-year high school students. One group was taught using concrete materials; the other was taught using transposition rules. The latter group became the more proficient equation solvers.

Moreover, Herscovics & Kieran (1980) point to the greater generality of arithmetic identities compared to the balance scale:

> "The physical limitations involved with the scale can be avoided by the use of arithmetic identities, for these are an arithmetic representation of the concept of equilibrium and not subject to physical restrictions." (p. 577)

They state that the balance model "does not lend itself readily to addition and subtraction of arbitrary rational numbers nor to the more complex operations of multiplication and division" (p. 577).

Dickson seems to prefer the flowchart approach - despite its limitation to equations with isolated constant terms - to the balance model. Her reasons are that the latter does "not readily represent situations involving subtractions" (p. 189); it leads students into talking in terms of "one side becoming lighter or heavier"; it ignores the actual mass of objects and containers; it cannot incorporate powers or roots; it sometimes bears no resemblance to the formal approach used by students; it promotes (through the 'getting rid of' actions) "an interpretation based on superficial symbol manipulation and a rote relationship to the 'change the side, change the sign' rules'… [and hence] a diminishing need to understanding the underlying rationale and process" (p. 190); and it is "less mathematically sound" than the flow diagram approach.

> "[The student was] perturbed by the positioning of the cans. She knew from practical experience that this would not work on a real balance because of the need to position weights very carefully to preserve equilibrium. This discrepancy distracted the pupil from the concept being developed and it required skilful handling by the teacher to remove the confusion." (Nolder, 1991, p. 110).

However, the researchers engaged in the "Children's Mathematical Frameworks" research (CMF), of which Dickson (1989) is a part, describe the students' success rate in learning formal methods for solving equations (by whatever instructional approach) as "extremely low" (p. 223) and mostly attribute this to the "difficulty that children have in linking the formal rule to that used prior to the teaching or in the pre-formalisation work" (p. 222). Moreover, Schliemann *et al.* (1992) note about the Leibniz theory of equality that "Mathematical axioms are sometimes so convincingly obvious that it is tempting to treat them as a priori truths known to anyone with the power of reason." (p. 298). My personal experience of observing students working on the balance puzzles from SMP (1983) suggests that many children never really appreciate the "necessary truth" of the Leibniz theory. If it is never really grasped, then not only will the link that CMF describes never be made, but the whole purpose of the balance approach is destroyed.

Kaput (1987) suggests that work using concrete models…

> "would have vastly different outcomes (1) if their concrete models had been instantiated in the computer medium, a medium much more congenial to variation and hence conceptual generalisation, and even more importantly, (2) if those models were then *actively linked to the associated algebraic formalisms*, so that transformations of a concrete model would have salient consequences in its formal counterpart, and vice-versa." (p. 352).

However, this assumes (as made clear earlier in Kaput's paper) that it is the making and meaningfulness of the links between different representational systems that should constitute the imperative of a substantive attempt to improve algebraic theories and concerns; and, moreover, that it is *generality* that is lacking from concrete models. (This is what Kaput calls the "representational perspective".) If, on the other hand, the imperative is a conjectural, problem-solving environment which allows students to grasp transformational and representational problems (regardless of the "inherent generality and abstractness" of the algebraic statements involved), the simultaneous display of transformations' effects on concrete models and symbolism might not be crucial. But in this case, some other sort of feedback is required, to support - as Edwards (1991) puts it - "children's 'debugging' of their own solutions, and of their own conceptual models of the mathematical entities instantiated in the environment.".

For example, Feurzeig (1986) and Roberts *et al.* (1989) provide a computerised balance model ("The Marble Bag Balance Laboratory") in which students can operate on one side or both. The balance tilts if the student does something that does not maintain equivalence. However, it does not attempt to overcome the physical limitations of the model, to demonstrate the value of the Leibniz strategy over informal strategies, or to transfer the strategy to new situations.

Can a software environment be created that help students to grasp the Leibniz method? Can the physical limitations of concrete models that Herscovics & Kieran, Filloy & Rojano, and Dickson describe be minimised by setting the problem situation inside such a software environment? Such a use of technology relates to the allusion of Kaput (1987) to "the need to focus research on the possible learning environments of the future rather than those of the past - to take an inherent difficulty… and then build and test new teaching and learning environments that respond to that difficulty." (p. 352). I suggest that the finding of an unknown weight on a balance scale may be slightly easier to make a concern than learning rules governing expressions, easier than finding a missing number in an arithmetic identity, and easier than using inverse operations in a flow diagram. I suggest that such a software environment could allow students to make the transition from informal to formal methods *for themselves*, and this would be done without "superficial" or "rote" symbol manipulation. And I suggest that generalisability to subtraction, powers and roots is irrelevant while such large numbers of children seem to spend much of their schooling worrying about simpler equations. However, if one were to follow through such suggestions, what would this software environment look like? What facilities would it need to provide? How exactly would it *help* students to learn the Leibniz method?

3.5.8 Automatic Equation Solvers

The work of Heid (1990) and her colleagues suggests that when students on Computer Intensive Algebra courses are taught paper-based transformation skills, it takes 6 to 8 weeks for them to be able to perform as well as "their counterparts in year-long traditional courses." (p. 655). Is this because of the more realistic contexts, or because CIA allows a focus on exploring (rather than acquiring) methods? However, while performance in word problems is undoubtedly enhanced by the use of an equation solver, is the ability to solve transformation problems maintained when the technology is withdrawn?

Hunter *et al.* (1995) carried out a study in which 14-15 year old students studied quadratic functions; some classes with Derive, parallel classes without. All classes plotted functions, found minima and intercepts, solved equations, and expanded and factorised expressions. Pre- and post-tests based on CSMS showed little improvement attributable to software usage. In an additional post-test based on the quadratic content, in which computers were not allowed, the control groups did better than the experimental groups. Those using computers were allowed to use them on this test two weeks later: their scores improved on the pencil and paper sitting; but in one school, the score was still lower than the control group.

Because automatic equation solvers like Derive remove the *choice* element from transformation, students do not have the opportunity to discover that the power of algebra is dependent on the existence of standard transformation rules. Without this discovery, the need for standardised representation does not become manifest. Moreover, the opportunity to learn those rules by conjectural means does not arise.

3.5.9 Interactive Manipulators

Fey (1989b) notes that "Use of computer symbol manipulation or computer algebra systems as tools for learning about symbol manipulation itself is an almost totally unstudied area. But imagine the discoveries that students could make if they could call on an algebraic assistant to test the effects of various operations on a planned series of example expressions." (p. 254).

The program GED devised by McArthur (1987) and McArthur, Stasz & Hotta (1987) provides such a focus on solution strategy. It allows students to enter the equation for each step in a solution procedure, to go back to a previous equation and try alternative procedure, and to see a record of all the solution paths in a "tree" display. The program can indicate the correctness and appropriateness of steps, suggest a new step, and elaborate the details of suggested steps. Thompson (1989) suggests that it is the screen record of the steps taken that is of most value, because the path can then be examined as an object in itself, fostering an appreciation of the relationships between the tactical choice of operator and the strategic goals.

Larkin (1989), meanwhile proposes a program to show the structure of equations using various levels of tiles (p. 132) during the solution process, which she suggests might not only improve transformation strategies, but also the use of hierarchical representation.

The "Algebra Workbench" (Roberts *et al.*, 1989), "Algebraland" (Brown, 1985) and "EXPRESSIONS" (Thompson & Thompson, 1987) allow students to enter an equation and then have the computer carry out whatever formal operations are requested. For example, solving equations such as $2 + x = x/17$ by hand involves transformations such as "subtract x from both sides" or "take the x over to the other side". Transformations such as $x/17 - x$ must then be known. However, in such a tutor the *choice* to initiate a transformation is made by the student, while the *performance* of that transformation can be executed by the computer (which is instantaneous, responsive and reliable, unlike a textbook, a busy teacher or a fellow student). This allows students to focus on the solution strategy rather than on the tactics of performing an operation. EXPRESSIONS converts each equation into a tree display; Algebraland shows solution paths in a tree; while the Algebra Workbench shows solution paths and can suggest a next step. APLUSIX (Nicaud, 1992) offers similar facilities for the factorising of polynomials.

A non-computer attempt to achieve the separation of choice and execution is described by Shavelson, Webb, Stasz & McArthur (1988), and is characterised by Kieran (1992) as the teacher reminding students after each tactical "lower-level transformation" of the role of the transformation in fulfilling strategic "higher-level reasoning" purposes. Nevertheless, a reminder from the teacher to pay attention to "global" features of an equation is unlikely to be very effective with students who are struggling with the "local" difficulties of arithmetic and performance of algebraic transformations, and are rather unconfident about what those global features might be. Hence, the computer is extremely useful in taking care of these local difficulties.

Such interactive manipulators can provide vital feedback: Sleeman (1986) concludes that "Few pupils have evolved mechanisms by which they can verify whether a proposed algorithm is feasible." (p. 52); but studies such as Lewis, Milson & Anderson (1987), Sleeman *et al.* (1989) and Anderson, Boyle, Corbett & Lewis (1990) seem to suggest that a learner can correct his or her strategic theories in the light of experience, as an AI system's "self-diagnostic routine" might attempt to "debug" its reasoning procedures.

The question also arises: can allowing students to focus on the strategic choice of operation actually improve *performance* of that operation? It would be interesting to examine the extent to which simplification theories are improved as a by-product of separating choice and performance of solution steps.

However, in all of these activities, the assumption is made that students already have a concern for equation-solving and at least tentative strategic theories for formal operations. They also rely on a teacher or textbook choosing appropriate tasks for the student to tackle within the environment. They do not therefore seem *by themselves* appropriate as a first introduction to algebra. Nevertheless, Thompson & Thompson show how the EXPRESSIONS program can be used to explore arithmetic identities before introducing letters, thus increasing the accessibility of the Leibniz method. The "Marble Bag Balance Laboratory" of Roberts *et al.* (1989) comes close, in that it provides a context accessible to novice algebra students, but does not appear to introduce the transformations progressively. When the laboratory was piloted, quantitative and qualitative evaluations "suggested that the sixth-grade students could learn algebra concepts" (p. 265) in the context of the program. Although any improvements could not be attributed to use of the laboratory, since other software and Logo programming also formed part of the curriculum, testing appeared to indicate, in any case, that children had difficulty in transferring what they had learned within the laboratory to the "more traditional content of an algebra curriculum" (p. 265).

3.5.10 Tutors

Lewis, Milson & Anderson (1987) and Anderson, Boyle, Corbett & Lewis (1990) provide a computer algebra tutor underlying which "is an ideal model of how students should solve the [transformation] problems and a model of how students err." (1990, p. 29). They also detail some of the tensions underlying the construction of such software. The tutor characterises skills acquisition in terms of symbolic manipulation goals (for example "collect constants", "clean-up", "distribute"); it is recursive in decomposing problems to primitive sub-goals; and lower-level goals are optional for the student if he or she can carry out higher-level goals. However, it should be noted that at this stage it does not seek to make the top level goals such as solving equations into concerns; and it does not introduce algebraic ideas *piecemeal*. Even so, its designers are clearly committed to increase the effectiveness of computer-based instruction from the typical "less than half of a standard deviation of improvement" (p. 43) towards the 2 standard deviations of improvement over standard classroom instruction that they cite as being common for human tutors. The success of such a mission would raise many issues; but from a cognitive point of view, it would be most interesting to find out in what ways the explicit models of student skills, errors and learning assumed by the software can be improved. The tutor does not allow students to pursue "non-optimal" problem-solving paths, for example.

Roberts *et al.* (1989) raise some further issues about such tutors:

> "Should the software *know* the best solution through a problem at any point along the way? Should the software allow students to make mistakes? How small should the steps be when the computer presents a solution (when multiplying both sides of an equation to eliminate fractions should the cancellation be explicit?)" (p. 263)

Moreover, how responsive and adaptive to students' thinking does a so-called "intelligent tutoring system" have to be? Sleeman *et al.* (1989) investigated "model-based remediation" in the context of learning algebra using human tutors. This is when a model of students' thinking is developed, errors are pointed out and the consequences of the errors are discussed. This may include *showing* why an error is an error, by comparing the strategy with a rule such as "whatever you do to one side you must do to the other" or "if you want to undo add, use subtract" or "try to get all the unknowns on one side"; *asking* why the error is an error, which seeks comparison with such rules; or *showing that* an error is an error, for example by substitution. It was found, against expectations, that model-based remediation seemed to have a similar effect to straight re-teaching, although there may have been methodological reasons for this result; in particular that the distinction between the specific model-based remediation and re-teaching employed in the experiment may not have been very great in practice.

However, Kaput (1987) suggests that all such tutors, even though they may enrich the experience of symbol manipulation by providing "history windows" and interactive feedback, accept the reduction of mathematics to learning compartmentalised skills, and that students respond by using "ever more superficial learning strategies". Alternatively, when Lesh & Kelly (1991) studied a "human simulation of computer tutors", they found that as the human tutors developed progressively greater expertise, they spent less time trying to diagnose or correct students' procedural errors, but instead emphasised multiple representations, and asked questions that encouraged manifold types and levels of interpretations and responses. But it is not clear from the report that such trends were independent of the nature of the problems that the students were tackling. Nor was the interaction between students and teachers in the study (via pre-prepared email messages and "simple graphics that the teacher could create at runtime") as great as the interaction possible between students and the interactive algebra tutors described earlier. Even so, the study's results question the implicit assumption of some "procedurally-oriented types of AI-based tutors" that an evermore precise model of the knowledge state of a student is crucial to the improvement of that student's theories.

3.5.11 Games

Edwards (1991) describes a slightly different model for learning using technology (in the context of "Green Blobs" - see Dugdale & Kibbey, 1986 - which links symbolic and graphical representations of equations):

- "Instead of directly teaching the properties of the mathematical entities, these entities are incorporated into a game-like situation, in which the learner must *use* them to solve a problem.
- "In order to use the entities effectively, the student must understand how they work.
- "This understanding is built through an iterative process of conceptual debugging: Students generate solution attempts based on their current model of how the entities work; if these attempts fail, they compare their internal model of the entities with what they see on the screen. In their next attempt, they refine their solution based on this visual feedback. The process continues until the learners have gained a sufficient understanding of the entities to succeed in the game." (p. 7)

Unlike algebra tutors and learning management programs, the learning that results from algebra games is (as far as the student is concerned) a by-product rather than an explicit aim. Of course, from the teacher's point of view there is no difference in aims.

On the other hand, Hoyles *et al.* (1991) raise questions about such programs:

> "Conventional wisdom asserts that by exploration and use of… computer tools and reflection upon computer feedback learners will come to understand the mathematical structures and relationships which have been planted according to a priori learning objectives. Our experience suggests however, that mathematical learning tends not to be unproblematic. Key issues of debate centre on the degree of explicitness and timing of pedagogical intervention whilst maintaining a climate of pupil decision-making and exploration." (p. 197)

3.6 Conclusions

As can be seen from this chapter, the reasons given earlier for gloom about the prospect of comparing activities can in fact provide criteria for comparisons between activities:

- *Different things work for different people*

But one can conjecture that, for a particular theory, activity A works for more students than activity B. While it would be a *safer* approach to assume that an appropriate choice of teaching strategy depends on the particularities of the student's existing theories, it can be worth attempting to investigate the limitations of generalising about activities.

- *Contexts can limit*

But one can conjecture that activity A leads to a grasping of theories of greater transferability than theories encouraged by activity B.

- *Understanding is always incomplete*

But one can conjecture that activity A enables a deeper understanding than activity B.

- *Learning cannot not timetabled*

But one can conjecture that activity A takes less time than activity B to produce the same engagement.

- *An activity does not depend on its rationale*

But supporting claims can sometimes be challenged by comparing outcomes from different activities.

- *Benefit from an activity is difficult to measure*

But one can conjecture that activity A is of more benefit that activity B for a particular individual. We could be wrong, but the whole idea of helping someone to learn depends on the fact that our guesses are sometimes good.

It may have been noted that a theme runs through this chapter. It has been suggested in section 3.4 that meta-algebraic theories may be potentially useful *by-products* of engagement in algebraic problems rather than "causes" of failure. It was also suggested that concern to use equations can be a *by-product* of engagement in word problems, especially when software is used to support representation as a tool rather than as an end in itself. It has now been suggested in section 3.5 that formal equation-solving theories can be *by-products* of engagement with a computer game employing a minimal level of diagnostic management, and in which a balance model and the separation of operation choice and execution might feature. If the representational and transformational aspects could be combined, it would appear that the reinterpretation of research into the learning of algebra using Popperian psychology has led to an instructional proposal.

Chapter 4
A Proposal for Improving Students' Equation Theories and Concerns

4.1 Introduction

This chapter draws together the arguments of the previous chapter into an argument for certain constraints on an activity that could productively test conjectures arising from the reinterpretation of the research literature from a Popperian perspective.

It has been maintained that finding ways to turn problems amenable to algebra into concerns is crucial for improving algebraic theories. It has also been concluded that learning equation theories may best be done by constructing an activity that enables students to develop a genuine concern for symbolic algebra.

A number of activities have been seen in chapter 3 for promoting representation:

- attention to meta-algebraic theories
- formalisation of method, using syncopation, recording of trial-and-improvement or tables
- step-by-step extension of existing representations
- word problem representation software: programming, spreadsheets, automatic equation solvers
- expressing generality via patterns or rules.

It has been argued that acquiring a concern for symbolic algebra should involve word problems, because the advantages of algebra over informal strategies can be appreciated. Formalising method and expressing generality may not be of such obvious value to a student (although several innovative counter-examples to this have been examined). Rather than being portrayed as the focal "causes" of failure, meta-algebraic theories can develop from engagement in problems in which symbolic algebra plays a mediating role. It was therefore suggested that a learning environment that provides syntactical and transformational feedback can effectively enable students to extend existing representation theories by tackling word problems.

However, such an environment may not be enough to enable students who have no prior knowledge of how to represent a situation using standard conventions to learn to represent situations. It may also not be enough to promote the Leibniz method for solving equations. The first section of this chapter looks at what advantages the balance model may offer in these regards. The second section formulates some conjectures that could be tested by computerising the balance model in certain ways. The third section describes the software environment that has been written as a result of this analysis.

4.2 Rationale for using the balance model

A number of activities for promoting transformation were seen in chapter 3:

- attention to meta-algebraic theories
- learning explicit rules for an expressions calculus, the Leibniz method or transposition
- trial-and-improvement substitution
- flowchart methods - undoing, inverse operations, cover-up
- concrete models - balance scales, areas, cups & beans, lines
- automatic equation solvers
- interactive manipulators, tutors, games

A rationale was reconstructed for each proposal, so that what these proposed activities might improve could be discussed and compared. The indicators of improvement used by these studies were considered critically; but although the test items came in for heavy criticism in chapter 2, it was not concluded that the items could in principle tell us nothing about children's algebraic theories and concerns. What was seen, however, was that many of the "suggested implications for teaching" and rationales would be challenged by an activity, with certain characteristics, that promoted analytic uses of equations, as detected by a collection of those very test items that were used to justify alternative proposals.

The balance model was found to have several advantages, in particular that it can promote the formal Leibniz method for solving equations by starting in a familiar, concrete situation. Moreover, if Lins (1992) is correct that the balance model is "one of the most popular didactic artefacts used to teach the solution of linear equations" (p. 208), then this in itself justifies an attempt to explore its deficiencies critically. Nevertheless, the model is seen to suffer from a number of severe limitations, including physical limitations, potential metaphor confusions, misleading letter interpretations, the use of unknowns not variables, ineffectiveness in promoting a cancellation strategy, and ineffectiveness in promoting conventional algebraic notation.

By combining the balance model with some of the features of interactive manipulators, tutors and games, it may be possible to overcome some of the limitations, and also give students an opportunity to create and test strategic theories.

4.3 Questions and Conjectures

The unifying influence of the Popperian psychological perspective enabled the construction of some conjectures from the vast research into activities for initial learning about equations. To test some of these conjectures, an activity with certain characteristics has to be constructed. The following table summarises the conjectures and the consequent implications for the activity.

Conjecture	**Characteristics**
The balance model cannot effectively promote a simplification strategy of formal operations for solving simple linear equations.	A balance context. Simple linear equations. No explicit Leibniz, transposition or expression rules. No encouragement for substitution or flowcharts.
The balance model cannot lead naturally into negative signs, negative solutions, decimal coefficients, multiplication and division.	Equations with negative signs, negative solutions, decimal coefficients, multiplication and division.
The balance model cannot effectively promote a strategy of using conventional algebraic notation.	Word problems. Conventional notation.
The balance model encourages the view of letters as objects rather than numbers.	Initial letters used to represent quantities (e.g. S for number of students)
The balance model does not establish the power of algebraic symbols, because students can only consider procedures for manipulating symbols as meaningful when generalised arithmetic has established the need for algebra.	Unknowns rather than variables. No expression of generality. No formalisation of method. Emphasis on finding a particular numerical value for x rather than developing general solution methods.

The desired outcomes of the activity are firstly, improved strategic theories and concern to use the Leibniz method (once students have this they can appreciate why algebraic transformation might be *useful for them*); secondly, improved strategic theories and concern to formulate equations using algebraic notation (once students have this they can appreciate why algebraic *representation* might be useful for them); thirdly, possibly, improved meta-algebraic theories about equations, expressions and letters. Whether or not such theories "underpin" knowledge, they can at the very least be useful as strategic theories in contexts other than the one in which they were created. On the other hand, one must remain sceptical about other proposed functions of such theories.

Note that the following are target strategic theories:

1. Taking the same quantity off each side of a balance makes it easier to work out an unknown weight.
2. Taking the same number (known or unknown) off each side of an equation makes it easier to work out an unknown number.
3. If you perform the same operation on each side of an equation, the answer is still the same.
4. Performing the same operation on each side of an equation can sometimes make it easier to work out an unknown number [simplification].
5. Finding an equation to represent a situation can sometimes make it easier to work out an unknown quantity [utilisation].

But the following are *not* target strategic theories:

1. If from two equal things the same quantity be taken away, the things will remain equal.
2. In a balance, if the sides weigh the same as each other, there will be no tilting.
3. In a balance, if there is no tilting, the sides weigh the same as each other.
4. An equation is like a balance in that the left hand side is equal in value to the right hand side.
5. In a balance or an equation, what you do to one side, you have to do to the other, otherwise equality will not be maintained.

One issue involved in computerising the model is how much the computer automates. If too *little* of the transformation strategy is carried out automatically for the students (as with solving equations without a computer), the students may get so bogged down in the details of transformation that firstly, they fail to realise that transformation aims to make finding a solution easier, rather than constituting an end in itself; and secondly, the problem of representing a situation does not become a concern, because a representation entails angst about transformation. On the other hand, if too *much* of the transformation strategy is carried out automatically (as with computer symbol manipulators), why should the problem of developing strategic theories for transformation become a concern?

One way out of this transformation dilemma is to use word problem representation software (spreadsheet, Logo, CARAPACE, etc.). The solution strategy is controlled by the student, but transformation is made easier by the computer carrying out the substitution of the requested number in the requested formula. However, in this case, the transformation uses trial-and-improvement rather than formal operations on equations. One solution is to construct an environment in which firstly, students are permitted a very limited range of transformation strategies; secondly, in which transformations chosen by the student are carried out by the computer; thirdly, in which it is obvious when certain transformation strategies result in a simpler situation; fourthly, in which the situation progressively becomes represented by conventional algebraic equations, and the transformations progressively become formal operations; and finally, in which the range of transformation strategies is gradually increased.

Although several IT innovations have been attempted for promoting operations on equations (the algebra tutor of McArthur *et al.*, 1987; the Algebraland software of Brown, 1985; the Balance Laboratory of Roberts *et al.*, 1989; the EXPRESSIONS software of Thompson & Thompson, 1987; and the algebra tutor of Anderson *et al.*, 1990), none of these has simultaneously incorporated:-

- the tackling of word problems in a challenging context
- accessible success criteria at *all* stages (i.e. it is necessary to find a number, rather than to use certified representations or to perceive "underlying structure")
- an option for transforming the situation which becomes increasingly necessary as the problems get harder and informal strategies become more difficult to implement
- the separation of the choice of operation to simplify from the execution of that operation (which obviates the need for arithmetic, and for accurate theories of operations on expressions and therefore allows students to focus on strategic operations)
- the piecemeal introduction of algebraic notation to represent the situation more conveniently
- a retreat from the context to extend the applicability of the representation
- the introduction of a course of solving suitably chosen situational problems that can be tackled using equations to demonstrate the power of the extended representation

A program incorporating all the above features (call it "EQUATION") could allow four research questions to be explored:

1. To what extent can EQUATION improve students' algebraic theories and concerns, as measured by the test instruments developed here?
2. How can these pre-post improvements be related to the students' experiences of the learning activity?
3. What are the limitations of the balance model?
4. How do these effects compare with those of related activities in the research literature?

The first research question relates to the possibility of transferable learning:

> "If pupils already have experience of using and manipulating symbols in computer-based environments, what effect could this have on their developing understandings within a paper-based algebra context?" (Sutherland, 1990, p. 171)

The analysis in chapter 3 would suggest that if students could develop a concern for symbolic algebra then they may be better able to construct strategic theories not just for solving equations and word problems *away* from the computer, but also (if what has been argued about the transferability of knowledge between problems is accurate) for certain representation and transformation items. Clearly, this conjecture can be explored. For example, with which students does EQUATION succeed and fail? With which problems, theories and concerns? Do students become aware that algebra is a tool? Cortes, Vergnaud & Kavafian state about a first attempt at using algebra in word problems that "The gap between the problem to be solved and the pupil's knowledge creates a paradox which can only be resolved through the teacher's tutorial activity." (p. 29). Clearly, then, this assertion is testable. Aside from solving word problems, do students seek to express generality using algebra? Are there any improvements in meta-algebraic theories of equation objects, processes and relations?

The second research question addresses the challenge of how to explain potential improvements identified in addressing the first question. One way of illustrating Popperian psychological analysis is to attempt the important task of reconciling pre-post testing with data gathered on students' learning experiences with EQUATION.

Conjectures and issues relating to the third research question have been detailed earlier in this section.

Finally, the fourth research question is not intended to refer to a direct empirical comparison of learning activities using matched samples or random allocation, but relies on comparing EQUATION's effects with those reported in the research literature for other initiatives. This is firstly because prima facie evidence that EQUATION can in fact improve theories and concerns in some way is required before such a direct comparison could be undertaken; and secondly because it is necessary to explore the effect of merely repeating the written test without overt learning activities having taken place. There are reasons for suspecting particular test items could show improvement merely because of a repeat sitting. The fourth question is raised rather because if EQUATION were to improve *representation* theories, this would potentially challenge the claims to pedagogical priority of activities such as:

- practising the formalisation of method
- expressing mathematical relationships in natural language before algebraic language
- explicit consideration of meta-algebraic theories
- practising the production of unclosed expressions as legitimate answers
- the introductory use of letters to represent generalised numbers
- lengthy experience making explicit "the transition from procedural to structural conceptions" (Kieran, 1992, p. 414)
- explicit teaching of conservation of equation.

It is also noted that EQUATION could, if it improved *transformation* theories, potentially challenge the priority claims of:

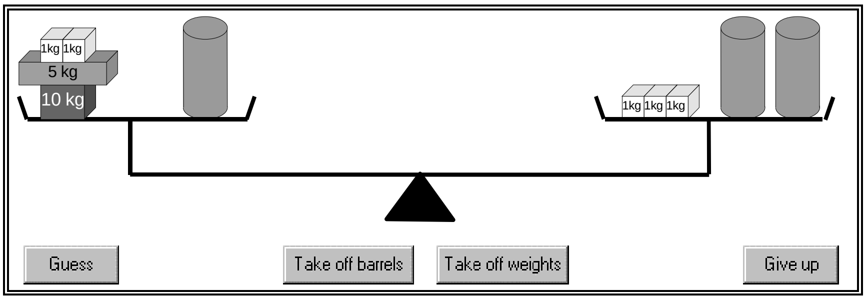
- learning explicit rules for an expressions calculus, the Leibniz method or transposition

- trial-and-improvement substitution
- flowchart methods
- waiting for “cognitive readiness”.

Note that the reinterpretation of the research literature using Popperian psychology has generated the conjectures and questions above, which place constraints on what the software does and looks like. It is, however, possible to analyse the design of the program in greater detail, again using Popperian psychology, and this is done in chapter 5 as part of the description of the development of the program. EQUATION constitutes an integral part of this research, therefore, for three reasons. Firstly, the major constraints outlined above on its specification illustrate Popperian psychology applied to the research literature. Secondly, *within* these constraints, the particular decisions made in its development illustrate Popperian psychology applied to the design of educational software. Thirdly, in using EQUATION in empirical research, the conjectures and questions generated from the reinterpretation of the literature are explored.

4.4 How does EQUATION work?

EQUATION has now been written for Windows 95 and Windows 3.1. The program starts with balance puzzles that are accessible to many students’ informal strategies. The interface is that of a computer game rather than a tutor program.

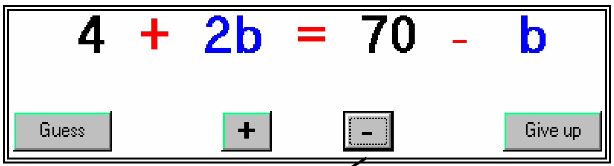


As the puzzles get increasingly difficult, concern to use the Take off… buttons grows. These buttons eventually lead to algebraic operations on equations. Once students have shown they can solve the puzzles at a particular level of difficulty, they are promoted to the next level. Note that the levels are those of a computer game, not the National Curriculum. Students also get a score which depends on their speed and skill. Clearly there are issues of when to promote, what assistance to provide and how to acknowledge failure. Piloting of the software (outlined in chapter 5) has helped to shape policies.

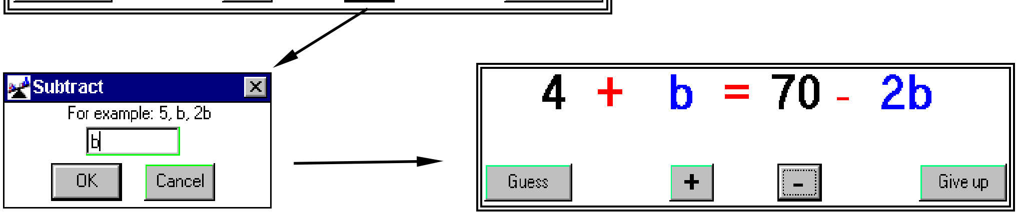
Level 1 involves up to 4 barrels, which weigh an integer number of kilograms. The puzzles are of the one-step form mb = c and b + c = d where b is the weight of a barrel, m is the number of barrels, and c and d are the total weight on one side of the balance. For Level 2, puzzles are of the two-step form mb + c = d or the one-step form mb + a = (m+1)b with a one-barrel difference between the sides (and potentially across the "cut" of Filloy & Rojano). For Levels 3, the puzzles are of the form $mb + c = (m+1)b + d$. The potential number of barrels and their weight increases at level 4, and halving is necessary in the three-step form $mb + c = (m+2)b + d$. Level 5 puzzles are of the form $mb + c = nb + d$ where n is the number of barrels on the right-hand side. Note that although division may be necessary, the answer is still ensured to be an integer. At Level 6, this restriction is eased to allow simple decimal or fractional answers (e.g. 0.25, 2/3, 0.7), and lifted completely at Level 7.

At Level 7, the weights are combined; and on Level 8 the barrels are labelled with a "b", and there is just one button, labelled $\boxed{-}$. Levels 9 and 10 replace the balance pictures with simple linear equations that could represent balance pictures. Levels 11 and 12 introduce negative answers and negative signs.

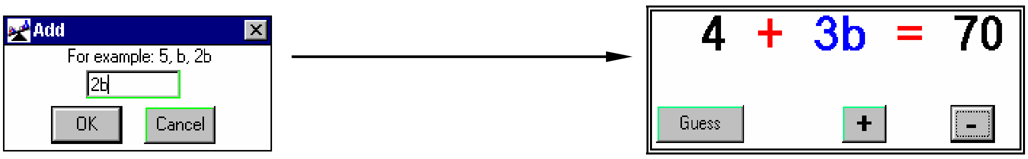
EQUATION offers a major advantage over paper-based exercises. Although the student chooses the operation to perform, the computer executes it. This means firstly that attention can be devoted to strategic simplification decisions, without worrying about arithmetic; and secondly that students see the effect of an operation instantly, thus preventing pages and pages of error-strewn workings:



Suppose the student decided here to try to subtract an unknown (a typical error).



It is clear to the student now that *subtracting* was not a useful strategy. It is necessary to *add*.



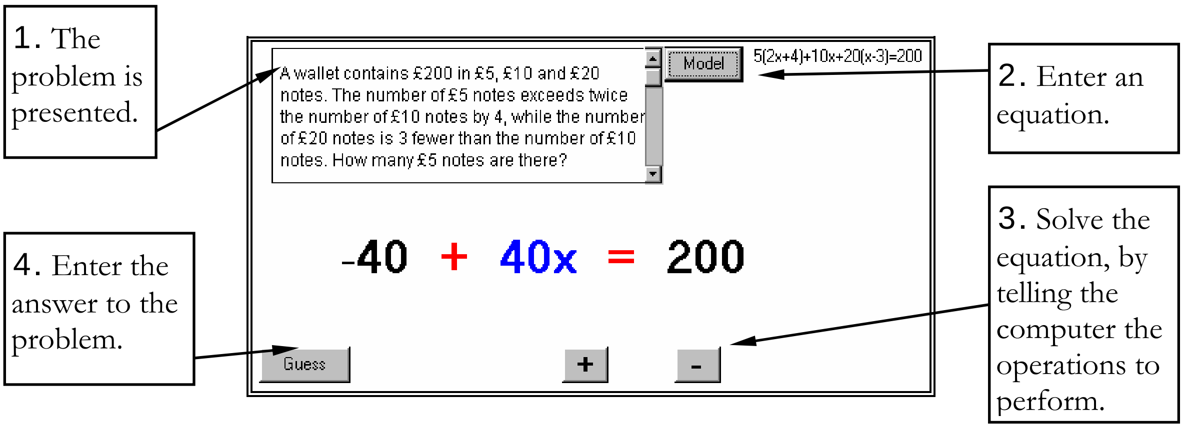
Note that this use of the balance picture as a "source domain" from which students can learn the "target domain" of the equation fits with the recommendations of Rumelhart & Norman (1981), based on their analysis of analogical learning:

1. "[The source domain] should be based on a domain with which the student is very knowledgeable and in which the student can reason readily.
2. The target domain and the source domain should differ by a minimum number of specifiable dimensions.

3. Operations that are natural within the target domain should also be natural within the source domain.
4. Operations inappropriate within the target domain should also be inappropriate within the source domain." (p. 358).

However, they go on to say "Typically, no single model will suffice for any reasonably complex subject matter." and they suggest the creation of several such models for the target domain. In introducing negative answers and negative signs within the equation domain, no new source domain is introduced. In this way, the limitations of the balance model are explored for research purposes, although ideally for teaching purposes alternative models should be introduced.

The final levels contain word problems, similar to those in the research literature. They can be represented as an equation by using the Model button, which gives them some feedback on syntax.



Level 13 involves a description in words of balance puzzles; Level 14 involves puzzles that may include negative signs; Levels 15 and 17 tends to be about combining ratios; and Levels 16 and 18 require the use of expressions. By graduating the problems in this way it is hoped that students will develop five or six standard problem-situation schema (Berger & Wilde, 1987).

EQUATION incorporates three further useful features. Firstly, students can enter their own equations into the computer. Secondly, teachers can prepare a standard file of puzzles for students. Thirdly, as in Thompson & Thompson (1987), each student's puzzles and entries into the computer can be recorded in a log file. This means that a student's problem-solving experiences can be re-played on-screen; the research is not limited by the number of video cameras or recorders available; and the data for a number of students can be easily analysed quantitatively if required (for example, to chart the abandonment of guessing as a strategy).

This design addresses most of the "conceptual difficulties" of Cortes, Vergnaud & Kavafian (1990): the particular conventions of equations are made apparent; the "concept of the unknown" is clear; the equals sign is used to indicate equivalence; "pre-requisite" arithmetic is bypassed; and "detour behaviour" (simplifying before attempting to calculate the unknown) is encouraged. The homogeneity of an equation is not, however, addressed explicitly. With regard to the categories of Berry, Graham & Watkins (1994) outlined in chapter 1, the program combines several functions to some degree. It is clearly an interactive tutor, in that it has specific learning targets; on the other hand it does not have extensive management capabilities, because the range of activities included is strictly limited (in line with the constraints outlined at the beginning of the previous section). It is a demonstration aid, in that an iconic display turns into an equation and into a natural language problem; however it makes no attempt to link multiple representations dynamically. It is clearly also an investigative environment for exploring the simplification of balance situations, operations on equations and the representation of word problems; on the other hand the range of options open to the student is limited. It is a problem solving assistant, in that it supports the development of algebraic strategies for solving word problems; and it can also act as a mathematical tool for solving equations and word problems given to the student by the teacher.

4.5 Final Remarks

Programming represents another way of interacting mathematically with a computer, and one that is, in many ways, of much greater potential benefit to a child than EQUATION. EQUATION is about the child "programming" the computer using algebra, but in a much more limited way than Logo, Boxer or Visual Basic. It is certainly about computers acting as "carriers of powerful ideas" (Papert, 1980, p. 4), but only in a modest way. It should also be pointed out that, as far as the students are concerned, algebra is introduced as a *game*; and then shown as a *linguistic tool*; and perhaps only later will they appreciate its other qualities. But it is primarily a research instrument rather than an idealised teaching activity, and in this sense it does have the potential to "challenge current beliefs about who can understand what and at what age.", as Papert puts it. More levels are clearly possible - for example simultaneous equations, fractions (× and ÷ buttons), quadratics and re-arranging formulae - but this initial set of problems is sufficient to test the conjectures that have been devised above.

In order to illustrate Popperian psychology, fieldwork will be carried out to explore the questions asked in this chapter. In particular, it is hoped to identify improvements in students' algebraic theories and concerns using the pre-post instruments generated in chapter 2; to relate these improvements to the algebraic problems contained in the software; and to compare the improvements with the initiatives described in chapter 3. The empirical research will be a success if it contributes something towards finding out about the extent to which a computerised balance model could enable students to use algebraic symbolism to explore problems.

Chapter 5
Fieldwork

5.1 Introduction

The previous chapter described a proposal to improve students' theories and concerns that would illustrate the theoretical arguments of this research and provide prima facie tests of conjectures arising from the reinterpretation of the research literature. That chapter therefore set the major constraints on the fieldwork. This chapter provides details of the development, planning and execution of the fieldwork.

As has been seen in the previous chapter, the fieldwork requires a *student activity* that can promote particular learning in particular ways; *research instruments* that can detect such learning and provide a comparison with other activities; and arrangements that can allow *data collection* on the relationship between the activity and the learning.

With regard to the activity (EQUATION), chapter 4 has examined the major constraints on its design - constraints set by the need to test the claims arising from the analysis of the research literature. This chapter summarises the development and piloting of EQUATION, including the many minor practical design decisions that had to be made, their reasons, and the pros and cons of alternatives.

With regard to the research instruments, chapters 2 and 3 provided a large number of test items that might detect learning while at the same time allowing comparison with other activities. This chapter summarises the development of a written test instrument that incorporates some of these items, and supplements them with questions designed to probe equation concerns and consideration independently of particular mathematical problems. This includes the use of interviews.

The major constraints on the student activity also tend to place limits on the arrangements for data collection. However, many minor decisions had to be made in practice, arising out of negotiation with the teachers collaborating in the research and constrained by the available resources. This chapter summarises these arrangements, from the earliest pilot work to the three main experiments. The nature and extent of data collected are indicated.

5.2 The Development of EQUATION

Popperian psychology can provide three criteria against which the potential of an activity for promoting certain target theories could be judged:

1. there are problems for which target theories offer advantages over existing theories;
2. there are mechanisms for these problems to become concerns;
3. there are opportunities for the target theories to be created and tested.

These are not "all or nothing" characteristics of the activity. For example, changing an aspect of the activity could change the problems and hence the relevant "existing theories" and consequently the "advantages" offered by the target theories. Similarly, changing an aspect of the activity could affect the mechanisms for turning problems into concerns, or the opportunities for target theories to be created and tested. Therefore, different options for changing the activity might be judged "better" or "worse" against these criteria. So what are being sought (as argued in chapters 1 to 4) are improvements to problems, mechanisms and opportunities for the target theories of operating on equations and formulating equations to find an unknown number in a situation.

Note, however, that in the early stages of EQUATION's development, the perspective was not fully elaborated, and so the construction of the program was rationalised in less constrained terms.

5.2.1 The BALANCE Program

If struggling with procedures gets in the way of understanding the strategy of simplification by operating on both sides of an equality, and if this strategy does not depend on algebraic symbols, could less abstract experiences be provided to promote the strategy prior to formalisation? The first version of EQUATION (called "BALANCE", and started in February 1996) was intended to correct some of the deficiencies identified by Soh (1995) in the pictorial textbook experience and the findings of Schliemann *et al.* (1992) that children rarely use such a strategy spontaneously to find unknown weights on physical balance scales.

BALANCE gave feedback on answers, allowed arithmetic to be bypassed in the way described in the previous chapter, allowed experimentation with strategies without the need for re-drawing or crossings out, allowed a variety of strategies, but visibly showed the benefit of simplification. It did not pretend to be "concrete", but provided a constrained environment in which the range of possible confusions was reduced because the range of actions was limited. At this stage there was no intention to make a link with algebraic symbolism, or to test the limitations of the balance model. The program started with barrels on both sides - barrels so that the weights of actual objects did not intrude and so that there could be a wide variation in the possible weights of the objects used. There were 5 levels, and it was possible to skip up to the next level if desired.

Equations could be entered to be turned into pictures, but there was no algebraic notation introduced in the program.

The basic rationale was simple. Since the balance puzzles are initially accessible to informal strategies - such as guessing, cover-up, matching or part-reasoning - more sophisticated strategies such as trial-and-improvement, graphs or inversing are not necessary. As weights and barrels appear on both sides, and the numbers involved get larger, informal strategies are harder to implement; but the Take Off… buttons allow the puzzles to be simplified without arithmetical error. It is usually obvious that the situation is simpler after using the buttons. By separating the task of *deciding what to do* from *doing it*, developing strategic theories for the latter can be left until later, leaving the student free to refine strategic theories for the former. Given this problem situation, mechanisms to turn the problem of simplification into a concern included: accessible success criteria (finding an unknown number in a concrete situation); a game-like interface in which success in solving the problems is rewarded by a score and a new challenge; the likelihood of early success; few penalties for failure and encouragement of experimentation; the usefulness of competition and collaboration with one's colleagues; the futility of giving up (not even a hint); colourful pictures; minimal reading and writing; minimal requirements for prior knowledge or help from external sources; potentially worrisome arithmetic is bypassed; tolerance for experimentation allows curiosity to explore what the Take Off… buttons do; and the puzzles are easy once the target strategic theory has been constructed.

