DEVELOPING AN UNDERSTANDING

OF ALGEBRAIC SYMBOLS

by

Cyril Ronald Edmund QUINLAN, fms,
B.A.(Syd.), B.Sc.Gen.(Hons.I) (Lond.),
Dip.Theol.(S.T.U.), M.Sc.(Hons.) (Macq.)

School of Education

Submitted in fulfilment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF TASMANIA

July, 1992

DECLARATION

I, Cyril R. E. Quinlan, the undersigned,

hereby declare that this thesis does not contain

any material which has been accepted for

the award of any other higher degree or graduate diploma in any tertiary institution,

and that, to the best of my knowledge and belief,

this thesis does not contain

any material previously published or written by another person,

except when due reference is made in the text of the thesis.

Signed: Cyril Quinlan

(Cyril Quinlan)

Dated: 16th July, 1992

ABSTRACT

The major objective of this research project was to investigate the difficulties that beginning algebra students experience in developing an understanding of the meaning and use of algebraic symbols.

Learning problems identified by relevant research projects during the previous two decades provided a starting point, and items used in these projects for written tests or interviews were valuable in the formation of a new test instrument. By incorporating aspects investigated by several other researchers, a broad-based approach was employed to extend their work of applying psychological understandings of cognition to the learning processes involved in early algebra. Investigations examined interrelationships between measures previously studied in separate projects.

Data were collected for analysis from a sample of 208 Year 7 secondary school students as they began their study of algebra in the form of generalized arithmetic. Methods of data collection were repeated written tests, interviews and lesson observations. To locate the responses of the beginning Year 7 students in the learning continuum about algebraic symbols as numerical variables, research data were also collected from another 309 Years 7 to 12 students.

Scales were established for measuring and reporting on the patterns of thinking revealed by the students' responses. The pool of research information about the learning of algebra was expanded by the frequency data for individual items and for scaled groups of items. Comparisons and contrasts with findings of earlier researchers were reported where possible.

Hierarchies of difficulty, as proposed by previous researchers for distinguishing levels of understanding of algebraic symbols, were tested for their applicability to the student sample and to see if they reflected any identifiable learning sequences. The most difficult challenge for students beginning their study of the algebra of generalized arithmetic was found to be the attainment of an understanding of algebraic symbols as representing numerical variables. Some Year 7 students made little progress towards this goal during the seven months of the study. The tendency to regard symbols as standing for objects or people was one focus of attention.

Evidence supported the view that the level of achievement on the algebraic tasks presented is related to the degree of progress towards understanding algebraic symbols as numerical variables. Empirical data were shown to agree with psychological reasons for arranging some of the tasks into hierarchical orders of difficulty and/or into sequential orders of learning. There was some elucidation of the key steps in learning which distinguish students likely to progress in algebra.

ACKNOWLEDGEMENTS

I owe a great debt of gratitude to my supervisor, Professor Kevin F. Collis, Head, Department of Educational Studies,University of Tasmania. He was of great assistance to me as I worked my way through the learning experience of planning the research, analyzing data and identifying the key issues for this thesis. He was readily available and was remarkably prompt in returning well-informed comments on any work I submitted to him. Above all, he was most patient during stages in which effort appeared to be expended fruitlessly and when the implications of the data were gradually becoming clear.

I am grateful to my assistant supervisor, Dr Jane Watson, Lecturer, Centre for Educational Studies, Universtiy of Tasmania, for her perceptive comments as the work progressed.

Deserving of particular thanks is Dr Ian McKay, Hobart, for his invaluable expertise in painstakingly checking the drafts.

I am indebted to Dr Glen McPherson for his suggestions and advice in his role as consultant statistician, University of Tasmania.

Mrs Joan Marr completed the demanding task of transcribing onto computer files the audio-tapes of interviews.

I record my thanks to colleagues and follow postgraduate students with whom I discussed my work whether at conferences where I presented aspects of my work (in Hobart, Launceston, Sydney, and Assisi), in the corridors at the Centre for Educational Studies, within St.Virgil's, Hobart, or Lavalla, Westmead.