Larkin (1989, p. 131) suggests some criteria for a display that attempts to represent a theory:

1. It should be isomorphic to the theory.
2. It should not imply spurious relations.
3. It should use the standard notation.
4. "… ordinarily difficult processes should be represented by easy perceptual processes".
5. It should be memorable.

Lewis, Milson & Anderson (1987) note in addition the interface should minimise keystrokes and mouse clicks, it should be highly interactive, it should be responsive to student errors and it should keep to a minimum the number of parts of the screen to be scrutinised. These principles capture something of the initial rationale for the interface, although - as will be seen in the next section - certain of these aspects needed to be improved.

A number of design decisions were made that seemed reasonable at the time, given the limited aims of this first attempt at the program, but which could now be questioned.

For example, the balance is never seen to tilt. Therefore the balance shown on the screen is largely iconic, and the point of removing items from *both* sides is not made. One way of making this point, suggested one teacher involved in the research, is to see the balance wobble when the items are removed; or the items could be removed from one side just before the other, and the balance starts to tilt before moving back. However, it could be argued – citing the psychological principle of learning occurring in response to concerns – that this feature would serve no useful purpose *as far as the student is concerned*; it has no impact at all on the actions that need to be carried out, and would add to background "noise" merely to try to satisfy those already in-the-know that some sort of attention was being paid to this aspect. The same would apply to showing - at the same time as the balance picture - an algebraic representation of the situation on the screen, or a line graph, the "ghosts" of removed items, or a record of actions and past balances,. No doubt more mathematically sophisticated users of the program would find much of interest in such representations, but the activity is intended for precisely those who do not have such prior concerns. A better solution to the problem of promoting the need for taking the *same* items from *both* sides would appear to be requiring separate buttons for each side; the student would then have to discover that taking the same items from both sides was the only way to achieve balance. Unfortunately, this has disadvantages too. If the program restricted the student to taking off only those items which have a corresponding item on the other side, one non-algebraic strategy would be to take off as many items as allowed from one side, and then do the same to the other. On the other hand, if the program allowed *all* the items to be taken off, the student could end up in a situation in which it was impossible to find the weight unless items could be put back on. In this case, the non-algebraic strategy of clearing off all objects except for one barrel and putting weights progressively on the other side would work. In any case, having to remove from both sides makes simplification twice as cumbersome, and so makes it less attractive than alternative informal methods. Most importantly, this feature detracts from the idea of simplification by subtraction. The problem of finding a way for the program to promote doing the *same* operation to both sides still remains, although dealing with negative signs might eventually address this.

Alternatively, when a wrong guess is made, the balance could tilt appropriately. Firstly, of course, this might seem counter-intuitive - after all, the barrel has a weight, independently of what is guessed, so why would the mere action of guessing affect the physical situation? Therefore, in order to provide a physical model for guessing causing tilting, the barrels could be seen to be replaced by the guessed weight before tilting, and then returned to equilibrium when the barrel is reinstated. More significantly, this guess-and-tilt feature would encourage a strategy of trial-and-improvement. Now this may be very desirable, as it could perhaps be turned into algebraic substitution. However, it is the idea of simplification by operating on the equality that is the target theory, rather than of substitution. In chapter 4 a number of theories that are *not* targets were described. In particular, the following were not target theories: "In a balance, if the sides weigh the same as each other, there will be no tilting"; "In a balance, if there is no tilting, the sides weigh the same as each other"; "In a balance or an equation, what you do to one side, you have to do to the other, otherwise equality will not be maintained". Nevertheless, a guess-and-tilt feature could be useful in a future version of the activity, perhaps as a separate series of levels, so as to avoid confusion with the promotion of simplification.

Another decision was that items are removed by clicking on a button to call up a "dialogue-box" in which the required number of barrels or kg can be entered, rather than the perhaps more natural method of dragging-and-dropping. This was because drag-and-drop again runs into the difficulty of more attractive non-algebraic strategies. Interestingly, one researcher when shown the program immediately attempted to transpose a barrel from one side to the other, because "getting the unknown by itself" by moving things between the sides is the algebraic strategy he uses. Another objection to drag-and-drop is that it makes weight-splitting seem unnatural: although the removal of "items" has been talked about, this is not quite the same as removing "objects". A 10kg weight can have 1kg removed if necessary. By experimenting with such operations, students can quickly discover that the actual weights that are present are not significant - it is their total. Automatic weight-splitting offers an advantage over the textbook approach, which can cause confusion on this point. Such confusion of course raises the question: why have different types of weights at all? There were a number of reasons behind this. Firstly, showing single weights would have made a visual matching strategy easier than removing weights, even when the numbers get large (because whole rows can be matched at a time); and it would – for those without a matching strategy – have necessitated lots of potentially error-prone counting. Secondly, Schliemann *et al.* (1992) show that children tend to use idiosyncratic methods based on the particular weights present; so showing totals from the start might have prematurely made the assumption that students were convinced that the particular weights could be ignored. Thirdly, showing different types of weights rather than totals would make a matching strategy easier when the numbers are small - and this was seen as a good thing because such a strategy would lead naturally onto removing matched objects completely when the numbers got larger. Finally, several "mathphobic" adults who tried the program commented that they found the solid, multi-coloured pictures more reassuring than a spartan total number would have been. At a higher level, the multiple weight pictures were replaced by a single total weight, once it was considered that the simplification strategy had been grasped, (and hopefully therefore that only the total was important).

The dialogue boxes can provide an example of the many minor interface improvements that took place, that do not greatly affect the promotion of mathematical strategies and so are not discussed here. Many students clicked on Take Off Barrels ; then entered a number; and then, rather than clicking OK (as is customary under the Windows interface), clicked Take Off Barrels again. This had the effect of recalling (and blanking) the dialogue box rather than removing the requested number of barrels (as might have been the student's intention). Now of course it would have been possible to make the Take Off Barrels button disappear once it was clicked, or provide a beep to indicate this was not a possible action (as is current Windows practice). Instead, because there was little confusion caused by such an action, re-clicking the button was changed to remove the barrels. Many such interface design issues occurred, and of course there were myriads of purely programming design decisions that had to be taken; but they are omitted from this account because their significance is considered small with respect to the research conjectures. Nevertheless, interface issues such as this one or programming issues such as the speed of operation can lessen the opportunities for concerns to be grasped and for strategic theories to be created and tested; and so can contribute to the learning environment that is under scrutiny.

Confusion can arise over the representation of the scale pans. Many students know from personal experience in the playground or from science experiments, that the distance of an object from the pivot of a seesaw affects how much the object makes the seesaw tilt. Nolder (1991) has described how a student was distracted by the positioning of the objects. The decision was taken to show the scale pans schematically on the screen. The evidence available from piloting this version of the program seemed to suggest that there was no confusion on this point.

Another decision was to "grey out" each $\boxed{\text{Take Off…}}$ button when it was no longer applicable. This was to draw attention to the fact that the button was no longer useful - which may go a little way towards emphasising that it must be possible to remove the same items from each side. This feature was also there so that greying out would act as a prompt on later levels involving negative answers and negative signs that a certain action was no longer helpful. However, it would have been perfectly possible to allow the entry of attempts to remove more items. Such attempts would, moreover, have provided an insight into students' appreciation of the possible actions. Nevertheless, it was considered important that in the early stages of introducing simplification, as little as possible should interfere in the efficient grasping of the strategy. The program should then, in later stages, ease this restriction, so that careful thought would be required as to whether the action was useful.

It was adversely noted by one teacher that attempting to take off impossible weights, or wrong guesses were not signalled. Although correct answers are acknowledged (with text, a noise, and a picture), no other feedback is given. This was to encourage experimentation - vital at this stage if the strategy of splitting weights is to be grasped, and important later when negative signs come in. However, there could be more feedback. What happens if the students get stuck? How are they to understand the reasons why what they wanted to do could not be done, or why the answer was wrong? How do they know that the computer has understood their entry? At this stage there was no real answer to these questions, and we return to this point later.

Another decision was to keep the order of weights and barrels constant. So although the program varies the sides - so you could get both "20kg weighs the same as 2kg plus a barrel" and "2kg plus a barrel weigh 20kg" - you couldn't get "A barrel plus 2kg weighs the same as 20kg". There were three reasons for this. The first reason was purely opportunistic: it was easier to program fixed positions, and - like the idea of having just barrels rather than a variety of objects of unknown weight – order did not seem crucial to grasping the strategy of simplification. Moreover, it was conjectured that there would be very few students who would fail to appreciate that the order of weights and barrels was not a relevant factor in either the value of the unknown weight or the simplification strategy for finding it. The second reason was slightly more subtle: keeping irrelevant variation initially fixed (against the spirit of Dienes' variability principles) was seen as a *good thing*, because it would allow students to focus on the more important aspects of the situation (such as the numbers of barrels and the total weights) rather than to be side-tracked by issues of lesser importance. The variation should, of course, be introduced at some point in the program, but an interesting question is at what point. The third reason for a fixed order is to do with this point, and makes sense within the context of the move to algebraic symbolism, and so is discussed later.

The decision to generate puzzles *randomly* (within the parameters of each level) was taken to avoid recall or copying being of use, to minimise the possibility of idiosyncratic strategies based on particular numbers, to provide a modicum of variety and to save the bother of creating puzzles by hand. However, the puzzles could have been fixed, so as to ensure a full variety of cases were tackled, to ensure progression within each level, and to provide standardised questions for comparing students. Another possibility was random generation from a bank of puzzles. The advantages of random generation were roughly considered to outweigh the advantages of "given" puzzles. At this stage, variety was to be achieved by the researcher, teacher or student choosing the moment for the move up to the next level; and by constructing the levels so that variety within a level would be only a minor factor in determining the difficulty of a level and in any comparison of students.

A decision that outraged one teacher was the apparent measurement of "weight" in kilograms rather than Newtons. This decision was angst-ridden. On the one hand, personal experience has taught that neither "mass" nor "Newtons" are as familiar from life outside school to students as "weight" and "kilograms"; on the other hand, science teachers have long complained about the confusion between the ideas of mass and weight, and between their units, that could perhaps be minimised by consistent usage of the terms in all classes. And in this artificial computer environment, surely measuring the weight in Newtons can be enforced without doing much damage to the student's grasp of the situation? Nevertheless, the overriding consideration here was making the situation as accessible as possible; even though guilt was only partially assuaged by the thought that usage in just this one program would be unlikely in itself to be enough to sway students either to or from the confusion.

It could be argued that the BALANCE program should be just one of a number of activities used if the general effects of IT on learning were to be studied. However this was not the intention of the research, nor could it have been a feasible objective within the constraints of time and resources. Moreover, in many ways having such a sharp focus would have the advantage that more attention could be given to identifying instantaneous learning outcomes in as much detail as possible, isolated as far as possible from other algebraic activities. The main disadvantages of using just one activity would be firstly a lack of breadth of experiences, which would limit the scope of the conjecture-testing; secondly a smaller time-scale and range of environments over which to assess learning; and thirdly a lack of comparison with other activities. In order to mitigate the second disadvantage, detailed evidence would have to be collected from the classwork and not just from a written instrument and interviews. The third disadvantage is partly addressed by selecting items that can be compared with activities studied in the literature.

5.2.2 The EQUATION Program – Pre-symbolic

So that this activity could be used in a wider range of school computer systems, with fewer installation obstacles and licensing issues, and so that certain interface features could be made more user-friendly, the program was re-written in Microsoft Visual Basic 4 (rather than VBA for Microsoft Excel). In order to emphasise the change in aims of the activity – from an attempt at correcting textbook deficiencies in one aspect of a "concept" to a research tool for exploring the improvement of certain student theories and concerns – the program name was changed from "BALANCE" to "EQUATION".

A primitive mechanism for promotion to the next level was also introduced – 5 puzzles solved correctly without error would enable promotion; several errors would mean extra puzzles might have to be solved. A score was displayed. Although such one-dimensional assessments of progress can be confidence-sapping, this score could only increase (as in many computer games). The dangers of competition – such as a concern to increase score rather than appreciating the learning "for its own sake", and diminished self-esteem for those with low scores – were carefully thought about, but were eventually accepted, given the motivational advantages from portraying the program as a game rather than a set of exercises.

The program was trialled with a variety of people, to improve the interface and functionality. In most cases, no more than half an hour was spent on the program.

Firstly, four pairs of students aged 10-13 at School B used the program. The school is a middle school on the outskirts of a town in the South Midlands. The two Year 8 students soon got the idea of simplifying the situation to solve problems, and rapidly attained the highest levels. This usage suggested that the program did not start, perhaps, at an unreasonably high level, and that progress through the program was not guaranteed to be patronisingly slow. The students had met equations before, but only one of the students seemed to appreciate the connection ("Have you done anything like this before?" "Yeah… like equations"). Their teacher commented that it was really simultaneous equations where their difficulties lay, rather than with linear equations in one unknown. However, no testing was carried out as part of this trial. It would be possible to incorporate some simultaneous equations levels in the program in the future, but these were not seen as vital aspects at this stage.

All four Year 7 students found the idea of subtracting barrels and weights an obvious thing to do, despite no experience of equations. However, the idea that if 2 barrels weigh 10kg, one barrel must weigh 5kg was not immediately obvious; and various guesses were tried. After two problems like this, they got the idea of dividing (although they found the arithmetic, even in cases like 12 ÷ 3, to be taxing) and once they were prompted to enter the calculation they wanted to do (i.e. a fraction) as an answer, they had no further difficulties with the program and attained the higher levels. This suggested that each level should contain at least 3 problems of the same type before moving on, to give the students a chance to confirm their strategies before the introduction of greater complexity. Moreover, the decision to allow fractional answers (such as "5/3" or "1 2/3") seemed to have been vindicated, in that were this not the case, progress with simplification would probably have stalled.

On the subject of fractions, it was decided that the program should not insist on lowest terms, because it is the solution process (rather than the form of the answer) that is being emphasised. This does mean, however, that students can get into the syndrome of entering a string of characters (such as 12/3) without necessarily appreciating that a simple numerical answer is available. It was suggested by an adviser that students should see the simplest answer displayed, when "Correct" is shown. But would this really address a student concern? A better method may be to introduce levels on which lowest terms were required, although this does not address an algebraic theory. An alternative, then, would be to insist on lowest terms on levels where a $\boxed{\div}$ button is available, so that the student ends up having to construct an algebraic theory.

Neither Year 6 student found the idea of subtracting barrels and weights at all obvious. They initially tried to solve the picture of $x + 13 = 20$ by trial-and-error. After prompting about the Take Off… buttons, one of the students suddenly got the idea, and would happily subtract as a first step. However, he would always subtract one barrel at a time, even when there was more than one barrel on each side. The other student struggled with the idea of subtraction. Whether this was because she considered it "cheating", whether she was intimidated by her colleague's confidence, or whether she just did not appreciate the value of subtracting is not clear. Nevertheless, these two students made it to level 4 of the program, with some prompting about fractions.

There was some enthusiastic comparison of scores, and three of the students apparently returned in their own time to attempt to beat the highest score reported. Although discussion between the students was limited (see later), the existence of a score did not seem to inhibit questioning - "Can you add weights?" asked one Year 6 child, in a situation something like 37kg + 4 barrels = 9kg + 5 barrels (adding 3 kg might have made the subtraction 37 – 9 easier). On the other hand, diminished self-esteem might be an unlikely characteristic to be clearly observed.

Minor improvements to the design of the program were made in the light of this trial. These included providing a background to fill the whole screen, thus removing superfluous windows from view, and reducing the chance of multiple loading of the program. Minor bugs were corrected. The promotion mechanism (which had functioned incorrectly) was made more sophisticated, and now used the time taken to solve each puzzle and the number of actions as factors. Because the Year 6 students found certain puzzles – particular those with barrels and weights on both sides – on the existing first level much harder than others, it was decided to insert two levels before this one (the levels were then re-numbered). This meant that the new first level introduced the ideas of subtracting weights and dividing weights separately (and mentally); the new second level then allowed puzzles in which subtraction and division might both occur, and also puzzles in which there was only a one-barrel difference between the sides. So those students who struggled with the old first level would now find the same puzzles on level 3 easier.

This gradation of the first level meant that the age for which the program is appropriate could be explored. A six-year-old was given the program to play with. Although she was shown how the interface worked by her parents, she continued with it in her own time, and determinedly stuck with it, by herself, without deliberate pressure from adults. (Judging by her grasp of language, she was a rather advanced six-year-old). She reached level 5, and then was distraught to discover that she could not save her work. A save option was left out not only because experience had shown that file-handling on school networks led to frustrations for students and teachers alike; but also because it was hoped to be able to see what happened when students returned to the program, in particular to what extent the strategies had been remembered. It was also seen as useful revision, and the new promotion mechanism meant that very little time was wasted in starting from Level

1. The Enter an Equation button would enable teachers or students to skip to a higher level, should this be desired. Alternatively, by this time an option for loading a file of pre-prepared equations had been added, so that each student could tackle the same problems, and so that higher levels could be introduced more quickly.

A small number of teachers did not like the word “guess” on the button allowing entry of the weight of a barrel. This, they suggested, would encourage precisely that - guessing - rather than a more algebraically sophisticated strategy. It has to be pointed out that this was precisely the point. Informal strategies should be encouraged, because students construct their concern to find better strategies only by their existing strategies being put under stress. After all, if informal strategies prove to be more attractive than simplification in the long run, then why bother with simplification? The evidence from trialling appeared to be that blind guessing as an initial strategy was used by only an extremely small proportion of students. Meanwhile, nearly all students seemed able to experiment with all the buttons on the screen - encouraging such experimentation was another aim of using the word “guess” to license failure.

In response to the earlier questions about providing more feedback, an option was incorporated into the program to give messages to indicate why certain actions were not possible, to indicate why certain entries were invalid, to indicate that the answer was wrong, and to suggest ways forward if the student appeared stuck. The presence of a Give Up button also became optional. Nevertheless, as had been expected, the message features were much less popular than the standard version of the program. The excitement that so many students exhibited when using the standard version was rarely found amongst the 4 or 5 people who tried the message version. One or two when questioned about this even pointed that “It is much more fun to find things out for yourself” than be patronised in this way. Rather than the curiosity engendered by a computer game, the message version encouraged the view that this is a “program that teaches you maths”, with all the negative connotations that that might inspire in some. Experimentation, commitment and tenacity were noticeably diminished. In reply to those earlier questions, then, it appears that in the absence of such messages, students do their best *not* to get stuck, or use the Give Up button as a license to tackle a fresh problem, or trade insights with their friends. They *work out for themselves* what might be going wrong or what the computer might intend, and (as we shall see), their explanations are something equally wrong, but these can be put to the test and the minimalist approach allows them to see failure as a step on the road to success.

One “bug” in the program that remained unfixed for some time was actually seen as a “feature”. If the weight on one side of the balance went over 385kg (as it could do, on some levels), there was the possibility that not all the weights could be seen. For the sake of interface simplicity an expanding window was not implemented; and the illegibility of shrunk weight pictures precluded that option. The easiest solution would have been to limit the total weight on a single side to 385kg. However, this was not done, because it was felt that students would gain an appreciation of the value of the simplification strategy if the only way they could solve the problem was by removing sufficient items from both sides so that the weights on the obscured scale pan could then be seen. However, there were puzzles in which even doing this would not make all the weights visible. Again, these puzzles could have been excluded from generation, but it was thought that the principle that no one method was sufficient for all problems would be promoted by the only feasible strategy - intelligent guessing to eliminate the possible values. However, many teachers and students found this occurrence to be annoying; and once the study of the research literature was complete, it was found that this insufficiency principle did not form part of the conjectures being tested. So the bug was fixed. However, entry of puzzles that would be only partly visible was later allowed, to give teachers the option of promoting the insufficiency principle, and to allow students to experiment with the larger range of puzzles.

Some adults who considered themselves "non-mathematical" enjoyed using the program very much. Several of these were, however, reluctant to use the Take Off… buttons, preferring to demonstrate their mental arithmetic skills. It was noticeable that complete simplification was rarely carried out even as far as Level 5. However, in situations where they were not confident about the results they used the Take Off… buttons rather than guess, and thereafter were happy to use the buttons on each puzzle. Although virtually all teachers who have tried the program have particularly praised the bypassing of arithmetic, some researchers have disliked this because students would not be able to practise their arithmetic and would, in exams, have to carry out the simplification without a computer. Nevertheless, several teachers have suggested that students would, in any case, want to bypass the Take Off… buttons if they could calculate the sums quicker in the head and thus gain a higher score. Moreover, the issue of "not being able to see the wood for the trees" (i.e. not being able to grasp the point of algebra because of arithmetical difficulties) was seen as a very real issue by various teachers.

5.2.3 The Move to Algebraic Symbols

When students have finished solving the balance puzzles, do they then understand the *algebraic* strategy of simplification? It seemed unlikely that the purpose of symbolic algebra would be automatically clear. So promotion of pre-symbolic simplification was not enough - further levels were required to introduce the algebraic symbolism. Moreover, it was conjectured, based on experience of variations of the balance situation, that the transition to equations with negative signs could be better made by moving away from the balance situation rather than trying to adapt it (e.g. by using balloons or "negative weights). Alternatively, new situations (as described in the previous chapter) could be introduced. Tackling the physical limitations of the balance model requires an attempt to introduce negative signs, and negative solutions; however decimal coefficients were left until the modelling levels - a decision that occurred by default, and that could now be questioned.

On the level after multiple weights were reduced to a single total weight on each side, a level was added labelling barrels with the letter "b" and turning the two Take Off… buttons into a single Subtract button. The next level replaced the balance picture with an algebraic equation, but returned to the use of two buttons for removing unknowns and constants; the level after re-introduced a single – button. The idea was to introduce algebraic notation gradually as a convenient abbreviation; while the de facto separation of operation decision from operation execution obviates the need for an initial requirement for accurate theories of operations on expressions. The letter “b” had to be chosen in order to maximise the potential for an “object view” of letters, as outlined in the constraints on the research activity in chapter 4. Levels with negative signs and negative answers were also added, thus breaking with the balance model. Finally one modelling level was added. The concern mechanisms are just as before, although it must be pointed out that the transition to conventional algebraic notation is intended to be as *unthreatening* as possible. Some adults trying the program found the jump worth noting by a worried noise such as “Mmm!”, “Ooh!” or “Oh no!”, but very few children did so - the closest some came was “Ah!”.

A further reason for a fixed order of weights and barrels was intimated earlier. When an equation appears on the screen for the first time, where a balance picture used to be, the intention is to encourage the view that the equation is merely a shorthand for the balance situation. This is in line with the idea of attempting to facilitate the transfer of the simplification strategy from balance puzzle to equation. By fixing the position on the screen of the weights and barrels, there can be an exact match between the order of weights and barrels and the order of constants and unknowns; and this match helps to encourage the shorthand view. However, there are two points that need to be made. Firstly, should the variation of this fixed order occur before or after the transfer to symbols? If *before*, the advantage is lost of matching weights to constants and barrels to unknowns; if *after*, the advantage is lost of understanding that the variation does not affect the solution or strategy in terms of moving objects around on a single scale pan. Since the variation was never implemented anyway, this question did not have to faced. But a compromise would be to introduce variation prior to the level on which the single Subtract button appears, return to fixed positions on that level and the symbolic one after, and then return to varied positions from then onwards. The second point is that the failure to introduce variation meant that the equations selected for solution on the written test were coincidentally in the “wrong” order, so those who did not grasp the unimportance of order would not have an adequate strategy to tackle the item, and thus potentially diminishing any improvement in facility. It would have been useful to have included some test items checking what the effect of varying order might be.

A feature that is suddenly of prime importance when there are negative signs is the role of feedback. By instantly replacing the equation when an operation is selected, the effect of that operation is seen straightaway, rather than misconceived ideas about what it does being incorporated into the method of solution, as is possible with pencil-and-paper. This speed of feedback makes it less likely that the student will lose the thread of what they are doing.

Changes were again made to the design of the program.

The effect of pressing the Give Up button was changed from reducing score and allowing another puzzle on the same level to be tackled to slipping back a level. This was because it was recognised that if students moved on too quickly to higher levels (for example, they were entering or loading equations that they could not solve, or had not acquired an appropriate strategy at the previous level), there had to be some means for them to return to a lower level.

The scoring was improved by being multiplied by a factor of 10. This not only made the scores appear more dramatic (and more like the scores on video games or pinball machines), but allowed greater diversity in score - which would go some way to countering the false impression that there was a fixed path through the program. For similar reasons, a bonus score was added for speedy solution. This also meant that the Year 8 students who had tried the program before would have been able to move up to higher levels more quickly. The loss of score on giving up was fixed to be roughly the equivalent of having to solve another 1½ puzzles; wrong guesses lost 1 point (about 5% of the standard score for solving a puzzle); and the speed bonuses were fixed so that it usually required at least 3 goes to get onto the next level.

The Enter an Equation facility was improved. Firstly, equations entered with solution 0, negative signs or decimal coefficients were now allowed, because with the introduction of symbols, there was no longer any need to reject equations that do not conform to the limited balance model. Secondly, entered equations were also now allowed to use letters other than "b". Unfortunately, none of the symbolic levels ever introduced letters other than "b". This was a serious failing, because students would be expected to use other letters on the modelling levels. Finally, a major change occurred that at first sight might appear entirely internal, and without implications for the visible operation of the program. A module that checked entered equations were legal and linear in one unknown was replaced by an all-purpose algebraic module written for another program. This other program (called *Noy*) was originally - before it was decided that one activity could provide sufficient data to test the research conjectures - destined not only to provide all the computer algebra system facilities (simplification, solution, factorisation, expansion, substitution and iteration) required by students tackling the activities, but also all the algebraically-useful facilities of a graphics calculator and spreadsheet. Noy had the advantage over its nearest rivals (*Derive*, *Excel* and the TI-82 calculator) that it was much easier to use (one-key simplification, solution, graphing, tabling, etc. rather than multiple menus and entries) and free (therefore not placing any licensing limits on the number of students or schools involved in the research). Although the user would see no difference in operation, the use of this all-purpose algebra module meant that modelling levels could be much more flexible in accepting equations. For example, brackets, division and multiplication could be entered.

The idea that the computer could record all the students' interactions in a "log file" for research purposes was important. This idea meant that it was possible to obtain data from every student using the program, no matter how many computers, researchers or videos were available. Of course such a log has advantages and disadvantages compared to other means of collecting data, and these are discussed later. But an immediate impact of this innovation was that in developing the program it was possible to reproduce exactly what the student did and saw.

The Subtract button was re-labelled as – so as to minimise the words that have to be understood, and to provide a standard subtraction button throughout. The Subtract unknowns and Subtract numbers buttons on the symbolic levels were also replaced by this – button, because few difficulties had been observed in using the button for either the balance pictures or the equations, and the move back to two separate buttons did not seem to offer much of an improvement.

On levels 9 to 11, there is no $\boxed{+}$ button, so the $\boxed{-}$ button is firstly disabled when the equation reaches the form ax = b (where a and b are positive); and secondly, when the equation has a negative answer (on level 11), the program prevents the user reaching a state where the $\boxed{+}$ button is required, e.g. -40 – 10x = 0. Students could of course always add 40 by subtracting -40, but the intention of the level with negative signs was ensure that the student realises as quickly as possible that the answer has to be negative if ax + b = 0 or ax = -b, rather than to explore the effect of subtracting negative numbers. On level 12 and the modelling levels, however, the buttons are never disabled, which forces students who are not yet confident of their strategy to consider the circumstances under which the operations are appropriate.

The program had also previously used whatever letter the student had used in his or her equations for all equations on levels 9-12. This was changed to "b" as standard, because some had experimented with entering equations before they had moved from pictures to symbols, and would therefore not have the advantages and disadvantages of "b for barrel" as a mnemonic.

One mathematics teacher suggested that for many Year 10 students, the re-arrangement of formulas presents great difficulties, especially the move from numerical coefficients to variable coefficients. The original intention was to add levels requiring ÷ and ×, followed by re-arrangement of formulae and simultaneous equations. This did not happen, however. Development time was perhaps the most significant constraint on the range of algebraic problems that could be incorporated in the program; but it was also realised (when the study of research literature was complete) that there were more than enough ideas to test in an activity computerising the balance model, moving to symbolism and introducing modelling.

One suggestion, from several people, for improving the program, was to keep each equivalent equation on the screen. This would make the program more like AlgebraLand (Brown, 1984), which shows a trace for each solution attempt. However, such a trace would, in this program, not help to introduce or address target concerns. Moreover, by increasing the apparent complexity of the situation each time an operation was attempted, the trace would interfere with the theory that the Leibniz method can make the situation simpler. On the hand, it must be admitted that such a trace is an attractive device that may clearly aid the understanding of more complicated operations than addition and subtraction (such as factorisation, or formula rearrangement), because comparing a resulting equivalent equation with the previous equation can be made easier if the previous equation can be seen rather than remembered. This would have an advantage over an "undo" facility, in that it would be possible to compare the equations visually term-by-term. Longer traces would later enable the comparison of different solution methods, perhaps made a concern by requiring the shortest solution path.

It was found at School D that students would enter answers such as "34/7b" or "34kg/7" rather than "34/7", and that these answers were accepted by the program. The first answer could in principle cause great confusion if transferred to paper; while the second is at best clumsy, and is counter to the principle of dealing with pure numbers rather than quantities in order to aid the transfer of strategies from balance puzzle to equation. The program was duly altered to allow only the third type of answer.

For some teachers, the issue of the level at which to start students weighed heavily. Although this has to be a matter of judgment, it should be noted that someone who has thoroughly grasped the appropriate strategies can take just 3 minutes to get to the first symbolic level. On the other hand, when the students at School D started their second lesson, they were encouraged to experiment with the Enter an Equation button as a challenge to find an equation as hard as the ones they were solving the previous lesson, and some of their entries are very interesting in terms of what they might indicate about their understanding. When the students at School E started their second and third lessons, they were given equations to enter. This was problematic for some, who may have benefited from revision of the early levels, but there were constraints of time (algebra was not strictly part of the Year 7 curriculum).

So for example:
$50 + 11b = 34 + 13b$ would take them to Level 4 (balance puzzles with a two-barrel difference);
$98 + 3b = 46 + 15b$ goes to Level 7 (balance puzzles with fractional answers);
$2.1 + 5x = 1.4 + 12x$ goes to Level 10 (symbolic notation);
$5 - 6x = 2 + 17x$ goes to Level 12 (equations involving negative signs).

Should they end up at too high a level, the Give Up button allows movement down the levels.

5.2.4 The Modelling Levels

A science teacher at School B thought at first glance the balance pictures were to introduce the idea of moments. This misconception inspired the inclusion of a wider range of puzzles via the modelling levels. The research literature yielded a large number of problems, but a decision had to be taken about which word problems were to go into EQUATION, and which into the test (see later). Contingent limitations of the symbolic algebra program module meant that problems for the computer would have to be considered solved by the student entering a single numerical answer. Concern mechanisms are similar to those listed earlier, with the addition that the continued availability of the – and + buttons (that have been so useful) is intended to bestow confidence.

Other decisions centred on the role of equations in the solution. Firstly, in order to conform to the principle that students must discover that their existing strategies are inadequate, it was not to be a requirement for the Model button to be used to solve the problem. This meant that some students initially used other methods (such as blind numerical trial-and-error, operational trial-and-error, trial-and-improvement, or parts reasoning). When they subsequently became stuck on higher levels, it was often necessary at School D not only to point out the Model button, but also to redirect them to a previous level in order to explore how it might work.

A second decision was that the computer would disallow *illegal* equations (giving feedback on incorrect syntax, e.g. displaying a message saying that a letter was required, or that there had to be an equals sign, that only one letter could be used, or that the equation was just not understood by the computer), but that *inadequate* equations would be allowed (i.e. the equation was syntactically correct, but would not obtain the correct solution to the problem when solved). Illegal equations were treated this way because the program would be sometimes used by those who had never formulated an equation before, and needed to be aware of the specific constraints on the form. But the acceptance of inadequate equations meant that when students solved the equation (a solution of which they would now be confident, because they were using the computer in the same way as previously), their attention would almost inevitably be pulled back to the original equation. They would then be able to improve the model. It is much harder to simulate this process when using pencil-and-paper; firstly, because there isn't usually any immediate feedback; secondly, because students are not confident that they have solved the equation correctly; and thirdly because they cannot easily see the effect of re-working the equation.

Why was it expected that those without prior algebraic modelling experience would be able to cope with these levels without teacher intervention? It was with this question in mind that at first some of the problems that were judged to be more difficult to formulate were moved to another level; and then the problems were graduated over four levels (the "types" of problems were listed in the previous chapter), with the first modelling level (13) containing only word descriptions of balance puzzles. Therefore, the students have been solving balance puzzles over Levels 1-8; they have, in effect, seen these balance puzzles being represented symbolically, for the first time on Level 9; they have solved a variety of equations, many of which could be interpreted as representing balance puzzles; and the first word problems they have to solve are descriptions of balance puzzles.

Some of the problems at higher levels (e.g. "Think of a number") started to cause difficulties for those students who earlier had had successful strategies by thinking in terms of combinations of objects, but who now had to construct each side using operations on an unknown number. It may have been preferable to find some way for this challenge to the object interpretation of letters to have come prior to modelling (which has its own complexities).

Moreover, it was found at School D that there were difficulties with problems for which the obvious choice of unknown quantity to be labelled by a letter was not identical with the unknown quantity requested in the word problem. For example, the driving problem caused great frustration for many students when it asked how far had to be driven *after* lunch rather than before. Not only did they have to contend with formulating an equation expressing a more complex relationship than earlier, but they also had to realise that if they constructed the equation $d + 7.9d = 445$, their answer would have to be multiplied by 7.9. Even if they noticed the discrepancy, most were tempted to try to re-formulate the equation rather than use the found quantity to carry out the simple calculation (in most problems) that would have obtained the quantity required. For this reason, the program was modified for School E so that the word problems were graduated over 6 levels rather than 4; and problems with a potentially problematic choice of unknown were reserved for the higher levels.

Another design decision was that once an equation is formulated, each side is simplified to standard fixed linear form so that it can be solved using the strategic theories already learned for formal operations. In this way, it is made easier to formulate and solve an equation than to use guessing, numerical trial-and-improvement, operation trial-and-error, part-reasoning, simple arithmetic, graphs or inversing. So for example, in the wallet problem, brackets can be introduced prior to any rules about how to simplify or expand them. Students did not seem to find this automatic simplification confusing on the whole, but one or two students (and David Hewitt, creator of *Grid Algebra*) noticed that, because of the fixed order for constants and unknowns, the program in effect rearranged their entered equations for no apparently good reason. Another criticism that could be made (and one that was made against computer algebra systems earlier, in the context of formal operations on equations) is that learning how to factorise and expand are not promoted as strategic theories because the computer does it for them. Nevertheless, there does not seem any great harm (and some advantages) in using brackets to represent situations before one knows anything about how to deal with them formally.

The choice of names of individuals featured in the word problems was tricky. Rather than attempting to stick to the original names found in the research (which may have been tied to particular circumstances) or rewriting the problems to be name-free (which would have made the problems more wordy and less attractive) - names were randomly selected from a bank of original names, and those students who piloted the program. This had the advantage of not favouring any one name for any particular problem, thus neutralising any potential for recall of strategy being based on recall of name. The equal likelihood of the names being male or female also minimised possible gender stereotyping. Although the actual scenarios might not be free of gender bias (scarcely yet "cultural bias"), an attempt was made to create a number of scenarios for each type of problem, and each level selected from an appropriate bank of problems. Of course names and scenarios could be added to the program - the teacher could even prepare a file of such problems, which the students then load using the Get a file of equations button. However, unlike the built-in problems, the program cannot randomly vary the parameters, or randomly select problems from the file.

The numbers in each built-in problem are randomly generated within sensible limits. This was in order to minimise numerical recall as a strategy, although little could be done about operational recall, beyond creating a number of similar problems involving slightly varying relationships.

When piloting the program, two adults who had previously expressed great dislike for word problems at school progressed very fast through the levels, and particularly enjoyed tackling the word problems using the program. They were soon building up expressions on Level 16. The feature, mentioned earlier, whereby no buttons ever become inactive on the modelling levels, was found to be extremely interesting to a number of those using the program, who took advantage of this to experiment with simple algebraic manipulations. One feature that was added after this was that the previous equation could be clicked to recall it for editing, thereby making it easier to improve on the model, because it did not have to be retyped.

Nearly all the teachers who tried the program expressed approval of EQUATION's approach to introducing algebra; on the other hand, several advisers who tried the program wanted to know what the prerequisites for the program might be. Perhaps the demand for "prerequisites" is symptomatic of a concern that students with certain existing strategic theories may have unfortunate experiences with the program; but since they consistently failed to specify what these theories might be, perhaps it is symptomatic of a view (which Vygotsky, 1978, characterises as typically Piagetian) that students should not face problems before they are "ready" in terms of cognitive maturity. Meanwhile some mathematics education researchers who tried the program independently pointed out that the program could not be considered to encompass all algebra. The issue as to whether the program "covers" all algebra was not one that was ever given any priority - after all, what could such a program look like?! Nevertheless, this comment could be aiming (quite fairly) to determine what aspects of algebra are *not* addressed by the program. The apparently obvious suggestion that it does not help with dealing with variables is, as has been discussed in the previous chapter, one of the conjectures that is being tested.

5.2.5 Further Piloting and Changes

The only strongly negative comments came from a teacher whose school had just purchased an Integrated Learning System, and she considered that using another computer program for mathematics would confuse students. In particular, the ILS introduces new topics "only when they're ready", and who is to tell whether they are ready for EQUATION? This would appear to rather diminish the learning-management role of the teacher. It should be noted that this teacher had not yet used the ILS with any of her classes, nor did she know how to select algebraic topics as opposed to any other.

One teacher proposed a printed certificate indicating students' progress, demonstrating that the students have not been wasting times in lessons. On the other hand, the level is a pretty good indicator of what has been grasped already, and the last two digits of the score give an indication of progress through the level. Another teacher suggested cross-referencing against the National Curriculum and providing a list to the teacher indicating what skills each level introduced. The most obvious relevance of EQUATION to the National Curriculum is at NC Level 6, which demands the ability to "formulate and solve linear equations with whole number coefficients". This then provides the foundation for simultaneous equations, re-arranging formulae, and solving inequalities at higher levels. However, it is a corollary of this thesis that EQUATION can in fact provide a purpose for symbols to describe situations and hence the use of letters to represent variables. This, then, has immediate relevance to the KS2 Programme of Study in Number: "Pupils should be given opportunities to… progress to interpreting, generalising and using simple mappings, e.g. C = 15*n for the cost of n articles at 15p*, relating to numerical, spatial or practical situations, expressed initially in words and then using letters as symbols".

The two 9-year-olds at School C (see later) were able - given an hour a week, for 4 weeks - to get to Level 14 of the program, with minimal intervention from the researcher. This confirmed that the program did not require major changes to improve accessibility. The students also showed great enthusiasm. On the final day, they showed such a determination with the program that they worked furiously for 75 minutes without a break.

When the program was used at School D (see later), it was found that the first lesson was a great success in terms of accessibility - all the students were on task for a whole hour, and 90% got to the modelling levels. The second lesson was less successful - from Level 14 onwards severe difficulties started to emerge, and off-task behaviour increased dramatically. Nobody managed to solve a level 16 puzzle unaided. The increased frustration that was noted by both researcher and teachers led directly to the subsequent further graduation of the word problems. An additional suggestion from one teacher was to have the modelling levels accessible through a separate button on the program, so that the student would have the satisfaction of knowing that they were now accomplished in solving simple linear equations, and any frustrations felt with modeling problems would not affect that feeling of satisfaction.

5.3 The Development of the Test

The choice of individual items for the test - and the knowledge they might indicate - has been largely dealt with in chapters 2 and 3. This section considers the test as a whole.

The test was intended to provide a snapshot of the algebraic problems that a student could solve; to supply clues as to the strategies that the student might be using; and hence to indicate potential improvements as a result of using EQUATION. However, as will be seen when the subtleties of individuals' work in class are compared with what they write down when attempting the test, such a "before and after" written test that can be completed in an hour and that attempts to test all of initial algebra - including equations, expressions and variables, although excluding functions and graphs - is only a crude instrument for detecting improvements in strategic theories. Why then use a before-and-after extensive, 45 minute written test?

It is clearly not because of some notion that the only appropriate research techniques for assessing student's understanding are quantitative measurements (Popper, 1934, 1963). The main reason is simply that of ease of comparison with other studies. The research issues, the conjectures, the interventions, the tests, the data collection arrangements, and so on, do not form themselves in a vacuum; nor are they predetermined by some sort of logical dissection of the meanings of words. They have arisen as a result of complex interactions between judgments about the best ways to illustrate the perspective, constraints on what is feasible in practice, negotiations with the teachers involved in the study, expectations about what is possible and expectations about what might happen; hence this chapter attempts to capture the most salient aspects of such interactions.

Since some conjectures related to the potential for transfer between different problems, an extensive test was devised in preference to testing in greater detail only those aspects of algebra that EQUATION ostensibly aimed to improve (i.e. equation-solving and modelling). This did, however, limit the extent to which item variations could be used to explore more precisely the limitations of the relevant strategic theories. The items used in the interviews were different versions of a small number of the written items. The interview schedule was designed with a eye to it being used as a working document for discussion about differences in responses in the post-interview.

There was limited room in a test that had to be completed in under 45 minutes; a time limit set so that it would fit into a single lesson and so not take up too much of the Year 7 students' time (for whom algebra was not an explicit part of that year's scheme of work), nor add too much to the already large test burden on the Year 10 students. So, what with the need for a clearly legible font, reasonably-sized diagrams and spaces for "working out" included on the test paper (to minimise student effort in moving between a question sheet and an answer sheet), the original 11 pages had to be reduced substantially. This was done by removing items that seemed similar (using judgment about the similarities of the strategic theories that were expected to solve the items), and splitting the test into two - the second test was intended to contain those items that only someone who was able to get most of the first test correct should tackle. In the event, the maximum score of any student in the study on the first test and first attempt was 74%; only around 10% of the students got over 60%; and those who were given the second test found little they could do; so it was not thought worth using the second test for comparisons.