Special thanks go to my confrères, Dr Pat Fahy, fms, and Dr Peter Codd, fms, for their advice and organization of computing and printing facilities at Westmead.

The school principals, staffs and students of the six schools involved have my gratitude.

For allocating time for this project, I am most grateful to the Provincial Council of the Sydney Province of the Marist Brothers and the authorities of Australian Catholic University - N.S.W.

CONTENTS

CHAPTER 1
RATIONALE AND OBJECTIVES

CHAPTER 2
THEORIES OF COGNITION AND PREVIOUS RESEARCH

CHAPTER 3
RESEARCH METHODOLOGY FOR
TRIALLING PROGRAM AND MAIN RESEARCH PROGRAM

CHAPTER 4

STUDENT RESPONSES: A GLOBAL VIEW

CHAPTER 5
ESTABLISHMENT OF SCALES

CHAPTER 6

FIRST STUDY: A STUDY OF RELATIONSHIPS WITHIN LEVELS OF UNDERSTANDING FOR SYMBOLS AND BETWEEN LEVELS AND ACHIEVEMENT ON ALGEBRAIC TASKS

CHAPTER 7

SECOND STUDY: NUMBERS VIEW AND OBJECTS VIEW

CHAPTER 8

THIRD STUDY: A STUDY OF HIERARCHIES:
CHANGES IN LEVELS OF UNDERSTANDING OF SYMBOLS

CHAPTER 9

FOURTH STUDY: A STUDY OF DIFFERENCES:

DIFFERENTIAL DEVELOPMENT RATES FOR UNDERSTANDING SYMBOLS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1

RATIONALE AND OBJECTIVES

Overview

The rationale for carrying out a research project on the teaching and learning of algebra is established in terms of the cultural and historical background to algebraic symbolization, the important place algebra holds in the field of mathematics, the practice of including the study of algebra in secondary school curricula, and the abundance of evidence indicating that many students find algebra difficult.

Psychological learning theories found relevant to the study are mentioned and a list is given of the research projects from the past two decades that formed its foundation.

A broad range of possible research agenda is considered in the context of the overall research effort and aspects selected for this project are identified, leading to the statement of objectives and a brief account of the procedures selected for attaining them.

The Importance of Research on Learning Algebra

At the outset, it would seem useful to consider why algebra is a particular aspect of mathematics learning in schools that warrants close attention. The next few pages present a brief consideration of the cultural background to the development of the actual symbols used in algebra and some aspects of the importance of algebra to mathematicians and mathematics educators. Evidence is then assembled to make the point that many students have difficulty in understanding algebra. The research project was undertaken in response to the knowledge that this important branch of mathematics has been found so difficult by so many students, and in the hope of contributing to some redress of that situation in the future.

Algebraic Symbols in Historical and Cultural Perspective

The use of alphabetic symbols is the most obvious characteristic to distinguish the algebra currently taught in many Junior Secondary schools from arithmetic. This, however, is but a superficial distinction (cf. Lins, 1990). Mathematics that is classified as algebra was in existence many centuries before the French mathematician François Vièta, in 1591, published his work *Isagoge in artem analyticam*, marking the first systematic effort to introduce the use of letters as algebraic symbols. He used vowels for unknowns and consonants for knowns (Eves, 1969, p. 223; Gittleman, 1975, p. 140). What identifies much earlier work as algebra is the incorporation of the concept of an unknown. In fact, the term "algebra" was derived from the name of a book, *Kitab al-jabr wa'l-muqabalah*, "The Book of Completion and Cancellation" or "The Book of Restoration and Balancing" (Arndt, 1983, p. 669) in which the author, Al-Khwārizmī, in 825 A.D., showed little tendency towards algebraic symbolism apart from calling the unknown "the thing" (Boyer, 1968, p. 257; Italey, 1969, p. 612): He "wrote everything out in words, including numbers" (Gittleman, 1975, p. 132). It is his use of the concept of variable in presenting generalizations, while giving a clear exposition of how to solve linear and quadratic equations, that classifies the book as algebra, despite the fact that algebraic symbols, so common today, were missing. The term *al-jabr* referred to the process of "transposing a quantity from one side of an equation to the other" (Hollingdale, 1989, p. 97)

As early as 2000 B.C., the Babylonians tabulated solutions to equations of second and third degree, and the ancient Egyptians referred to unknown quantities as "heaps" (Prohhorov, 1973, p. 244; Gittleman, 1975, p. 8). Later, the Greeks expressed algebraic ideas in the language of geometry (Eves, 1969, p. 61).

Algebra was developed for many centuries without alphabetic symbols. In *rhetorical* algebra, everything was written out in grammatical sentences. Archimedes (287 - 212 B.C.) used *rhetoric* without the use of alphabetic symbols (Hollingdale, 1989, p. 73) to express his algebraic reasoning on topics as varied as spirals, quadrature of the parabola, spheres, cylinders, conoids and spheroids. In *syncopated* algebra, which followed as a later stage of development, abbreviations were used for commonly recurring quantities and operations. For example, the first letter was used in place of a term such as *plus* or *minus* (Ball, 1960, p. 239). Diophantus (around A.D. 250) used the Greek letter sigma, ς, as a symbol for an unknown quantity. "This symbol in verbal description he [Diophantus] calls ... 'the number', *i.e.*, by implication, the number *par excellence* of the problem in question" (Heath, 1910, p. 32). However, the symbol was "more of the nature of an abbreviation than an algebraical symbol like our x " (Heath, 1910, p. 34). Diophantus, "compelled by limitation of notation" (Heath, 1910, p. 51), kept to one unknown per problem and dealt with indeterminate situations involving more than one unknown by giving arbitrary values to all but one of the unknowns.

The development of *symbolic* algebra marked the middle years of the Renaissance (c.1450 - 1637). Apart from Vièta's introduction of letters, the symbols '+' for addition and '-' for subtraction were established about the fifteenth century. They were used, for instance, by Johann Widman in 1489 (Scott, 1969, p. 94). The sign 'x' for multiplication was introduced by Englishman Oughtred in 1631 (Scott, 1969, p. 103), although several other signs were used, such as the two that are still accepted, even in Junior Secondary classrooms: a dot as used by Englishman Harriot in 1631, and juxtaposition or conjoining as used by Frenchman Descartes in 1637 (Ball, 1960, p. 241). The use of the symbol '=' for equality was not introduced until 1557, by Englishman Record (Eves, 1969, p. 215). Vièta and others had earlier employed the same symbol "written between two quantities to denote the difference between them" (Ball, 1960, p. 240). Parentheses appeared in 1544 (Kline, 1972, p. 260) and brackets were first used by the Dutch mathematician Girard in 1629 (Ball, 1960, p. 242).

The now familiar symbols of Arabic numeration went through many stages of development (Ifrah, 1988, pp. 480, 482) and were brought into general use (Boyer, 1968, p. 280; Ball, 1960, p. 168) by the 1202 publication, *Algebra et almuchabala*, or *Liber Abaci*, by Leonardo of Pisa, who was also known as Fibonacci (Eves, 1969, p. 209). The need for *zero* as a place-value arose when figures were written rather than recorded on a counting board but its modern meaning of "numeral" was not acquired until the 1491 publication of *De Arithmetica Opusculum* by Italian Philippi Calabdra (Ifrah, 1988, pp. 481, 483). Writing negative numbers was not accepted for some time: Harriott (1560 - 1621) was "one of the first algebraists" to do so occasionally (Kline, 1972, p. 252). People like Descartes and Vièta allowed letters to stand for positive numbers only, and it was John Hudde (1633 - 1704) who first used a letter for both positive and negative numbers (Kline, 1972, p. 262).

Algebraic symbols provide a concise and precise way of communicating mathematical concepts and so can be vehicles for efficient and productive thought. For instance, the expression '$(10 - x).(24 + x)$' in algebraic symbols is much more concise than the five lines or so of Latin needed by Fibonacci to describe it rhetorically (Resnikoff & Wells, 1984, p. 204). The introduction of symbolism to algebra can be considered to have empowered the masses with algebraic thinking and problem-solving ability by routinizing the processes for manipulating numerical variables.