Piloting of the test with a variety of individuals suggested that 45 minutes was a reasonable length of time to tackle the test. However, when the final version was used as part of a regular school lesson of an hour's length at School D, great variations in working speed were found. Some students ran out of time, and speed of response was not intended to be a factor here. The number of items (43) is also deceptive. The first section (Patterns) consists of 2 groups of linked items (7 in total); the second section (Modelling) contains only 3 items, but may each require a much longer solution time than other items; the third section (Representation) contains 12 items, many of which involve short symbolic answers, but also reading a description or studying a diagram of a situation; the fourth section (Transformation) contains 21 items, but many can be answered quickly with only a moment's thought. The number of items could probably have been reduced to a smaller set, like the items used for interview; but Year 7 at School E was given the same test as Year 10 at School D so that comparison between the two age-groups would be possible.

Could a more compact test have sufficed? Looking at the quantitative results obtained for these four classes, little would have been lost by omitting Patterns items A1(i), A1(ii), A2(i), A2(ii) and A2(iii); Modelling items B1 and B3; Representation items C1(i), C1(ii), C1(iii), C2(i), C7 and C8; and Transformation items D2, D4(i), D4(ii), D4(v), D6, D7, D8, D9, D11 and D12. These 23 items were largely unchanged between test sittings, and the correlation between the test without them (20 items) and the test *with* them (43 items) is about 0.98 (and not less the 0.9 for any one class). Qualitatively, however, there are interesting features about the lack of change shown in these items, and about strategic theories and concerns evidenced by responses to them.

Although all the Year 10 students would have nominally covered all the topics addressed by the paper, only a few of the Year 7 students would have met any of the topics. The order of items was therefore important to avoid students accidentally missing out items that they might be able to attempt, because of worries induced by so many alien questions. The Patterns and Modelling items - with the exception of A2(iv) - do not require algebra in the item text or solutions; so these sections therefore had to come first. Since the Patterns items required fewer steps, these were put on the first page, and Modelling on the second. Although the Representation items involved more text than the Transformation items, they were considered marginally easier; and so came next. In each section, an attempt was made to put the items in ascending order of expected facility, to minimise student anxiety; although in the case of linked items (A1, A2, C1, C2, D1, D3, D4) the ascending order was internal. On the evidence of the Year 7 and Year 10 classes, this judgment was about right, although C5 and C7 were found to be harder than C8, and the order for D4 onwards was mistaken.

The choice of word problems for the test was a difficult one. A1(iii) is a variant of the patterns questions suggested by Küchemann (1981); the sequence B1 is from Sutherland (1993); B2 is from Lins (1992); while B3 is a "Think of a Number" problem that can be modelled by an equation of the form $ax + b = cx + d$. B2 is found in EQUATION; B1 and A1(ii) are not; while there are variants of B3 in the program. There were many other alternatives choices here.

One error has already been mentioned - the accident by which the order of unknown and constant in D3 was the reverse of that arbitrarily fixed in EQUATION. Another error related to the creation of a post-test. On the whole, the wording and the structure of diagrams in the post-test stayed the same as the pre-test; but numbers and symbols were changed so as to minimise the possibility of recall. One error in preparing the post-test was that the rectangle whose area had to be found in C1(iii) changed from of a width $3b$ and length $4b$ to a width of $4c$ and a length of $3d$. This erroneous change clearly resulted in improvements in this item's facility in all classes. There is always the possibility that other changes can cause more subtle, undetected effects.

The test items are not identical to those found in other studies. Wordings and diagrams were changed in relatively minor ways to minimise potential linguistic confusions. For example, compare B2 with Lins' original formulation which included arrows pointing downwards, referred to "Sam plus bricks" with 189kg just below, did not include scale pans but did include a second diagram showing the balanced situation. Finally, it should be pointed out that the test results were presented to the students' teachers as a means of formative assessment.

5.4 The Development of the Questionnaire

The first phase of the fieldwork at School A aimed to see what was realistically possible to find out about the theories that relatively successful students (i.e. A-level students) might have about what sort of thing an equation is; and to construct an interview schedule for accessing such theories. 8 students (chosen by the teacher for an interesting cross-section) from two Year 12 classes were interviewed in pairs about their "consideration" (perceptions, experiences, rationalisations, opinions) of objects, processes and relationships, reasons for changes in knowledge. So, for example: "What sort of thing is an equation?", "What verb do you connect with equations?", "What does the equals sign tell you?", "What do the letters mean?", "Is an unknown called that because it doesn't have a value, or because you don't know it?", "Do you think there are equations to which no-one knows the answer?", "How are equations and graphs connected?", "How do you like equations? Have you always regarded them this way?", "What are the most interesting things you have done involving equations?", "What do you think about the view 'Students should not use calculators or computers until they have understood and learnt the ideas thoroughly.'?".

Interviewing pairs was found to be helpful in that there was less pressure on each student, and the disagreements and clarifications of each other's responses (even when one student did much more of the talking) were very informative. This pilot work was also helpful in improving interview techniques.

Yet as has already been indicated, later work became increasingly sceptical about the role of such abstract rationalisations in students' mathematical understanding. So while the data which resulted from the interviews could be analysed from many possible perspectives and with many research questions in mind, when analysis was limited to considering just the questions of immediate relevance to this research, the data proved less than complete. It was only teachers' comments, for instance, that provided the insight that the students did not necessarily connect the end results of trial-and-error, simplification, factorisation, crossing the x-axis and numerical methods. The relationship between consideration of equations and the theories used to handle equation-related problems was simply not uncovered by the interviews. Moreover, the evidence did not seem to support the view that students had a well-established and consistent perception of an equation. Students also found it difficult to recall clearly improvements in their own knowledge and to relate them coherently to potential causes of improvement. These negative results seem to suggest that the proper context for examining algebraic understanding should be that of tackling algebraic problems rather than that of recollecting past experience. Students' theories are, it has been argued, far more accessible through observing their active strategy-making than through relying chiefly on their rather faltering and often contradictory responses to interview questions about the meanings of words

Nevertheless, some of the interview questions were adapted for use in interviews with students before and after using EQUATION. But rather than seeking stable images that could "explain" or "underpin" understanding, these later interviews sought theories *about* equations primarily as insights into concerns. The intention was not to find out how and why "students' views about what sort of thing an equation is" change, linking up with mathematician's views about what sort of thing mathematics is, and educators' views of what education is; but to identify the concerns that such views address, and their relationship to the strategic theories that enable participation in mathematical activities.

On the other hand, this work demonstrated that relatively unstructured audio-taped interviews - of around half an hour - with pairs of students on decontextualised issues of consideration can reveal some features not only of theories about algebraic objects, but also of algebraic concerns. For example: none of the students even referred to terms, operations or expressions. Only one student used the word "solve" to refer to equations, despite the interviewer's unregulated use of the word. No-one referred to defining, assigning, equating or expressing. Proving only arose when initiated by the interviewer, and one teacher (there were two formal interviews with the students' teachers) suggested that students tend to be unconcerned about the proofs they are shown. Expectations of the usefulness of equations appear to be generally low: other than seeking "formulas" in investigations or coefficients for standard curves in Physics experiments, students did not seem concerned to pose problems solvable using equations or to form equations. A teacher ventured that students say they have never made up equations because they see investigations as describing a pre-existing rule. As she put it, the students "just want to get the right answer" and lack a "need to want to know". On the other hand, students were not observed in situations where their control over the problem agenda would have been sufficient to allow them to demonstrate such concerns.

Moreover, since the need to get beyond clichés and truisms into "underlying" views of what the students "really" thought an equation to be was no longer quite so pressing, the obligation on the part of the interviewer to be proactive in helping the students' towards introspection (yet without unnecessarily initiating ideas) decreased; so that it was eventually possible to countenance a written questionnaire that might obtain data from a less narrow range of students, although data of admittedly less richness. Questionnaires were trialled with the A-Level students. As expected, the richness of the answers was reduced, but much of the diversity and commonality seemed to be still apparent. These questions were refined in the light of the questionnaire; and several questions were discarded or combined because vital questions were missed out when the interviews overran. The interview questions were simplified for the Year 6 students at School C; while the questionnaire was simplified and shortened for Schools D and E. More in-depth interviews with some students during the later fieldwork allowed distinctions to be made such as that about "method" - a method for finding answers in a practical situation, a method for solving an equation, a method as a set of arithmetic operations on numbers or letters standing for numbers, etc. As will be suggested below, however, interviews and questionnaires are not perhaps very informative when given to students who have had little explicit algebra teaching. Nevertheless, subsequent changes after using EQUATION can be interesting.

5.5 Data Collection

5.5.1 School A

Because the point of operating on an equation - making solution easier - often seemed to be lost in the struggle to remember algebraic rules and deal with arithmetic, the early pilot work for this research asked: what "images" do reasonably successful equation-solvers have of an equation that enable them to make sense of algebra? It might then be possible to provide a variety of experiences to suit these different images. It was expected that "balance" would be one image, perhaps "function machine", "graph" and "statement" would be others.

The earliest study involved interviewing A-Level students at an 11-18 comprehensive in a town in Oxfordshire, and observing their lessons. The 4 pairs of students were chosen by their teachers as exhibiting a variety of approaches to learning. The interviews were followed by observation of two Year 12 classes. The students undertook coursework over three to four weeks into numerical methods for solving equations. They used graphics calculators (to explore graphs, and produce tables of values), Excel (to speed up calculations in the various methods) and Derive (to differentiate functions and draw cobweb diagrams), including worksheets and suggestions. I sat in on the lessons, talking to students about their work, helping them with any technological problems, recording their conversations about equations and related topics, and taking notes. Video-recording was briefly considered as an option, but would perhaps have added little, and may have acted as a distraction over such a small time-scale. Audio-recording of interviews was used, but proved to be of limited value in the classroom. The teachers gave brief comments about individuals' progress after each lesson. There were then brief post-interviews with students, and longer teacher interviews.

However, the results suggested that although a case could be made that the anticipated images (and some others) were "there" in some sense, they did not seem to play the expected role of underpinning thought processes. It seemed perfectly possible to be unable to explain clearly what sort of thing an equation is, or to rationalise experiences using incommensurable images, and yet *still* be able to do all the mathematical things with equations that one might expect from somebody with "understanding". When even quite reflective and eloquent adult mathematicians were unable to match up to the exacting standards apparently required for "relational understanding", it was realised that such images were not after all pivotal in knowledge or learning. If learning could not be portrayed in terms of the grasping of images, there would be fewer cognitive insights into struggling beginners gained from observing A-Level students using technology for numerical methods than from observing Year 8 students learning about equations for the first time.

Moreover, it would be important to gain insight from students' talk as they worked on problems, otherwise there would be little evidence of success with regard to detecting improvements in theories and concerns. It would be important for both teachers and researcher to *tell* students less, and instead discuss with them what is being asked, and what they might do. Steps would be taken to promote more individual problem-solving, without lengthy introductions from the teacher - hence the development of a specific activity. Observation might also perhaps have gained greater insight into exactly how, when and why theories change by continuously recording the interactions between selected individuals, rather than attempting an observation of the whole class.

5.5.2 School B

So when EQUATION was trialled at School B, the data collection was of this style. The students worked in pairs, and their talk while solving the problems was audio-recorded; but the computers were sited in the library, and the silence of the room may have intimidated them somewhat. They were more communicative when there was background noise, but the recorder consequently failed to pick up this discussion. It may be preferable, therefore, to have the students work on the problems in their normal surroundings, and use personal microphones. An atmosphere of informality in other trials of EQUATION seemed conducive to discussion. The 4 pairs of School B students (one pair from Year 8, two from Year 7, one from Year 6) were selected by their teachers without negotiation - probably on the basis that they were seen as able students, although it would have been preferable to have a wider range of students.

5.5.3 School C

The first substantial study was at School C in June 1997. As with Schools A and B, this primary school in a town in Oxfordshire was chosen purely because it was local, taught children of the appropriate age, and had enthusiastic teachers who wished to make better use of computer facilities.

The study involved just two Year 6 students, selected by their teacher as likely to appreciate the greater challenges that EQUATION might afford them over and above their work in class. After pre-interviews, they used the program for about an hour a week for 4 weeks. They were then re-interviewed to find out how their views and responses to specific items had changed. They were given items to do by hand, and also with the aid of the computer.

This was the first opportunity for the log file to be used. The idea of the computer recording in a log file all the students' interactions with EQUATION was a valuable data collection technique. It meant firstly that there was at least some indication of every student's experiences, no matter how many students were in the class; secondly that detailed observational notes did not have to made on what was visible on the screen at a particular time; thirdly that the study was not limited by the number of observers, audio-recorders or video-recorders; and fourthly that it was possible to replay, at any speed (including real-time or in synchronisation with a tape counter), the puzzles faced by a student or pair, the actions they took, and what they saw on the screen. Jumping to any level, puzzle, step, time or counter value is allowed, as is searching for particular actions or types of problem. Other tools allow the creation of a "script" that summarises these interactions at various degrees of detail and also gives the times, a chart for showing how long each puzzle took, and a chart for showing how long each level took.

Of course there are disadvantages of such arrangements. The impression might be gained that only what is logged is of interest; whereas the discussions, mutterings, glances and gestures that are not logged may be very valuable in conjecturing the status of the problems being addressed. Although the discussions and mutterings of the two students at School C would have been picked up on the audio-recorder, video might well have been seen as intrusive (in a way that the audio-recorder was not), and so glances and gestures were observed by the researcher. However, inconspicuous, unthreatening note-taking turned out to be difficult to do, while at the same time attempting to keep up with the almost continuous banter of conjecturing. Video would have allowed repeated viewing.

The data, then, consists of logfiles and audio-recordings of the two students working at the computer, audio-recordings of pre- and post-interviews, and field-notes of observations and items attempted.

5.5.4 School D

The second substantial study was at School D - a 13-18 comprehensive in a large town in Oxfordshire. This school was chosen because a particular teacher working there was already collaborating on a number of other research projects, and had expressed a particular interest in helping his students with algebra.

The study involved two classes of Year 10 students. One class was to use EQUATION, the other would act as a control group that would do no algebra at all. This was because it was suspected (and later confirmed) that some items on the test would show improvement merely because of a repeat sitting. Moreover, the questionnaire results may have been affected by the sitting of the test. However, because of the difficulty of obtaining two “parallel” groups able to be part of the study, at a time convenient for observation, the control group was taught by a different teacher.

Having two groups does not solve any of the “internal validity” problems of Campbell and Stanley (1963), but it does enable one to be a little more circumspect about such issues as history and maturation. Nevertheless, given the variety of different activities available in the research literature, and the flexibility that the theories-concerns instrument provides in comparing learning outcomes, it is a task for further research to compare directly the EQUATION group with groups using different activities - either activities studied in the literature or the algebraic introductory activities that a class would usually have. It is only once a prima facie case of improvement has been made that such a task would be appropriate.

Two friendship pairs of students were selected from the EQUATION group by their teacher, on the basis that they offered “interesting ways of thinking” but were also able to articulate their thinking. They were interviewed for about half an hour; the first part of which centred on selected consideration questions; the second part on selected test items. Both classes then completed the questionnaire, followed by the test. This order was so that anxiety induced by the test did not place undue influence on questionnaire responses. Certain students were also given the second test (see above) if they had completed as much as they could do of the first test, if they asked for it, and if the teacher considered this a serious request.

The EQUATION group used the program for two lessons; the control group tackled questions from past GCSE papers that did not involve any algebra, and were mostly to do with the "shape and space" part of the syllabus. The students in the EQUATION group were encouraged to work in the same pairs for each lesson. The four selected students were audio-taped working at the computer, and encouraged to talk about their thinking; while all the students were logged. The researcher and teacher observed a self-imposed restriction that they would give minimal specific help to the students, beyond asking the same questions back, or (as a last resort) pointing out buttons or sentences on the screen that may have been overlooked. This ordinance broke down somewhat in the first lesson, when it was found necessary to intervene to point out that fractions could be entered, and when the students came across negative solutions and it was then decided to ask "What could you do now?" if they were faced with equations like $3x + 12 = 0$. The ordinance also broke down in the second lesson for three types of events: when some students had to be directed to work through an entire level rather than continuing to try to find an equation that could be entered to get them past level 12; when some students were redirected back to level 13 to work out how the Model button worked (see above); and when a complete impasse was reached on problems on Level 15 or 16.

The selected students were then re-interviewed, also drawing on their responses in the pre-interviews where there appeared to be change. Both classes then sat the post-test, some two weeks after the pre-test, in order to identify improvements in strategic theories and concerns. The questionnaire was not given because it was thought too soon after the last one to make any long-term difference, and to avoid overloading the students. Instead, the questionnaire and a delayed post-test (which was the same test as for the pre-test) were given 6 months later. However, this decision was reconsidered for School E, as it was realised that so many other things may have gone on in the meantime that it would be difficult to ascribe differences to EQUATION. Moreover, the EQUATION classwork took place in June, the delayed tests were in December, but the students had a change of teacher in September. The conditions for the delayed test were far from ideal, owing to an error in counting the number of tests required. The delayed test results were therefore not used in any statistical analysis, although they are quoted in the qualitative analysis where relevant, and with the appropriate caveats.

Although on the network at School D the program could save the log file in the user's directory (which makes identifying the students easier), it was noted after the fieldwork, that absenteeism on one of the days would mean that it could not be guaranteed that the same pairs were working together, or using the same user ID as previously, and so the identity of the other member of the pair would be in doubt. It may not also be immediately apparent when use of the program has been by students in other classes. The program was therefore adapted to ask in future for the class and student's name.

Some of the students in the EQUATION group also made use of the OurQuestions Internet project to compare students' concern to pose and critique problems solvable using algebra.

The data for School D, then, consists of logfiles of around 25 students working in pairs at the computer for two one-hour lessons; audio-recordings of two pairs of students working at the computer; audio-recordings and field-notes of pre-post interviews with those two pairs, of a post-interview with one other student, and of a post-interview with the teacher; pre-post written tests and questionnaires for each student in the class; a supplementary written test attempted by about ten students; a delayed post-test for most of the class; and problems posed by seven students.

5.5.5 School E

The third substantial study was at School E - another 11-18 Oxfordshire comprehensive in March 1998. The study involved two classes of Year 7 students, taught by the same teacher. There were no formal interviews or audio-taping for this study, because the aim was to see whether the remarkable test findings with the School D Year 10 class would be replicated with a younger age group.

After the questionnaire was handed out, some of the students in the control group revealed that they had never come across the word "equation" before. Rather than leave this as the questionnaire response, a decision was made to indicate to the students to what the word might refer, given that they may have come across equations without calling them that. Arithmetic identities were put on the board as an example e.g. $2 \times 3 = 5 + 1$; followed by equations with one unknown, e.g. $4x + 6 = 18$; and an example of an equation with letters, e.g. $y = mx + c$. Some students expressed realisation that they now knew to what the word might be referring. The questionnaire then took about 10 minutes to complete. Several other words in the questionnaire caused difficulties, including "confident" and "statement". The pre-test was tackled in the remaining time (45 minutes). Nearly all the students had completed as much as they could well before the end of the time available. When the EQUATION group was asked how many had come across equations before, there was a similar lack of avowal, and so identical examples were given. As at School D, the students were allowed to use calculators, but they were strongly encouraged to show how they worked out their answers. They were also told that there would be quite a few problems that they had not been taught how to solve, but that they should have a go anyway. The word "algebra" was never used, not only so that any method of obtaining an answer was seen as valid, but also so that those who had not done any algebra would not worry unduly about a topic which is sometimes trailed as "very difficult" by older siblings or parents.

The EQUATION group used the program for three lessons; while the control group also used computers for their lessons, but to do non-algebraic work on databases. The classwork was followed by the same questionnaire and then a post-test.

The data for School E, then, consists of logfiles of around 25 students working in pairs at the computer for three one-hour lessons; field-notes from observations in class and interviews with the teacher; and pre-post written tests and questionnaires for each student in the class.

Chapter 6
Analysis

6.1 Introduction

This chapter analyses the fieldwork in basic terms, giving interpretations of what happened during the classwork and in the pre-post testing. The agenda for the first part is to identify from the classwork what contextual equation strategies might have improved, and to provide some clues as to when and why. The agenda for the second part is to identify from the testing and interviewing some algebraic problems in which there may have been improvements. The final chapter will attempt to use the analysis presented here to reconcile these two parts, to discuss implications for objections to the balance model and hence perhaps to illustrate the Popperian psychological perspective.

6.2 What strategies improved during the classwork?

In order to analyse the learning that took place during the activity, annotated logs are provided in this section, on the left-hand side of the page. The logs relate to a pair of students, working together - Rebecca and Nicola in the Year 10 class using EQUATION (10EQ). This pair was one of those whose conversations were audio-taped, and who were interviewed. In the analysis that follows, the text on the right-hand side of the page provides details about the level, an interpretation of the students' actions shown in the log, and supplementary evidence from the audio-recorded conversations. The audio-recording (still less its transcript) makes little sense without a simultaneous presentation of the puzzles being tackled. The puzzles are numbered for ease of reference; and the times taken to accomplish each action are given in seconds.

6.2.1 Simplifying Balance Puzzles

Levels 1-7 of EQUATION are all balance puzzles, and aim to promote the use of a simplification strategy.

Level 1: Balance Puzzles E + b = F & Kb = E

Rebecca & Nicola

#	Time	Puzzle
1	32	2 + b = 20
	23	*Guess: 18*
	9	*Continue*
2	21	18 = 2b
	16	*Guess: 9*
	5	*Continue*
3	10	5b = 15
	8	*Guess: 3*
	2	*Continue*
4	13	3 + b = 13
	12	*Guess: 10*
	1	*Continue*
5	11	10 = 5b
	9	*Guess: 2*
	2	*Continue*

The simple Level 1 balance puzzles are of the form E + b = F and Kb = E, where "E" means a weight - anything up to about 20kg at this level; "b" means one barrel; and "Kb" means a number of barrels - up to four at this level. These puzzles have integer answers <20. So arithmetically there is only a single calculation to do (either F – E or E ÷ K); although of course "counting on", recall of number facts and matching of objects are among the many informal strategies that might be used.

Rebecca and Nicola work out very quickly that the problem is to find the weight of the barrel; begin to solve the puzzles with little difficulty; and are therefore quickly promoted to Level 2. From the tape it appears that Rebecca is doing all the work - she's describing the operations and saying the answers; Nicola agrees or seeks clarification, but does not initiate solution. On the other hand, the microphone is attached to Rebecca and does not pick up Nicola's voice very well.

Another pair - Rajiv and Seb - note after four Level 1 puzzles that "we're good at this". After five, Rajiv comments "not very hard, is it?". They solve balance puzzles like 3b = 15 by finding the arithmetic calculation, which here would be 15 ÷ 3.

Level 2: Balance Puzzles E + K*b* = F & E + K*b* = (K + 1)*b*

Rebecca & Nicola

#	Time	Puzzle
7	29	19 = 1 + 2b
	27	*Guess: 9*
	2	*Continue*
8	15	15 = 7 + 2b
	13	*Guess: 4*
	2	*Continue*
9	20	2 + b = 2b
	17	*Guess: 2*
	3	*Continue*
10	14	6 + 2b = 3b
	12	*Guess: 6*
	2	*Continue*

Puzzle #7 requires two arithmetical steps, and so it is not surprising that it takes them a little longer; but further puzzles at this level are found quickly. It is perhaps a little surprising that the puzzles with *barrels* on both sides are solved as easily as the puzzles with *weights* on both sides, but maybe this reflects the accessibility of the context.

At the moment the puzzles are accessible to informal strategies such as guessing, cover-up, matching or part-reasoning. There is little indication which strategies Rebecca and Nicola are using, which is unsurprising given that all these strategies are intuitive and mental. Other than barely audible mutterings, the same is true of the other 10EQ and 7EQ students. Of course had there not been a requirement of the research to avoid as far as possible drawing explicit attention to method, it would have been possible to ask them how they solved particular puzzles. Nevertheless, the existence of these informal strategies is not in doubt - and several were revealed in the pilot work for this research.

The two buttons labelled Take off weights and Take off barrels appear on this level, but most students ignore them for the moment because they are having little difficulty with the puzzles. Rajiv, for example, says "Take off weights" when the first Level 2 puzzle appears; but does not use this button until Level 5 - the subtractions are easy enough for mental calculation until then. However, the first puzzle they face with barrels on both sides ($2b = 8 + b$) causes them to use the Take off barrels button; which they continue to use thereafter. It is to be expected that the more successful the informal strategy, the later the regular use of the Take off… buttons. All but two 10EQ students start to use both buttons by Level 5. Debbie and Liam (working separately) use the Take off weights button on Level 3, but do not use the Take off barrels button in any sustained way. Liam uses $\boxed{-}$ to subtract barrels on Level 8; and Debbie uses $\boxed{-}$ to subtract unknowns on Level 9. Harry, on the other hand, uses both buttons from Level 2 onwards.

Level 3: Balance Puzzles E + Kb = F + (K+1)b

Rebecca & Nicola

#	Time	Puzzle
11	123	10 + 3b = 9 + 4b
	18	*Take off barrels: 1*
	18	*Take off weights: ESC*
	53	*Guess: 9*
	23	*Guess: 1*
	10	*Continue*
12	141	19 + 4b = 12 + 5b
	26	*Take off barrels: 1*
	12	*Take off barrels: 1*
	17	*Take off barrels: 1*
	56	*Take off barrels: 1*
	25	*Guess: 7*
	4	*Continue*
13	37	13 + 4b = 11 + 5b
	8	*Take off barrels: 3*
	10	*Take off barrels: 1*
	11	*Guess: 2*
	8	*Continue*
14	32	12 + 2b = 14 + b
	14	*Take off barrels: 1*
	16	*Guess: 2*
	2	*Continue*
15	57	5 + 3b = 14 + 2b
	6	*Take off barrels: 2*
	17	*Guess: 7*
	16	*Guess: 8*
	16	*Guess: 9*
	2	*Continue*
16	33	9 + 4b = 6 + 5b
	13	*Take off barrels: 4*
	14	*Guess: 3*
	6	*Continue*

This is the first level with barrels and weights on both sides.

The first puzzle presents difficulties - two minutes is spent on it. They successfully take off a barrel, without apparently realising the benefit of taking off any more. They then consider taking off weights but Rebecca points out "No, you've got to take it off of the same side that, one on each side, haven't you?". They guess 9 for no apparent reason, although the fact that the answers to the previous two puzzles have both been (coincidentally) the same as the weights on one side may be the inspiration for this. When this fails they decide to take off one barrel (Rebecca says she wants to take off two) but Nicola selects Guess by mistake. So when she enters "1", they find the correct answer by accident. Theory development has therefore stalled on this puzzle (a fact which they recognise: "Oh well...") and so will not assist them on the next puzzle, which is of identical form but again takes over 2 minutes.

Rebecca starts puzzle #12 confidently "I reckon you take off barrels until you get one each", and Nicola interprets this as meaning that they should take off a barrel. Rebecca then says "Yeah, but in equations you've got to do the same on each side, haven't you?". Nicola takes off a barrel and Rebecca concludes "Yeah, it's done it the same on each side.". She tells Nicola to take another off, which she does, and then they take off another. At this point the picture is 19 + b = 12 + 2b. Instead of taking off the final barrel, Rebecca now wants to know "What's 38 ÷ 2?". This fits however with her theory "take off barrels until [you] get one each". Nicola asks where she gets 38 from, and Rebecca replies "Dunno". A boy passes by and says "Take off one barrel.". "Yeah", says Rebecca in confident tones "take off one barrel.". There's a pause and then she asks him "What... have you done this one before?". This suggests that Rebecca still doesn't appreciate the need for the barrel to go. But when she sees the new picture (19 = 12 + b) she exclaims "Ah that's obvious now… One barrel's 7". They both cheer when "Correct" is shown. The question is now - will Rebecca continue to take off barrels piecemeal, or will she modify the strategy to take off in one go as many barrels as there are on the side with the smaller number of barrels? In fact she then says "Just take off barrels until you get there.", which might be interpreted as something like this theory.

For puzzle #13, Rebecca says "Take off 3 barrels", followed by "Take off 1 barrel". In other words, she has the theory "take off as many barrels as you can rather than "take off as many barrels as there are on the side with the smaller number of barrels". Nicola's "Why?" when Rebecca gets the correct answer indicates that she has not yet grasped the same theory. When faced with $9 + 4b = 6 + 5b$ (puzzle #16) Rebecca says "Take off fo... four barr... three barrels... no four. Four barrels.". Thereafter the more efficient theory seems to be preferred.

Rebecca and Nicola are typical in demonstrating such strategic improvements. Rajiv and Seb, for example, faced with their first puzzle on Level 3 ($18 + 5b = 19 + 4b$), choose to take off only 1 barrel (at Rajiv's suggestion). Seb then notices that 19 + 3 = 22 and 18 + 4 = 22, so the answer must be 1kg. But he remarks to Rajiv after the answer is checked, "You weren't sure about that were you?". On the next puzzle with more than one barrel on each side ($1 + 4b = 18 + 3b$), they take first 2 barrels and then one more. Thereafter they always take off the maximum number of barrels in one go, with just one slip - on Level 5, when they attempt to take off the larger number of barrels.

Level 4: Balance Puzzles E + K*b* = F + (K+2)*b*

Rebecca & Nicola

#	**Time**	**Puzzle**
17	56	24 + 14b = 50 + 12b
	11	*Take off barrels: 13*
	15	*Take off barrels: 11*
	8	*Take off barrels: 1*
	20	*Guess: 13*
	2	*Continue*
18	60	49 + 13b = 19 + 15b
	12	*Take off barrels: 13*
	45	*Guess: 15*
	3	*Continue*
19	32	36 + 9b = 46 + 7b
	7	*Take off barrels: 7*
	21	*Guess: 5*
	4	*Continue*
20	37	8 + 5b = 36 + 3b
	5	*Take off barrels: 3*
	17	*Guess: 16*
	13	*Guess: 14*
	2	*Continue*
21	18	7 + 7b = 5 + 9b
	7	*Take off barrels: 7*
	9	*Guess: 1*

The maximum number of barrels increases to 15 and the answer can now be less than 50. There is also a two-barrel difference between the sides - there are therefore three arithmetical steps involved.

Difficulties with Level 4 seem to stem from counting and arithmetic, although it might be possible that Rebecca is still using some sort of "1 less than" rule to determine the maximum number of barrels that can be subtracted. Having to subtract the weights, notice the two-barrel difference, and divide by two does not appear to hold Rebecca up at all. Nicola, however, is still struggling to grasp Rebecca's strategy. For example, faced with puzzle #18, Rebecca has no hesitation in taking off 13 barrels and doing $(49 - 19) \div 2$; but Nicola asks "How do you do that?". Rebecca's initial explanation may or may not be useful to her:

Rebecca: OK, look. Here is 49, right?
Nicola: Yeah.
Rebecca: So you minus 10 off 49.
Nicola: 39

Rebecca: 5...
Nicola: (*giggles, nervously*)
Rebecca: 9, OK? 9 off 30...
Nicola: Umm... Thir... (*giggles*) This isn't...
Rebecca: ...is 30.
Nicola: Yeah...
Rebecca: So it's balanced out, we've taken all that off... and 30 off there. So it's only, um, it's only 30 left.

But then Nicola asks the question that has really been bothering her, and thereby reveals that it is the simplification strategy (as opposed to the arithmetic strategy) that she is trying to grasp.

Nicola: Why are you minusing it?
Rebecca: (*pause*) Because then you get it a balanced equation and then you just divide it by the last two barrels left. ... 15.

So even though Rebecca has a good strategy for solving the puzzle, her rationale makes no mention at all of the idea of simplifying a situation to make it easier. In any case, Nicola is quite happy using subtraction on the very next puzzle - further evidence that grasping a strategic theory does not always depend on having a coherent rationale for it. "It's well easy, isn't it?" says Rebecca.

Meanwhile Seb, having simplified a puzzle down to $25 = 1 + 2b$, attempts to convince Rajiv that the answer is 12 by saying "twice 12 plus 1 is 25".

Rebecca & Nicola

#	Time	Puzzle
22	338	323 + 5b = 295 + 9b
	7	*Take off barrels: 5*
	169	*Guess:*
	76	*Guess: 18*
	38	*Guess: ESC*
	13	*Take off weights: 200*
	8	*Take off weights: 50*
	13	*Take off weights: 45*
	8	*Guess:*
	3	*Guess: 7*
	3	*Continue*
23	57	297 + 5b = 87 + 15b
	9	*Take off barrels: 5*
	16	*Take off weights: 87*
	29	*Guess: 210/10*
	3	*Continue*
24	32	338 + 3b = 129 + 14b
	6	*Take off barrels: 3*
	8	*Take off weights: 129*
	15	*Guess: 209/11*
	3	*Continue*
25	39	130 + 9b = 202 + b
	5	*Take off barrels: 1*
	4	*Take off weights: 202*
	11	*Take off weights: 130*
	15	*Guess: 72/8*
	4	*Continue*
26	43	321 = 27 + 6b
	10	*Guess:*
	1	*Guess:*
	6	*Take off weights: 27*
	23	*Guess: 294/6*
	3	*Continue*

Again it appears to be the size of the numbers involved that causes Rebecca and Nicola most difficulty. They start to use the coloured weight pictures to help them in the subtraction 323 – 295 in the first puzzle. But Rebecca eventually concedes defeat as far as the mental arithmetic is concerned - “Need some paper. Can’t think like this.”. They seek confirmation from their teacher that they are allowed to use paper. About 3 minutes is spent working on this one subtraction, via possible answers such as 18 and 172 (95 – 23 + 100). At this point, I intervened to check what they were doing and to point out the Take Off Weights button that had so far been unused and could be a way of saving some effort. Rebecca initially suggests taking off 300 but then sees that “you haven’t got 300” and so starts by taking off 200. A prompt from me that “you can take off more if you like” was probably unnecessary, but Rebecca then takes off 50 rather than 95, which would be the obvious choice if she had grasped the idea that one could take off (analogously to the barrels) as much weight as on the side with the smallest number of kg displayed. In other words, the theory for barrels does not automatically get employed for weights. She then suggests taking off 40, before realising that 45 can be taken off. By the next puzzle she has generated the more efficient weight strategy.

They were also shown at this point that “210/10” was an acceptable answer to the computer. They use both Take Off… buttons on the next puzzle and enter their answer as “209/11”. The bypassing of arithmetic seems to be appreciated by Rebecca & Nicola, who might otherwise have missed the point of the simplification strategy because of their fear and loathing of arithmetic. Nicola says “We’re so thick. We could have done this so much quicker”. They attempt to take off the larger number of kg on the next puzzle, and it is Nicola who points out the mistake. Their speed on the remaining puzzles on level 5 shows their confidence in using the strategy.

Some 10EQ students (such as Harry, Jack and Jane) take off the maximum number virtually from the start; others (such as Debbie, Tracy and Cedric) seem to struggle a little more before appreciating the value of simplifying as much as possible. For many students, Level 5 appears to be a watershed - the requirements to subtract barrels, subtract large weights and divide unfriendly numbers by the end of this level have led to nearly all students exploiting the $\boxed{\text{Take Off…}}$ buttons.

Rajiv and Seb have a rather different experience to that of Rebecca and Nicola. Their first Level 5 puzzle is $457 = 377 + 4b$. Rajiv initially wants to take off weights, but changes his mind. Is this because (like some of the adults in the pilot study) he prefers to test his mental arithmetic, rather than use the $\boxed{\text{Take Off Weights}}$ button? Unfortunately they appear to have difficulty in keeping both the large numbers in mind simultaneously. Despite correctly adding up the weights on each side, their similarity leads them to misremember 377 as 357, and so they conclude there is a 100kg difference rather than 80kg. When they discover their mistake, they have to keep adding up the weights on each side to remind themselves of the calculation they have to do. They resort to a calculator; and soon after they use the $\boxed{\text{Take Off Weights}}$ button, and are shown how to enter fractions. But as for Rebecca and Nicola, when Rajiv and Seb use the $\boxed{\text{Take Off Weights}}$ button, they do not automatically take off the maximum number in one go, as they have been doing for barrels. This does not mean, of course, that they do not realise that this is the more efficient strategy in terms of actions - they might reason that adding up all the weights on each side takes time, and so it may be more efficient in terms of time to take off obviously matching weights first. For example, for $321 = 202 + 7b$, they take off 200kg; for $406 + b = 66 + 11b$, they take off only 6kg (there are no other matching weights). However, even after they reach the single-weight pictures of Level 7, they continue to take off weights piecemeal. It is only on the symbolic levels that they begin to take off the maximum in one go.

Oddly in the latter puzzle, when they are left with $400 = 60 + 10b$ their initial guess is $400 \div 60$. This is despite their earlier success, and despite Seb saying "400 divided by 10" before he is overruled by Rajiv. They then guess 0.666, and when this fails Rajiv suggests taking off weights. Seb is reluctant and then guesses 46, saying "400 and 60, divided by 10". Somehow the talk about division has diverted them from their previously robust strategy. Rajiv then says "No, wait. 400… and we need 400 there…" and then enters 34. Seb asks "How did you get that?". Rajiv says "You minus 60.". So this looks like a version of the cover-up strategy: 400 balances some weights and some barrels; the weights are 60; so the barrels must be 340; 10 barrels, so 1 barrel is 34.

Their next puzzle is $7 + 8b = 497 + b$, but the two 2kg of the 497kg has disappeared off the top of the screen. Rajiv begins by yet again (for the fourth time) asking to use the $\boxed{\text{Take Off Weights}}$ button. Seb this time agrees, but insists that they should "take away 40", even after Rajiv has asked "Are you sure?". Once they have taken off one barrel and 7kg however, they quickly get the answer; and by the end of Level 5, they seem to have a more coherent simplification strategy worked out.

Level 6: Balance Puzzles with simple decimal answers

Rebecca & Nicola

#	Time	Puzzle
27	23	48 = 6 + 10b
	8	*Take off weights: 6*
	12	*Guess: 42/10*
	3	*Continue*
28	32	2 + 14b = 45 + 10b
	11	*Take off barrels: 10*
	8	*Take off weights: 2*
	10	*Guess: 43/4*
	3	*Continue*
29	23	42 + 2b = 7 + 12b
	4	*Guess:*
	3	*Take off barrels: 2*
	5	*Take off weights: 7*
	9	*Guess: 35/10*
	2	*Continue*
30	23	15b = 49 + 10b
	11	*Take off barrels: 10*
	10	*Guess: 49/5*
	2	*Continue*

Up until now the weight of the barrel has been a whole number. Level 6 involves fractional answers.

Each puzzle on this level takes Rebecca and Nicola around half a minute to solve, the refined strategy now consisting off: take off maximum barrels, take off maximum weight (always in that order) and enter the answer “Remaining weight ÷ Number of remaining barrels”. Nicola seems able to carry out the strategy without prompting from Rebecca, and can be heard on the tape saying the relevant numbers. Puzzles like $48 = 6 + 10b$ (#27) - in which taking off barrels is inappropriate - or like $15b = 49 + 10b$ (#30) - in which taking off weights is inappropriate - do not cause difficulties. In the latter puzzle, Rebecca says “OK then, 15... Oh no, we take it off both the sides.” before taking off 10 barrels.

Level 7: Balance Puzzles with fractional answers and a single weight picture

Rebecca & Nicola

#	**Time**	**Puzzle**
31	28	44 + 7b = 25 + 10b
	15	*Take off barrels: 7*
	5	*Take off weights: 25*
	6	*Guess: 19/3*
	2	*Continue*
32	26	40 = 1 + 10b
	9	*Take off weights: 40*
	5	*Take off weights: 1*
	10	*Guess: 39/10*
	2	*Continue*
33	27	47 = 1 + 12b
	8	*Take off weights: 1*
	8	*Guess: 46/22*
	9	*Guess: 46/12*
	2	*Continue*

Many students have realised by this level that the visible weight pictures are not as important as the total. For example, the strategy and speed of Rebecca & Nicola are unaffected by the replacement of the multi-coloured weight pictures by a single grey weight picture (and similarly for Rajiv and Seb); which might not have been the case had they been continuing to use the pictures to help them subtract. In fact at one point Rebecca says "This is so tedious". The excitement of strategy development that was so evident on level 5 and before has now been replaced by boredom with executing the same strategy again and again. Curiosity about what might be different on the next level is perhaps the only source of interest. On the second puzzle Rebecca attempts to take 40 from 40 = 1 + 10*b*, and Nicola divides by 22 instead of 12 on the third puzzle, but these are probably just momentary slips.

Rajiv has a habit of saying, in equations like 49 = 2 + 12*b*, "49 ÷ 2" rather than "49 – 2". Once or twice he attempts to go ahead with the division, then realises that subtraction is necessary, and finally remembers to involve the 12. However, more often, Seb corrects his "49 ÷ 2" with "49 – 2" so the error is not followed through.

6.2.2 Simplifying Balance-Like Equations

Levels 8-10 introduce algebraic notation. They aim to transfer the simplification strategy to equations that could represent balance puzzles.