The symbols system required for today's high school algebra, and for all the tasks used in the present research program, was established by the end of the seventeenth century (Resnikoff & Wells, 1984, p. 206). Present-day algebra students have the advantage of working with the highly-structured system of algebraic symbols which took mankind so long to develop. Students face the challenge of understanding this system if they are to use it meaningfully in its varied applications. The research reported in this thesis was undertaken to elucidate precisely this challenge, the challenge of developing an understanding of the meaning and use of algebraic symbols.

Algebra as Mathematics

The type of thinking involved in algebra is essentially different from that in arithmetic. Arithmetic deals with specific numbers and methods of obtaining numerical answers to various computations with number. Early high school algebra deals with the processes involved in such computations, and with generalizations about arithmetic relationships: "The frontier between arithmetic and algebra is considered to have been crossed when the learner adopts the usage of the letter as a 'numerical variable' " (Harper, 1979, p. 3). Booth moves beyond this frontier: "In algebra the focus is on the derivation of procedures and relationships and the expression of these in general, simplified form" (1988, p. 21). This requires the movement away from specific numbers, beyond the idea of generalized numbers, to the notion of *variables*. The latter is a quite abstract notion: A numerical variable is able to represent simultaneously any number from a domain of possible numbers (Harper, 1979, p. 242).

The notion of variable is central to algebra and accounts for its power and importance within the science of mathematics.

> The concept of variable is central to mathematics teaching and learning in junior and senior high school. Understanding the concept provides the basis for the transition from arithmetic to algebra and is necessary for the meaningful use of all advanced mathematics.
> (Schoenfeld & Arcavi, 1988, p. 420)

Nunn wrote that the invention of variables could be "the most important event in human evolution" (1919, p. 8). He recognised that the idea of variables is used in common language, but saw algebra as making a special use of the notion.

The advent of algebraic variables made it possible to deal with relationships between mathematical processes so that generalizations could be clearly stated. Algebraic generalizations make use of symbols for variables in order to state inferences based on the properties of mathematical processes. For example, the statement '$a(x + b) = ax + ab$' can be interpreted as a relationship which is true for any particular arbitrary set of number-values for the symbols a, x, and b, or true for all sets of number-values simultaneously. This fluidity of interpretation, moving from the particular to the general and back again, is characteristic of generalizations in algebra. Moreover, the power to form such generalizations is an essential characteristic of algebra. Traditional algebra "can perhaps best be defined as the study of generalizations" (Norris, 1973, p. 1).

"One of the important functions of algebra is to permit the concise representation of general relationships and procedures" (Booth, 1986, p. 2).

Algebra has brought to the science of mathematics the pervasive power of forming generalizations, and a structure for analysing and comparing such generalizations. To be able to make use of this power, a person needs to understand the concept of variable. The challenges involved in developing this concept in the minds of beginning algebra students is the main focus of this study.

Once this concept is in place, the way is open for progress into higher mathematical pursuits:

> To view pronumerals not only as generalized numbers but also as variables and to be able to express relationships among variables are necessary concepts for an appreciation of abstract mathematical systems.
> (Collis & Biggs, 1979, p. 97)

Algebra in Current School Curricula

As we have seen, the power of algebra makes it of great importance to mathematicians. It is no surprise, then, that it is included in the curriculum for secondary school students around the world.

Convenient and appropriate data on how firm a place algebra has in school curricula are provided by the First and Second International Mathematics Studies, the first conducted across 12 educational systems in the years 1960 to 1964 (Husén, 1967), and the second conducted in the years 1980 to 1982 across 20 educational systems (Robitaille & Garden, 1987, 1989). Algebra was regarded as holding an important place in the mathematics curricula of the countries surveyed:

> Algebra was also considered important in both studies, and it is the major focus of the Junior Level curriculum in many systems. Its importance has increased in most systems.
> (Robitaille & Garden, 1989, p. 175)

The publication, Standards for School Mathematics (NCTM, 1989), pinpoints the following reasons for judging algebra to be important in the school curriculum for the 1990s:

> Algebra is the language through which most of mathematics is communicated. It also provides a means of operating with concepts at an abstract level and then applying them, a process that often fosters generalizations and insights beyond the original context. (p. 150)

The expectation expressed in this Standards document (p. 150) is that, prior to Grade 9, algebra would be developed as a generalization of arithmetic, extending in Grades 9 to 12 by focusing on algebra's own logical framework and consistency. This could lead to the more sophisticated use of algebraic symbols for representing objects such as polynomials rather than numbers. Once this level is achieved, algebra is seen as having still further importance: "This more sophisticated understanding of algebraic representation is a prerequisite to further formal work in virtually all mathematical subjects" (NCTM, 1989, p. 150).

Although, as Eves (1969) wrote, "it is probably correct to say that mathematicians have ... studied well over 200 [such] algebraic structures" (p. 367), mostly in this century, the present study concentrates on the basic form of algebra which may be described as generalized arithmetic because the only operations involved are arithmetic, such as addition and multiplication, and it makes use of alphabetic symbols to represent numbers only. This form can build upon the experience of arithmetic which the students have had during their primary grades. It is the form which is currently taught in the early years of secondary school, in line with the views referred to above from the Standards document (NCTM, 1989). Algebraic symbols in this study are restricted to representations for numbers because the focus is on the early stages of learning algebra and not on the more advanced stages.

The National Statement on Mathematics for Australian Schools (Australian Education Council, 1990) promoted the tradition of teaching algebra in schools by pointing out that algebra provides notation for efficient work with mathematical generalizations. The Statement also highlights the importance of developing an understanding for such notation, and, by implication, indicates the relevance of the present study:

> Mathematics brings to the study of patterns an efficient and powerful notation for representing generality and variability, and for reducing complexity - algebra. In order to master this notation, the conceptual understandings that underpin it must be developed. (p. 187)

There are copious data that students have problems with this basic form of algebra. Some of these data are considered in the following section.

Students Have Problems with Early Algebra

To establish the fact that algebra has been a source of difficulty for students from many countries, some of the findings are now summarized from four surveys, namely, The Second International Mathematics Study, The Fourth Assessment of Mathematics in U.S.A., Concepts in Secondary Mathematics and Science in U.K., and the 1989 N.S.W. Reference Test in Mathematics.

1. Performance statistics from twenty countries. The Second International Mathematics Study, conducted in the years 1980 to 1982 across 20 educational systems, produced performance statistics for a group referred to as Population A, namely, 79 677 students in the grade (year level) where the majority have attained the age of 13 years to 13 years 11 months by the middle of the school year. The algebra

items used with this group were described as "dealing with topics such as integers, order of operations, simplification of algebraic expressions, evaluation of expressions and formulae, and solutions of linear equations and inequalities" (Robitaille, 1989, p. 110). The form of algebra tested was generalized arithmetic in which any variables used were numerical variables. The test data showed that more than half the population were incorrect on more than half of the algebra items, and less than one-quarter of the population answered more than 60% of them correctly. The quartile scores were: 25th Percentile: 39%, Median: 43%, 75th Percentile: 57% (McKnight, Travers, Crosswhite, & Swafford, 1985, p. 24).

The following three questions, Items 086, 118 and 151, drew special comment from the evaluators. Success rates are given as percentages.

<u>Item 086</u>: If $\frac{4x}{12} = 0$ then x is equal to

A. 0 B. 3 C. 8 D. 12 E. 16.

Mean 42%. Israel 57% (highest). Nigeria 19% (lowest).

<u>Item 118</u>: $\frac{x}{2} < 7$, then

A. $x < \frac{7}{2}$ B. $x < 5$ C. $x < 14$ D. $x > 5$ E. $x > 14$.

Mean 35%. Japan 45% (highest). Luxembourg 16% (lowest).

<u>Item 151</u>: If $5x + 4 = 4x - 31$, then x is equal to

A. -35 B. -27 C. 3 D. 27 E. 35

Mean 26%. Japan 58% (highest). Swaziland 9% (lowest).