Level 8: Balance Puzzles with $\boxed{-}$

Rebecca & Nicola

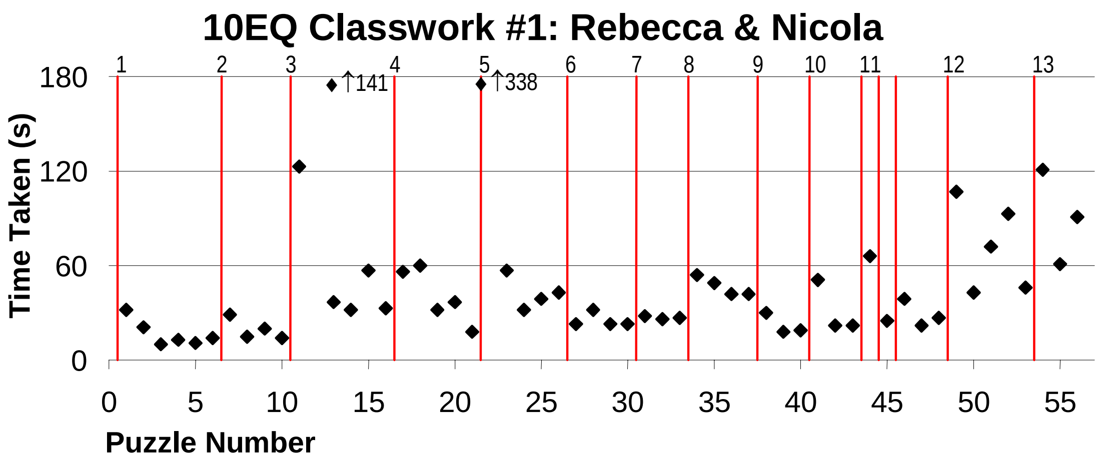
#	**Time**	**Puzzle**
34	54	269 + 6b = 129 + 26b
	12	*Guess:*
	1	*Guess:*
	21	*Subtract: 6b*
	9	*Subtract: 129*
	9	*Guess: 140/20*
	2	*Continue*
35	49	34 + 24b = 233 + 12b
	5	*Guess:*
	8	*Subtract: 233*
	15	*Subtract: 34*
	8	*Subtract: 12b*
	11	*Guess: 199/12*
	2	*Continue*
36	42	7 + 31b = 30 + 14b
	5	*Guess:*
	6	*Subtract: 7*
	11	*Subtract: 14b*
	18	*Guess: 23/17*
	2	*Continue*
37	42	71 + 30b = 209 + 16b
	4	*Guess:*
	6	*Subtract: 71*
	9	*Subtract: 16*
	9	*Subtract: 16b*
	13	*Guess: 138/14*
	1	*Continue*

On Level 8, the barrels are labelled with the letter 'b', and the two "Take Off" buttons are replaced by a single $\boxed{-}$ button. All the students soon work out how to use it. Up until this point Rajiv and Seb have been referring to "taking off" weights and barrels; as the calculations became harder they started (around Level 5) talking about "taking *away*". The few times they used the word "subtract" were when they were talking about a specific arithmetical calculation which always involved subtracting the total weight on one side from the total weight on the other; but there was no hint of removing objects on these occasions. However, as soon as the $\boxed{-}$ button appears they say things like "minus 106kg" (meaning, in context, "remove 106kg from the balance"); and "minus 16*b*" (meaning "remove those 16 barrels from the balance by clicking the $\boxed{-}$ button and entering '16*b*'."). They continue to refer to "taking away" and "taking off", but since they are now taking off maximum weights and barrels, they no longer need to refer to "subtracting".

Rebecca & Nicola		
#	**Time**	**Puzzle**
38	30	11 + 25b = 51 + 15b
	14	*Subtract: 11*
	8	*Subtract: 15b*
	7	*Guess: 4*
	1	*Continue*
39	18	37 + 9b = 87 + 4b
	8	*Subtract: 4b*
	5	*Subtract: 37*
	4	*Guess: 10*
	1	*Continue*
40	19	37 + 15b = 13 + 18b
	5	*Subtract: 15b*
	2	*Subtract: 13*
	11	*Guess: 8*
	1	*Continue*

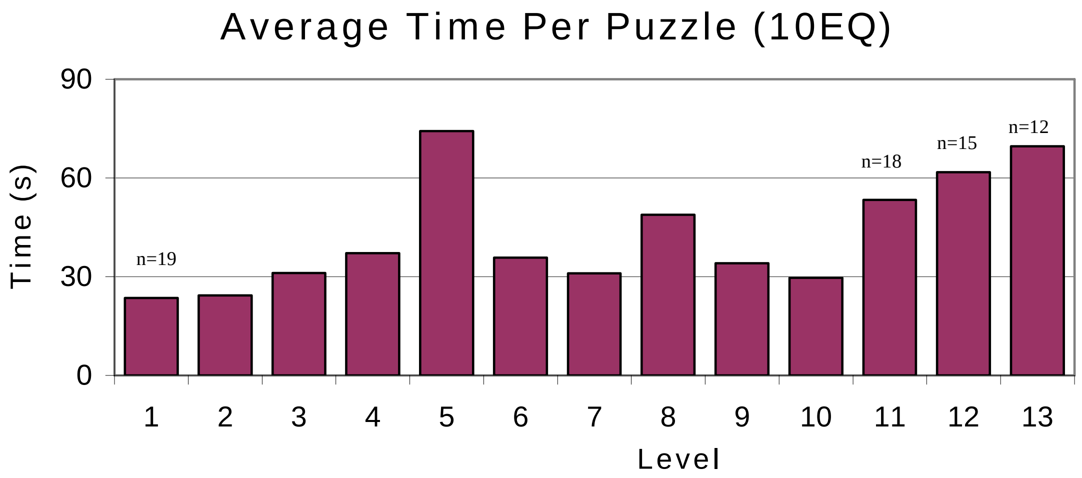
On Levels 9 and 10, the balance pictures are replaced by symbolic notation. The move from pictures to symbols does not cause comment from Rebecca & Nicola; Seb just says "hmmm"; and they all continue with the existing strategy.

When the first equation is simplified to $10b = 40$ Nicola's comment "Er... 4, maybe?" is in a tone that suggests that this is rather an easy problem. Of course the whole point of the program is that the problem didn't appear at *all* easy at first glance - it was only the use of the strategy that made it easy. Similarly, Rajiv sees $12 = 6b$ and says "That's 2… 6 two's are 12.". Is his pre-existing algebraic knowledge coming into play, or is he noting that if 6 barrels weigh 12 then each barrel is 2? We shall see below that breaking with the balance model causes ructions, which suggests that the knowledge being used has been generated during the activity. Moreover, there are no attempts by anyone to re-arrange equations by dragging, or to solve them on paper, or to use inverse operations. And the strategies continue just as before. In fact, the students are even quicker than with the pictures because they no longer have to count barrels. The following graph illustrates this remarkable result:



Note the improvement in strategies occurring over the course of each level up to 8, in response to changes in the nature of the problem at the start of the level; yet the times for Level 9 continue the downward descent of Level 8. After solving the next equation at speed, Seb again says "We're good at this".

This pattern for Rebecca and Nicola is typical of other students:



(The figure for each level includes only students who completed that level)

The figures here treat each pair of students as one. Each pair tackled in excess of 50 puzzles in total.

Level 10: Balance Equations with fractional answers

Rebecca & Nicola

#	**Time**	**Puzzle**
41	51	176 + 20b = 183 + 6b
	4	*Subtract: 6b*
	5	*Subtract:*
	7	*Subtract: 176*
	3	*Guess: 2*
	7	*Guess: 1*
	13	*Guess: -2*
	10	*Guess: 7/14*
	2	*Continue*
42	22	77 + 14b = 27 + 19b
	5	*Subtract: 14b*
	1	*Subtract:*
	4	*Subtract: 27*
	4	*Guess:*
	1	*Guess: 5*
	6	*Guess: 50/5*
	1	*Continue*
43	22	162 + 13b = 269 + 8b
	6	*Subtract: 8b*
	5	*Subtract: 162*
	10	*Guess: 107/5*
	1	*Continue*

Solving the first equation of Level 10 goes well until it gets to $14b = 7$. Rebecca confidently answers 2. This isn't shown as correct. "What?!" she says, and then after a pause guesses 1. "Why can't you do it?" she asks plaintively. Nicola says "Maybe it's minus 2.". This doesn't work either, so Nicola suggests entering 14 ÷ 7. This suggestion provides the clue that it should be 7 ÷ 14. Now they have happily solved, on Levels 7 and 8, balance pictures of the form *several barrels* = *weight*, and with non-integer solutions. However, this is the first time the answer is less than one. Nicola and Rebecca could therefore be using the strategy "Divide the larger number by the smaller number" rather than "Divide the constant by the number of unknowns". Interestingly, none of the other Year 10 students make this "larger ÷ smaller" error. However, Rebecca and Nicola are unusual in not having met an answer less than 1 on Levels 7 or 8, in which the error would involve dividing the number of barrels by the weight (rather than vice-versa).

This equation has taken a minute to solve, but the next two equations are as quick as Level 9 (around 20s). However, it is interesting that when they are faced with $50 = 5b$ they answer "50/5" rather than

"10". More equations with answers less than 1 would perhaps be advisable if the strategy is to be reinforced. It would not be surprising if such an equation causes difficulties later.

This delay makes Rebecca and Nicola untypical with regard to Level 10: even though Level 9 has integer answers, and Level 10 does not, nearly all the students are faster on Level 10 than on Level 9. This would fit with the typical strategy being one of simplification rather than informal, number-specific methods. Those few slower on Level 10 (e.g. Tracy and Jocelyn) were often unaware of the facility for entering fractions. The single exception is Craig, who did not attempt to use the $\boxed{-}$ button on Levels 8 or 9, even though he had been removing weights and barrels on earlier levels. However, he was not especially slow at these levels, at least in part because he did not spend time, like the other students, working out how to use the button until Level 10.

Some students attempt to solve equations like $42 = 9b$ by entering $42/9b$. Unfortunately the version of the program used with 10EQ allowed this as an answer, thus possibly permitting a new error to flourish.

6.2.3 Simplifying Equations that are Unlike Balances

Levels 11 and 12 break with the balance model by introducing negative answers and negative signs. They aim to expand the strategy of simplifying a situation by removing items into a strategy of operating on an equation.

Level 11: Linear Equations with negative answers

Rebecca & Nicola

#	Time	Puzzle
44	66	72 + 18b = 33 + 15b
	7	*Subtract: 15b*
	3	*Subtract: 33*
	11	*Subtract: 15b*
	5	*Subtract:*
	13	*Subtract: ESC*
	25	*Guess: ESC*
	2	*Give up*
		Level 10
45	25	176 + 30b = 329 + 13b
	12	*Subtract: 13b*
	4	*Subtract: 176*
	8	*Guess: 153/17*
	1	*Continue*
46	39	44 + 11b = 50 + 32b
	7	*Subtract: 40*
	8	*Subtract: 11b*
	4	*Subtract: 4*
	4	*Subtract: 6*
	14	*Guess: -6/21*
	2	*Continue*
47	22	36 + 27b = 16 + 7b
	5	*Subtract: 16*
	4	*Subtract: 7b*
	5	*Subtract: 20*
	7	*Guess: -20/20*
	1	*Continue*
48	27	55 + 7b = 20 + 2b
	5	*Subtract: 2b*
	4	*Subtract: 20*
	3	*Subtract: 35*
	14	*Guess: -35/5*
	1	*Continue*

Having simplified the first equation down to $39 + 3b = 0$ Rebecca & Nicola are stuck - "You can't do that can you?" says Rebecca. They attempt to take off a further 15b without result. They select "Give Up", and this takes them back to a Level 10 equation which they solve without difficulty.

During the class, there were many requests for assistance as students came onto this level and faced an apparently "impossible" balance. Even though they should have been reasonably familiar with negative numbers, many did not catch on until there was an explicit negative sign. For example, Rebecca and Nicola simplify equation #46 to $0 = 6 + 21b$, from which they then subtract 6 and only then guess -6/21. This is a very common strategy; although Jocelyn gives up on 6 equations before success. The fact that someone like Debbie fails to see the "obvious" solution to $0 = 20 + 10b$ - or that Craig is so unbelieving that the answer to $39 + 13b = 0$ is *not* 3 that he has to also guess 39/13 and 0.3 just to be on the safe side - must mean that many students are still thinking in terms of positive rather than signed numbers.

Rajiv and Seb, on the other hand, do not hesitate to take off 85 from $85 + 20b = 45$, rather than the 45 that the "Take off the smaller constant term" strategy might suggest, and that they consider very briefly. Are they guided here by a pre-existing "Get the unknown by itself" rule, or are they simply noting that they need the standard form Kb = E that has been a feature of virtually every puzzle since Level 4? They are rare in not taking the 45 off first, followed by the 40.

Cedric provides another interesting example: having simplified his first Level 11 equation down to $31 + 15b = 0$, he guesses 31/15, attempts to subtract 0, guesses 15b/31, subtracts 31, guesses 15/-31, and finally guesses -31/15b. Thereafter he has no difficulties with this level. Dick similarly inverts the division when dividing the constant by the coefficient of *b* fails (because the negative sign is omitted). Does this indicate a deep-rooted worry about which way to divide?

Note how in Rebecca and Nicola's equation #47, the answer to $20b = -20$ is given as $-20/20$ rather than -1, suggesting that the strategy developed for positive answers is continuing to be followed, without paying any attention to the numerical answer. Many other students do the same; which would suggest at one extreme that they do not see the simple numerical answer, and at the other that they see no reason to look for simpler answers in this context.

Rebecca and Nicola solve the final two equations on this level almost as fast as the balance puzzles on Level 2, a fact which provides evidence that they are as confident about executing the simplification strategy as they are about solving simple balance puzzles. Of course whether this grasping of the simplification strategy remains a reliable member of their theoretical tool-kit remains to be seen.

For example, Jocelyn was the only student to complete Level 11 in the first lesson, but did so only after giving up on 6 occasions. Having then apparently worked out that negative numbers were required - he successfully solved three Level 11 equations - one might imagine that he would have few difficulties in the second lesson. Yet at the start of the second lesson he again struggles with Level 11 equations, giving up on 6 more occasions. The nature of his difficulty is the same each time: he is able to subtract constants and unknowns so that he is left with something like $3b = -7$; but cannot then enter the answer -7/3. He seems convinced that the answer should be -21. Even $5 = -5b$ is beyond him: he guesses -25. When the coefficient of *b* is 1, there is no difficulty. Yet he is able to solve every Level 10 equation, such as $4b = 108$. Interestingly, faced with equations such as $9b = 18$ or $17b = 17$, he always enters the numerical answer - unlike many other others who enter 18/9 or 17/17. Finally, after 25 minutes of giving up and retrying on Level 11, he is presented with five Level 10 equations in a row, and then gives the correct answer to $-43 = 8b$. He is then able to solve the rest of Level 11 and 12 quickly.

Level 12: Linear Equations with negative signs and $\boxed{+}$

Rebecca & Nicola

#	Time	Puzzle
49	107	-27 + 24b = -22 + 7b
	5	*Add: 7b*
	3	*Add: 22*
	5	*Add: 14b*
	8	*Add: 28b*
	13	*Add: -56*
	9	*Subtract: 56*
	7	*Subtract: -112*
	5	*Subtract: 56b*
	6	*Subtract: -5*
	9	*Subtract: 5/17*
	16	*Subtract:*
	1	*Subtract:*
	18	*Guess: 5/17*
50	43	-18 = 5 + 3b
	11	*Add:*
	3	*Add:*
	5	*Subtract: 5*
	4	*Subtract:*
	18	*Guess: -23/3*
51	72	-22 -14b = -12 + 4b
	5	*Add: ESC*
	6	*Subtract: 12*
	10	*Subtract: 4b*
	11	*Add: 24*
	11	*Add: 10*
	5	*Add:*
	14	*Guess: -18/10*
	9	*Guess: 10/-18*
52	93	-1 -24b = 8 -7b
	6	*Add: 1*
	7	*Subtract: 7*
	6	*Subtract: 7*
	5	*Subtract: 2*
	5	*Add: 16*
	5	*Add: 7*
	5	*Add:*
	6	*Subtract: 7*
	5	*Subtract: 7b*
	5	*Subtract: 9*
	4	*Subtract:*
	4	*Add: 9*
	5	*Add: 14b*
	3	*Add:*
	20	*Guess: 9/-17*

Rebecca & Nicola have difficulties with the first equation on Level 12 that are probably more connected with having failed to noticed that the appearance of the $\boxed{+}$ button means that the shortcut keys they have been using to enter an answer or to subtract no longer work as they intend. There will always be a question (whether using audio-tape logs, direct observation or video) as to whether it was the choice of operation or the selection of the button that was at fault. One must be careful in equating what was done with what was intended.

For equation #51, Rebecca and Nicola initially select $\boxed{+}$ (by mistake again?), and then cancel the input. They then subtract 12 (ignoring the negative sign?). They then subtract 4*b*. Nicola now realises that one must use $\boxed{+}$ rather than $\boxed{-}$ to eliminate the constant:

Nicola: You should've *plussed* 12. (*pause*) Should be *plus* 24. 'cos you want to get rid of it, so you want to get it to nought, so you plus it.

Rebecca: Do you?

Nicola: Yeah.

Rebecca: Oh yeah.

If Rebecca was previously using a purely object-based strategic theory (subtracting 17 to get rid of the object '17' or subtracting -8 to get rid of the object '-8') then it would appear to have been superseded by the arithmetic strategy of adding. They add 10, and then select $\boxed{+}$ again before realising that they can now enter an answer because this is a form they recognise. Their guess –18/10 suggests that they are still unsure about the correct order of the division.

Note how the feedback to Action 1 allows them to debug their strategy. Nicola says "You should've *plussed* 12."; but Rebecca isn't convinced, until she tries it.

One might think that the theory that $\boxed{+}$ can be used to eliminate subtracted quantities is now grasped. Not at all - equation #52 raises more questions: Action 1 (suggested by

Nicola) seems entirely reasonable, but what about Action 2 (from Rebecca alone)? It is difficult to see, if we assume (given their experiences so far) that they appreciate by now that the 7 and the *b* are wedded to each other more firmly than subtraction can sunder, what might be the intention of subtracting 7. So we have to try to guess what simple, accidental, forgetful slip has been made. Perhaps she meant to subtract 7*b*, in which case they might have the theory that the minus sign "goes with" the 8 rather than the 7*b*; she might not have noticed the minus sign, or not considered it important; she might think that adding to get eliminate subtracted quantities applies only to constants and not to unknowns.

Perhaps she meant to subtract -7*b*; or to add 7*b*. But these latter two scenarios require two slips and, moreover, the very next action is exactly the same: "Subtract 7". It could be then, that she has forgotten the requirement to include the *b*. Alternatively, it might be the 7 on the left-hand side of the new equation that is now the object of attention. There are many possibilities, and Action 4 is especially difficult to explain.

One scenario is that Rebecca carries out Action 1 without really appreciating why it is done. She then subtracts 7 (Action 2) to try to eliminate the 7b, not noticing the minus sign and forgetting to put the *b*. Nicola says "It was minus 7.", pointing out to Rebecca therefore that there is a minus sign in front of the 7b, and so subtraction was not appropriate. Rebecca replies "Whoops!". However, she interprets the comment as referring either to a wrongly entered minus sign in the input box; i.e. "You put a minus in front of the 7 when you shouldn't have."; or to the desired operation not being carried out; i.e. "We wanted minus (subtract) 7, but you entered add 7 instead.". In either case Action 3 (Subtract 7) could be seen as undoing the effect of the first. Alternatively, in subtracting 7 for the second time, Rebecca could be responding to the new equation rather than to Nicola's comment - she might be attempting to get rid of the -7 by subtraction, forgetting the minus sign as before. Just as Rebecca is carrying out Action 3, Nicola says "Then subtract 2."; but at that moment Rebecca presses Enter and the new equation appears: $-14 - 24b = -5 - 7b$. Responding to the new equation, Nicola says "and then add 14.". However, Rebecca concentrating on subtracting 7, has only just heard "Add 2"; so this is what she does (Action 4). There are other possible interpretations throughout this - for example Rebecca might have looked at the screen after Action 2, decided she wanted to subtract 7*b* and 2, left off the *b* when subtracting 7*b* in Action 3, and not looked at the resulting equation until after subtracting 2 in Action 4. In Action 6, it is surely 7*<u>b</u>* that was intended rather than 7 - but when this has been undone (by subtracting 7) it is *subtract* 7b that is entered. Attention now turns to the sole remaining constant - the 9 on the right-hand side. This is subtracted off only to be added on again when this is seen not to help. Finally, 14b is added on to reduce the equation to a known form.

But we are interested not so much in producing a definitive version of Rebecca's cognitive history, as in seeing - in Actions 1, 5, 6, 12 and 13 - the gradual development of strategies for dealing with negative signs in equations. The final equation on this level shows the new strategies in action:

Rebecca & Nicola		
53	46	-35 -11b = 3 + 8b
	10	*Add:*
	4	*Subtract: 3*
	5	*Subtract: 8b*
	5	*Subtract:*
	3	*Subtract:*
	4	*Add: 38*
	14	*Guess: 38/-19*
	1	*Continue*

There are many examples of students struggling to develop such strategies. Rajiv and Seb's first equation on this level, for example, is $-4 - 8b = -4 - 29b$, which they manage to "simplify" to $128 - 128b = 128 - 149b$ along the way to discovering that $b = 0$. Their next equation takes almost 20 actions, as they work out the effect of adding negative numbers to positive numbers, subtracting positive numbers from negative numbers, and so on. By the last equation they can solve them in 3 actions, but there have to be doubts about the extent to which they will remember the strategies they have developed here given the number of permutations.

For example, when there is a negative sign in front of the unknown term, Lisa tends to attempt to subtract rather than add: she tackles $37 + 13b = 6 - 3b$ with "subtract $3b$". She corrects this strategy, but it recurs: in particular when tackling the first equation of the second lesson. Many students have to remind themselves of what happens to the equation according to the sign of the terms in the equation, whether the operation is add or subtract, and the sign of the number being added or subtracted.

6.2.4 Posing Linear Equations

Although 80% of 10EQ reached Level 13 during their first lesson, details of that work will be included in the next section on modelling word problems. At the beginning of the second lesson, however, students were encouraged to enter their own equations to get back up to the higher levels.

Rebecca and Nicola entered the following equations, and received certain messages from the program:

Equation entered	Message
$x^2 + 25 = y + 100$	This program cannot handle more than one letter
$x^2 + 25 = 2x + 5$	Left-hand side is not linear
$x^2 + 25 = x$	Left-hand side is not linear
$x^2 + 25 = x = 5$	An equation must have only one equations sign
$x^2 + 25 = x + 5$	Left-hand side is not linear
$x + 48 - 22 = 12 - 49$	

They then successfully solved the latter equation (which had been automatically simplified to $26 + x = -37$), but gave the answer $-63/x$, then 63, and finally -63. They then successfully completed Level 12.

Billy was more ambitious, in trying to find complicated equations that the program either could not handle or dramatically simplified. He posed $(x \times 2.89) - (y/342) = (z + 34)/(x \times 99)$, and then $(x \times 2.89) - (x/342) = 32$. A wide variety of equations were posed to get to as high a level as possible - some tried to arrange for negative answers; others quickly discovered that a negative sign was the best indicator of a high level.

Rajiv and Seb, meanwhile, posed the equations $876b = 45/56 + 2$, $34 + 17b = 56 + 12$, $32 + 78b = 67/65b$, $32 + 78b = 67 + 76b$ and $23b + -15 = 18b + 30$. In the post-interview, Rajiv and Seb said they enjoyed the equation-posing part of the activity, and would like more - perhaps also posing equations for others to try to solve.

6.2.5 Modelling Word Problems

Level 13: Word Problems with balances

Rebecca & Nicola

#	**Time**	**Puzzle**
54	121	4 oranges plus 227g weighs the same as 9 oranges plus 17g. What is the weight of an orange?
	29	*Guess: 4+227=9*
	36	*Guess: 4+227=9+17*
	53	*Guess: 210/5*
	2	*Continue*
55	61	12 cups plus 417g weighs the same as 18 cups plus 129g. What is the weight of a cup?
	58	*Guess: 48*
	3	*Continue*
56	91	'17 pencils plus 390g weighs the same as 23 pencils plus 132g. What is the weight of a pencil?'
	90	*Guess: 258/6*
	1	*Continue*
	19	*Quit*

On Level 13, each problem is a description of a balance puzzle. For Rebecca & Nicola, the script for the modelling puzzles does not look promising. Faced with the orange problem, their initial "guess" of 4 + 227 = 9, apparently shows little awareness of the relevance of unknowns or that such an expression cannot constitute an answer to the problem. But in each puzzle on Level 13 they obtain a correct answer and the first and the third are in the fractional form that one might expect had they solved an equation. The tape reveals that, after the abortive attempt to represent the first problem situation as an equation, Rebecca says "Take 4 oranges off each side.". It appears, then, as though Rebecca is using a simplification strategy based on the objects in the balance situation rather than the unknowns in an equation. She then says "Take 17 off each side.". Note the form of words: if she were using a typical arithmetic strategy she might say something like "If 4 oranges plus 227g weighs the same as 9 oranges plus 17g, then 5 oranges weights 227 – 17.", but instead she says "Take... off each side".

Rose also entered an equation like 4 + 227 = 9 + 17, but found another way to solve the Level 13 balance problems. Another common formulation amongst students for the Level 13 balance problems is something like 4 + 227g = 9 + 17g, which suggests that the point of using a letter to stand for an unknown number has been lost. It is interesting that several students were able to get the correct answer using this equation - their grasp on the strategy for solving 210g = 5 not being very strong perhaps.

Of course having appreciated from the earlier balance puzzles that dividing the weight difference between the sides by the difference in number of objects gives the answer, and being confident in one's arithmetic (or having a calculator) one may not attempt to be tempted to solve the problem using algebra. Indeed, several students did just that. For example, Tracy solved the Level 13 balance problems and Level 14 seesaw problems without using the Model button. However, some students who do this then attempted to find by trial-and-error the sequence of operations that would solve problems on higher levels. This strategy very rarely worked, and so they were redirected back by the teacher or researcher to Level 13 to find a way of using the $\boxed{\text{Model}}$ button to solve the problem.

However, several students worked out effective ways of modelling. Seb, for example, happily entered an appropriate equation to solve their Level 13 problems, using the letter *a* or *x* in each case, whatever the objects in the problem. Jack, Harry, Kirsty, Lisa, Mike, Dick, Jane and Debbie successfully solved the Level 13 balance problems using the $\boxed{\text{Model}}$ button. On one problem, Harry omits one of the unknowns, but when his answer is rejected re-enters the equation correctly.

Level 14: Word Problems with subtractions

Jack

Charlotte has a secret number. If she multiplies it by 5 and subtracts the result from 155, she gets 240. What is her number?

Time	Action
58	*Model: (5x)-155=240*
9	*Subtract: 240*
11	*Add: 10*
7	*Subtract: ESC*
8	*Add: 385*
7	*Guess: 395/5*
11	*Guess: 395/5*
10	*Guess: 395/5x*
8	*Guess: 5/395*
61	*Model: 155-(5x)=240*
12	*Subtract: 155*
13	*Guess: 85/-5*
1	*Continue*

Some of these problems started to cause problems for those who had met early success using a strategy of treating expressions as combinations of objects. A slow-down in students' success rate and enjoyment was noted in all classes who worked this far. This would tend to support the proposal to challenge the letter-as-object strategy prior to modelling.

Jack's first Level 14 problem is interesting (see left). He is among several students who algebraically translate the TOAN phrase "subtracts the result from" as "subtracts", but then re-models the equation when the answer is not correct. He then successfully uses models the remaining Level 14 problems, although he initially attempts to model the seesaw problem by using 4 variables - two for the people's weights and two for the bricks' weights - and two equations. He again attempts simultaneous equations for the tickets problem, before creating an accurate single equation.

For the Level 14 CD problem, Rose divided the difference in price by the difference in number of CDs. For the Level 14 seesaw problem, she entered the equation $48n - 5n = 39n - 2n$.

Craig obtained correct answers for Level 13 without using the $\boxed{\text{Model}}$ button; attempted to model the Level 14 seesaw problems without negative signs - a fairly common strategy - but found the answer another way; successfully modelled (and solved) the Level 14 TOAN problems; and successfully modelled (and solved) a Level 15 cinema problem. Grace was similar in not using $\boxed{\text{Model}}$ until TOAN. Lisa initially attempted to model the Level 14 CD problem without negative signs, but thereafter has no difficulties with that level.

Level 15 & 17: Word Problems with ratios

As noted earlier, in the version of the program used with Year 10, this level did not discriminate between problems in which the quantity requested happened to be the obvious choice for x and problems in which the quantity requested was *not* the obvious choice for x. When the (correct) solution to the equation was rejected as a solution to the problem, many students attempted to re-formulate the equation rather than carry out the simple calculation using the equation's solution that would have obtained the correct solution to the problem.

On the other hand, when Dick successfully modelled Levels 13 to 15, the only difficulty was the first Level 15 cinema problem, which took him 6 minutes to find an appropriate equation; and it took only another minute to deal with the "hard x" obstacle.

Level 16 & 18: Word Problems with expressions

Again, in retrospect, the word problems could have been graduated better - being asked to find a quantity that was not the obvious choice to be represented by a letter was a bigger hurdle than anticipated. However, much of the frustration felt by students at this stage could have been avoided simply by introducing this subtlety in an already familiar word problem. The value of this frustration is debatable. Nevertheless, it was amazing not so much that students chose to formulate their own equations and that they were able to; but that once the equation appeared on the screen, students said things like "Ah, now I can do it!" and "It's easy now!". In other words, the equation had become for them a powerful problem-solving tool that they were confident about using. Enjoyment in using algebra was, for many of the students in the study, a new experience.

6.3 Which problems improved according to pre-post testing?

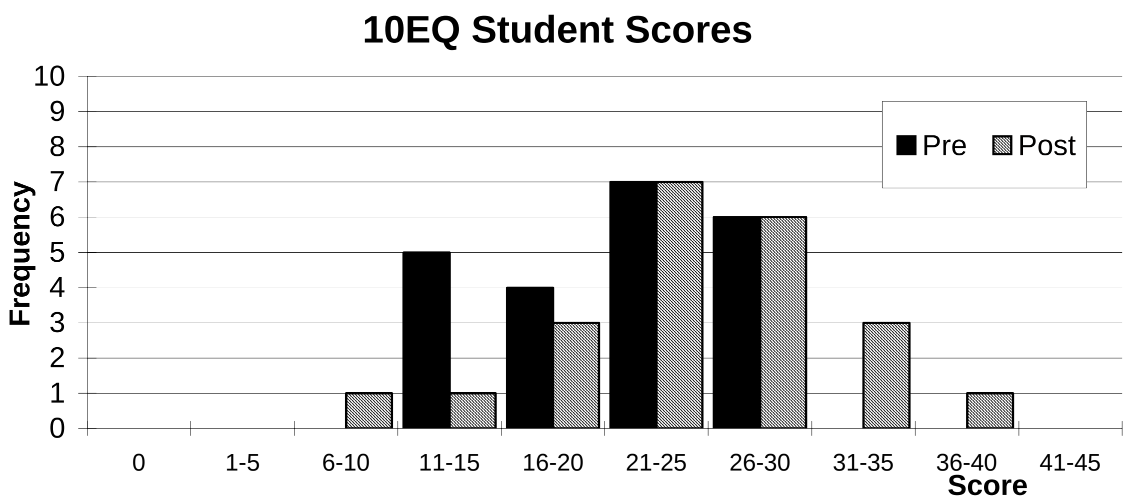
This section uses evidence obtained before and after the activity to attempt to identify the problems in which there were improvements attributable to EQUATION. Evidence obtained *during* the activity will *not* be considered here, because the task of objectively identifying improvements that are unexpected from other theoretical perspectives (a task for which the most appropriate research instruments are perhaps those that do not depend on the specific activity; for example - with regards to strategic theories - the written test, interviews and problem-posing) must not to be conflated with the task of relating any such improvements to particular aspects of the specific activity (a task which of course relies heavily on evidence from the activity).

6.3.1 All Tested Algebraic Problems

Class	n	Pre	Post	Imp	Wor	p
7EQ	26	11	14	72 / 992	43 / 126	0.012
7CON	21	10	12	43 / 815	25 / 88	
10EQ	22	50	58	149 / 470	79 / 476	0.006
10CON	24	50	53	153 / 518	121 / 514	

NB "7EQ" and "10EQ" refer to the Year 7 and Year 10 classes (respectively) using EQUATION. "7CON" and "10CON" refer to the control classes - i.e. those *not* using EQUATION. "n" refers to the number of students. "Pre" and "Post" refer to the mean test percentage scores. "Imp" refers to the number of items that improved, compared to the number of items where improvement was possible. "Wor" refers to the number of items that worsened, compared to the number of items where worsening was possible. "p" refers to the p-value of the t-test with matched samples, i.e. the significance of the change from pre-test to post-test for that particular class. Recall that the EQUATION groups and control groups are not matched.

The test results suggest that at least some problems were improved by the use of EQUATION. Even though there were a large number of items in the test that were not expected to show improvements, both EQUATION groups showed small but significant increases for the test as a whole. There were small increases for the control groups, but these were not significant.



Even allowing for enthusiasm generated by being involved in research, this apparent success disguises the fact that, as will be seen, some facilities improved and some worsened. Looking at both Year 10 groups, approximately a third of the items where improvement was possible did improve; and a fifth of the items where decline was possible did decline. The situation for Year 7 was very different. Only around one in twenty items where improvement was possible did improve; and a third of the items where decline was possible did decline. The differences in scores *between years* were much larger than between the EQUATION and control groups.

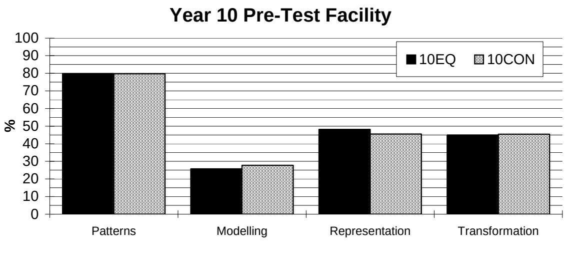
Facilities for Year 10 exceeded those for Year 7 in virtually every item (and the exceptions were trivial).

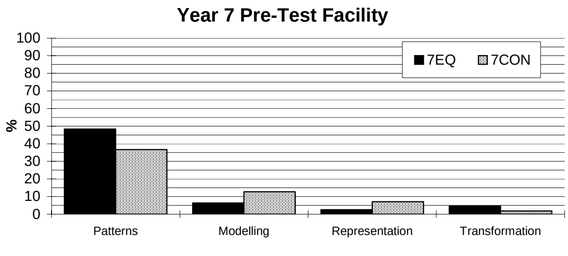
One particularly unfortunate occurrence with the Year 10 students for which there is some evidence (a cluster of "unusual" answers from adjacent students) is of copying, although this mostly applies to the control group. So what with this and the small samples (22 in 10EQ, 24 in 10CON, 26 in 7EQ, 21 in 7CON), any statistical analysis here has to be treated with some scepticism. Even with these provisos, it seems reasonable to suggest that EQUATION does not by any means provide a panacea for algebra - the post-test results for the EQUATION groups are not significantly higher than those of the control groups, for the whole test and for the individual sections. Moreover, because gains tended to be on a continuum, it is difficult to clearly identify students who particularly gained from use of the program.

The highest mark in 7EQ was 44% (Basil on the post-test), with the next mark being 26% (Melissa on the post-test); while the vast majority scored under 20% (for either test), and one student scored 0 both times. Many of the students increased their score on the post-test by small amounts, the only serious exception being Joe (from 23% to 14%); the best improvements were made by Margaret, Melissa, Jessica, Basil, Catherine and Eddie. Success in *modelling*, *representation* and *transformation* for the class as a whole remained minimal (below 10% correct).

The range in marks for 10EQ was 84% (Dick on the post-test) to 23% (Grace on the post-test), but there are no large gaps in the scores. Nearly all the students increased their score on the post-test; with Grace (down from 44% to 23%), Judy and Arthur (both dropping about 10%) being the only serious exceptions; the best improvements were made by Bruce, Rajiv, Dick, Cedric, Debbie, and May (all up about 20%). The extent of change for each student does not bear much relation to the pre-test score. As for 7EQ, the number of improved items for 10EQ was almost double the number of declined items.

The control groups were fairly similar to the EQUATION groups on the pre-test, as can be seen from the following charts:





The highest mark in 7CON was 60% (Deirdre), with the next mark being 30% (Stephen); over half the students scored under 5%. Over a third of the total number of improvements (43) were obtained by Deirdre; and she did not decline in any items. The other students were fairly evenly divided between those who improved their score, those who stayed the same, and those who declined; but the size of change was small.

The range in marks for 10CON was 74% (Robin on the post-test; Elizabeth on the pre-test) to 9% (Thomas on the post-test). There were roughly equal numbers of students increasing and decreasing with their score; the extent of change for each student is largest for those with the smallest pre-test score. The largest increase was for Reuben (up from 23% to 58%); while the largest decrease was for Thomas (down from 33% to 9%).

For both control groups, the number of item improvements is only slightly higher than the number of item declines; and, taking both control groups together, if the three most improving students are removed as outliers, these two indicators are virtually identical. If the same is done for the EQUATION groups, however, the number of improvements is still some 60% higher than the number of declines.

6.3.2 Comparison of the Problem Types

Representation

Class	Pre	Post	Imp	Wor	p
7EQ	3	8	22 / 304	4 / 8	0.002
7CON	7	9	5 / 234	1 / 18	
10EQ	48	57	45 / 137	22 / 127	0.017
10CON	45	50	49 / 157	37 / 131	

The results suggest that *representation* scores were improved by the use of EQUATION. The improvement in representation items for 7EQ was significant; while 10EQ increased by about 10%; neither control group increased significantly.

Transformation

Class	Pre	Post	Imp	Wor	p
7EQ	5	5	16 / 521	15 / 25	
7CON	2	5	16 / 433	4 / 8	
10EQ	45	52	66 / 253	37 / 209	0.016
10CON	45	48	71 / 275	57 / 229	

The results also suggest that using EQUATION improved *transformation* scores for 10EQ students; but not for 7EQ. The control groups' scores each increased by about 3%, but this was not significant.

Modelling

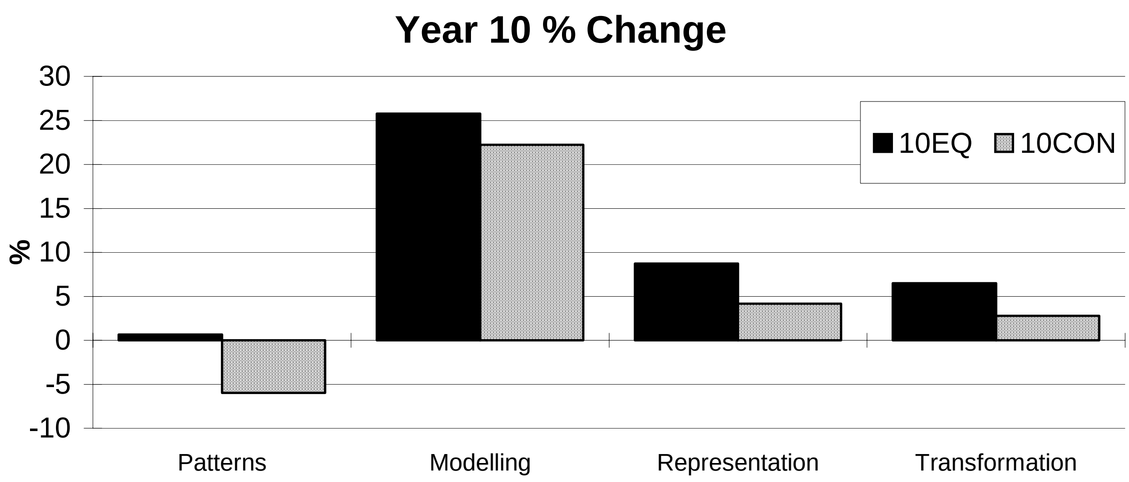
Class	Pre	Post	Imp	Wor	p
7EQ	6	8	6 / 73	5 / 5	
7CON	13	17	6 / 55	3 / 8	
10EQ	26	52	21 / 49	4 / 17	<0.001
10CON	28	50	16 / 52	0 / 20	<0.001

Neither of the Year 7 groups improved their *modelling* score significantly; whereas both the Year 10 groups did so, the average facility moving from about a quarter to about half.

Patterns

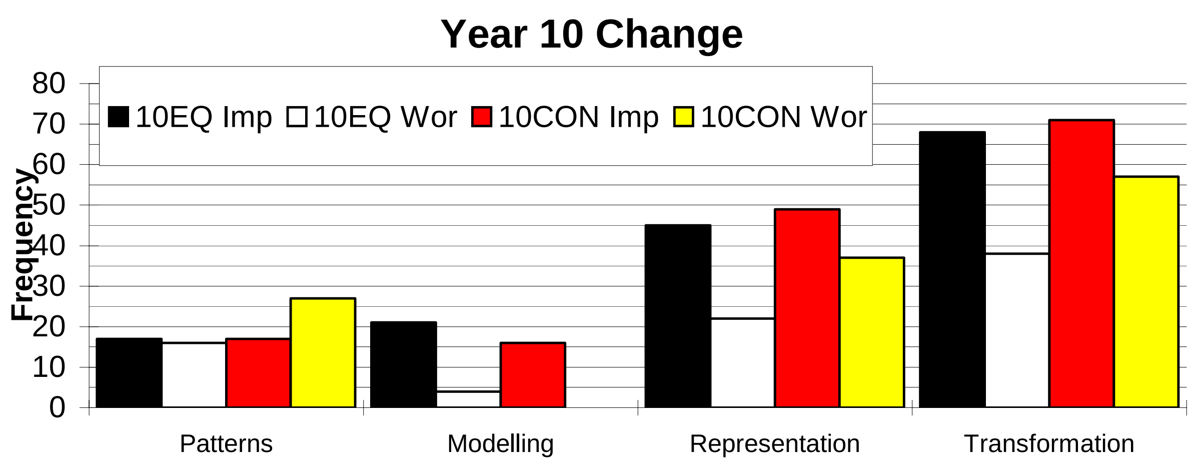
Class	Pre	Post	Imp	Wor	p
7EQ	48	53	28 / 94	19 / 88	
7CON	37	36	16 / 93	17 / 54	
10EQ	80	81	17 / 31	16 / 123	
10CON	80	74	17 / 34	27 / 134	

The improvement in the EQUATION groups' *patterns* scores was not statistically significant; while both control groups declined slightly.



Caution is required in comparing pre- and post-tests for Year 10, with particular regard to *transformation*, as some students appeared to run out of time before reaching some questions in that section. Some students were also given the option by their teachers during the pre-test of attempting additional (harder) questions, which may have distracted them from the problem types under discussion here.

The following chart shows the proportion of those students who improved compared to those who could improve; and the proportion of those students who worsened compared to those who could worsen.



It is interesting to note that there seems to be little correlation between score in one problem type and score in another; and that the same seems to be the case for improvement in score. On the other hand, *patterns* and *modelling* consist of fewer items than *representation* and *transformation*. It would be difficult, in any case, to conclude much about strategic improvements without looking at the individual problems being tackled, unless one had a conjecture about strategies that work at the level of problem types. The following sections attempt to identify problems for which there is prima facie evidence (from any source, but mostly from the written test) of improvement as a result of using EQUATION.