(Robitaille, 1989, p. 114; Robitaille & Garden, 1987, Appendix E)

The evaluator's comment on these three questions summed up their deep concern so well that it is quoted in full as follows:

> None of these items (086, 118, 151) is particularly difficult, and at least the two involving equations were considered to be in the Intended Curriculum in almost every system. Yet performance levels are relatively weak on all three, and particularly on Item 151. If such items are appropriate to the curriculum, if teachers say they have taught this material, yet performance levels are so low, it may be a case that this material is beyond the capability of many Population A students. In fact, 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 quite poor.
> (Robitaille, 1989, p. 114)

The evaluators referred to these items as "straightforward" items. In all three the form of algebra being tested is clearly generalized arithmetic, as 'x' in each case stands for a number or a range of numbers. In Item 086, 'x' represents one particular number which could be termed the "discovered content" for 'x' and students were expected to consider 'x' equal to zero as a possibility for the value of this discovered content. The outcome was that less than half succeeded. Item 118 moved beyond the notion that 'x' had some particular value, and required students to allow 'x' to stand for a host of possible values simultaneously. Thus the item measured their ability to think in terms of a numerical variable more so than Item 086, which could account for the mean success rate being lower. However, the mean success rate for Item 151 was even lower, and this question asked for the discovered content for 'x'. The most probable reason for the increased difficulty level is the fact that Item 151 presented an equation which could not be readily solved by trial and error, even in the multiple-choice format (not many, presumably, could easily substitute '-35' as a trial value). The equation demanded a series of steps for its formal algebraic solution.

Other studies have recorded similar reasons for uneasiness. Brief extracts from data for three countries suffice to further emphasise the fact that many students perform poorly on algebra in the early secondary grades.

2. Performance statistics from U.S.A. The National Assessment of Educational Progress conducted the Fourth Assessment of Mathematics during the 1985 - 1986 school year across the United States of America. Below are some of the conclusions, as given by Swafford and Brown (1989), from the performances by 31 938 students in a sample of eleventh-graders on "approximately fifty items intended to measure students' ability to work with mathematical variables and algebraic expressions" (p. 55), each item being given to about 2 000 students. Some Grade 7 students were also tested.

1. "Little of algebra is intuitive. Students must study algebra to learn algebra" (p. 55). With this point in mind, data were collected in the present study from students *before* as well as after they began their classroom study of algebra.

2. "Algebra students had difficulty with technical vocabulary, graphic representations, and all but the most routine problem situations" (p. 56).

3. "Algebra students were generally successful at simplifying linear algebraic expressions" (p. 56). For example, of students who had done Algebra 1,

86% correctly simplified $3x + 2y + 5x$,

74% correctly simplified $9(1 + 5x) + 13$, and

51% correctly simplified $2 - 4(5 - x)$.

The fact that the same students revealed poor understanding in other items indicated the possibility that they had developed manipulative skills without necessarily understanding what the alphabetic symbols represented.

4. "About two-thirds of the Algebra 2 students and half of the Algebra 1 students could correctly translate word problems into algebraic expressions" (p. 56).

However, "only half of those with two years of algebra chose the correct equation to describe the situation 'the number of chairs (*C*) is twice the number of students (*S*)' " (p. 60). Such a question focuses on the interpretation of algebraic symbols as representing numbers of objects as distinct from objects themselves or abbreviations for them.

5. "Students at both grade levels [i.e., Grades 7 and 11], with and without algebra, had difficulty interpreting or manipulating formulas" (p. 56).

6. Over 70% of students without algebra were correct, as were over 90% of algebra students, on giving the value of '$a + 7$' when $a = 5$ but when the same question was asked using function notation, the percentages dropped to about 30% and 70% respectively (p. 62). This outcome highlights the difficulty students can have with formal notation. Function notation was excluded from the present study.

7. Performance on the tasks of interpreting and substituting values into a formula was low, e.g., fewer than 30% of algebra students, given the formula relating Fahrenheit and Centigrade temperatures, could give the increase in degrees Fahrenheit for one degree Centigrade increase, and only 40% of Algebra 2 students and 30% of Algebra 1 students could correctly substitute into the formula (p. 62).

The following two summary comments show that many of these students were not really coping with algebra.