6.3.3 Patterns Problems

For the *patterns* section of the test, the overall change in facility, the changes in facility for each item, and the numbers of students improving or worsening suggest little change for any class. However, there are four anomalous changes for 7CON at the 5% level. (Facilities for each item are in the Appendix.) Nevertheless, there are roughly equal numbers of improvements and declines for this class, and a quarter of the declines are by Deirdre. Although there were slightly more improvements than declines for 7EQ, a quarter of the improvements are by Hugh and Tony. For 10EQ, there are again roughly equal numbers of improvements and declines; while for 10CON there are more declines than improvements, but over a third of the declines are by Christine, Thomas and Isabelle.

Moreover, looking at the individual scripts, there is little evidence of systematic strategic change, for any group. One exception resulted from the fact that the usually used rule for A1 in the pre-test was $3n + 1$; whereas in the post-test there were two: $4n - 1$ and $4(n - 1) + 3$. (Josh in 10EQ, for example). One student used the post-test rule $3n + (n - 1)$. Rajiv in 10EQ used a distinctive (but unknown) strategy to get an answer on the pre-test; but divided by 4 on the post-test and rounded up. Quite a few students were able to get the correct answer by dividing and then checking, rather than subtracting (or adding) and then dividing. A couple of students solved A1(i) in the pre-test by using the pile number and by using the pile difference in the post-test. The most important result for strategies is a negative one: nobody used algebra to tackle A1(iii) in either test.

In A2(iv), it is odd that all 14 10CON females were correct in the pre-test and then 6 declined, whereas 7 of the 10 males were correct on the pre-test and the other three then improved on the post-test (one declined).

6.3.4 Modelling

Although neither Year 7 group improved in the *modelling* section, it is clear from the fact that *both* Year 10 groups improved their scores significantly that attempting the pre-test can assist in the post-test (assuming that 10CON did not do anything between the tests that could have contributed to this improvement). Why this should be - and just for the modelling section - is interesting, and will require a closer look at the strategies used for individual items.

Note that attribution of an algebraic strategy requires more than just the presence of letters - there must be evidence of an attempt to represent the situation using some sort of notation.

6.3.5 B1: The Field Problem

Many students divided the perimeter by six, and we are given no further clue as to how they arrived at this calculation; although whole-parts reasoning can be conjectured. Harriet in 10CON, for example, wrote “2-length, 2-width; 4-length, 2-width = 6” to get the correct answer. Bruce in 10EQ divided by 3 in the pre-test (the two widths are equivalent to one length, perhaps?) to get the correct answer. The simplicity of the problem worked against much elucidation of strategy; so for example Debbie in 10EQ divided the perimeter by 4 in the pre-test, but changed this to 6 in the post-test, but no indication is given why. A popular strategy was to divide the perimeter by 4, and then add this number onto half itself. Letters did not appear to play any role in solution for 10CON, and only for a few students in 10EQ.

For 10CON, 9 out of 14 females were correct on the pre-test, compared with only 2 (out of 10) males. 1 female and 3 males improved in the post-test. What advantage did these four students gain? One tried trial-and-improvement rather than the "divide by 4, and add half on as much again" strategy; and one started by dividing by 6 rather than by 4; and two showed no working.

Potential strategic improvement for 10EQ is perhaps best exemplified by Darren, who used the "divide by 4 and add half on as much again" strategy in the pre-test; but on the post-test there is not only evidence of an initial attempt to divide the perimeter by 4, but also a calculation involving halving the perimeter, dividing this number by 3, and then doubling. Beside this are two sketches of the field: the first with the width marked as x and the length as $2x$; the second with the width marked (correctly) as 19 and the length as 38. Of course one interpretation of his script is that he was just playing with whole-number combinations in an attempt to find a number which worked. But there is a checksum on the post-test that these numbers produce the required perimeter; but not on the pre-test. Were the letters in the diagram standing in for "parts"? Did the diagram help him realise that half the perimeter would be $3x$? Is that why he halved (to find $3x$), divided by 3 (to find x) and then multiplied by 2 (to get $2x$)? His delayed-test is almost identical to the post-test, but without the sketches. This scripts might indicate that the x was being used to assist in the execution of a whole-part strategy. If so, such a use of algebra would be a remarkable product of using the program. In Lins' terminology, Darren would be showing evidence of a whole-part, non-internal, but analytical approach.

However, there is also evidence to suggest that EQUATION hindered students' use of algebra to represent situations. For example, Judy used the equation $102 = x + 2x$ in the pre-test (forgetting that four sides make up the perimeter, and so obtaining an answer twice as big as the correct one); but in the pre-test she wrote the equation $114 = x \times 2$, which does not suggest a good grasp of why algebra might be useful here. Jack did exactly the same as Judy on both tests. Jane, who correctly used a whole-part strategy in the pre-test, used the unhelpful equation $L = 2p$ in the post-test. Yet all three students scored above average for the class on the test as a whole, and (apart from Judy's confusion between area and perimeter in C2 on the post-test) were reasonably successful in the representation section. Evidence from Jack's script from the advanced pre-test also shows that he is a competent equation-solver of simple linear equations in one unknown, with or without negative signs.

The only really successful use of literal algebra on this item (in any class) was by Liam in 10EQ, who left the question blank on the pre-test, but apparently made use of the equation $2a + 2(2a) = 114$ on the post-test, after having first attempted to divide by 8. However, the mechanism by which the equation provides the answer is not clear: this equation is followed by "$= a + 2a = 114$" on the same line, and then $3a = 114$ below, which would provide the length correctly as $a = 38$ (which he then checks produces the required perimeter). But how was $a + 2a = 114$ obtained? One interpretation would be that he realised that he could replace $2a$ by a in the equation if a were chosen to represent the length rather than the width. Unfortunately, the symbols $-a$ appear below the $2a$ and $2(2a)$ terms in the first equation, suggesting that he found the second equation by subtracting a from each of the two terms on the LHS of the equation without altering the RHS. On the other hand, this interpretation would require that when he starts to check that a width of 38 produces the required perimeter, he discovers that it is too large to be the width, and so decides it is the length. There is no evidence for this, and it may be that the $-a$ indicates the replacing of the $2a$ by a rather than the subtraction of a from both sides of the equation. Whatever the interpretation, it is clear that Liam is attempting to use an equation to help him solve the problem - he seems to have an algebraic strategy that was not there before, albeit his use of algebra is non-standard. His delayed-test does not provide much further insight: the initial equation is written as $2(2w)w + 2w = 102$, but it is simplified to $4w(1) + 2w = 102$, then $6w = 102$, which is then solved correctly.

In contrast to Year 10, there is not only little evidence that Year 7 improved strategies for this item, but there are also no indications of algebraic activity. Three students in 7EQ were correct on the pre-test (Margaret, Charlotte & Basil). Three different students were correct on the post-test (Melissa, Mickey & Eva). Margaret's method on the pre-test was "I kept try new numbers"; whereas she divided the perimeter by 4 on the post-test, labelled each side of a diagram of the field with this figure (28.5), then drew another diagram with the sides labelled 57, 28.5, 57, 28.5, and answered 57m. Melissa left the item blank on the pre-test, and just wrote down the answer on the post-test. Charlotte gave no working on the pre-test, and used an unknown strategy on the post-test. Mickey showed no working on either test. Basil used an unknown strategy on the pre-test; and on the post-test appeared to use the perimeter divided by 8 as the width (14.25) and then found the length by subtracting twice the width from the perimeter and dividing by 2, without noticing that the length was therefore 3 times the width, rather than twice the width. Eva left the item blank on the pre-test, and divided the perimeter by 6 on the post-test to find the width. There were similarly few indications of strategic improvement for 7CON.

6.3.6 B2: The 4× Seesaw Problem

Lins (1992) found a facility for this item of 22%. The overall pre-test facility here was 15%. But there are important differences in the experimental arrangements: the most significant being that Lins's students had to tackle a very much smaller number and range of problems. So they probably felt they had more time not only to spend on the items, but to elaborate their reasoning. However, his research did not take account of possible improvements on subsequent occasions, which is relevant here because for 10EQ at least, B2 is the item with the most improvements.

Rebecca in 10EQ provides an interesting before-and-after case, in that there are examples of at least 4 different strategies in the pre-test, which are given here in the conjectured order.
(1) "Difference ÷ 4" - Calculation of the difference in weight between the sides; and then division by 4. This she checked by finding the final weight on each side, and she concluded that the

strategy didn't work ("no").
(2) "Difference ÷ 2 – 2" - Calculation of the difference between the sides, and then assignment to George of 4kg more than Sam. Although she wrote down the principle of this strategy ("George has to get rid of 4kg + the amount Sam throws away") she quickly realised that this wasn't going to balance the seesaw.
(3) Solution of the equation $189 - x = 273 - 4x$. Unfortunately, when adding x onto the equation she thought that it should be $5x$ (instead of $3x$) on the RHS. She checked this solution and concluded "wrong doesn't fit".
(4) Trial-and-error - repeated trials based on the checking process ("tried other numbers in my book") before finding the correct answer.

In the post-test Rebecca used only one strategy - the solution of the equation, which this time she carries out successfully ($-4x + x$ becomes $-3x$) and then checks ("144 4"). Hence we appear to have prima facie evidence of strategic improvement. Incidentally, in the delayed test six months later she again goes straight for an algebraic strategy, which is again carried out accurately without the $-4x + x = -5x$ error.

Rebecca was the only student who attempted the equation strategy in the pre-test; but there were 14 students (3 from 7EQ, the rest from 10EQ) who used the equation strategy in the post-test. 7 of these (May, Melvyn, Jack, Dick, Jane, Judy and Harry) who had in the pre-test either left the question blank or used difference division, successfully used the equation strategy in the post-test. However Judy wrote down how much Pat threw away rather than Sarah; while Harry accidentally transposed 248 to 284. Jocelyn wrote "can't do it" in the pre-test, and used the equation strategy in the post-test; but he made a slip ($-B + 4B$ becoming $4B$ rather than $3B$). The choice of letter is interesting - is it B for "weight of bricks" or B for "bricks" or B because that's what EQUATION used (for barrels)? Liam and Melissa did virtually the same as Jocelyn. Grace used trial-and-error in the pre-test; and the equation strategy in the post-test; but she made a slip connected with negative numbers. Jennifer and Mickey both started to represent the situation, but gave up after finding it difficult to write down the right-hand-side of the equation.

For the control groups, on the other hand, only 1 student attempted (half-heartedly) an algebraic strategy, with the others evenly divided between trial-and-error, "Difference ÷ 4 or 5" (variants of which featured in Lins' study), unknown strategy or blank. All the correct answers with identifiable strategies were obtained by trial-and-error. Trial-and-error was a far more popular strategy that in Lins's groups.

Interestingly, there is a marked difference between the males and females on this item for Year 10. Only one male (out of 13) got it correct in the pre-test, and 7 in the post-test. The females improved from 4 (out of 9) to 6. For 10CON, only two males (out of 10) were correct for each test; whereas the females improved from 5 (out of 14) to 8.

B2	10EQ			
	Pre		Post	
Name	✗/✓	Strategy	✗/✓	Strategy
Debbie	✗	-	✗	D
Rebecca	✓	T	✓	A
Judy	✗	-	✗	A
Arthur	✗	?	✗	?
Samuel	✗	?	✗	-
Jocelyn	✗	-	✗	A
Liam	✗	-	✗	A
Harry	✗	D	✗	A
Bruce	✓	T	✗	?
Darren	✗	D	✓	?
Tracy	✓	?	✓	?
Josh	✗	D	✓	?
May	✗	-	✓	A
Rajiv	✗	-	✓	?
Dick	✗	D	✓	A
Cedric	✗	-	✓	?
Jane	✗	D	✓	A
Grace	✓	T	✗	A
Joanna	✓	?	✓	?
Jack	✗	-	✓	A
Rose	✗	D	✓	T
Melvyn	✗	-	✓	A

B2	10CON			
	Pre		Post	
Name	✗/✓	Strategy	✗/✓	Strategy
Christopher	✗	D	✗	D
Duncan	✗	-	✗	-
Fatima	✗	D	✗	D
George	✗	D	✗	-
Chelsea	✗	D	✗	?
Ashleigh	✗	D	✗	T
Guy	✗	D	✗	O
Robin	✓	?	✓	T
Reuben	✗	-	✗	-
Harriet	✓	T	✓	T
Heather	✗	-	✓	?
Rhiannon	✗	?	✓	T
Helen	✓	T	✓	T
Mathew	✓	?	✓	T
Jerry	✗	?	✗	A
Nick	✗	D	✗	-
Lauren	✗	T	✓	?
Christine	✓	?	✓	?
Elizabeth	✓	?	✓	T
Isabelle	✓	T	✓	T
Maria	✗	?	✗	-
Thomas	✗	-	✗	-
Hannah	✗	-	✗	-
Alice	✗	D	✗	D

Strategies: A = algebraic strategy, T = trial-and-error, D = difference division, ? = unknown strategy, O = other strategy, - = item left blank, ✓ = answer correct, ✗ = answer incorrect

It seems reasonable to suggest, then, that for about half the Year 10 students, EQUATION encouraged the development of an algebraic strategy for the seesaw puzzle; and that about half of these had difficulties with some aspect of executing the strategy. Year 7 were not generally helped to develop an algebraic strategy.

6.3.7 B3: Secret Number ($2x + 6 = 3x - 70$)

On the pre-test, 10EQ mostly either left the question blank (9) or tried an algebraic approach (8, of which 3 successful); 3 used an unknown strategy (2 successful); and 2 tried trial-and-error (0 successful). Interestingly, half the females and none of the males were correct.

On the post-test, those in 10EQ who had tried an algebraic approach on the pre-test tried it again on the post-test, with the exception of two who left the item blank; and 6 more students tried it. Those who were successful before were successful again; and two who were unsuccessful with the algebraic strategy the first time corrected their errors the second time. However, half of the 12 in total who used an algebraic strategy were unsuccessful. There were 5 unknown strategies (all but one successful), 1 trial-and-error (successful), and 4 blanks. Errors in the algebraic strategy centred around difficulties with the negative sign. 5 males improved, and only 1 female; yet proportionally 2/3 of the females were correct and just over 1/3 of the males.

B3	**10EQ**			
	Pre	**Post**		
Name	✗/✓	Strategy	✗/✓	Strategy
Debbie	✗	-	✗	-
Rebecca	✓	A	✓	A
Judy	✗	-	✗	A
Arthur	✗	-	✗	?
Samuel	✗	-	✗	A
Jocelyn	✗	A	✗	A
Liam	✗	A	✓	A
Harry	✗	?	✗	A
Bruce	✗	T	✓	T
Darren	✗	A	✗	-
Tracy	✓	?	✓	?
Josh	✗	A	✗	-
May	✗	-	✓	?
Rajiv	✗	-	✓	?
Dick	✗	A	✓	A
Cedric	✗	-	✗	-
Jane	✓	A	✓	A
Grace	✗	T	✗	A
Joanna	✓	?	✓	?
Jack	✗	-	✗	A
Rose	✓	A	✓	A
Melvyn	✗	-	✓	A

B3	**10CON**			
	Pre	**Post**		
Name	✗/✓	Strategy	✗/✓	Strategy
Christopher	✗	-	✓	?
Duncan	✗	-	✗	-
Fatima	✗	-	✓	?
George	✗	-	✓	T
Chelsea	✗	-	✗	-
Ashleigh	✗	-	✗	-
Guy	✗	A	✗	A
Robin	✗	-	✗	-
Reuben	✗	-	✗	-
Harriet	✓	T	✓	T
Heather	✗	A	✓	?
Rhiannon	✗	T	✓	T
Helen	✗	O	✓	T
Mathew	✗	?	✗	T
Jerry	✓	A	✗	A
Nick	✗	-	✗	-
Lauren	✗	T	✓	?
Christine	✗	A	✗	A
Elizabeth	✗	T	✗	-
Isabelle	✗	A	✓	?
Maria	✗	?	✓	?
Thomas	✗	-	✗	?
Hannah	✗	?	✗	?
Alice	✗	-	✗	T

A = Algebraic strategy, T = Trial-and-error, ? = unknown strategy, O = Other strategy, - = question left blank

5 students in 10CON, meanwhile, used an algebraic strategy on the pre-test (one of whom was successful); and only 3 used one on the post-test (none of whom were successful). Trial-and-error appeared to be effective, however there were 6 who were successful with unknown strategies. This made B3 the most improved item for that class. Although there was no male-female discrepancy in the pre-test, 7 females improved (out of a possible 13), compared with only 2 males (out of a possible 9).

There is little evidence of Year 7 improving their strategies. The Year 6 students, meanwhile, were unable to solve word problems outside the context of EQUATION.

For B3, as for B2, there is a marked difference between the males and females on this item for Year 10. None of the 13 males were correct in the pre-test, and 5 were correct in the post-test. The females improved from 5 (out of 9) to 6. For 10CON, only 1 male (out of 10) was correct in the pre-test, and 3 in the post-test; whereas the females improved from 1 (out of 14) to 8. Putting the scores for B2 and B3 together, it can be seen that the 10EQ males (from 1 to 12, out of 26)and 10CON females (from 6 to 16, out of 28) appeared to gain most by repeating the test; whereas the 10EQ females (from 9 to 12, out of 18) and 10CON males (from 3 to 5, out of 20) improved only slightly. Nevertheless, for virtually every class on each sitting, females outperformed males on the word problems as a whole.

Comparing B2 and B3 for 10EQ, it can be seen that those who attempted an algebraic strategy on the post-test of B2 mostly also tried one on B3. Moreover, B3 seemed to induce more algebraic strategies than B2, on both tests. It is striking that nobody in 7CON apparently attempted an algebraic strategy for any of the modelling items, and well over half failed to get any correct answers at all. For 7EQ, there were only three apparent attempts at an algebraic strategy (all unsuccessful), and over three-quarters failed to get any correct answers at all.

6.3.8 Posing Word Problems

Because the students were not asked to pose their own algebraic problems at the start of the research, it would be presumptuous to claim that the problems posed at the end were as a direct result of using EQUATION. Nevertheless, it is possible to use the problem-posing evidence to place limits on students' prima facie ability to devise problem situations in which an algebraic strategy might be an effective way of solving the problem. For example, Kirsty and May posed a TOAN puzzle, which might involve the use of bracketed expressions:

> "I think of a number and double it then I add 6, I then halve it and minus 3. I end up with 80, so what was my original number?"

Yet in their solution, they suggested a non-algebraic reversal strategy: "To get the answer you have to plus 3, then multiply by 2, minus 6 and then halve it.". In their own critique of the problem, they acknowledged the lack of challenge, but failed to recognise the value of taking steps to ensure unknowns on both sides of the equation - they did not go beyond suggesting that a "harder formula" be used. Nor did they note the ambiguity of the second "it".

Dick posed a problem which is reminiscent of the album sale, but which is typical of introductory examples for simultaneous equations (which did not appear in EQUATION):

> "John buys 4 tapes and 3 CDs for £60. Alfonz buys 1 tape and 2 CDs for £35. How much does (a) a CD and (b) a tape cost?"

Rajiv and Seb similarly adopted the language of an EQUATION problem, but without managing to pose a problem for which algebra would be appropriate:

> "Sarah and Claire are sitting on a seesaw and Sarah is carrying 195lb and Claire is carrying 390lb and they are not balanced at all. If they both throw away ¼ of their weight, find the ratio of the weight they are carrying compared to each other."

Nor (it is clear from their solution) did they realise that the existence of the seesaw or the throwing away of ¼ of the weight do not affect the ratio, even given the potential ambiguities of the problem wording. They suggested that the problem is "complicated"; and even though there are "not enough variables", the mathematics is "about right".

Similarly, there is rather less to Samuel's problem than the age problems in EQUATION (presumably his father is *always* 30 years older):

> "My father was 30 years older than me 3 years ago, and I will be 16 in 2 years. How old is my father?"

Lisa, meanwhile, posed a more elaborate problem which uses algebraic notation explicitly yet somewhat artificially, and fails to avoid ambiguities that may be responsible for the problem being insoluble:

> "My dad owns a car, a Rover 3600 made in 1967. It cost £1500 when new. Now it is worth my dad's age times 80. My dad was born 1500 add 22 minus 1482 years ago. My dad also says I can drive it when I am seventeen. At the present moment I am x divided by 2 minus 140 add 5 years old. x is the cost of the car in 1967 minus the present value, divided by 10, times 2 and minus 40.
>
> a)How much is my dad's car worth now?
>
> b)How old is my dad?
>
> c)How old am I at the moment?
>
> d)What is x?"

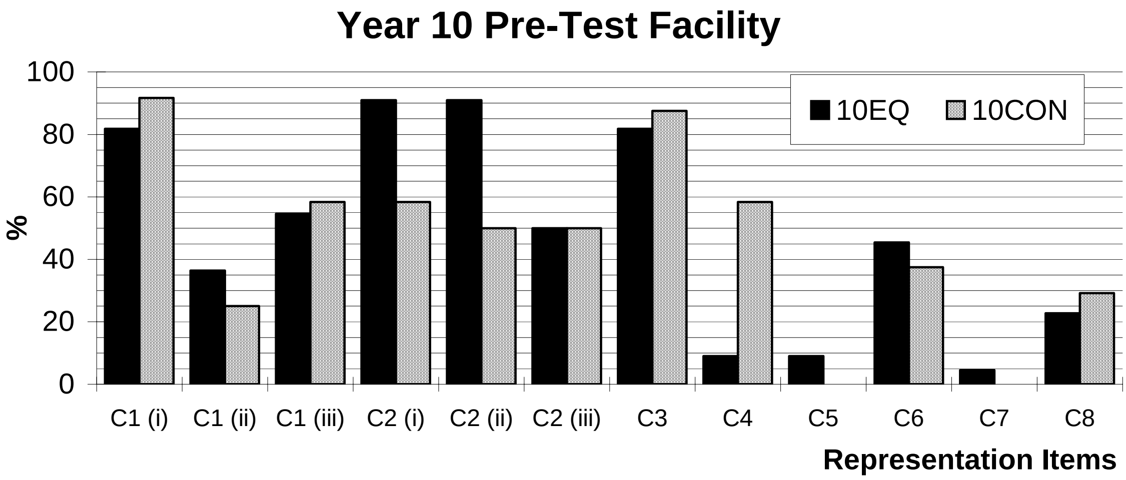
But then there are problems like Debbie's, that are creative improvements on problems in EQUATION, and that are sufficiently challenging for algebra to be useful:

> "If you take my age and triple it, it is my father's age. My mother is my father's age divided by two, multiplied by three and you then subtract 25. My mother's age is twice mine plus 10. What are our ages?"

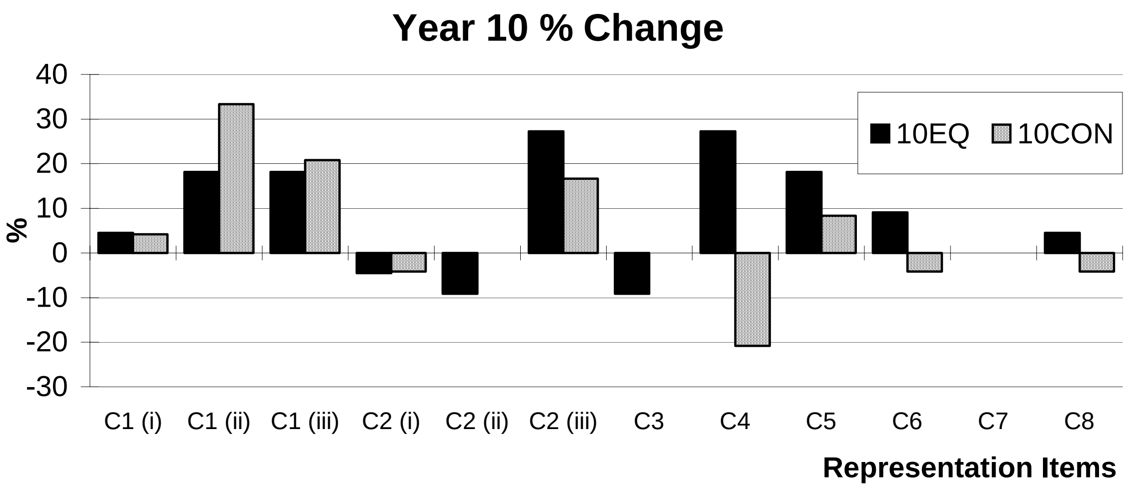
In item C4 of the written test, the students were asked to find a situation in which $x = 4c$ might help. Although this is not a direct request to pose a word problem (which it perhaps might have more profitably have been), it does demand the ability to recall or create a situation in which modelling might be useful. 7 students in 10EQ improved on this item; while 10CON declined.

6.3.9 Representation

As the chart below shows, 10EQ and 10CON are similar for most representation items. However, there is clearly some discrepancy on C4, but it is not really a representation problem (see above). The reasons for differences on items C2(i) and (ii) - which require the finding of a perimeter - are not so clear. One popular 10CON answer for C2(i) was $4h \times t$, so maybe confusion over perimeter is to blame.



The following chart shows the changes in these items:



C1(ii), C1(iii), C2(iii), and C5 showed improvement for both classes. Yet overall for the representation items, 10EQ made significant progress whereas 10CON did not. The interpretation item C4 shows a large divergence: 10EQ improved from 9% to 36%; 10CON declined from 58% to 38%; both are statistically significant. C1(iii) is a test aberration (see the previous chapter), but explaining the other 10CON improvements is difficult. C1(ii) on the pre-test required the area of 5 × (e + 2) rectangle; while the post-test involved a 7 × (k + 3) rectangle; this seems to be an innocuous enough difference. Wrong answers on the pre-test included the obvious 5 × e + 2, but also 5e + 2, 5e + 4, 5 × e + 4, 5 × e, 5 × e5, and 5 × e × 2. It is possible (though perhaps implausible) that the reason for improvement is that D7 on the pre-test assisted C1(ii) on the post-test. The same argument might then apply for D8 and C2(iii). To add to the puzzle, the males remained fairly static for C1(ii) (around half correct); whereas the females improved from 2 correct (out of 14) to 9. Meanwhile, all 5 of the males who got C2(iii) wrong on the pre-test were correct on the post-test, compared with 3 out of the 6 females (one decline for each). In any case, as for modelling, the value of having a control group is again demonstrated by these unexpected improvements, as a caution against hasty interpretation of the experimental group's results. In particular, Hart *et al.* (1981) found a facility of 24% for Year 8, rising to 41% for Year 10. The results here would suggest that minimal algebraic experience (such as for these Year 7 classes) makes this a virtually impossible item; while although the pre-test facility for Year 10 corroborates the CSMS result, merely giving the test again can apparently result in a quarter of the students improving. Similar considerations apply to D1(ii) below.

It is difficult to identify strategies from the items in this section, because there is little need for intermediate written steps between question and answer. However, some error analysis of the student-professor problem and penny-dime problem is possible for 10EQ:

10EQ	**C6**		**C7**	
Name	**Pre**	**Post**	**Pre**	**Post**
Debbie	✓	✓	e	e
Rebecca	r	r	e	e
Judy	✓	✓	-	e
Arthur	r	r	o	o
Samuel	-	r	-	✓
Jocelyn	r	✓	r	e
Liam	✓	✓	e	e
Harry	r	r	e	e
Bruce	✓	r	o	e
Darren	o	✓	✓	r
Tracy	✓	o	o	o
Josh	o	✓	r	r
May	r	✓	-	e
Rajiv	r	✓	r	e
Dick	✓	✓	r	r
Cedric	o	✓	e	r
Jane	r	r	r	r
Grace	r	r	e	e
Joanna	✓	o	e	o
Jack	✓	✓	r	r
Rose	✓	r	r	e
Melvyn	✓	✓	e	e

✓ = *correct (e.g.* $S = 15T$ *for C6,* $P = 5F$ *for C7), r = reversal error (e.g.* $15S = T$ *for C6,* $F = 5P$ *for C7),- = item left blank, e = equality (e.g.* $P = F$*), o = other error*

C6 is unusual in that around half of Year 10 changed their answer on the post-test; it is one of the most volatile items on the test.

C8 is odd in that none of the 9 females in 10EQ got it correct in either test; yet 5 out of 13 males were correct on the pre-test, 3 improved on the post-test, and 2 declined.

Comparing performance on problems C6-C8 with problem A2 reveals a rather interesting result. In A2 students were given two simply-related sequences of numbers, and asked to predict a number in one sequence given the corresponding number in the other sequence. The facility for algebraic representation was around 90% for Year 10, even though they had to find the relation before they could represent it. Yet when, in C6-C8, they were given a relation in words, the facility for algebraic representation was no better than 45%.

One major reason for carrying out the Year 7 fieldwork was to provide corroboration for the (potentially astonishing) improvement in representation problems. The Year 7 results do indeed provide this corroboration to some extent; however, the facilities are much smaller, and the number of items showing improvement is also smaller. Only 4 (out of 26) students in 7EQ got any representation items correct in the pre-test (Basil 4, Joe 2, Melissa 1 and Eva 1). There were hints that the difficulty lay in understanding what was being asked. For example, Jordan asked "Is there an alphabet of numbers?"; Margaret gave numerical answers to C1 to C3; while Hugh effectively offered numerical answers in the guise of algebraic expressions for C5 ("4r and 13b"), C6 ("6T 90S") and C7 ("4F 20P"). However, Scott and Kevin did at least manage to provide a reversed equation in C6. In the post-test, 12 got at least one item correct; and the number of correct items went up from 8 to 26. The most improved item was C3 (from 3 students correct to 10), followed by C6 (from 1 student correct to 7). No other item was correctly answered by more than 2 students. There was little indication of improvement in C2(ii) or C2(iii) - the only items featuring just one letter - only Basil improved in these two. There were still responses to C3 such as "J = 10, P = 16, 26 marbles altogether" (Anna and Tony); but even those who failed to make much improvement in score seemed from their responses to have a better grasp of what sort of answer would be appropriate. For example, Susan did at least abandon numerical answers:

Item	**Pre**	**Post**
C3		J + P = m
C5	b = 6 R = 10	
C6	S =	15 × S ÷ T
C7	P = 30 F = 30	F/P =

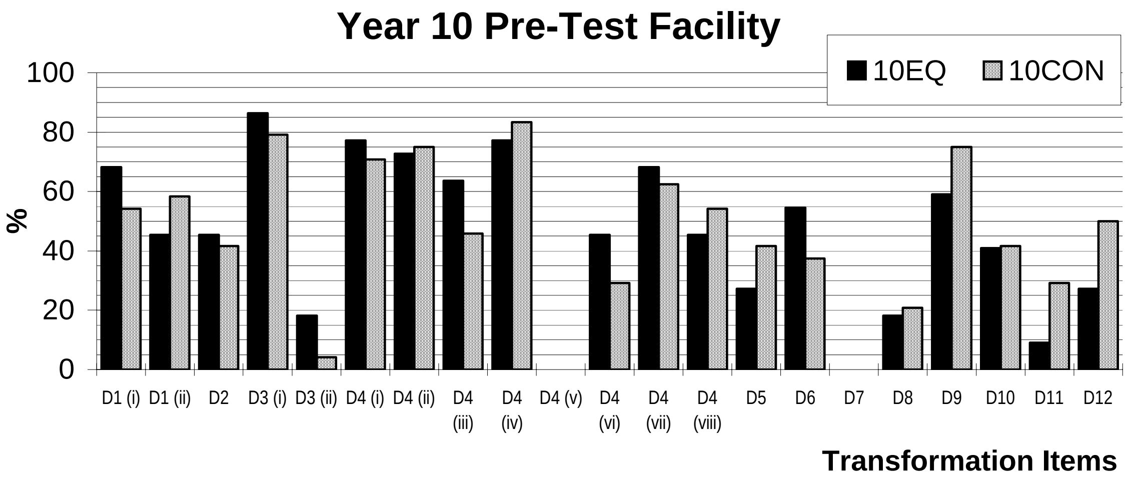
Similarly, Jessica on C6 moved from "T = 15 S = 225" to "S = T × 15". Meanwhile, Eva recognised in C6 the need for equations (S ÷ 15 = T on the post-test) rather than expressions (15S × T on the pre-test).

For 7CON, on the other hand, not only did the number of students getting at least one item correct move up by only 1 (from 7), but there were only 5 improvements (and 1 decline), of which Deirdre accounted for 3. Most of the responses here are numerical, but show little consistency with the pre-test.

The two Year 6 students from School C similarly showed little improvement in representing the perimeter of a shape labelled with letters and numbers - their expectation appeared to be that the answer should be a number (as opposed to an expression). Ways of achieving this including assuming "standard" values for letters (e.g. $a = 1$, $b = 2$, etc.), assuming symmetry or attributing an arbitrary role to a letter (e.g. $3n$ means multiply 3 by itself). If a letter appears more than once in a diagram, it can stand for something different in each case.

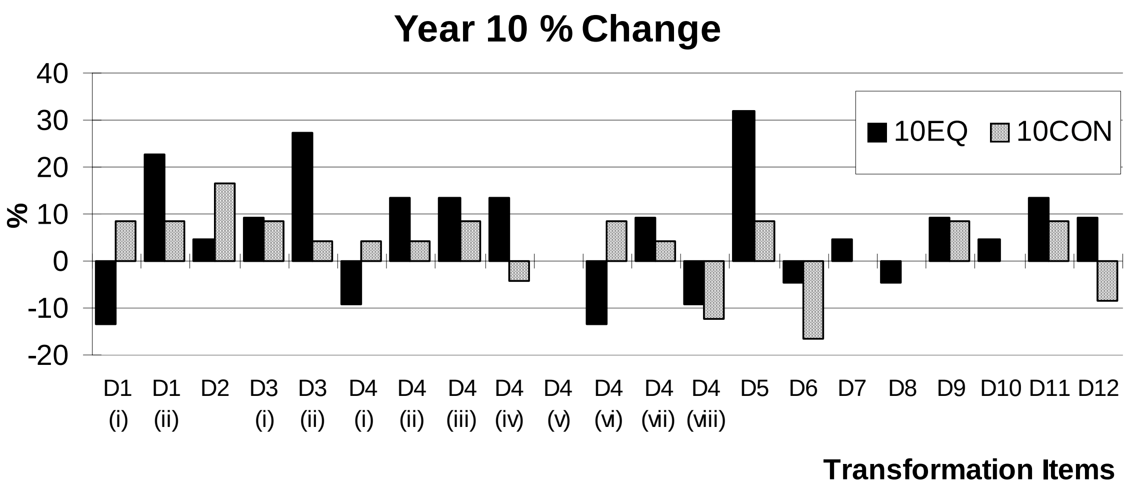
6.3.10 Transformation

The following chart shows the facility of the 21 transformation items.



It is clear that there are pre-test differences between 10EQ and 10CON. 10CON seem to be better in items D9, D1(ii), D12, D5 and D11. 10EQ seem to be better in items D1(i), D4(iii), D6, D4(vi) and D3(ii). Although these correspond to small differences in frequency (around 3 students), it is interesting that 10CON seem to do better on items requiring substitution and 10EQ seem to do better on items for which treating letters as objects (as in Küchemann, 1981) can help get a right answer. It is puzzling that while all the females in 10EQ were correct in D4(ii), around the half the males were incorrect.

The post-test changes are interesting:



As can be seen, change in many of the items was fairly small. Of the equation-solving items, D3(ii) improved for 10EQ quite dramatically (8 students improving, 2 declining); while D3(i) moved up only slightly given that the facility was already quite high. There are dramatic changes for D5; and the overall change in the representation section is significant for 10EQ even if D3 is excluded from consideration.

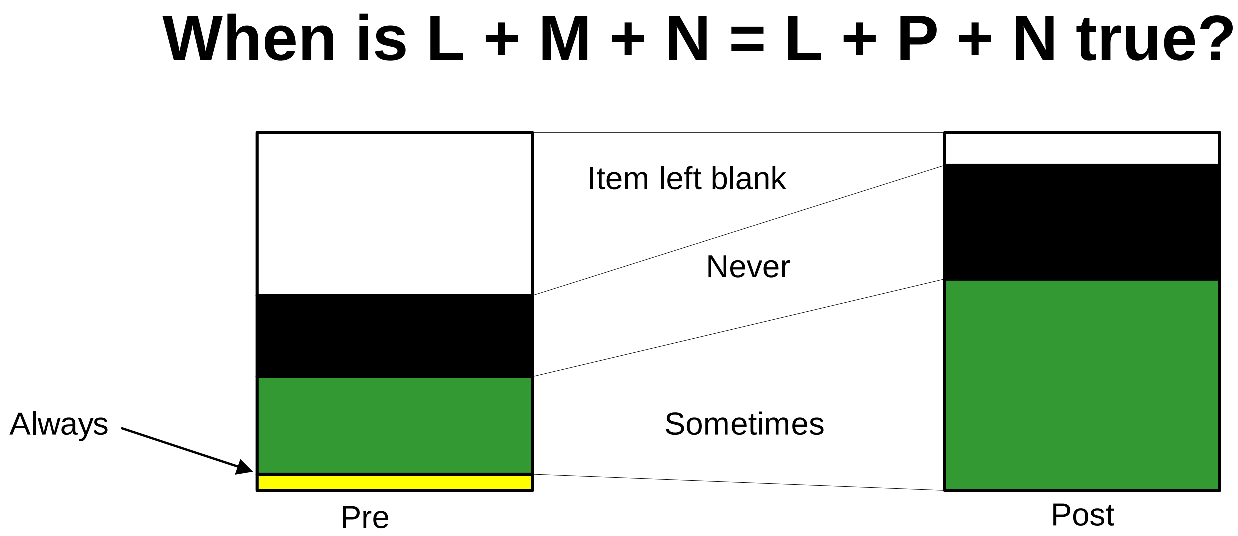
The facilities hide changes in D9. For the EQ groups there were 8 improvements and 5 declines in total; whereas for the control groups, there were 6 improvements and 2 declines. Also hidden is the fact that for 10EQ, the facility of D12 for the 13 males improved from 4 to 7; whereas the 9 females declined from 2 to 1. These results also raise questions about the validity of the CSMS item D1(ii) - see C2(iii) above - although of course such a tiny sample size (100 compared to 3000) mitigates against further conclusions in this regard.

For 10EQ, one student (Bruce) improved on 8 items; while another (Dick) improved on 7, although he declined on 3 others; Darren improved in 6 and declined in 1; Grace declined in 6 without improvement. However, it is difficult to claim that these resulted from the use of EQUATION: 4 students in 10CON improved in between 6 and 8 items.

Similarly to the representation section, transformation strategies are difficult to detect in the scripts. However, in D1(ii) the simplification $4 + 3n = 7n$ is evident:

10EQ	**D1(ii)**		**D5**	
Name	**Pre**	**Post**	**Pre**	**Post**
Debbie	7	7	-	S
Rebecca	✓	✓	N	N
Judy	O	-	-	S
Arthur	7	7	S	S
Samuel	-	7	A	N
Jocelyn	✓	✓	S	S
Liam	✓	✓	S	S
Harry	O	✓	S	S
Bruce	O	✓	-	N
Darren	O	✓	N	S
Tracy	O	✓	-	N
Josh	7	✓	-	N
May	✓	✓	-	-
Rajiv	✓	✓	-	S
Dick	✓	✓	-	S
Cedric	7	7	N	-
Jane	✓	✓	S	S
Grace	7	-	N	N
Joanna	7	✓	N	N
Jack	✓	-	-	S
Rose	✓	✓	S	S
Melvyn	✓	✓	-	S

D1(ii): ✓ *= correct (e.g.* $4 = 3n$*), - = question left blank,* 7 = 7*n written,* O = *other error*
D5: A = Always, S = Sometimes, N = Never



Reasons given for “Never” include “Because both sides of the equation don’t equal the same”, “Because N isn’t in the first equation on the LHS”, “Can’t have all different letters = same thing”, “P would have to be M”. Darren ticked Never on the pre-test because “it’s not balanced” and Sometimes on the post-test because “I don’t know what L, P, N equal”. However, his reason changed to “if P = M” on the delayed-test.

Even though there were only two equations to be solved in the standard test, the post-test facilities clearly demonstrate severe limitations on Year 10’s skills in this regard. Although over 80% could solve D3(i) $3x + 5 = 17$, fewer than 50% in 10EQ could solve D3(ii) $3x + 6 = 2x - 30$ (an equation with a negative sign, a letter on each side, and a negative answer); and below 10% for 10CON. The advanced pre-test and the interviews provide a little more detail, showing that students cannot be easily categorised by the type of equations that they can solve. Cedric, for example, was able to solve equations with a letter on each side, even though he could not solve D3(ii). Darren was able to solve equations in which the negative signs could be eradicated on each side independently, but ran into difficulties with equations involving brackets, and similarly with all the other items containing brackets - D1(ii), D4(iii), D4(vi) and D7. Bruce continued in the post-interview to treat the minus sign as applying to the term to the left of it, rather than to the right of it. For example, he solved $8 - x = 2x + 5$ by adding 8, and getting $x = 2x + 13$. Jocelyn failed to solve D3(ii) on both pre and post-tests because of (different) difficulties with negative signs, but he was able to solve $5y - 11 = 2y + 1$ and $63 - 5n = 28$ because those particular difficulties were not applicable to these equations. Judy, on the other hand, had difficulty with any of the equations involving a negative sign, and there was an apparent knock-on effect for those without negative signs: $5x + 12 = 3x + 24$ was simplified to $8x = 36$. Lisa, meanwhile, simplified it to $8x = 12$, and appeared to have a policy of always adding the unknowns. Yet she was also able to solve with ease $561/x = 22$, equations with brackets, and simultaneous equations.

As in previous sections, success for Year 7 was rare. Over a third of 7EQ failed to get a single transformation item correct; Kevin got 4 (out of 21), the rest got 1 or 2. Some pre-test answers showed a classic lack of understanding of the questions being posed. For example, both Kevin and Scott thought that $2a + 7b$ could be written more simply as $7c$. To D8 ("Which is larger, $2n$ or $n + 2$?"), Shannon wrote "I'm not sure as the 2n could be a two figure number. Whereas n+2 may not be.". Susan illustrated this with a numerical example: "$2n$ because $2 + 8 = 10$ but you could have 28". Margaret, on the other hand wrote "Neither there the same!!!!". Dylan replied to D6 that $7w + 22 = 109$ and $7n + 22 = 109$ *sometimes* have different solutions, "because the numbers aren't usually the same letter".

On the post-test for 7EQ, no items showed significant improvement, and a similar proportion of students scored nothing. D1(i) and D5 remained the easiest items. Basil got 5 items correct; and Catherine got 4. There was actually a decline in D3(i): 3 of the 4 students who got D3(i) correct on the pre-test got it wrong on the post-test. Catherine responded to this item (which asked "If $3x + 5 = 17$ then what is x?") with "x is a symbol to signify a thing". On D8 she wrote that $3n$ and $n + 3$ are the same because "if you add $n + 3$ you get $3n$". Melissa thought they were the same for a different reason: "they use the same numbers and letters". Basil, meanwhile thought that "$3n$ means $3 \times n$ which is more than $3 + n$". Jordan wrote that $7w + 22 = 109$ and $7n + 22 = 109$ *always* have different solutions because "they have different letters". Kevin's answers to D4 are interesting: $7a - 3a$ simplifies to $7 + 3 = 10a$; $(a + b) + a$ simplifies to $a + b + a = p$; $a + a + a \times 2 = 2c$; $(a - b) + b = c$; $5a - b + a = n$; $4a - 2 + 7a + 1 = x$. Charlotte responded quite reasonably - given C6, that is - to D11 ("What can you say about r if $r = s + t$ and $r + s + t = 40$?") with "t × 15=s". The only apparent ray of hope was in the responses of Jennifer and Shannon, whose pre-test answers to D1-2 were numerical, but whose post-test answers were at least expressions.

The pre-test scores and responses of 7CON were similar to those of 7EQ. However, on the post-test, although both groups had 16 improvements, 7EQ had 15 declines and 7CON had 4 declines. For 7CON, Deirdre accounted for 8 of these improvements, and Stuart accounted for 3. The obvious implication would be that Deirdre received some sort of tuition between the tests. She wrote, however, that $a + b + c = a + f + c$ was never true (D5) because "f doesn't = b"; and that $3n$ is larger than $3 + n$ because "multiplying makes things bigger than adding does". Graham though that if $e + f = 8$ (D10) then $e + f + g$ is $5 + 6 + 7$.

The two Year 6 students were able to solve mentally equations with whole number solutions (e.g. $5x + 6 = 21$) with greater speed. But other equations caused difficulties - for example when the equation was changed to $5x + 6 = 20$, the students re-interpreted $5x$ as $5 + x$ in order to allow the possibility of a whole number solution, because it was recognised that otherwise the answer would be "2 point something". Jacob suggested that if he had the computer he would take 5 away - and when the equation was turned into a balance picture, he realised that 6 should be subtracted. They were unable to solve such equations outside the context of EQUATION.

6.3.11 Meta-algebraic Aspects

One striking fact about the questionnaire results is that while the students in both groups produced a wide variety of responses, what each student in the control group wrote in December is remarkably similar to what he or she wrote in June. This is in spite of a change of mathematics teacher in September, and also, presumably, more focus on exam technique than understanding. These would have constituted good reasons for scepticism about ascribing any changes to a particular intervention. Yet when the EQUATION students were asked why they thought some people found equations difficult, there was a shift from *operational* reasons such as “because they do not understand how to do them” to more *structural* reasons such as “the use of letters to stand for numbers confuses people”. From around half the students emphasising structure in June, almost three-quarters did so in December (the control group remained at 40%).

When they were asked how they would describe equations to others in their own words, there was a richer set of responses from the EQUATION group, as they tried to cover not just the syntactical and operational aspects in their answers, but also the semantic and purposive aspects.

For example “numbers put together like a sum” seems to be (if we’re generous), trying to capture the objects, structure and operations of an equation, and so addressing the syntactic aspect. But it does not emphasise a *sense of equality* (the semantic aspect), because, in “sums”, if the equals sign is not completely bypassed using the notation of standard algorithms, as was noted by Behr *et al.* (1976), it nearly always means “makes” rather than “is the same value as”. Nor does “numbers put together like a sum” appear to attempt to give either an idea for the sort of ways in which one might *deal with* an equation (the operational aspect) or a rationale for *being concerned with* equations (the purposive aspect).

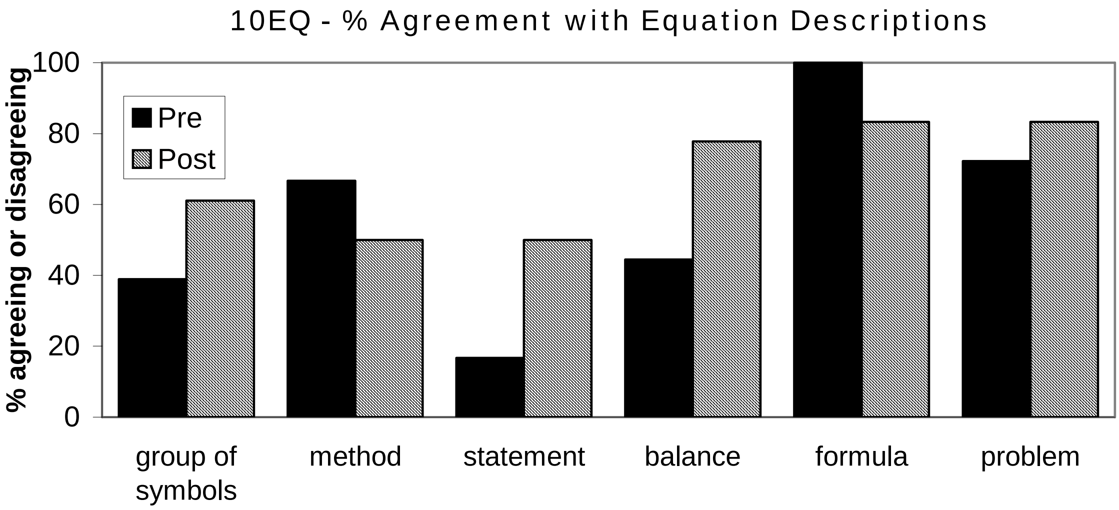
On the other hand, “... like a set of balancing weighing scales and you have to eliminate parts to find the anonymous value” seems not to be concerned so much with syntax as with a metaphor for equality and with what operational procedures might be appropriate.

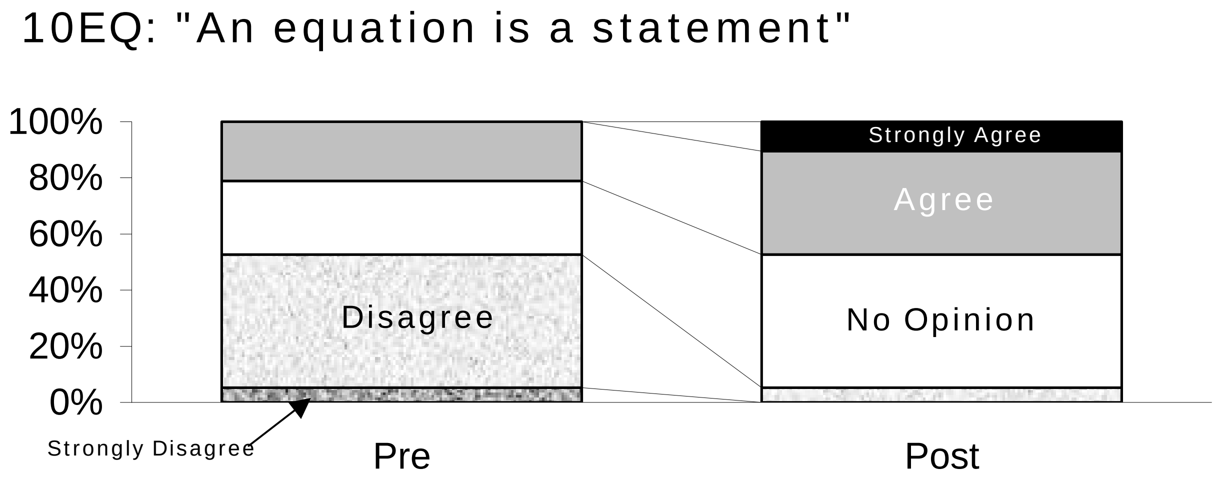
“... something that helps you solve a problem” clearly does not specify very much that is unique to algebra, but it does at least acknowledge the question of why one might be concerned about equations.

A formulation that involves syntax, operation and purpose is: “An arrangement of numbers and letters that can be simplified to work out a problem. The letters in the equation replace numbers.”. But this does not address the semantic aspect.

In explaining equations to others, syntax continued to be viewed as significant, but awareness of operational factors grew. The control group did not show the greater richness of responses between June and December that the EQUATION group showed.

When the students were asked to select from given descriptions, there was an increase in agreement with “statement” as a description for 10EQ, whereas 10CON remain static.





The diversity of responses is similar to that obtained when A-Level students were interviewed, and said things like “[An equation is] sort of a way of finding unknowns”, “a rule that determines”, “Balance, because each side is equal, so it’s balanced”, “Creates a picture if you draw it.”, “Thing with *x*’s in.”, “It’s more to me as a formula… because you’ve got something that is equal to something else. And you could put values in to find out what other values are going to be. … The formula has to balance, doesn’t it? For it to *be* a formula.”. Again, for the Year 6 students: “Stuff like 4*x*…”, “sometimes they have another question underneath [an equation] that you have to work out the answer”. But note the lack of statement-related comments. Moreover, the value of equations was seen either in terms of schoolwork or science-related occupations.

For 10EQ, confidence about algebra remained roughly the same. Although 6 people indicated higher confidence in EQUATION, compared with only two in the control group, this is hardly indicative of any sea-change in attitudes. This is in sharp contrast with the comments made after the first EQUATION lesson, and, to a lesser extent, the second lesson. Comments such as “It’s really good.”, “It makes things really easy.”, “I get it now.”, and “It really helps you understand.” suggested that a more positive attitude to algebra might ensue. However, both the teacher and researcher noted that the second lesson was more frustrating for some students, as they struggled to create equations to model the word problems; and the work in the summer term on inequalities and in the autumn term on quadratics and simultaneous equations will have reminded students that there is more to algebra than formulating and solving simple linear equations. Moreover, this question of confidence is surely more prone than most to feelings about the last piece of algebra work attempted.

Chapter 7
Conclusions

7.1 Introduction

The previous chapter identified some student strategies that are conjectured to have improved while using EQUATION in class, and some problems in the pre-post testing for which there appear to have been improvements. In this final chapter, these improvements are first discussed in order to elucidate coherently the main empirical claims. Then the analysis is used in three ways:

1. to reconcile the classwork and pre-post testing;
2. to challenge claims about the balance model;
3. to compare EQUATION with other initiatives.

The thesis concludes with some suggestions for further research.

7.2 Discussion: Classwork Improvements

7.2.1 Simplifying Balance Puzzles

The logs of student usage recorded by the program appear to show the active creation and improvement of strategies. This is especially convincing when listening to the audio-tape of the students' conversations while the program replays on the screen what the students saw and did. Many of their strategies are clear - some are more obscure - but students developed for themselves a simplification strategy for the balance puzzles. There is no evidence that they used prior algebraic knowledge at all; and although some students (such as Rebecca and one of the Year 8 students at School B) perhaps saw parallels with simplification of equations, there is no evidence that they appreciated any implications.

The graphs of times per puzzle and times per level for each student suggest that the strategies were developed piecemeal, in response to changes in the objective problem situation that the students faced - mostly at the start of each level. The developed strategies replaced informal strategies, at varying points in the program for different students. This strategy development in response to critical situations constitutes an indication that the concern to find an unknown number in a situation was well formed. Students' "images" or rationales for strategies seemed to play a minimal role in this improvement, however. Moreover, the strategies had the character of expectations about what aspects of the problem situation were important, and how they might need changing. Recall "I reckon you take off barrels until you get one each.", "Yeah, it's done it the same on each side." and "Why can't you do it?".

The strategies developed include:

- Match barrels and weights [possible source of error: counting and subtraction].
- If a certain weight plus an unknown weight weigh the same as another known weight, subtract knowns to find the unknown [possible source of error: larger – smaller].
- If a certain number of barrels weighs a certain amount, divide the amount by the number to find the weight of each barrel [possible source of error: larger ÷ smaller].
- Remove weights and barrels from both sides to simplify a situation [possible source of error: forgetting to remove from *both* sides].
- Take off as many barrels and weights as you can.
- Take off as many barrels as there are on the side with the smaller number of barrels.
- Take off as much weight as there is on the side with the smaller known weight.
- The particular visible weight pictures can be ignored - it is the total weight that determines the answer.

However, were the students just subtracting *objects*, or were they also subtracting *quantities*? The move to single weight pictures on Level 7 may have made object-subtraction less amenable, but it is difficult to tell much more without moving away from the concrete situation. Nevertheless, there is some evidence that the bypassing of arithmetic was crucial in enabling students to focus on simplification decisions.

7.2.2 Simplifying Balance-Like Equations

I suggest that what we have seen here from Level 9 onwards is the *continued* application of the simplification strategy. It is, of course, possible that students used their pre-existing algebraic knowledge to solve these equations, or that they developed strategies anew without reference to the balance puzzles. But there is no evidence for either of these scenarios. Moreover, the fact that the strategies continued just as before (even a little faster), and that the later break with the balance model caused ructions, suggests that the strategies developed during the balance puzzles were transferred to the balance-like equations.

Although this means that concern to find an unknown number survived the transfer to symbolism, it also means that the balance pictures are not far from the surface: at one extreme, $5b$ means "5 barrels", rather than "5 multiplied by the number that b represents". One way of breaking with the balance model at this point - if one were not exploring the limitations of the balance model - would have been to introduce a new problem situation (such as TOAN) so that students can realise that their simplification strategy does not need objects to work.

7.2.3 Simplifying Equations that are Unlike Balances

It is claimed that the break with the balance model (in this case, equations with negative numbers and negative signs) caused the *extension* of the strategies developed earlier for balances to equations that are unlike balances.

For Rebecca & Nicola, the break with the balance model on Levels 11 and 12 was less traumatic than might be the case with paper-and-pencil exercises. I suggest this is because a very limited range of transformation strategies was involved, because this range was increased only gradually, and because they were able to develop strategies for coping with negative answers and negative signs by exploring algebraic form. With paper-and-pencil, students not only have to choose the operations to perform, but they also have to carry them out correctly. With EQUATION they have the opportunity to experiment with algebra, to try hunches and make mistakes that do not lead to having to start again.

The role of feedback here is crucial with respect to an important theory that has not hitherto been discussed. Rebecca and Nicola seem to be well aware that the choice of an unhelpful operation is not fatal to the solution process, because it can be undone - the solution will be unaffected because the operations are carried out on both sides of the equation. Recall "You should've *plussed* 12… Should be *plus* 24. 'cos you want to get rid of it, so you want to get it to nought, so you plus it.". Of course a degree of feedback can be provided by a teacher, a textbook or a fellow student; but EQUATION has a speed advantage, which makes it less likely that the student will lose the thread.

However, the graph of puzzle times suggests that Rebecca and Nicola needed more practice in equations with negative signs, to consolidate the variety of possible permutations. Moreover, the transfer from EQUATION-based solution to paper-and-pencil should not be taken for granted. The de facto separation of operation decision from operation execution obviates the need for an initial requirement for accurate strategies for operations on expressions; but 100% error-free execution without EQUATION is hardly likely.

The evident satisfaction derived from the program kept classes on task for an hour's lesson; but this apparent enthusiasm for solving the equations should not be assumed to be maintained away from EQUATION, the classroom or a research project.

There is no classwork evidence that students improved their strategies in manipulating expressions and variables, other than learning about the effects of adding and subtracting terms like 7 and $7x$. There is also no evidence that such manipulation became a concern; nor, in fact, that simplifying an equation became a concern *in itself* (as opposed to for finding an unknown number).

The classwork differences between students appear small. No groups of students - other than those who missed lessons - stood out as being especially fast, slow, inspired or bored. There was a tendency for speed to increase with age, but this should not be interpreted as indicating "cognitive maturation". Rather, as indicated especially by the research at School B, and also by comparing Year 7 with Year 10, the number of strategies needing development tended to be higher for younger children. This would be an obvious consequence of differential experience with certain types of problem situation.

7.2.4 Modelling Word Problems

The scripts show an increase in effective usage of the Model button. Early attempts to represent the situation either fail to show awareness of the relevance of unknowns, or bypass the Model button in favour of numerical trial-and-improvement or operational trial-and-error. This difficulty in formulating equations in order to represent particular situations is in spite of students' success in posing their own linear equations in order to get to a high level.

The TOAN word problems made another break with the strategy of thinking in terms of combinations of objects rather than operations on unknown numbers; yet it is equations rather than inversing that students appeared to use. Also, the equals sign was used repeatedly as an equivalence relation, rather than purely to indicate a result. However this does not mean that students were necessarily first seeking a relationship in the situation which they could then represent algebraically - they could also have been seeking *operations* in the situation which could be represented algebraically and legitimately linked by an equals sign.

An equation did not represent an endpoint for students - success was only indicated onscreen when the correct numerical answer was obtained; but comments such as "Ah, now I can do it!" when the equation they entered was automatically simplified suggest that algebra had become for them a problem-solving tool. There is no evidence that algebra was seen as a generalisation of arithmetic; nor that students could express generality using algebra. No indication was given as to whether students could represent situations using equations away from EQUATION; nor that they would want to. On the other hand, there were no comments to the effect of "Why are we doing this?" which might be expected were handling the notation seen as ritualistic.

7.3 Discussion: Pre-Post Improvements

The Year 10 class had already been taught much algebra - including simplifying expressions, solving linear, quadratic and simultaneous equations, and functions. Yet the tests showed that many of the students struggled with algebra, and the questionnaires suggested that many saw algebra as a pointless ritual. Most of the Year 7 class, on the other hand, had been taught no literal algebra.

This section focuses on the problems identified in the last chapter for which there is clear *post hoc* evidence of improvement attributable to EQUATION, to try to identify the strategies that may be responsible for improvement.

Firstly, of course, equation-solving generally improved; but Year 10 students made uneven progress with various types of linear equation; and younger students made little progress at all. Robitaille (1989) reported that only a quarter of UK 14-15 year-olds students could solve $5x + 4 = 4x - 31$. 10EQ moved from 18% to 45%.

Similarly, Year 10 tended to adopt algebraic strategies for some word problems (but not all); while younger students did not. Lins (1992) suggested that "for the Brazilian 7th graders the 'default' approach is non-algebraic, and for the 8th graders it is an algebraic one, namely the use of equations." (p. 199); while there were only 3 successful attempts at using an algebraic strategy by the English students in his study. This led him to suggest that "the development of *algebraic thinking* is a process much more akin to *cultural processes* than to age-related stages of intellectual development." (p. 228-9). It is interesting here, then, that very few students from any group appeared to use an algebraic strategy on the pre-test; yet, for 10EQ on the post-test, equation usage was promoted over trial-and-improvement, whole-parts reasoning, and guessed direct calculations. There was also evidence of improvement in the posing of word problems.

Turning to representation and transformation, there are clear improvements for all ages. There are, however, few indications what particular strategies might be improving on many items, particularly as it is difficult to find plausible explanations for some Year 7 responses. For example, Rajiv wrote for D4(i) that $7a - 3a$ simplifies to $a^2 = 7 - 3 = 5a^2$; Darren wrote for D1(ii) that 4 added to $3n$ becomes $16n$; and Scott wrote for D8 that $3 + n$ is larger than $3n$ because "if n is nose it would be nose + 3" (at least this perhaps fits with his simplification of $7a + 2b$ to "7 apples + 2 bananas").

Note also that comparison with the original CSMS results is not necessarily helpful, because the curriculum has changed a great deal in the last twenty years, and because there were demonstration items given in the original CSMS items that were not given in the test used here. For example: it was explained that n multiplied by 4 could be written as $4n$, and that $a + 3a$ could be simplified to $4a$. Examples of perimeters were also given.

But we can conjecture firstly that younger students began to appreciate that answers need not be numerical: Year 7 answers, for example, start to involve expressions and equations (there are, incidentally, few indications of the oft-noted phenomenon of students assuming that all algebraic expressions should be equations).

Secondly, the idea that the plus sign joining two letters is optional starts to be questioned (particularly for items C3, C8, D1(ii), D7 and D8). For example, Debbie in 10EQ wrote for D8 on the pre-test that $2n$ and $n + 2$ "are the same because $2n$ is simplifying $n + 2$"; whereas on the post-test she indicated that $3n$ means $3 \times n$. However, Judy, May and Jack - also in 10EQ - gave the answer "JP" for C3 on both tests. Jocelyn even gave the "simplification", in his post-test answer, of "J + P = JP". On the other hand, For "Add 4 onto $3n$" (D1(ii)), 10EQ obtained 45% on the pre-test - much the same as CSMS - yet improved to 68% on the post-test.

Merely repeating the test caused around half the Year 10 students to reconsider their answer to the student-professor problem C6. So although an overall facility of 30% on both tests appears to corroborate Philipp's (1992) study of high school students - compared with 40% for Kaput and Clement's (1979) older social science students and 60% for Rosnick's (1981) engineering students - this figure hides substantial changes between the tests. It also hides the fact that a quarter of 7EQ improved, without any declines, compared with no change for 7CON.

It is difficult to conjecture what an improved strategy might be for the equation-truth problem D5 ("When is L + M + N = L + P + N true?"), because the Year 10 improvement mostly comes from 6 students who left the item blank the first time. However, the improvement could indicate a new concern: "When is the equation true?". So although the pre-test facilities (about 15% for Year 7; 35% for Year 10) appear to roughly correspond to Hart's (1981) finding of 11% for Year 8 and 27% for Year 10, the fact that 10EQ improved by over a third on the post-test may raise doubts about the role of some sort of cognitive maturation. On the other hand, 7EQ showed little change.

The item with blue and red pencils (C5) saw 10EQ improving from 9% to 27%. This compares with CSMS Year 10 obtaining 13%. This result directly contradicts the claim of Herscovics (1989) that working with equations in one unknown does not assist in constructing equations in two variables (p. 63). The CSMS team suggested that answers such as b + r = 90, b + r, 6b + 10r = 90 and so on indicate treating letters as objects. No students in the EQUATION groups made any additional errors on this item in the post-test. On the other hand, no Year 7 students got it right. The CSMS team also suggested that certain errors in D2 and the C1(ii) are likely to be caused by students ignoring letters. There is no evidence of these errors for the EQUATION groups appearing in the post-test. However, both Year 10 classes improved on C1(ii) and there is little evidence of improvement in D2.

The CSMS analysis suggested that errors in D10, C2(iii) and D11 could result from evaluating letters. There is no evidence that EQUATION groups made notably greater errors on the post-test for any of these items.

CSMS items with a danger of "premature closure" include C2(i), C2(ii), D4(ii), D4(iv), D5, D8 and D12. There is evidence of improvement in D4(iv) and D5, but not the others. There is little evidence of decline.

Finally, with regards to meta-algebraic aspects, EQUATION students placed greater emphasis on the structural and semantic aspects, and showed greater awareness of the role of an equation as a statement.

In discussing all these improvements however, it ought to be pointed out that various provisos were pointed out in the previous chapter; that sustainability of improvements is not claimed; and, moreover, that the conjecture that such improvements are more widely replicable has to be tested further. It should also be noted that there is no evidence of improvement in appreciating algebraic variation - there was little change in items D8 ("Which is larger, $3n$ or $n + 3$?") and D12 ("What can you say about c if $c + d = 8$ and c is less that d?").

7.4 Can the classwork and pre-post testing be reconciled?

7.4.1 The conundrums

Equation-solving: If the analysis is sound, then it can be claimed that the classwork results show students developing for themselves strategies for simplifying and solving balance puzzles; the transfer of these strategies to balance-like equations; the extension of these strategies to equations that are unlike balances; and the development of the theory that unhelpful operations can be undone. There was no paper-and-pencil work during the classwork, yet the pre-post testing suggests progress in simplifying and solving linear equations using paper-and-pencil. The main differences between students in the classwork seem to be linked to the number of pre-existing informal strategies, yet the pre-post testing shows that success varies from student to student and equation to equation. The use of the balance permits the possibility of metaphor confusions, yet little pre-post evidence has been found that the metaphor confusions are crucial obstacles to learning.

Modelling: The classwork results would appear to show the development of strategies for representing situations using equations. There was no paper-and-pencil work during the classwork, yet the pre-post testing suggests progress with older students in using equations as a technique for solving word problems. The main classwork differences between students are linked to prior experience of representing operations on numbers algebraically; but the pre-post testing shows that success is dependent on the student and on the type of problem. There is also some progress in posing word problems.

Representation and transformation: The classwork shows no evidence of development of skills in manipulating variables, or in representing situations involving more than one letter. Yet the pre-post testing suggests some progress in transformation and representation items that require dealing with variables, in particular the student-professor problem and the equation-truth problem. There is also some progress with younger students in appreciating that expressions and equations may be required as answers to representation and transformation items, and that the plus sign joining two letters cannot be omitted. On some levels of the program, letters can be treated as objects or as representing concrete objects, as opposed to representing numbers. Yet the dangers of interpreting letters as objects only seem to have been exhibited by a very small number of students.

Meta-algebraic aspects: The classwork shows no evidence of reflection on the nature of equations. Yet the pre-post testing suggests some progress in considering the structural, semantic and purposive aspects of equations.

How then can the classwork and pre-post testing be reconciled?

7.4.2 Note: Distinguishing strategic theories and concerns

One important point should be made before addressing these conundrums. Concerns are intimately associated with strategic theories; this is a natural consequence of the recursive nature of understanding. For example, a *concern* to use algebra is, from observing classroom behaviour in *this* study, very often indistinguishable in practice from a *strategy* of using algebra. So although neither the theoretical work nor the empirical analysis are possible without the distinction between strategic theories and concerns, no great play has been made here of the distinction between them *with respect to classwork outcomes*. Of course further studies could be set up to dissect the recursion between strategic theories and concerns more precisely; but such an aim would have to be central to the research, because it would entail rather different experimental arrangements. In particular, there would have to be greater opportunities for students to select the problems they wanted to solve, and to justify their choice. This would consequently lessen the extent to which it would be possible to explore the limitations of the balance model - an aim that is central to *this* research.

7.4.3 Equation-Solving

From the Popperian psychological perspective, the rationale for EQUATION already anticipated improvements in simplification and solution. However, some elaboration may be helpful.

The students' simplification and solution strategies constitute theories. When these are developed in response to the concerns raised in EQUATION, they coexist with a varying number of informal strategies of varying scope - hence the variations in time before the target strategic theories are developed, which are especially significant when comparing Year 7 and Year 10. However, other classwork differences between students are small because the target strategic theories are sharply circumscribed and immediately effective. When the students tackle the post-test equations, many recognise that the strategic theories they developed are applicable. Some students however failed to develop theories for the requested simplifications that EQUATION executed - or they cannot recall the theories. Time spent on Level 12 (negative signs) would be particularly relevant here, as this is a level on which such theories are most important. Of the 10EQ students for whom the appropriate data is available, none of the four who spent most time on Level 12 and none of the four who spent least time on Level 12 got D3(ii) correct on the post-test. All but one of the remaining nine students got the item correct. The lack of post-test success for 7EQ in equation-solving could be attributed partly to forgetting theories for simplification and solution (which students without prior experience of algebra may not have attributed importance to remembering); partly perhaps to the reversal on the test of

EQUATION's fixed E + K*b* order (which students without prior experience of algebra may not have appreciated was equivalent to K*b* + E) and partly to forgetting or never having developed theories for the requested simplifications that EQUATION executed (again which students without prior experience of algebra may not have attributed importance to remembering; but also which were developed in response to Level 12 - a level on which many 7EQ students turned out not to have spent much time).

It should be noted, nevertheless, that the range of equations that students end up solving as a result of using EQUATION is limited. Limitations include: the restriction on term order, the lack of brackets, the single constant term on each side, the single unknown term on each side, and the lack of multiplication or division within the equation. In principle, though, there is nothing to prevent expansion of the program to encompass these features; and then direct comparison with, for example, Payne & Squibb (1990) would then be possible.

There is no evidence of any "didactic cut" here (Filloy & Rojano, 1989; Herscovics & Linchevski, 1991 & 1994). This raises questions about the claim of Sfard & Linchevski (1994) that the cut is the inevitable consequence of a demarcation between operational and structural conceptions (p. 106). Or does this indicate that the balance model has "structural tendencies" (Boulton-Lewis *et al.*, 1997)?

Of course this empirical improvement in equation-solving does not at all "refute" alternative psychological perspectives. It is also explicable from the point of view that characterises students' understanding as sense-making rather than strategic theories and concerns; or as imagining; or as re-enactment. It is claimed, however, that the characterisation of "understanding as problem-solving" enabled a coherent argument to be made for EQUATION in the first place, using concerns as a learning force; and that it is unlikely that the program would have been developed in quite this way if one had had meanings, images, empathy, decontextualised cognitive structures or modes of thought as goals rather than strategic theories.

7.4.4 Modelling

Just as for equation-solving, from the Popperian psychological perspective the rationale for EQUATION already anticipated improvements in formulation. But some of the links between classwork and pre-post results can be elaborated.

The transfer from balance puzzles to balance-like equations meant that some students developed the initial theory that equations were abbreviations for balance puzzles. So, for example, Year 6 students asked in interview what $4 + 2b = 3 + 5b$ might mean started talking in terms of weights and barrels rather than numbers and unknowns. Of course the older students would also be aware that such symbols also have a more abstract interpretation, but such a contextualised interpretation would not be in contradiction. Thus many students were being introduced (albeit fairly surreptitiously) to a standardised way of representing a pictorial balance puzzle. So when on Level 13 the idea arose to represent a verbal balance puzzle, a "natural" way to do this was to adopt the standardised symbolism. Indeed the Model button did not allow much variation from this. Therefore students had moved from *seeing* how balance puzzles could be represented to *choosing* to represent them in this way. We thus have the beginnings of an algebraic modelling strategy, and the potential for the posing of problems for which an algebraic strategy is appropriate, even though only one lesson was spent working on such problems. Of course transferring this strategy to paper-and-pencil relies not only on confidence in recalling conventional syntax and confidence in executing solution strategies, but also on recognition of problems for which an algebraic strategy is appropriate. The post-test items B2 and B3 are very similar in wording to problems found in EQUATION; whereas A1(iii) and B1 are not; and this may be why algebra was not attempted in these latter items. This supports the results of Berger & Wilde (1987). For the younger students, algebra may not have been used in B2 or B3 because of the issues of confidence, given that they had more simplification and solution strategies to develop than the older students. For Rebecca in 10EQ, however, it would appear that EQUATION helped her firstly to debug her equation solving method; and secondly to conclude that using an equation would be the best way to solve the problem.

Note also that the evidence from 10CON that a second attempt on certain problems can lead to a significant improvement vindicates the questioning of the attribution of algebraic knowledge to the influence of a "mode of thinking" that is in some sense "underlying" knowledge and difficulties. For example, Lins found a facility for B2 of 22%, which is similar to 23% for 10EQ and 29% for 10CON. Yet on the post-test, 10EQ rose to 59% and 10CON rose to 42%. There are also dramatic rises for both classes for B3, but a direct comparison with Lins' study is difficult because he gave the rule governing the secret number in syncopated form. That errors centred on the negative sign supports the teacher's suspicion that not enough time had been spent consolidating this aspect of the strategy, and emphasises the significance of EQUATION executing the operations.

7.4.5 Representation and Transformation

The thesis is that promoting simplification and utilisation can provide a purpose for algebraic symbols. If this were correct, then it would be expected that there might be improvements in transformation and representation problems. These results corroborate the inference.

EQUATION does not instruct students in tackling representation or transformation problems - it does not give examples, explain things, probe their understanding, assign them activities appropriate to their current knowledge, or provide any of the other learning experiences that one might expect from being instructed. It pays no attention to identities, multiple letters or variables, and seems to positively encourage the "*b* for barrel" misconception.

Yet the pre-post testing suggests that 10EQ made widespread improvements across a range of items; that 7EQ started to achieve Collis' so-called "acceptance of lack of closure"; that the "letter as object" interpretation played little role; and that there were significant improvements in the student-professor and equation-truth problems - considered by some to border on the intractable. So what would appear to be happening is, as conjectured a priori, that EQUATION's easing of the algebraic strategy of simplification of a situation - combined with practice in tackling word problems - has led to prima facie evidence of improved theories for representation and transformation. In other words, the results are consistent with transference from utilisation to representation and transformation.

A greater tendency to treat letters as objects would lead to improvement on the items C2(i), C2(ii), C5 and D4(i); and there is little evidence of this. Therefore, it seems reasonable to suggest that at least some of the significant improvements made by students on representation items result from the grasping of the theory that treating letters as numbers can often help solve a problem.

Of course it is not an entirely black-and-white issue - there are items where there is conflicting evidence of manipulation of expressions, and confusion over letters as objects. There is little evidence of progress in the Patterns items. Not all students progressed equally. Moreover, EQUATION could not possibly ever be expected to provide *all* the manipulation skills with which to handle variables.

Yet there are a number of viewpoints discussed in chapter 3 from which there would be no reason to expect *any* improvement in these items: among them, the view from the neo-Piagetian perspective, the view that it is only practice in transformation that can improve transformation skills; the view that a fundamental lack of understanding of the notion of variable lies at the heart of manipulation difficulties; and the view that "letter as object" constitutes an underlying cognitive conception. On the other hand, the idea that some sort of algebraic "mode of thought" has been developed that can now be applied relatively unproblematically in other algebraic items is a viable explanation of these improvements, albeit one that this research rejects for reasons given in chapter 2. From the Popperian psychological perspective, the context-specific nature of knowledge overrides such general thinking, and so this transfer has to be explained in other ways. The arguments in Chapters 4 and 5 outline how the strategic theories developed when using the program are robust enough for use in other contexts.

But if transfer needs to be explained, so does lack of transfer. Perhaps the fewer improvements for Year 7 shows that those with already greater knowledge have a better chance of using EQUATION to improve their strategic theories. There are also a larger number of more subtle theories that perhaps younger students are having to develop in the time: for example that $4 + 3n$ cannot be simplified unless one knows what n stands for; or that there can be two letters in an equation; or that the same letter can represent different numbers in different contexts; or that different letters can represent the same number; or that every time a letter appears in an equation it stands for the same number.

If Stacey & MacGregor (1993) are right that word-order matching is not a complete explanation for the reversal error (and such a strategy would, in any case, obtain the correct answer in C8), none of Clement's (1982) strategies seem to be able to explain the discrepancy between item A2 (facility around 90% for Year 10) and items C6-C8 (facility less than 45%). "Operative pattern" and "substitution followed by operative" should get the right answer in C6-C8 - just as they would in A2; and "static comparison" - although not perhaps applicable in C7 - would gain as many wrong answers in A2 as C6. But an analysis of students' mathematical knowledge in terms of theories and concerns can provide an explanation.

A2 is designed in such a way as to encourage the student to seek an operation. The very giving of a table prompts pattern-seekers to find operations obtaining one number from another *before* this is expressed as a symbolic relation; A2(i) and A2(ii) reinforce this prompt; and then A2(iii) puts this prompt explicitly: "Describe in words *how you would find* Q if you were told what L is" (emphasis added). When it comes to writing an algebraic rule, what is being represented (I would suggest) is an *operation*, not a relation. Therefore, the problem expectation (i.e. concern) is to tell someone else how to calculate one quantity from another. But in C6-C8 the wording of the questions is such that students conjecture that they are to represent the given relation. However, without the prompts to find an operation, a concern to ensure that the representation conforms to arithmetic conventions (as opposed to ratio conventions) does not arise. Hence we have at best a reversal, and at worst an idiosyncratic representation. Here is a hypothesis to act as a test of this explanation: change the wording of C6 from "Write an equation to represent the statement: 'At this school there are 15 times as many students as teachers.'. Use S for the number of students and T for the number of teachers." to "At this school there are 15 times as many students as teachers. Use algebra to describe how to find the number of students from the number of teachers. Use S for the number of students and T for the number of teachers." and I predict the facility should shoot up.

If this is correct, it could be that a way of curing the reversal error is to encourage the seeking of an operation (rather than a relation). In EQUATION this was largely done by Level 15 problems such as "At a cinema there are 7.1 times as many cheap seats as expensive seats. Altogether there are 810 seats. How many expensive seats are there?"; and such problems might be expected to lead to the use of something like "7.1e" to represent the number of cheap seats. However, many students in 10EQ struggled with level 15 (often because the quantity asked for was not always the quantity represented by a letter), and few progressed beyond these problems without assistance. It would be interesting to see if longer time spent with these multiplicative-relation problems could assist further with the student-professor problem. Similarly, although there was no sign of progress at all in relation to the penny-dimes problem, and the use of letters in EQUATION was restricted to one letter (not two) and unknowns (not variables), I believe it would be still worth testing whether prolonged interaction with Level 15 problems made a difference to this item.

7.4.6 Meta-algebraic aspects

If students now have a purpose for algebraic symbols, it is inevitable that this purpose will be reflected in their views of the nature of equations (as opposed to their views of the nature of equations "underpinning" their understanding). Having "played a game" with equations, they then used them to represent situations. In everyday language, *statements* tell us something about situations. That is not to say, of course, that an equation is not also a group of symbols, a balance (an equality and capable of representing a balance), a method (because it constitutes two sets of operations, because it can be used as a way of solving word problems, and because it is subject to solution techniques), a formula and a problem.

Initial reasons for finding equations difficult almost inevitably (for Year 10) focused on the tactical operations that one has to learn. Yet those using EQUATION quickly realised that also of great importance is the relationship between the symbolism and whatever is being represented - and hence they have concerns with more structural aspects. In describing an equation to others, then, the idea of equality has to be emphasised.

For the reasons given in the last chapter, not too much should be read into the anecdotal examples of more positive attitudes towards equations. But I would certainly want to suggest that in the long-run, valuing equations as tools rather than obstacles would lead to greater confidence about algebra.

7.5 How are claims about the balance model affected?

We have seen earlier that the balance model had advantages, as a non-symbolic example of equality; as a concrete experience for promoting simplification; and as a solid metaphor for an equation. Yet there were severe limitations, some of which the empirical analysis suggests EQUATION may go some way to overcome:

Physical limitations: By leaving the balance model behind, Levels 11 and 12 were able to introduce negative signs and negative solutions. These breaks with the model were noted by many students, yet having done so they then seemed happy to accept the situation. There was no evidence of the "major cognitive difficulties" that Linchevski & Herscovics (1996) found could result from physical limitations. Decimal coefficients were introduced on the modelling levels, but could easily have been introduced earlier to address Lins' recognition that different equations within the same algebraic form can be treated in very different ways. Little was also done here to address the need for division and multiplication, as demanded by Herscovics and Kieran (1980). Even so, I suggest that these deficiencies could be corrected to demonstrate that the balance model *can* naturally lead onto decimal coefficients, multiplication and division. In any case, it would be difficult now to maintain that the use of a balance model actively conspires against algebraic strategies.

Ineffectiveness in promoting an algebraic strategy: The analysis has suggested this is no longer a valid objection. It is true that the research corroborates the finding of Schliemann *et al.* (1992) that children rarely use a spontaneous cancellation strategy. Yet although there was minimal teacher intervention, no issuing of Leibniz, transposition or expression rules, and no encouragement for substitution or flowchart methods, students were encouraged to develop a cancellation strategy by means of the Take off… buttons, and a careful graduating of puzzles such that informal strategies progressively break down. This cancellation strategy led naturally onto formal operations for solving simple linear equations. This result overturns the findings of Dickson (1989) with respect to the balance model. Moreover, the examples found by Filloy & Rojano (1989) of the semantics of the balance model delaying construction of an algebraic syntax would not appear to be applicable in EQUATION, because use of a conventional algebraic syntax for modelling was encouraged. Similar considerations apply to Booth's (1987) warnings about students failing to appreciate the connections between concrete or ideographic approaches and formal procedures.

Metaphor confusions: As noted earlier, little evidence has been found that the metaphor confusions are crucial obstacles to learning, although students' rationales were sometimes incoherent.

Misleading letter interpretations: Similarly, the judgement of Dickson (1989) and Booker (1987) that letters tend to label objects rather than unknown numbers has been seen as less dire than might be expected.

Unknowns not variables: The above analysis also raises doubts about Booker's demand (along with Arcavi, 1994) for generalised arithmetic to come prior to any other encounter with algebraic symbols. Indeed, I would argue that EQUATION itself has "established the need for and power of algebraic symbols" (p. 279), but independently of generalised arithmetic. Of course generality is important, and EQUATION is limited in what it can promote in this respect; but would it still be fair to conclude with Kaput (1987) that "the inherent particularity of [concrete] models... runs entirely opposite to the inherent generality and abstractness of algebraic statements."?

However, there is one further objection that, in the light of this research, could be usefully added:

Ineffectiveness in promoting equality principles: The analysis suggests that the computerised balance model enabled many students to develop theories such as "Taking the same number (known or unknown) off each side of an equation can make it easier to work out an unknown number." and "If you perform the same operation on each side of an equation, the answer is still the same.". Nevertheless, some subtly different theories that were not targeted may not have been grasped. For example: "If from two equal things the same quantity be taken away, the things will remain equal", "In a balance, if the sides weigh the same as each other, there will be no tilting", "In a balance, if there is no tilting, the sides weigh the same as each other." and "In a balance or an equation, what you do to one side, you have to do to the other, otherwise equality will not be maintained.". There is no doubt that these theories appear important principles; but even so, the extent to which they actively affect practical strategies is debatable.

In conclusion, this study suggests that the same objections that might be applied to a physical balance or a balance picture in a textbook do not necessarily apply to the interactive, game-like, computer-based balance model in EQUATION.

7.6 How does EQUATION compare with other activities?

7.6.1 Priority Claims

There are a number of views outlined in Chapters 3 and 4 that would suggest that EQUATION could not improve representation and transformation items because it does not incorporate certain activities that are deemed by some to be essential for learning in algebra. The empirical results would therefore raise questions about the priority of activities such as using letters to represent generalised numbers, expressing mathematical relationships in natural language prior to algebraic symbolism, making explicit the transition from procedural to structural conceptions and emphasising conservation of equation.

Nevertheless, it could be argued perhaps that EQUATION *indirectly* encourages formalisation of method, construction of meaning or imagery, and production of unclosed expressions as legitimate answers. What would follow, though, would be the rejection of the idea that such activities are *essential*, and perhaps that making them more explicit would improve EQUATION's efficacy. Even so, the hypothesis that those who struggle with algorithms *require* explicit consideration of meaning in a variable-centred approach to algebra would appear to be challenged by the apparently greater understanding of the role of algebraic representation demonstrated by these students. The claim that students' algebraic difficulties *centre* on deficient "construction of meaning" for algebraic objects and processes is not perhaps the truism that might be assumed.

However, test conditions, students' informal discussions between the tests, and the excitement of being involved in research might prove to be better explanations for statistical significance than EQUATION itself. Moreover, there are no guarantees that these are long-term improvements. Therefore, given that this empirical work was primarily to illustrate a theoretical analysis of learning processes, rather than to achieve a large, rigidly controlled experiment with random sampling and allocation, further research is required to substantiate such priority claims.

Moreover, although the results are consistent with the conjecture that transformation, representation and interpretation problems are less accessible than word problems, there has been little testing of this conjecture. Even if "the central feature of algebra is that it is an ideal medium through which one can see and express general statements" (Mason *et al.*, 1985, p. 1), we have seen that tackling word problems can lead to concerns for representation, interpretation and transformation. But little has been found out about the converse.

Note that this research has not attempted directly to compare the computerised balance model *empirically* with other activities for promoting operations on equations and representation as an equation. Rather, the empirical work has focused on exploring whether there is prima facie evidence of at least some improvement. The analysis then used this evidence to explore whether that improvement can be related to the classwork, to explore proposed limitations of the balance model, and (below) to investigate connections between the improvements with EQUATION and improvements with other activities reported in the research literature. Given that this prima facie evidence exists, the next step would be to conduct such comparative research.

7.6.2 Usual Curriculum

The Year 7 students' usual introduction to algebra that would normally consist (in Year 8) largely of practice in transformation problems (including substitution), followed by a small selection of representation and word problems. Flowcharts and inverse operations would be used to introduce operations on equations. Informal strategies like the cover-up method and trial-and-improvement substitution may also be taught.

The expert opinion of the Year 7 teacher was that EQUATION was a useful addition to the range of algebraic experiences her classes could have. Enjoyment of algebra, having a purpose for the symbolism, and confidence with operating on equations were all characteristics that she identified from the three lessons that would usefully supplement the curriculum. The higher modelling levels she considered to be an ideal way of practising formulation of equations from the beginning of Year 9, once the initial practice in transformation items is complete.

The focus for the Year 10 students returned, prior to using the program, to transformation items - this time involving multiple variables, quadratic terms, brackets and inequality symbols, and representation items involving graphs. Memorising formal operational, transposition and expression rules and the conventions of algebra would play a large role in this.

The Year 10 teacher noted that this memorisation was easier for some students than others; and that EQUATION not only made the Leibniz method more accessible for students struggling with memorisation, but it also motivated all students towards seeing algebra as something more than school exercises. He considered the modelling levels to be an excellent way of extending to algebra his pedagogic strategy of asking students to pose their own mathematical problems. Although this strategy was possible without the program, many students were usually so unconfident about their equation solution strategies, that they were loathe to spend any more time than the absolute minimum considering the form of the equation that might model a situation. The students were attracted by the idea that "good problems" might be incorporated in future versions of the program.

7.6.3 Other initiatives

EQUATION tends to compare rather unfavourably with other initiatives if all the items in the pre-post test are considered. This seems a slightly unfair comparison, given that the test was intended to cover a range of algebraic items, even ones for which there was no expectation of improvement. There is also no doubt that virtually all other studies allowed students considerably more than the two hours that most students had for EQUATION. Hunter *et al.* (1995), for example, involved three weeks' worth of lessons, and Booth (1984) about two weeks' worth.

Moreover, given the evident learning displayed (for example) by Year 7 that failed to be recognised in the relatively decontextualised environment of a paper-and-pencil test, there must be doubt about the extent to which pre-post testing can be used alone to compare studies. Some of the interview episodes reported in Sutherland (1995), for example, are remarkably similar to those in this research, in that students who started out with little idea what the question "If John has J marbles and Peter has P marbles, what could you write for the number of marbles they have altogether?" might be asking (unless J is 10 and P is 16 because A = 1, B = 2, etc.) ended up being able to talk about J + P being "any number".

Unfortunately none of the substitution studies carried out by Thomas and Tall into how substitution and programming can improve performance on CSMS items lists both the pre and post facilities of individual items for either control or experimental groups; the experimental and control groups for the substitution studies were matched, whereas they were not in the EQUATION study; the control groups in the substitution studies were usually taught by "traditional methods", whereas the control groups in the EQUATION study did no algebra; the substitution studies took place over 3 weeks, whereas the EQUATION study took place over less than one.

So direct quantitative comparison is not possible. However, if one compares experimental and control post-tests for the L + M + N = L + P + N item in the published figures from Graham & Thomas (1997), Tall & Thomas (1991) and Thomas & Tall (1988), they at best indicate an improvement of 15% for using graphic calculators, which compares with an improvement of over double this for 10EQ. By these same measures, the substitution and EQUATION studies produced similar results for the remaining items in common. On the other hand, 7EQ showed little change in these items.

In comparison with the "arithmetic identities" of Kieran & Herscovics (1980), there is a commonality in the fact that students could give examples of equation with operations on both sides only after the respective initiatives. It can be argued that in both studies, students' knowledge was "transformed gradually so that they can build for themselves the notion of an algebraic equation." (p. 573), albeit in very different ways. However, the students in the EQUATION study were able to solve such equations using formal methods, rather than having to rely on recall or trial-and-error. Also, by tackling the highest level of word problems, 10EQ students were able to make a start on developing strategies for dealing with issues of brackets and order of operations.

7.7 Suggestions for Further Research

EQUATION allowed students to develop strategies and concerns for solving and formulating equations; and it consequently provided a purpose for algebraic symbols that resulted in improved theories for representation, transformation and consideration. Yet the program was not intended as an ideal learning resource, but as a focused research tool to examine the claimed priority of certain activities, to question the pedagogical trend to focus on variables as the expense of unknowns, and to explore the limitations of the balance model. Nevertheless, in the light of the critique of the program in chapter 5, and the conclusions in this chapter, a number of improvements to EQUATION could be made, to better address the algebraic needs of students.

Interspersed Levels: The following new levels, interspersed among the existing levels, would diminish the reliance of strategic theories on particularities of the given balance situation, without fundamental changes to the program:

- variation of order of barrels and weights, unknowns and constants;
- simple decimal weights, constants and coefficients of the unknown;
- letters other than "b".

New Symbolic Levels: The following levels, added to the end of the symbolic levels, would augment existing algebraic strategic theories, without fundamental changes to the program:

- fractional coefficients, requiring the use of multiplication;
- a "lowest terms" constraint on fractions (to emphasise that answers are numerical) requiring the use of division;
- large parameters, so that answers involving fractions would not fit in the space provided, requiring the use of division;
- equations without numbers, requiring rearrangement.

Balance Puzzle Variants: The following are variants of the existing balance model, and could lead onto different strategic theories for equation transformation. They would require both balance levels and symbolic levels, and so could be placed prior to the existing symbolic levels. However, they do not fit easily into the existing structure of the program, in that they may interfere in the process of transition of simplification strategic theories from pictures to symbols, and in the process of transition of representation strategic theories from pictures to word problems. The variants might therefore be best implemented by placing them after the initial transitions.

- a tilting balance, which could lead onto decimal search, quadratics, square roots, indices, logarithms, etc.;
- dragging of objects, which could lead onto transposition;
- grouping of barrels, which could lead onto expansion of brackets, factorising and identities;
- two sorts of barrels, which could lead onto simultaneous equations.

Other Pre-Symbolic Scenarios: The following are alternatives to the balance model, and could not only lead onto different strategic theories for equation representation and transformation, but may in some cases also provide different scenarios through which to construct some of the *same* theories as in the balance scenario. Again, although they could be interspersed among the existing levels, placing these scenarios after the completion of the balance scenario could avoid interference with the existing transition processes.

- Cups & Beans

- Areas
- Think of a Number
- Flowcharts
- Line Graphs
- Arithmetic Identities
- Number Patterns

Other Features:

- a save option;
- audible puzzles (for those with reading or sight difficulties);
- a trace (when more complicated operations are involved);
- easier entry of students' own word problems;
- a gradual diminution of automatic simplification.

Now that there is prima facie evidence for the claim that EQUATION can promote a range of representation and transformation strategic theories, a direct empirical comparison between EQUATION and other initiatives would be desirable. The strengths and weaknesses, similarities and differences of the various initiatives could be explored in detail, qualitatively and quantitatively. The result that certain items can improve merely on repetition, for specific reasons, will assist in the design of future research instruments.

Moreover, although some of the CSMS items used here would suggest that EQUATION helped students in the transformation and representation of variables, it is to be expected that students are still likely to lack a dynamic appreciation of variables. As English & Halford (1995) point out: "It is all too easy to assume that students have a grasp of variables when they solve generalized arithmetic tasks. Frequently, these can be solved without an understanding of variables, that is, without considering the degree to which one set of values varies as a result of changes in another set." (p. 226). Further research is needed into what further aspects of understanding variables need to be promoted, and into how these aspects can be identified.

It is also not possible to tell clearly on the basis of this research if students' knowledge of equivalence of equations is affected in any way by EQUATION. It would be worthwhile repeating the study of Steinberg, Sleeman and Ktorza (1991) to see what happens - in particular for the third of the students using computed solutions as a primary strategy, and the high proportion using faulty reasoning.

Moreover, these results are not necessarily applicable to all students of algebra. Since this research was aiming to illustrate the Popperian approach in detail rather than to test the generalisability of claims about learning, the samples chosen were small and with little regard to the representativeness of students. Now that a number of items have been developed for identifying potential learning gains, a larger number of students could be involved. Further research with a larger sample could also usefully try to determine who gains, and what characterises the prior strategic theories and concerns of those who do not gain.

Another important point is that this research has focused on the individual student's encounter with World 3 with respect to introductory symbolic algebra, and primarily through interaction with a computer and with a fellow student, over a small time-scale. It has not chosen to examine the roles of the teacher, of textbooks, of the classroom, and of the wider social environment in shaping the longer-term encounter. There is nothing in the Popperian perspective that would preclude such an extended investigation.

In addition, there is no suggestion that what has been learned over two or three lessons is necessarily remembered and plays a significant part in future algebraic learning. In particular, this research has not linked all aspects of student's classwork, test and questionnaire responses to prior experiences. Longitudinal data would allow more scope to attempt this, and the analysis of such data might allow a greater mapping of algebraic knowledge using the theoretical tools developed here. Moreover, the research instruments here included a wide range of algebraic items, in order to test the hypothesis that improvement would occur only in certain problems. In-depth interviews in later research would be able to probe in much greater detail just those items where improvement is expected.

Finally, another strand of research should focus on the Popperian psychological perspective that has been developed here. This perspective has been used to combine the results of a number of research studies, to develop arguments for and against features of a computerised balance model, and to elucidate learning mechanisms. If the perspective is to be of any use, it clearly requires further critical scrutiny, both theoretical and empirical.

Glossary

algebraic knowledge	(in World 2) the theories and concerns relating to algebra
algebraic thinking	a hypothetical specific mode of thought involving, for example, "thinking arithmetically, thinking internally and thinking analytically" (Lins, 1992); or "handling the as-yet-unknown, inverting and reversing operations, seeing the general in the particular." (Love, 1986)
balance model of an equation	the use of balance scales (physically, pictorially, or metaphorically) to teach about equations
balance-like equations	equations that could represent a simple balance scale
BVSR	Blind Variation and Selective Retention - a World 2 Darwinian process of creative learning involving the generation and checking of thought-trials under selection pressures
concerns	problems of special interest to an individual; more precisely - any World 2 construction that exerts a selection pressure on the formation of theories; concerns include desires, motivations and fears; a concern incorporates background theories; concern to use a theory may vary
conjectural knowledge	refers to the view that knowledge growth occurs through error-elimination rather than through derivation from foundational knowledge
context	the problem situation
didactic cut	a hypothetical cognitive jump required when unknowns on both sides are introduced into linear equations
error-elimination	the critical scrutiny of theories, using logical arguments or empirical testing; also includes the putting of theories into forms so that they can be more easily scrutinised
fallibilism	the view that all knowledge is open to critical scrutiny, because certainty can never be attained
intersubjectivity	the testing of theories by a research community
intuition	prior theories applied in a new problem situation (as opposed to "immediate" knowledge)
learning	the growth of knowledge; learning occurs through trial-and-improvement of theories in response to concerns, rather than through the development of context-free modes of thought

Leibniz strategic theory	operating on both sides of an equality in order to simplify
meta-algebraic theories	theories about algebraic theories, strategies and concerns; in particular the theoretical consequences of attempts to consider the nature of World 2 or World 3 algebraic objects, properties, relationships and processes, or to rationalise practices
objective knowledge	refers to the view that some knowledge can have a degree of autonomy from the people who generated it; it is "real" in the sense of being able to conflict with other knowledge
objectivity	the scientific ideal that knowledge is open to criticism by anybody
Popperian epistemology	a theory of fallibilist, conjectural, objective knowledge, dependent on intersubjective error-elimination
Popperian psychology	a conjectural, fallibilist theory of subjective knowledge, dependent on the BVSR of theories in response to concerns, in which "understanding" is viewed as problem-solving, and in which recontextualised meta-algebraic theories are preferred to slowly maturing "underlying" algebraic cognitive structures
recontextualisation	the process of separating a theory from its originating concerns, often so as to examine the properties and relationships of a theory
rigour	the warranting of knowledge by systematic error-elimination
strategy	a plan for solving a problem
theory	a construction of reality - can be objective or subjective; some theories are strategic
transfer	the adaptation of theories generated in one problem situation to another
understanding	problem-solving (rather than sense-making, imagining or re-enactment); understanding a theory is understanding the problem it solves, and why other theories do not solve the problem; understanding a problem is understanding something of its background theories and why it might be a problem
World 1	the physical world
World 2	the subjective, mental world of conscious experiences
World 3	the theoretical world of objective contents of thought, especially of scientific and poetic thoughts and of works of art

Bibliography

Abidin, B. (1997) *Computer based representations for aiding problem solving in mathematics*, PhD Thesis, University of Leeds

Ainley, J. (1995a) "Re-viewing Graphing: Traditional and Intuitive Approaches", *For the Learning of Mathematics*, Vol. 15, No. 2, pp. 10-16

Ainley, J. (1995b) "Reasons to be formal: contextualising formal notation in a spreadsheet environment", in Meira, L. & Carraher, D. (Eds.), *Proceedings of PME 19*, Vol. 2, pp. 26-33, Recife, Brazil

Ainley, J. (1996) "Purposeful Contexts for Formal Notation in a Spreadsheet Environment", *Journal of Mathematical Behaviour*, Vol. 15

Anderson, J. R., Boyle, C. F., Corbett, A. T. & Lewis, M. W. (1990) "Cognitive modelling and intelligent tutoring", *Artificial Intelligence*, 42, p. 7-49

Arcavi, A. (1994) "Symbol Sense: Informal Sense-making in Formal Mathematics", *For the Learning of Mathematics*, Vol. 14, No. 3, pp. 24-35

Arzarello, F. (1991) "Procedural and relational aspects of algebraic thinking", in Furinghetti, F. (Ed.), *Proceedings of PME 15*, Vol. 1, pp. 80-87, Assisi, Italy

Austin, J. D. & Vollrath, H.-J. (1989) "Representing, Solving and Using Algebraic Equations", *Mathematics Teacher*, Vol. 82, pp. 608-612

Aziz, N. A. (1996) *The Use of Belief Revision to Model Contradictions and Entrenched Misconceptions*, PhD Thesis, University of Edinburgh

Ball, S. J. (1990) "Self-doubt and soft data: social and technical trajectories in ethnographic fieldwork", *International Journal of Qualitative Studies in Education*, Vol. 3, No. 2, pp. 157-71

Becker, G. (1988) "A classification of students' errors in secondary level algebra", *Proceedings of PME 12*, Veszprém, Hungary

Bednarz, N., Kieran, C. & Lee, L. (eds.) (1996) *Approaches to Algebra: Perspectives for Research and Teaching*, Dordrecht, Netherlands, Kluwer

Bednarz N., Radford L., Janvier B. & Leparge A. (1992) "Arithmetical and Algebraic Thinking in Problem-Solving", *Proceedings of PME 16*, Vol. 1, pp. 65-72, Durham, USA, University of New Hampshire

Behr, M., Erlwanger, S. & Nichols, E. (1980) "How Children View the Equals Sign", *Mathematics Teaching*, 92, 13-15

Bell, A. (1988) "Algebra - Choices in Curriculum Design", in Borbas, A. (Ed.), *Proceedings of PME 12*, Vol. 1, pp. 147-153, Veszprém, Hungary

Bell, A. W., Costello, J. & Küchemann, D. E. (1983) *A Review of Research in Mathematical Education. Part A: Learning and Teaching*. Windsor: NFER

Bell, A., Malone, J. & Taylor, P. (1988) "Algebra: An exploratory teaching experiment", Nottingham, UK, Shell Centre for Mathematical Education

Berger, D. E. & Wilde, J. M. (1987) "A Task Analysis of Algebra Word Problems", in Berger, D. E., Pezdek, K. & Banks, W. P. (Eds.), *Applications of Cognitive Psychology*, pp. 123-137, Hillsdale, NJ, Lawrence Erlbaum

Bernard, J. E. & Cohen, M. P. (1988) "An integration of equation-solving methods into a developmental sequence", in Coxford, A. F. (Ed.), *The ideas of algebra, K-12 (1988 Yearbook)*, pp. 97-111, Reston, VA, US, NCTM

Berry, J., Graham, E., & Watkins, A. (1993) *Learning Mathematics through DERIVE*, Chichester, UK, Ellis Horwood

Berry, J., Graham, E., & Watkins, A. (1994) "Integrating the DERIVE program into the teaching of mathematics", in Berry, John (Ed.), *International DERIVE Journal*, Vol. 1, No. 1, pp. 83-96, Hemel Hempstead, UK, Ellis Horwood

Bibby, N. (1991) "Wherefore 'plug and chug'?", *Mathematical Gazette*, No. 75, p. 40-8

Birenbaum, M., Kelly, A. E., & Tatsuoka, K. K. (1993) "Diagnosing Knowledge States in Algebra Using the Rule-Space Model", *Journal for Research in Mathematics Education*, Vol. 24, No. 5, pp. 442-459

Blume, G. W. & Schoen, H. L. (1988) "Mathematical problem-solving performances of eighth-grade programmers and non-programmers", *Journal for Research in Mathematics Education*, Vol. 19, pp. 142-156

Bobrow, D. G. (1968) "Natural-language input for a computer problem-solving system", in Minsky, M. (Ed.), *Semantic information processing*, pp. 135-215, Cambridge, MA, MIT Press

Bolter, J. D. (1984) *Turing's Man*, North Carolina University Press

Booker, G. (1987) "Conceptual Obstacles to the Development of Algebraic Thinking", in Bergeron J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 275-281, Montreal, Canada

Booth L. (1984) *Algebra: Children's Strategies and Errors*, Windsor: NFER-Nelson

Booth, L. (1987) "Equations Revisited", in Bergeron J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 282-288, Montreal

Booth, L. (1989a) "A Question of Structure", in Wagner, S. & Kieran, C. (Eds.), *Research Issues in the Learning and Teaching of Algebra*, pp. 57-59, Reston, VA, NCTM

Booth, L. (1989b) "Grade 8 students' understanding of structural properties in mathematics", in Vergnaud, G., Rogalski, J. & Artigue, M. (Eds.), *Proceedings of PME 13*, pp. 141-148, Paris

Boulton-Lewis, G. M., Cooper, T. J., Atweh, B., Pillay, H., Wilss, L. & Mutch, S. (1997) "The Transition from Arithmetic to Algebra: A Cognitive Perspective", in Pehkonen, E. (Ed.), *Proceedings of PME 21*, Vol. 2, pp. 185-192, Lahti, Finland

Brown, J. S. & Burton, R. B. (1978) "Diagnostic models for procedural bugs in basic mathematical skills", *Cognitive Science*, Vol. 2, pp. 155-192

Brown, J. S. (1985) "Process versus product: A perspective on tools for communal and informal electronic learning", *Journal of Educational Computing Research*, Volume 1, p. 179-201

Brownell, W. A. (1935) "Psychological considerations in the learning and teaching of arithmetic", in Reeve, W. D. (Ed.), *The teaching of arithmetic: The 10th Yearbook of the NCTM*, pp. 1-31, New York, US, Teachers College, Columbia University

Brownell, W. A. (1945) "When is arithmetic meaningful?", *Journal of Educational Research*, Vol. 38, No. 3, pp. 481-498

Bruce, M., Chanter, B., Cudmore, D., MacDonald, Z. (1993) "Introductory Algebra Resource Guide"

Bruner, J. (1960) The Process of Education, Cambridge, MA, US, Harvard University Press

Burton, M. B. (1988) "A Linguistic Basis for Student Difficulties with Algebra", *For the Learning of Mathematics*, Vol. 8, No. 1, pp. 2-7

Campbell, D. T. & Stanley, J. C. (1963) "Experimental and Quasi-Experimental Designs for Research on Teaching", in Gage, N. L. (Ed.), *Handbook of Research on Teaching*, pp. 171-246, Chicago, Illinois, US, Rand McNally

Campbell, D. T. (1960) "Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes", *Psychological Review*, Vol. 67, No. 6, pp. 380-400

Carr, W. & Kemmis, S. (1986) *Becoming Critical: Education, Knowledge and Action Research*, London, Falmer

Carraher, T. N. & Schliemann, A. D. (1987) "Manipulating equivalences in the market and in maths", in Bergeron J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 289-294, Montreal, Canada

Carry, L. R., Lewis, C. & Bernard, J. (1980) "Psychology of equation solving: An information processing study", Austin: University of Texas at Austin, Department of Curriculum and Instruction

Chaiklin, S. (1989) "Cognitive studies of algebra problem solving and learning", in Wagner, S. & Kieran, C. (Eds.), *Research Issues in the Learning and Teaching of Algebra*, pp. 93-114, Reston, VA, NCTM

Chalouh, L. & Herscovics, N. (1988) "Teaching algebraic expressions in a meaningful way", in Coxford, A. F. (Ed.), *The Ideas of Algebra (NCTM Yearbook)*, p. 33-42, Reston, VA, US, NCTM

Chazan, D. (1993) "F(X) = G(X)?: an Approach to Modelling with Algebra", *For the Learning of Mathematics*, Vol. 13, No. 3, pp. 22-26

Chevellard, Y. (1984) "Le Passage de l'arithmetique à algebre dans l'enseignment des mathematiques au college", Petit X, 5, IREM de Grenbole, pp. 51-95

Clement, J. & Kaput, J. (1979) "The Roots of a Common Reversal Error", *Journal of Children's Mathematical Behaviour*, Vol. 2, p. 208

Clement, J. (1982) "Algebra word problem solutions: Thought processes underlying a common misconception", *Journal for Research in Mathematics Education*, Vol. 13, No. 1, p. 16-30

Cobb, P., Yackel, E. & Wood, T. (1992) "A constructivist alternative to the representational view of mind in mathematics education", *Journal for Research in Mathematics Education*, Vol. 23, No. 1, pp. 2-33

Cockcroft, W. H. (1982) *Mathematics Counts: Report on the Committee of Inquiry in the Teaching of Mathematics*, London, HMSO

Collis, K. F. (1975) *A study of concrete and formal operations in school mathematics*, Hawthorn, Victoria, Australia, Australian Council for Educational Research

Cooper, M. (1984) "The mathematical 'reversal error' and attempts to correct it", in Southwell, B., Eyland, R., Cooper, M., Conroy, J. & Collis, K. (Eds.), *Proceedings of PME 8*, pp. 162-171, Sydney, Australia

Cooper, M. (1986) "The dependence of multiplicative reversal on equation format", *Journal of Mathematical Behaviour*, Vol. 9, p. 115-120

Cooper, T. J., Boulton-Lewis, G. M., Atweh, B., Pillay, H., Wilss, L. & Mutch, S. (1997) "The Transition from Arithmetic to Algebra: Initial Understanding of Equals, Operations and Variable", in Pehkonen, E. (Ed.), *Proceedings of PME 21*, Vol. 2, pp. 89-96, Lahti, Finland

Cortés, A. (1998) "Implicit Cognitive Work in Putting Word Problems into Equation Form", in Olivier, A. & Newstead, K. (Eds.), *Proceedings of PME 22*, Vol. 2, pp. 208-215, Stellenbosch, South Africa

Cortés, A., Vergnaud, G. & Kavafian, N. (1990) "From Arithmetic to Algebra: Negotiating a Jump in the Learning Process", in Booker, G. , Cobb, P., & de Mendicuti, T. N. (Eds.), *Proceedings of PME 14*, Vol. 2, pp. 27-34, Oaxtepec, Mexico

Coxford, A. F. (Ed.) (1988) *The ideas of algebra K-12* (1988 NCTM Yearbook), Reston, VA, NCTM

Davis, P. J. & Hersh, R. (1980) *The Mathematical Experience*, Penguin

Davis, R. B. (1985) "ICME-5 report: Algebraic thinking in the early grades", *Journal of Mathematical Behaviour*, Vol. 4, pp. 195-208

Davis, R. B. (1986a) *Learning Mathematics: The Cognitive Science Approach to Mathematics Education*, Norwood, NJ: Ablex

Davis, R. B. (1986b) "What 'Algebra' Should Students Learn?", *Journal of Mathematical Behaviour*, Vol. 5, pp. 21-24

Davis, R. B. (1989) "Research Studies in How Humans Think About Algebra", in Wagner, S. & Kieran, C. (Eds.), *Research Issues in the Learning and Teaching of Algebra*, pp. 266-274, Reston, VA, NCTM

Davis, R. B. (1992) "Understanding 'Understanding'", *Journal of Mathematical Behaviour*, Vol. 11, pp. 225-241

Davis, R. B., Jockusch, E. & McKnight, C. (1978) "Cognitive processes in learning algebra", *Journal of Children's Mathematical Behaviour*, Vol. 2, No. 1, p. 10-32

Davydov, V. V. (1962) "An experiment in introducing elements of algebra in elementary school", *Soviet Education*, V(1), p. 27-37

Demana, F. & Leitzel, J. (1988) "Establishing fundamental concepts through numerical problem solving", in Coxford, A. F. (Ed.), *The ideas of algebra, K-12 (1988 Yearbook)*, pp. 61-68, Reston, VA, US, NCTM

Denmark, T., Barco, E. & Voran, J. (1976) "Final report: A teaching experiment on equality", Project for the Mathematical Development of Children (PMDC) Technical Report No. 6, (ERIC Document Reproduction Service No. ED144805), US, Florida State University

Dennett, D. C. (1993) *Consciousness Explained*, London, Penguin

Dettori, G., Garuti, R., Lemut, E. & Netchitailova, L. (1995) "An Analysis of the Relationship between Spreadsheet and Algebra", in Burton, L. & Jaworski, B. (Eds.), *Technology in Mathematics Teaching - a bridge between teaching and learning*, Bromley, UK, Chartwell-Bratt

Dickson, L. (1989) "Algebra", in Johnson, D. C. (Ed.), *Children's Mathematical Frameworks 8-13: A Study of Classroom Teaching*, Windsor, UK, NFER-Nelson

Dienes, Z. P. (1960) *Building up mathematics*, London, UK, Hutchinson Education

Dienes, Z. P. (1963) *An experimental study of mathematics learning*, New York, US, Hutchinson & Co., Ltd.

diSessa, A. (1995) "Collaborating via Boxer", in Burton, L. & Jaworski, B. (Eds.), *Technology in Mathematics Teaching - a bridge between teaching and learning*, Bromley, UK, Chartwell-Bratt

Donaldson, M. (1978) Children's Minds, Hammersmith, Fontana Press

Drouhard, J-P. & Sackur, C. (1997) "Triple Approach: A Theoretical Frame to Interpret Students' Activity in Algebra", in Pehkonen, E. (Ed.), *Proceedings of PME21*, Vol. 2, pp. 225-232, Lahti, Finland

Dugdale, S. & Kibbey, D. (1986) *Green globs and graphing equations*, Pleasantville, NY, US, Sunburst

Edwards, L. D. & Núñez, R. E. (1995) "Cognitive Science and Mathematics Education: A Non-Objectivist View", in Meira, L. & Carraher, D. (Eds.), *Proceedings of PME 19*, 2, 240-247, Recife, Brazil

Edwards, L. D. (1991) "A Comparison of Children's Learning in Two Interactive Computer Environments", in Furinghetti, F. (Ed.), *Proceedings of PME 15*, Vol. 2, pp. 1-8, Assisi, Italy

Eisenhart, M. A. (1988) "The Ethnographic Research Tradition and Mathematics Education Research", *Journal for Research in Mathematics Education*, Vol. 19, No. 2, pp. 99-114

English, L. & Halford, G. (1995) *Mathematics Education: Models and Processes*, NJ: Erlbaum

English, L. D. & Sharry, P. V. (1996) "Analogical Reasoning and the Development of Algebraic Abstraction", *Educational Studies in Mathematics*, Vol. 30, pp. 135-157

Ernest, P. (1991) *The Philosophy of Mathematics Education*, Basingstoke: Falmer Press.

Feurzig, W. (1986) "Algebra slaves and agents in Logo-based mathematics curriculum", Instructional Science, Vol. 14, P. 229-254

Fey, J. (1989a) "School Algebra for the year 2000", in Wagner, S. & Kieran, C. (eds.), *Research issues in the learning and teaching of algebra*, Reston, VA, US, NCTM

Fey, J. (1989b) "Technology and Mathematics Education: A survey of recent developments and important problems", *Educational Studies in Mathematics*, Vol. 20, pp. 237-272

Filloy, E. & Rojano, T. (1989) "Solving Equations: The Transition from Arithmetic to Algebra", *For the Learning of Mathematics*, 9 (2), p. 19-25

Fisher, K. M. (1988) "The students-and-professors problem revisited", *Journal for Research in Mathematics Education*, Vol. 19, No. 3, p. 260-262

Foxman, D. (1992) *Learning Mathematics and Science: The Second International Assessment of Educational Progress in England*, Slough, UK, NFER

Foxman, D., Ruddock, G. & McCallum, I. (1991) *APU Mathematics Monitoring 1984-1988 (Phase 2) Assessment Matters 3*, London, SEAC

Fullan, M. (1982) The Meaning of Educational Change, New York, Teachers College Press

Gallardo, A. & Rojano, T. (1987) "Common difficulties in the learning of algebra among children displaying low and medium pre-algebraic proficiency levels", in Bergeron J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 301-307, Montreal, Canada

Garançon, M., Kieran, C. & Boileau, A. (1990) "Introducing Algebra: A functional approach in a computer environment", in Booker, G. , Cobb, P., & de Mendicuti, T. N. (Eds.), *Proceedings of PME 14*, Vol. 2, pp. 51-58, Oaxtepec, Mexico, PME Program Committee

Gardner, J., Morrison, H. & Jarman, R. (1993) "The impact of high access to computers on learning", *Journal of Computer Assisted Learning*, Vol. 9, No. 1, pp. 2-16

Geary, D. C. (1994) *Children's Mathematical Development*, Washington, DC, American Psychological Association

Gibbs, W. & Orton, J. (1994) "Language and Mathematics", in Orton, A. & Wain, G. (Eds.), *Issues in Teaching Mathematics*, London, Cassell

Glaser, B. G. & Strauss, A. L. (1967) *The discovery of grounded theory: Strategies for qualitative research*, Chicago, US, Aldine

Glasersfeld, E. von (1995) *Radical Constructivism: A Way of Knowing and Learning*, London: Falmer

Graham, A. & Thomas, M. (1997) "Tapping into Algebraic Variables Through the Graphic Calculator", in Pehkonen, E. (Ed.), *Proceedings of PME 21*, Vol. 3, pp. 9-16, Lahti, Finland

Gray, E. & Tall, D. (1993) "Success and Failure in Mathematics: the flexible meaning of symbols as process and concept", *Mathematics Teaching*, No. 142, pp. 6-10

Greeno, J. G. (1982) "A cognitive learning analysis of algebra", Paper presented at the annual meeting of the American Educational Research Association, Boston, MA, US

Griffin, P. & Hirst, S. (1989) *Project Mathematics Update: Preparing to Teach Equations*, Milton Keynes: Open University

Hammersley, M. (1995) *The Politics of Social Research*, London: Sage

Hammond, M. (1994) "Measuring the impact of IT on learning", *Journal of Computer Assisted Learning*, Vol. 10, pp. 251-260

Harper, E. (1981) "Psychological changes attending a transition from arithmetical to algebraic thought", *Proceedings of PME 5*, pp. 171-176, Grenoble, France

Harper, E. (1987) "Ghosts of Diophantus", *Educational Studies in Mathematics*, Vol. 18, pp. 75-90

Hart, K. (Ed., 1981) *Children's Understanding of Mathematics 11-16*, London: John Murray

Hawkes et al. (1993) "A simple learning environment improves mathematical reasoning", in Nwana, H. (Ed.), *Mathematical Intelligent Learning Environments*, pp. 162-184, Oxford, UK, Intellect

Healy, L. & Sutherland, R. (1991) *Exploring Mathematics with Spreadsheets*, Oxford: Blackwell

Heid, M. K. & Kunkle, D. (1988) "Computer-generated tables: Tools for concept development in elementary algebra", in Coxford, A. F. (Ed.), *The ideas of algebra, K-12 (1988 Yearbook)*, pp. 170-177, Reston, VA, US, NCTM

Heid, M. K. & Zbiek, R. M. (1995) "A Technology-Intensive Approach to Algebra", *Mathematics Teacher*, Vol. 88, No. 8, pp. 650-656

Heid, M. K. (1990) "Uses of technology in pre-algebra and beginning algebra", *Mathematics Teacher*, Vol. 84, pp. 194-198

Hembree, R. & Dessart, D. J. (1986) "Effects of hand-held calculators in pre-college mathematics education: A meta-analysis", *Journal for Research in Mathematics Education*, Vol. 17, pp. 83-99

Herscovics, N. (1989) "Cognitive obstacles encountered in the learning of algebra", in Wagner, S. & Kieran, C. (Eds.), *Research issues in the learning and teaching of algebra*, pp. 60-86, Reston, VA, NCTM

Herscovics, N. & Kieran, C. (1980) "Constructing meaning for the concept of equation", *Mathematics Teacher*, Vol. 73, No. 8, pp. 572-80

Herscovics, N. & Linchevski, L. (1991) "Pre-algebraic thinking: range of equations and informal solution processes used by seventh graders prior to any instruction", in Furinghetti, F. (Ed.), *Proceedings of PME 15*, Vol. 2, pp. 173-181, Assisi, Italy

Herscovics, N. & Linchevski, L. (1992) "Cancellation within-the-equation as a solution procedure", in Geeslin, W. & Graham, K. (Eds.), *Proceedings of PME 16*, Vol. 1, pp. 265-272, Durham, US, University of New Hampshire

Herscovics, N. & Linchevski, L. (1994) "A cognitive gap between arithmetic and algebra", *Educational Studies in Mathematics*, Vol. 27, pp. 59-78

Hewitt, D. (1985) "Equations", *Mathematics Teaching*, Vol. 111, pp. 15-16

Hinsley, D. A., Hayes, J. R. & Simon, H. A. (1977) "From words to equations: Meaning and representation in algebra word problems", in Just, M. A. & Carpenter, P. A. (Eds.), *Cognitive Processes in Comprehension*, pp. 89-106, Hillsdale, NJ, Lawrence Erlbaum

Holt, J. (1969) *How children fail*, Harmondsworth, UK, Penguin

Hoyles, C. & Sutherland, R. (1989) *Logo Mathematics in the Classroom*. London & New York: Routledge

Hoyles, C., Sutherland, R., Noss, R. (1991) "Evaluating a computer-based microworld: what do pupils learn and why", in Furinghetti, F. (Ed.), *Proceedings of PME 15*, Vol. 2, pp. 197-204, Assisi, Italy

Hunter, M., Marshall, P., Monaghan, J. & Roper, T. (1995) "Using a Computer Algebra System with 14-15 year old Students", in Burton, L. & Jaworski, B. (Eds.), *Technology in Mathematics Teaching - a bridge between teaching and learning*, Bromley, UK, Chartwell-Bratt

Hurd, M. (1993) "Student Learning with a Computer Algebra System and a Spreadsheet" in Monaghan, J. & Etchells, T. (eds.) *Computer Algebra Systems in the Classroom*, p. 54-67. University of Leeds

Inhelder, B., Hermine, S. & Bovet, M. (1974) "Learning and the development of cognition", translated by Wedgwood, S., London, Routledge and Kegan Paul

Johnson, D. C. (Ed.) (1989) *Children's Mathematical Frameworks 8-13: A Study of Classroom Teaching*, Windsor, UK, NFER-Nelson

Johnson, D. C., Cox, M. J. & Watson, D. M. (1994) "Evaluating the impact of IT on pupils' achievements", *Journal of Computer Assisted Learning*, Vol. 10, No. 3, pp. 138-156

Johnson, M. (1987) *The body in the mind: the bodily basis of meaning, imagination and reason*, Chicago, University of Chicago Press

Kaput, J. (1987) "PME 11 Algebra Papers: A Representational Framework", in Bergeron J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 345-354, Montreal, Canada

Kaput, J. (1989) "Linking representations in the symbol systems of algebra", in Wagner, S. & Kieran, C. (Eds.), *Research Issues in the learning and teaching of algebra*, Hillsdale, NJ, Lawrence Erlbaum

Kaput, J. (1992) "Technology and Mathematics Education", in Grouws, D. A. (Ed.), *Handbook of Research on Mathematics Teaching and Learning*, pp. 515-556, New York, Macmillan

Kaput, J. & Sims-Knight, J. (1984) "Errors in translations to algebraic equations: Roots and implications", in Behr, M. & Bright, G. (Eds.), *Focus on Learning Problems in Mathematics*, Vol. 5, pp. 63-78

Kaput, J. & Thompson, P. W. (1994) "Technology in Mathematics Education Research: The First 25 Years in the JRME", *Journal for Research in Mathematics Education*, Vol. 25, No. 6, pp. 676-684

Keys, W., Harris, S. & Fernandes, C. (1996) *TIMSS: First National Report Part 1*, Windsor: NFER

Kieran, C. (1981) "Concepts associated with the equality symbol", *Educational Studies in Mathematics*, Vol. 12, pp. 317-326

Kieran, C. (1988a) "Learning the structure of algebraic expressions and equations", in Borbas, A. (Ed.), *Proceedings of PME 12*, pp. 433-440, Veszprém, Hungary

Kieran, C. (1988b) "Two different approaches among algebra learners", in Coxford, A. F. (Ed.), *The ideas of algebra, K-12 (NCTM Yearbook)*, pp. 91-96, Reston, VA, US, NCTM

Kieran, C. (1989a) "The early learning of algebra: A structural perspective", in Wagner, S. & Kieran, C. (Eds.), *Research issues in the learning and teaching of algebra*, pp. 33-56, Reston, VA, NCTM

Kieran, C. (1989b) "A perspective on algebraic thinking", *Proceedings of PME 13*, Vol. 2, pp. 163-171, Paris, France

Kieran, C. (1992) "The learning and teaching of school algebra", in Grouws, D. A. (Eds.), *Handbook of Research on Mathematics Teaching and Learning*, pp. 390-419, New York, US, Macmillan

Kieran, C. (1996) "The Changing Face of School Algebra", Paper presented at ICME8, Seville, Spain

Kieran, C. & Wagner, S. (1989) The Research Agenda Conference on Algebra: Background and Issues, in Wagner, S. & Kieran, C. (Eds.), *Research Issues in the Learning and Teaching of Algebra*, pp. 1-10, Reston, VA, NCTM

Kieran, C., Boileau, A. & Garançon, M. (1989) "Processes of mathematization in algebra problem solving within a computer environment: A functional approach", in Maher, C. A., Goldin, G. A. & Davis, R. B. (Eds.), *Proceedings of PME-NA 11*, pp. 26-34, New Brunswick, NJ, US, PME-NA Program Committee

Kirshner, D. (1987a) "The grammar of symbolic elementary algebra", Unpublished doctoral dissertation, Vancouver, Canada, University of British Columbia

Kirshner, D. (1987b) "The Myth about Binary Representation in Algebra", in Bergeron J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 308-315, Montreal, Canada

Kirshner, D. (1989) "The visual syntax of algebra", *Journal for Research in Mathematics Education*, Vol. 20, pp. 274-287

Kirshner, D. (1990) "Acquisition of Algebraic Grammar", in Booker, G. , Cobb, P., & de Mendicuti, T. N. (Eds.), *Proceedings of PME 14*, Vol. 2, pp. 75-84, Oaxtepec, Mexico

Körner, S. (1960) *The Philosophy of Mathematics*, London: Hutchinson.

Küchemann, D. (1981) "Algebra", in Hart, K. M. (Ed.), *Children's Understanding of Mathematics: 11-16*, 102-119, London: Murray

Kuhn, T. (1962) *The Structure of Scientific Revolutions*, Chicago University Press

Kutscher, B. & Linchevski, L. (1997) "Number Instantations as Mediators in Solving Word Problems", in Pehkonen, E. (Ed.), *Proceedings of PME 21*, Vol. 3, pp. 168-175, Lahti, Finland

Laborde, C. (1995) "Designing Tasks for Learning Geometry in a Computer-based Environment" in Burton, L. & Jaworski, B. (Eds.) (1995) *Technology in Mathematics Teaching - a bridge between teaching and learning*, Bromley, Chartwell-Bratt

Lakatos, I. (1976) *Proofs and Refutations*, UK, Cambridge University Press

Lakoff, G. (1987) *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*, Chicago, University of Chicago Press

Lamon, S. J. (1998) "Algebra: Meaning through Modelling", in Olivier, A. & Newstead, K. (Eds.), *Proceedings of PME 22*, Vol. 3, pp. 167-174, Stellenbosch, South Africa

Larkin, J. H. (1989) "Robust Performance in Algebra: The Role of the Problem Representation", in Wagner, S. & Kieran, C. (Eds.), *Research Issues in the Learning and Teaching of Algebra*, pp. 120-134, Reston, VA, NCTM

LaTorre, D. R. (1995) "Using Supercalculators to Affect Change in Linear Algebra", in Burton, L. & Jaworski, B. (Eds.), *Technology in Mathematics Teaching - a bridge between teaching and learning*, Bromley, UK, Chartwell-Bratt

Lee, L. (1987) "The status and understanding of generalised algebraic statements", in Bergeron J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 316-323, Montreal

Lee, L. & Wheeler, D. (1987) "Algebraic thinking in high school students: Their conceptions of generalisation and justification", Research Report, Montréal, Canada, Concordia University

Lee, L. & Wheeler, D. (1989) "The arithmetic connection", *Educational Studies in Mathematics*, 20, p. 41-54

Leibniz, G. W. (1702) "Necessary truths and the natural light within us", in Edwards, P. & Pap, A. (Eds.) (1957) *A Modern Introduction to Philosophy: Readings from Classical and contemporary Sources*, New York: The Free Press

Leitzel, J. R. (1989) "Critical Considerations for the Future of Algebra Instruction", in Wagner, S. & Kieran, C. (Eds.), *Research Issues in the Learning and Teaching of Algebra*, pp. 25-32, Reston, VA, NCTM

Lepper, M. (1985) "Microcomputers in education: Motivation and social issues", *American Psychologist*, Vol. 40, pp. 1-18

Lesh, R. & Kelly, A. E. (1991) "Human simulation of computer tutors: Lessons learned in a ten-week study of twenty human mathematics tutors", in Furinghetti, F. (Ed.), *Proceedings of PME 15*, Vol. 2, pp. 301-308, Assisi, Italy

Lewis, C. (1980) "Kinds of knowledge in algebra", Paper presented at the annual meeting of the American Educational Research Association, Boston, MA, US

Lewis, C. (1981) "Skill in algebra", in Anderson, J. R. (Ed.), *Cognitive skills and their acquisition*, Hillsdale, NJ, US, Lawrence Erlbaum

Lewis, M. W., Milson, R. & Anderson, J. R. (1987) "Designing an intelligent authoring system for high school mathematics ICAI: The Teacher's Apprentice Project", in Kearsley, G. P. (Ed.), *Artificial intelligence and instruction: Applications and methods*, Reading, MA, US, Addison-Wesley

Linchevski, L. & Herscovics, N. (1996) "Crossing the cognitive gap between arithmetic and algebra: Operating on the unknown in the context of equations", *Educational Studies in Mathematics*, Vol. 30, pp. 39-65

Linchevski, L. & Sfard, A. (1991) "Rules without reasons as processes without objects - the case of equations and inequalities", in Furinghetti (Ed.), *Proceedings of PME 15*, Vol. 2, pp. 317-325, Assisi, Italy

Lincoln, Y. S. & Guba, E. G. (1985) *Naturalistic Enquiry*, Beverly Hills, CA, Sage

Lins, R. (1990) "A framework for understanding what algebraic thinking is", in Booker, G. , Cobb, P., & de Mendicuti, T. N. (Eds.), *Proceedings of PME 14*, Vol. 2, pp. 93-100, Oaxtepec, Mexico

Lins, R. (1992) *A Framework for Understanding what Algebra Thinking is*, PhD Thesis, Shell Centre for Mathematics Education, Nottingham University

Lochhead, J. & Mestre, J. P. (1988) "From words to algebra: Mending misconceptions", in Coxford, A. F. & Shuttle, A. P. (Eds.), *The Ideas of Algebra. K-12 1988 Yearbook*, pp. 127-135, 1988 Yearbook of the National Council of Teachers of Mathematics, Reston, VA, US, NCTM

Love, E. (1986) "What is algebra", *Mathematics Teaching*, Vol. 117, pp. 48-50

Love, E. (1995) "Software for Mathematics Education", in Burton, L. & Jaworski, B. (Eds.), *Technology in Mathematics Teaching - a bridge between teaching and learning*, Bromley, UK, Chartwell-Bratt

MacGregor, M. & Stacey, K. (1993a) "Seeing a pattern and writing a rule", in Hirabayashi, I., Nohda, N. Shigematsu, K. & Lin, F.-L. (Eds.), *Proceedings of PME 17*, Vol. 1, pp. 181-188, Tsukuba, Japan

MacGregor, M. & Stacey, K. (1993b) "Cognitive models underlying students' formulation of simple linear equations", *Journal for Research in Mathematics Education*, 24 (3), 217-232

Malone, T. & Lepper, M. (1987) "Making Learning Fun: A taxonomy of intrinsic motivations for learning", in Snow, R. E. & Farr, M. J. (Eds.), *Aptitude, learning and instruction: Volume 3, Cognitive and Affective Process Analyses*, pp. 223-253, Hillsdale, NJ, US, Lawrence Erlbaum

Margolinas, C. (1991) "Interrelations between different levels of didactic analysis about elementary algebra", in Furinghetti, F. (Ed.), *Proceedings of PME 15*, Vol. 2, pp. 381-388, Assisi, Italy

Mason, J. & Pimm, D. (1984) "Generic examples: seeing the general in the particular", *Educational Studies in Mathematics*, Vol. 15, pp. 277-289

Mason, J., Graham, A., Pimm, D. & Gowar, N. (1985) *Routes to / roots of algebra*, Milton Keynes, UK, Open University Press

Mathematical Association (1934) *The Teaching of Algebra in Schools*, London, G. Bell & Sons

Mathematical Association (1992) *Computers in the mathematics classroom*, Leicester, UK, Mathematical Association

Matz, M. (1982) "Towards a process model for high school algebra errors", in Sleeman, D. & Brown, J. S. (Eds.), *Intelligent Tutoring Systems*, New York, Academic Press

Mayer, R. E. (1981) "Frequency norms and structural analysis of algebraic story problems into families, categories and templates", *Instructional Science*, pp. 135-174

Mayer, R. E. (1982a) "Different problem solving categories for algebra word and equation problems", *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 8, pp. 448-462

Mayer, R. E. (1982b) "Memory for algebra story problems", *Journal of Educational Psychology*, Vol. 74, pp. 199-216

Mayer, R. E. (1985) "Implications of cognitive psychology for instruction in mathematical problem solving", in Silver, E. A. (Ed.), *Teaching and Learning Mathematical Problem Solving: Multiple Research Perspectives*, pp. 123-138, Hillsdale, NJ, Lawrence Erlbaum

McArthur, D. (1985) "Developing computer tools to support performing and learning complex skills", in Berger, D., Pedzel, K. & Ganks, W. (Eds.), *Applications of cognitive psychology*, Hillsdale, NJ, US, Lawrence Erlbaum

McArthur, D., Stasz, C. & Hotta, J. Y. (1987) "Learning problem-solving skills in algebra", *Journal of Educational Technology Systems*, Vol. 15, pp. 303-324

Meira, L. L. (1990) "Developing knowledge of functions through manipulation of a physical device", in Booker, G. , Cobb, P., & de Mendicuti, T. N. (Eds.), *Proceedings of PME 14*, Vol. 2, pp. 101-108, Oaxtepec, Mexico

Mestre, J. P. & Lochhead, J. (1983) "The variable-reversal error among five cultural groups", *Proceedings of PME 7*, pp. 180-188, Montreal, Canada

Monaghan, J. (1994) "New Technology and Mathematics Education: New Secondary Directions" in Orton & Wain (eds.) "*Issues in Teaching Mathematics*", London: Cassell

Moncur, D. R. S. (1994) *Students' understanding of literal algebraic equations and formulae*, PhD Thesis, University of Leeds

Mordant, I. (1993) "Psychodynamics of Mathematics Texts", *For the Learning of Mathematics*, Vol. 13, No. 1, pp. 20-23

Morelli, L. (1992) "A visual approach to algebra concepts", *The Mathematics Teacher*, Vol. 85, No. 6, pp. 434-7

Nathan, M. J., Kintsch, W. & Young, E. (1992) "A theory of algebra word problems comprehension and its implications for the design of learning environments", *Cognition and Instruction*, Vol. 9, No. 4, pp. 329-389

NCET (1994a) *Mathematics and IT - a pupil's entitlement*, Coventry, UK

NCET (1994b) *Integrated Learning Systems - A report of the pilot evaluation of ILS in the UK*, Coventry, UK, NCET

Nicaud, J.-F. (1992) "A General Model of Algebraic Problem Solving for the Design of Interactive Learning Environments", in Ponte, J. P., Matos, J. F., Matos, J. M. & Fernandes, D. (Eds.), *Mathematical Problem Solving and New Information Technologies*, NATO ASI Series F, Vol. 89, Springer

Niemiec, R. P. & Walburg, H. J. (1992) "The effects of computers on learning", *International Journal of Educational Research*, Vol. 17, No. 1, pp. 99-108

Noddings, N. (1990) "Constructivism in mathematics education", in Davis, R. B., Maher, C. A. & Noddings, N. (eds.), *Constructivist views on the teaching and learning of mathematics*, pp. 7-18, (Journal for Research in Mathematics Education: Monograph Number 4), Reston, VA, US, NCTM

Nolder, R. (1991) "Mixing metaphor in mathematics in the secondary classroom", in K. Durkin and B. Shire (Eds.), *Language in Mathematics Education: Research and Practice*, Milton Keynes, UK, Open University Press

O'Brien, D. J. (1980) "Children's understanding of algebraic equations and their solution", Unpublished working paper, Nottingham, UK, Shell Centre for Mathematical Education, University of Nottingham

O'Reilly, D. (1990) "Hierarchies in Mathematics: A Critique of the CSMS Study", in Dowling, P. & Noss, R. (Eds.), *Mathematics versus the National Curriculum*, pp. 77-97, Basingstoke, UK, Falmer

Oldknow, A. (1995) "Personal Technology and New Horizons", in Burton, L. & Jaworski, B. (Eds.), *Technology in Mathematics Teaching - a bridge between teaching and learning*, Bromley, UK, Chartwell-Bratt

Oldknow, A. J. & Smith, D. V. (1983) *Learning Mathematics with Micros*, Chichester, UK, Ellis Horwood

Olds, H., Schwartz, J. & Willie, J. (1980) *People and Computers: Who teachers whom?*, Newton, MA, US, Educational Development Center

Olivier, A. (1988) "The Construction of an Algebraic Concept through Conflict", in Borbas, A. (Ed.), *Proceedings of PME 12*, Vol. 2, pp. 511-518, Veszprém, Hungary

Oprea, J. M. (1988) "Computer programming and mathematical thinking", Journal of Mathematical Behaviour, Vol. 7, pp. 175-190

Ormell, C. (1991) "How Ordinary Meaning Underpins the Meaning of Mathematics", *For the Learning of Mathematics*, Vol. 11, No. 1, pp. 25-30

Paige, J. M. & Simon, H. A. (1966) "Cognitive processes in solving algebra word problems", in Kleinmuntz, B. (Ed.), *Problem solving: Research method and theory*, pp. 51-119, New York, Wiley

Palmiter, J. R. (1991) "Effects of computer algebra systems on concept and skill acquisition in calculus", *Journal for Research in Mathematics Education*, Vol. 22, No. 2, p. 151-6

Papert, S. (1980) *Mindstorms*, Hertfordshire, UK: Harvester

Pawley, D. & Cooper, M. (1997) "What can be done to overcome the multiplicative reversal error?", in Pehkonen, E. (Ed.), *Proceedings of PME 21*, Vol. 3, pp. 320-327, Lahti, Finland

Payne, S. J. & Squibb, H. R. (1990) "Algebra mal-rules and cognitive accounts of error", *Cognitive Science*, Vol. 14, pp. 445-481

Pea, R. (1987) "Cognitive Technologies for Mathematics Education", in Schoenfeld, A. (Ed.), *Cognitive Science and Mathematical Education*, pp. 89-122, Hillsdale, NJ,, LEA publishers

Pereira-Mendoza, L. (1987) "Error patterns and strategies in algebraic simplification", in Bergeron J. C., Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 331-337, Montreal

Petitto, A. (1979) "The role of formal and non-formal thinking in doing algebra", *Journal of Children's Mathematical Behaviour*, Vol. 2, No. 2, p. 69-82

Philipp, R. A. (1992) "A study of algebraic variables: Beyond the student-professor problem", *Journal of Mathematical Behaviour*, Vol. 11, pp. 161-176

Phillips, D. C. (1989) "Subjectivity and objectivity: an objective inquiry", in Eisner, E. W. & Peshkin, A. (Eds.), *Qualitative Inquiry in Education: The Continuing Debate*, Read in Hammersley (1993), New York, US, Teachers College Press

Piaget, J., (1971) *Genetic epistemology*, translated by Duckworth, E., New York, Columbia University Press

Pimm, D. (1995) *Symbols and Meanings in School Mathematics*, London, Routledge

Pirie, S. & Kieren, T. (1989) "A Recursive Theory of Mathematical Understanding", *For the Learning of Mathematics*, 9 (3), 7-11

Pirie, S. & Kieren, T. (1994) "Growth in Mathematical Understanding: How can we characterise it and how can we represent it?", *Educational Studies in Mathematics*, Vol. 26, pp. 165-190

Poincaré, H. (1913) *The Foundations of Science*, New York, Science Press

Popper, K. (1934) *The Logic of Scientific Discovery*, Read in the 1959 translation, Hutchinson

Popper, K. (1945) *The Open Society and Its Enemies*, London: Routledge and Kegan Paul

Popper, K. (1963) *Conjectures and Refutations*, New York: Harper and Row

Popper, K. (1972) *Objective Knowledge*, Oxford University Press; read in the 1979 revised edition.

Popper, K. & Eccles, J. (1977) *The Self and its Brain*, Berlin, Springer

Radford, L. (1995) "Before the Other Unknowns were Invented: Didactic Inquiries on the Methods and Problems of Mediaeval Italian Algebra", *For the Learning of Mathematics*, Vol. 15, No. 3, pp. 28-38

Reed, S. K. (1984) "Estimating answers to algebra word problems", *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 10, pp. 778-790

Reed, S. K. (1987) "A structure-mapping model for word problems", *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 13, pp. 124-139

Reed, S. K., Dempster, A. & Ettinger, M. (1985) "Usefulness of analogous solutions for solving algebra word problems", *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 11, pp. 106-125

Roberts, N., Carter, R., Davis, F., Feurzeig, W. (1989) "Power tools for algebra problem solving", *Journal of Mathematical Behaviour*, Vol. 8, pp. 251-265

Robertson, S. I. (1994) *Problem solving from textbook examples*, PhD Thesis, Open University

Robitaille, D. F. (1989) "Students' Achievements: Population A", in Robitaille, D. F. & Garden, R. A. (Eds.), *The IEA Study of Mathematics 2: Contexts and Outcomes of School Mathematics*

Rojano, T. & Sutherland, R. (1991) "Symbolizing and solving algebra word problems: The potential of a spreadsheet environment", in Furinghetti, F. (Ed.), *Proceedings of PME 15*, Vol. 3, pp. 207-213, Assisi, Italy, PME Program Committee

Rojano, T. & Sutherland, R. (1997) "Pupils' Strategies and the Cartesian Method for Solving Problems: The Role of Spreadsheets", in Pehkonen, E. (Ed.), *Proceedings of PME 21*, Vol. 4, pp. 72-79, Lahti, Finland

Rorty, R. (1980) *Philosophy and the Mirror of Nature*, Oxford: Blackwell

Rosnick, P. (1981) "Some misconceptions concerning the concept of variable", *The Mathematics Teacher*, Vol. 74, No. 6 (Sep), pp. 418-420

Rosnick, P. & Clement, J. (1980) "Learning without understanding: The effect of tutoring strategies on algebra misconceptions", *Journal of Mathematical Behaviour*, Vol. 3, No. 1, pp. 3-27

Royal Society / JMC Working Group (1997) *Teaching and Learning Algebra pre-19*

Rubio, G. (1990) "Algebra Word Problems: a numerical approach for its solution (a teaching experiment in the classroom)", in Booker, G. , Cobb, P., & de Mendicuti, T. N. (Eds.), *Proceedings of PME 14*, Vol. 2, pp. 125-132, Oaxtepec, Mexico

Rumelhart, D. E. & Norman, D. A. (1981) "Analogical Processes in Learning", in Anderson, J. R. (Ed.), *Cognitive Skills and Their Acquisition*, pp. 335-359

Ruthven, K. (1995) "Pressing On", in Burton, L. & Jaworski, B. (Eds.), *Technology in Mathematics Teaching - a bridge between teaching and learning*, Bromley, UK, Chartwell-Bratt

Sakonidis, H. & Bliss, J. (1990) "Children's writing about the idea of variable in the context of a formula", in Booker, G., Cobb, P. & de Mendicuti, T. N. (Eds.), *Proceedings of PME 14*, Vol. 2, pp. 133-140, Mexico

Sammons, P. (1995) "Gender, Ethnic and Socio-economic Differences in Attainment and Progress: a longitudinal analysis of student achievement over 9 years", *British Educational Research Journal*, Vol. 21, No. 4, pp. 465-485

SCAA (1996) *Standards in Public Examinations 1975 to 1995*

Schliemann, A. D., Lima, A. P. de A. B., Santiago, M. M. L. (1992) "Understanding Equivalences Through Balance Scales", *Proceedings of PME 16*, Vol. 2, p. 298-305

Schoenfeld, A. & Arcavi, A. (1988) "On the Meaning of variable", *Mathematics Teacher*, Vol. 81, pp. 420-427

Schwartz, J. L. (1987) "The Representation of Function in the Algebraic Proposer", in Bergeron J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 235-240, Montreal, Canada

Seeger, F. (1990) "Observations on the 'reversal error' in algebra tasks", in Booker, G. , Cobb, P., & de Mendicuti, T. N. (Eds.), *Proceedings of PME 14*, Vol. 2, pp. 141-148, Oaxtepec, Mexico

Sfard, A. (1994a) "Mathematical Practices, Anomalies and Classroom Communication Problems" in Ernest, P. (Ed.) (1994) *Constructing Mathematical Knowledge: Epistemology and Mathematical Education*, Falmer, London

Sfard, A. (1994b) "Reification as the Birth of Metaphor", *For the Learning of Mathematics*, Vol. 14, No. 1, pp. 44-55

Sfard, A. & Linchevski, L. (1994) "The gains and the pitfalls of reification - the case of algebra", *Educational Studies in Mathematics*, 26, 191-228

Shavelson, R. J., Webb, N. M., Stasz, C. & McArthur, D. (1988) "Teaching mathematical problem solving: Insights from teachers and tutors", in Charles, R. I. & Silver, E. A. (Eds.), *The teaching and assessing of mathematical problem solving*, pp. 203-231, Reston, VA, US, NCTM

Shumway, R. (1988) "Graphics calculators: skill versus concepts", in de Lange, J. & Doorman, M. (Eds.), *Senior Secondary School Mathematics Education*, Utrecht, Netherlands, OW & OC

Sierpinska, A. (1994) *Understanding Mathematics*, London, Falmer

Sims-Knight, J. E & Kaput, J. (1983) "Exploring difficulties in transforming between natural language and image-based representations and abstract symbol systems of mathematics", in Rogers D. & Sloboda J. A. (Eds.), *The acquisition of symbolic skills*, pp. 561-570, New York, US, Plenum

Singley, M. K. & Anderson, J. R. (1989) *The Transfer of Cognitive Skill*

Skemp, R. (1971) *The Psychology of Learning Mathematics*, Harmondsworth: Penguin

Sleeman, D. (1984) "An attempt to understand students' understanding of basic algebra", *Cognitive Science*, Vol. 8, pp. 387-412

Sleeman, D. (1986) "Introductory algebra: A case study of student misconceptions", *Journal of Mathematical Behaviour*, Vol. 5, pp. 25-52

Sleeman, D., Kelly, A. E., Martinak, R., Ward, R. D. & Moore, J. L. (1989) "Studies of diagnosis and remediation with high school algebra students", *Cognitive Science*, Vol. 13, pp. 551-568

SMP (1978) *Algebra 2: Equations, Formulas and Graphs*, Cambridge University Press

SMP (1983) SMP 11-16: Materials for Years 7 & 8, Cambridge University Press

Soh, L. (1995) "Two pupils' growth of understanding of the concept of equation through an individualised learning scheme", Paper presented at the First European Research Conference on the Psychology of Mathematics Education, Osnabrueck, Germany

Soloway, E., Lochhead, J. & Clement, J. (1982) "Does computer programming enhance problem solving ability? Some positive evidence on algebra word problems", in Seidel, R. J., Anderson, R. E. & Hunter, B. (Eds.), *Computer Literacy*, New York, Academic Press

Sproule, S. (1998) "The Role of Time and Efficiency in Students' Algebraic Reasoning and Symbolic Manipulation", in Olivier, A. & Newstead, K. (Eds.), *Proceedings of PME 22*, Vol. 4, p. 309, Stellenbosch, South Africa

Stacey, K. & MacGregor, M. (1997) "Multiple Referents and Shifting Meanings of Unknowns in Students' Use of Algebra", in Pehkonen, E. (Ed.), *Proceedings of PME 21*, Vol. 4, pp. 190-197, Lahti, Finland

Steinberg, R. M., Sleeman, D. H., & Ktorza, D. (1991) "Algebra students' knowledge of equivalence of equations", *Journal for Research in Mathematics Education*, Vol. 22, No. 2, pp. 112-121

Sutherland, R. (1987) "A study of the use and understanding of algebra related concepts within a Logo environment", in Bergeron, J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 241-247, Montreal, Canada

Sutherland, R. (1988) *A longitudinal study of the development of pupils' algebraic thinking in a Logo environment*, PhD Thesis, Institute of Education, University of London

Sutherland, R. (1989a) "Providing a computer-based framework for algebraic thinking", *Educational Studies in Mathematics*, Vol. 20, No. 3, pp. 317-344

Sutherland, R. (1989b) "Developing algebraic understanding: the potential of a computer-based environment", in Vergnaud, G., Rogalski, J. & Artigue, M. (Eds.), *Proceedings of PME 13*, Vol. 3, pp. 205-212, Paris

Sutherland, R. (1990) "The Changing Role of Algebra in School Mathematics: the potential of computer-based environments" in Dowling, P. & Noss, R. (Eds.) *Mathematics Versus the National Curriculum*, London: Falmer

Sutherland, R. (1991) "Some Unanswered Research Questions on the Teaching and Learning of Algebra", *For the Learning of Mathematics*, Vol. 11, No. 3, pp. 40-46

Sutherland, R. (1993a) "Consciousness of the Unknown", *For the Learning of Mathematics*, Vol. 13, No. 1, pp. 43-46

Sutherland, R. (1993b) *Project AnA - The Gap Between Arithmetical and Algebraic Thinking*, Final Report to the ESRC, Institute of Education, University of London

Sutherland, R. (1995) "Algebraic Thinking: the role of the computer", in Burton, L. & Jaworski, B. (Eds.), *Technology in Mathematics Teaching - a bridge between teaching and learning*, Bromley, UK, Chartwell-Bratt

Sutherland, R. & Rojano, T. (1993) "A Spreadsheet Approach to Solving Algebra Problems", *Journal of Mathematical Behaviour*

Sweller, J., Mawer, R. F. & Ward, M. R. (1983) "Development of expertise in mathematical problem solving", *Journal of Experimental Psychology*, 112, p. 639-661

Tall, D. (1989) "Different Cognitive Obstacles in a Technological Paradigm", in Wagner, S. & Kieran, C. (Eds.), *Research Issues in the Learning and Teaching of Algebra*, Hillsdale, NJ, US, Lawrence Erlbaum

Tall, D. & Thomas, M. (1991) "Encouraging versatile thinking in algebra using the computer", *Educational Studies in Mathematics*, Vol. 22, pp. 125-47

Thomas, M. O. J. & Tall, D. O. (1988) "Longer-Term Conceptual Benefits from Using a Computer in Algebra Teaching", *Proceedings of PME 12*, Vol. 2, pp. 601-608, Budapest, Hungary

Thomas, M. O. J. & Tall, D. O. (1989) "Verbal Evidence for Versatile Understanding of Variables in a Computer Environment", *Proceedings of PME 13*, Paris, France

Thomas, T. & Tall, D. (1986) "The value of the computer in learning algebra concepts", *Proceedings of PME 10*, pp. 313-318, London, Institute of Education

Thompson, P. (1989) "Artificial intelligence, advanced technology and learning and teaching algebra", in Wagner, S. & Kieran, C. (Eds.), *Research Issues in the Learning and Teaching of Algebra*, Hillsdale, NJ, US, Lawrence Erlbaum

Thompson, P. W. & Thompson, A. G. (1987) "Computer presentations of structure in algebra", in Bergeron J. C. , Herscovics N. & Kieran C. (Eds.), *Proceedings of PME 11*, Vol. 1, pp. 248-251, Montreal, Canada

Thorndike, E. L. (1922) *The psychology of arithmetic*, New York, US, Macmillan

Tomlinson, P. (1989) "Having it Both Ways: hierarchical focusing as a research interview method", *British Educational Research Journal*, Vol. 15, No. 2, p. 155-176

Ursini, S. (1990) "Generalisation process in elementary algebra: interpretation and symbolisation", in Booker, G. , Cobb, P., & de Mendicuti, T. N. (Eds.), *Proceedings of PME 14*, Vol. 2, pp. 149-156, Oaxtepec, Mexico

Ursini, S. (1991) "First steps in generalisation processes in algebra", in Furinghetti, F. (ed.), *Proceedings of PME 15*, Vol. 3, pp. 316-323, Assisi, Italy

Vygotsky, L. S. (1962) *Thought and Language*, MIT Press

Vygotsky, L. S. (1978) *Mind in Society*, Harvard University Press

Wagner, S. (1977) "Conservation of Equation and Function and Its Relationship to Formal Operational Thought", Paper presented at the American Educational Research Association, New York

Wagner, S. (1981) "Conservation of Equation and Function under Transformations of Variable", *Journal for Research in Mathematics Education*, Vol. 12, pp. 107-118

Wagner, S. (1983) "What are these things called variables?", *Mathematics Teacher*, Vol. 76, No. 6, pp. 474-9

Watson, D. M. (Ed.) (1993) *The ImpacT Report: An evaluation of the impact of information technology on children's achievements in primary and secondary schools*, London, King's College London (for the Department for Education)

Wenger, R. H. (1987) "Cognitive science and algebra learning", in Schoenfeld, A. H. (Ed.), *Cognitive science and mathematics education*, pp. 217-251, Hillsdale, NJ, Lawrence Erlbaum

Wenger, E. (1987) *Artificial Intelligence and Tutoring Systems*, Los Altos, California, Morgan Kaufmann

Wertheimer, M. (1959) *Productive Thinking*, New York, US, Harper & Row

Wheeler, R. H. (1935) "The new psychology of learning", in Reeve, W. D. (Ed.), *The teaching of arithmetic: The 10th Yearbook of the NCTM*, pp. 233-250, New York, US, Teachers College, Columbia University

Whitman, B. S. (1976) "Intuitive equation solving skills and the effects on them of formal techniques of equation solving", Doctoral Dissertation, Florida State University

Wollman, W. (1983) "Determining the sources of error in a translation from sentence to equation", *Journal for Research in Mathematics Education*, Vol. 14, No. 3, pp. 169-181

Zehavi, N. (1988) "Substitutions leading to reasoning", in Borbas, A. (Ed.), *Proceedings of PME 12*, Vol. 2, pp. 665-672, Veszprém, Hungary

Appendix

A Written Test

B Questionnaire

C Year 10 Test Facilities

D Year 7 Test Facilities

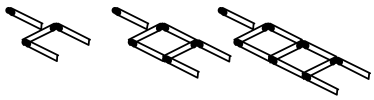
Appendix A: Written Test

Name: ______________________________ Class:__________

Solve each problem any way you can, but explain how you did it. Calculators allowed.

Section A: Patterns

1. Here are some piles of matches:



and so on…

Pile 1 Pile 2 Pile 3

(i) One of the piles contains 922 matches. How many matches are in the pile after?

Answer:

(ii) How many matches are in the 100^{th} pile?

Answer:

(iii) What is the number of the pile which has 568 matches in it?

Answer:

2. The results of an experiment that measured two quantities L and Q were:

L	3	5	9	21
Q	9	15	27	63

(i) What would you expect Q to be when L is 30? ____________________

(ii) What would L be when Q is 99? ____________________

(iii) Describe in words how you would find Q if you were told what L is:

(iv) Use algebra to write a rule connecting L and Q:

TC (written) J. C. Aczel, May 1997

Section B: Modelling to find an unknown

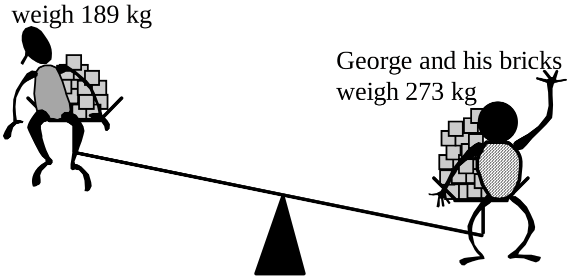
1. The perimeter of a field is 102m. The length is twice the width. What is the length?

Answer:

2.

Sam and his bricks weigh 189 kg

George and his bricks weigh 273 kg



Sam throws away some bricks and George throws away four times as many. Now they are balanced. How much weight did Sam throw away?

Answer:

3. Charlotte thinks of a number. If she doubles it and adds 6, she gets the same result as if she multiplies it by three and subtracts 70. What was her number?

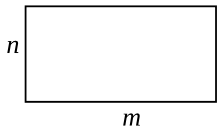
Answer:

TC (written) J. C. Aczel, May 1997

Section C: Representation

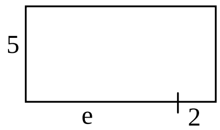
1. Find the area of each rectangle:

(i)



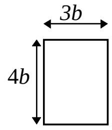
Area =

(ii)



Area =

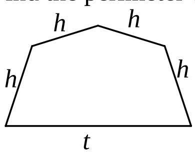
(iii)



Area =

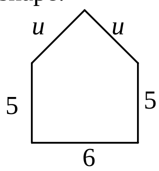
2. Find the perimeter of each shape:

(i)



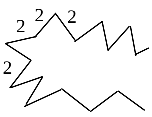
Perimeter =

(ii)



Perimeter =

(iii)



This shape has *n* sides, each of length 2, but only a part of it has been drawn.

Perimeter =

3. If John has J marbles and Peter has P marbles, what could you write for the number of marbles they have altogether?

4. Describe a situation in which $x = 4c$ could help you or tell you something.

____

5. Blue pencils cost 5p each and red pencils cost 6p each. I buy some blue and some red pencils and altogether it costs me 90p. If b is the number of blue pencils bought and r is the number of red pencils bought, what can you write down about b and r?

6. Write an equation to represent the statement:

'At this school there are 15 times as many students as teachers.'

Use S for the number of students and T for the number of teachers.

7. You have a pile of pennies and a pile of five pence pieces. The value of the pile of pennies is the same as the value of the pile of fives. Write this as an equation, using P for the number of pennies and F for the number of fives.

8. I have £x and you have £y. I have £6 more than you. Which of the following equations must be true? Tick *every* one you think is correct:

- ☐ $x = 6y$
- ☐ $6 + x = y$
- ☐ $6x = y$
- ☐ $x = 6 - y$
- ☐ $x = 6 + y$
- ☐ None correct

TC (written) J. C. Aczel, May 1997

Section D: Transformation

1. (i) Multiply $3n$ by 4 (ii) Add 4 onto $3n$

2. Multiply $n + 5$ by 4

3. (i) If $2x + 4 = 10$ then what is x? (ii) Solve $5x + 4 = 4x - 31$

______________________ ______________________

4. Write these more simply, where possible:

(i) $5a - 2a$ ______________
(ii) $2a + 5b$ ______________
(iii) $(a + b) + a$ ______________
(iv) $2a + 5b + a$
(v) $a + a + a \times 2$ ______________
(vi) $(a - b) + b$ ______________
(vii) $3a - b + a$ ______________
(viii) $4a - 2 + 7a + 1$

5. L + M + N = L + P + N is true…

☐ Always ☐ Sometimes ☐ Never (Tick one)

Why?

6. $7w + 22 = 109$ and $7n + 22 = 109$ have different solutions…

☐ Always ☐ Sometimes ☐ Never (Tick one)

Why?

7. Which of these can you write for $e + 2$ multiplied by 3?
Tick *every* one you think is correct:

☐ $e + 6$ ☐ $3(e + 2)$
☐ $3 \times (e + 2)$ ☐ $3e + 6$
☐ $3 \times e2$ ☐ $e + 2 \times 3$

8. Which is larger, $2n$ or $n + 2$? Explain.

____

9. What can you say about u if $u = v + 3$ and $v = 1$?

10. If $e + f = 8$, what is $e + f + g$?

11. What can you say about r if $r = s + t$ and $r + s + t = 30$?

12. What can you say about c if $c + d = 10$ and c is less than d?

TC (written) J. C. Aczel, May 1997

Appendix B: Questionnaire

Name: ____________________________________ Class:___________

Equations

1. Why do you think some people find equations difficult?

____

____

2. If you had to explain to somebody younger what sort of thing an equation is, how would you describe it?

____

____

3. Have you ever made up an equation? When and why?

____

____

4. How confident are you about algebra?

☐ Not at all ☐ Not very often ☐ Sometimes ☐ Mostly ☐ Always

Why? (Tick one)

____

5. How good are these descriptions of an equation?

	strongly disagree	disagree	no opinion	agree	strongly agree
group of symbols					
method					
statement					
balance					
formula					
problem					

6. How useful do you think calculators or computers are in learning about equations?

☐ Not at all ☐ Not much ☐ A little bit ☐ Very

Why? (Tick one)

____

TC (written) J. C. Aczel, May 1997

Appendix C: Year 10 Facilities

	10EQ					10CON				
Item	**Pre**	**Post**	**Imp**	**Wor**	**p**	**Pre**	**Post**	**Imp**	**Wor**	**p**
A1 (i)	100	86	0 / 0	3 / 22		96	83	1 / 1	4 / 23	
A1 (ii)	45	55	5 / 12	3 / 10		54	71	7 / 11	3 / 13	
A1 (iii)	50	50	5 / 11	5 / 11		54	46	2 / 11	4 / 13	
A2 (i)	95	100	1 / 1	0 / 21		96	92	0 / 1	1 / 23	
A2 (ii)	91	95	2 / 2	1 / 20		96	83	0 / 1	3 / 23	
A2 (iii)	86	91	2 / 3	1 / 19		75	71	4 / 6	5 / 18	
A2 (iv)	91	86	2 / 2	3 / 20		88	71	3 / 3	7 / 21	
B1	32	45	5 / 15	2 / 7		46	63	4 / 13	0 / 11	0.021
B2	23	59	10 / 17	2 / 5	0.009	29	42	3 / 17	0 / 7	
B3	23	50	6 / 17	0 / 5	0.005	8	46	9 / 22	0 / 2	0.001
C1 (i)	82	86	2 / 4	1 / 18		92	96	2 / 2	1 / 22	
C1 (ii)	36	55	6 / 14	2 / 8		25	58	9 / 18	1 / 6	0.004
C1 (iii)	55	73	5 / 10	1 / 12		58	79	8 / 10	3 / 14	
C2 (i)	91	86	2 / 2	3 / 20		58	54	4 / 10	5 / 14	
C2 (ii)	91	82	1 / 2	3 / 20		50	50	5 / 12	5 / 12	
C2 (iii)	50	77	8 / 11	2 / 11	0.028	50	67	6 / 12	2 / 12	
C3	82	73	0 / 4	2 / 18		88	88	2 / 3	2 / 21	
C4	9	36	7 / 20	1 / 2	0.015	58	38	2 / 10	7 / 14	
C5	9	27	4 / 20	0 / 2	0.021	0	8	2 / 24	0 / 0	
C6	45	55	6 / 12	4 / 10		38	33	6 / 15	7 / 9	
C7	5	5	1 / 21	1 / 1		0	0	0 / 24	0 / 0	
C8	23	27	3 / 17	2 / 5		29	25	3 / 17	4 / 7	
D1 (i)	68	55	2 / 7	5 / 15		54	63	4 / 11	2 / 13	
D1 (ii)	45	68	6 / 12	1 / 10	0.028	58	67	5 / 10	3 / 14	
D2	45	50	4 / 12	3 / 10		42	58	5 / 14	1 / 10	
D3 (i)	86	95	3 / 3	1 / 19		79	88	3 / 5	1 / 19	
D3 (ii)	18	45	8 / 18	2 / 4	0.028	4	8	2 / 23	1 / 1	
D4 (i)	77	68	0 / 5	2 / 17		71	75	4 / 7	3 / 17	
D4 (ii)	73	86	4 / 6	1 / 16		75	79	5 / 6	4 / 18	
D4 (iii)	64	77	4 / 8	1 / 14		46	54	5 / 13	3 / 11	
D4 (iv)	77	91	3 / 5	0 / 17		83	79	3 / 4	4 / 20	
D4 (v)	0	0	0 / 22	0 / 0		0	0	0 / 24	0 / 0	
D4 (vi)	45	32	0 / 12	3 / 10		29	38	4 / 17	2 / 7	
D4 (vii)	68	77	3 / 7	1 / 15		63	67	5 / 9	4 / 15	
D4 (viii)	45	36	1 / 12	3 / 10		54	42	3 / 11	6 / 13	
D5	27	59	7 / 16	0 / 6	0.003	42	50	5 / 14	3 / 10	
D6	55	50	2 / 10	3 / 12		38	21	1 / 15	5 / 9	
D7	0	5	1 / 22	0 / 0		0	0	0 / 24	0 / 0	
D8	18	14	1 / 18	2 / 4		21	21	1 / 19	1 / 5	
D9	59	68	5 / 9	3 / 13		75	83	3 / 6	1 / 18	
D10	41	45	3 / 13	2 / 9		42	42	4 / 14	4 / 10	
D11	9	23	5 / 20	2 / 2		29	38	5 / 17	3 / 7	
D12	27	36	4 / 16	2 / 6		50	42	4 / 12	6 / 12	

"**10EQ**" - Year 10 EQUATION group; "**10CON**" - Year 10 Control group;
"**Pre**" and "**Post**" - mean percentage score for the pre-test and post-test respectively;
"**Imp**" - number of items that improved, compared to the number of items where improvement was possible;
"**Wor**" - number of items that worsened, compared to the number of items where worsening was possible;
"**p**" - p-value of t-test with matched samples, for change from pre-test to post-test for a particular class

Appendix D: Year 7 Facilities

	7EQ					7CON				
Item	**Pre**	**Post**	**Imp**	**Wor**	**p**	**Pre**	**Post**	**Imp**	**Wor**	**p**
A1 (i)	77	85	4 / 6	2 / 20		81	67	1 / 4	4 / 17	
A1 (ii)	15	31	5 / 22	1 / 4		48	19	1 / 11	7 / 10	0.015
A1 (iii)	15	23	4 / 22	2 / 4		29	10	0 / 15	4 / 6	
A2 (i)	69	65	2 / 8	3 / 18		33	38	2 / 14	1 / 7	
A2 (ii)	65	65	3 / 9	3 / 17		24	48	5 / 16	0 / 5	0.011
A2 (iii)	62	62	4 / 10	4 / 16		24	38	4 / 16	1 / 5	
A2 (iv)	35	42	6 / 17	4 / 9		19	33	3 / 17	0 / 4	
B1	12	12	3 / 23	3 / 3		38	33	2 / 13	3 / 8	
B2	8	8	2 / 24	2 / 2		0	14	3 / 21	0 / 0	0.041
B3	0	4	1 / 26	0 / 0		0	5	1 / 21	0 / 0	
C1 (i)	8	4	0 / 24	1 / 2		0	5	1 / 21	0 / 0	
C1 (ii)	0	0	0 / 26	0 / 0		0	5	1 / 21	0 / 0	
C1 (iii)	0	8	2 / 26	0 / 0		0	5	1 / 21	0 / 0	
C2 (i)	0	4	1 / 26	0 / 0		5	5	0 / 20	0 / 1	
C2 (ii)	0	4	1 / 26	0 / 0		5	5	0 / 20	0 / 1	
C2 (iii)	0	4	1 / 26	0 / 0		5	5	0 / 20	0 / 1	
C3	12	38	8 / 23	1 / 3	0.008	24	33	2 / 16	0 / 5	
C4	4	0	0 / 25	1 / 1		5	0	0 / 20	1 / 1	
C5	0	0	0 / 26	0 / 0		0	0	0 / 21	0 / 0	
C6	4	27	6 / 25	0 / 1	0.006	24	24	0 / 16	0 / 5	
C7	0	4	1 / 26	0 / 0		5	5	0 / 20	0 / 1	
C8	4	8	2 / 25	1 / 1		14	14	0 / 18	0 / 3	
D1 (i)	19	31	5 / 21	2 / 5		0	10	2 / 21	0 / 0	
D1 (ii)	0	0	0 / 26	0 / 0		0	0	0 / 21	0 / 0	
D2	0	0	0 / 26	0 / 0		0	0	0 / 21	0 / 0	
D3 (i)	15	4	0 / 22	3 / 4		14	5	0 / 18	2 / 3	
D3 (ii)	0	0	0 / 26	0 / 0		0	0	0 / 21	0 / 0	
D4 (i)	12	4	1 / 23	3 / 3		0	5	1 / 21	0 / 0	
D4 (ii)	0	0	0 / 26	0 / 0		0	5	1 / 21	0 / 0	
D4 (iii)	0	4	1 / 26	0 / 0		0	10	2 / 21	0 / 0	
D4 (iv)	0	0	0 / 26	0 / 0		0	5	1 / 21	0 / 0	
D4 (v)	0	0	0 / 26	0 / 0		0	0	0 / 21	0 / 0	
D4 (vi)	0	0	0 / 26	0 / 0		5	10	1 / 20	0 / 1	
D4 (vii)	0	0	0 / 26	0 / 0		0	5	1 / 21	0 / 0	
D4 (viii)	0	0	0 / 26	0 / 0		0	5	1 / 21	0 / 0	
D5	23	23	3 / 20	3 / 6		5	10	1 / 20	0 / 1	
D6	12	8	1 / 23	2 / 3		0	0	0 / 21	0 / 0	
D7	0	0	0 / 26	0 / 0		0	0	0 / 21	0 / 0	
D8	0	0	0 / 26	0 / 0		0	0	0 / 21	0 / 0	
D9	12	15	3 / 23	2 / 3		10	19	3 / 19	1 / 2	
D10	0	4	1 / 26	0 / 0		0	0	0 / 21	0 / 0	
D11	4	4	0 / 25	0 / 1		5	5	1 / 20	1 / 1	
D12	0	4	1 / 26	0 / 0		0	5	1 / 21	0 / 0	

"**7EQ**" - Year 7 EQUATION group; "**7CON**" - Year 7 Control group;
"**Pre**" and "**Post**" - mean percentage score for the pre-test and post-test respectively;
"**Imp**" - number of items that improved, compared to the number of items where improvement was possible;
"**Wor**" - number of items that worsened, compared to the number of items where worsening was possible;
"**p**" - p-value of t-test with matched samples, for change from pre-test to post-test for a particular class

Name Index

A

B

C

D

E

F

G

H

J

K

L

M

N

O

P

R

S

T

U

V

W

Y

Z

Subject Index

D

E

F

G

I

K

L

M

N

O

P

R

S

T

U

V

W