> The performance of those students who had taken algebra indicated that, although they had learned symbol manipulation, they were unable to use variables and relations except in the most routine problem situations.
> (Swafford & Brown, 1989, p. 63)

> The majority of 11th-grade students who had completed one or two years of algebra could perform the symbolic manipulations involved in solving equations or simplifying expressions ... They could manipulate the variables, but they did not understand what they represented.
> (Carpenter & Lindquist, 1989, p. 165)

3. Performance statistics from England. The research program, Concepts in Secondary Mathematics and Science (CSMS) (Hart, 1981a), included testing in algebra of three thousand children in 1976 and another sample in 1977. One hundred children completed a longitudinal algebra study over the years 1976 to 1978. They were across classes from the 2nd Year Secondary (approximate age 13 years) to the 4th Year Secondary (approximate age 15 years). Some of the findings are summarized below, grouped in terms of Küchemann's (1981) classifications of children's differing interpretations of letters. Percentage success rates are given for 13-, 14-, and 15-year-olds.

1. Letter evaluated:

   "What can you say about a if $a + 5 = 8$?"

   86% (13) 92% (14) 93% (15)

   "What can you say about m if $m = 3n + 1$ and $n = 4$?"

   44% (13) 62% (14) 67% (15)

2. Letter not used

   "If $e + f = 8$, $e + f + g = ...$"

   25% (13) 41% (14) 50% (15)

3. Letter as object (shorthand for objects or objects in their own right)

   "$2a + 5a = ...$ "

   77% (13) 86% (14) 87% (15)

   "$2a + 5b + a = ...$ "

   40% (13) 60% (14) 66% (15)

   "$3a - b + a = ...$ "

   27% (13) 47% (14) 56% (15)

4. Letters as specific unknowns

   "Add 4 onto $3n$"

   22% (13) 36% (14) 41% (15)

   "Multiply $n + 5$ by 4"

   8% (13) 17% (14) 25% (15)

5. Letter as generalized number

   "What can you say about c if $c + d = 10$ and c is less than d ?"

   21% (13) 30% (14) 35% (15)

   and "$L + M + N = L + P + N$ is Always Sometimes (say when) Never true"

   11% (13) 25% (14) 27% (15)

6. Letter as variable

   "Which is larger $2n$ or $n+ 2$. Explain."

   4% (13) 5% (14) 8% (15)

(Küchemann, 1981, pp. 102 - 116.)

In the summary section, under the heading Implications for teaching, Küchemann (1981) wrote:

> On the algebra test the majority of 13, 14 and 15 year olds were ... not able to cope consistently with items that can properly be called algebra at all, i.e., items where the use of letters as unknown numbers cannot be avoided. (p. 118)

When Küchemann's research is treated in Chapter 2, the implications of these findings in relation to the development of an understanding of algebraic symbols are discussed.

4. Performance statistics from Australia. Each year in New South Wales all Year 10 students sit for a State-wide Reference Test in Mathematics. Results are quoted below for the 1989 Reference Tests. The plural is used here because testing is conducted at three levels: About 20% (or 11 thousand) sat for General, the lowest level, while about 50% (or 27 thousand) sat for Intermediate, the middle level, and about 30% (or 17 thousand) sat for Advanced, the highest level. The tests are used for norm referencing across the State in order to guide schools regarding the allocation of grades in the School Certificate awarded at the end of Year 10, and so the test designers favour questions with good discrimination qualities. However, the tests still provide a record of achievement on the questions used for this large population of students. They had started on algebra three years previously, in Year 7, and their ages were about 15 or 16 years. Outcomes are given below. All candidates were allowed to use calculators.

A sample of the success rates for these candidates on some of the multiple choice questions in these papers is as follows:

1. Manipulation of symbols:

   "Simplify $2x + 5y - x + 3y$" : 49% for the General candidates

   and 80% for Intermediate;

   "$2(y + 5) + 3(y + 1) =$" : 24% for General.

2. Substitution into equations:

   For a question requiring the substitution $D = 30$ in the equation $T = \frac{750}{D} - 5$: 47% for General;

   "If $f(x) = 3x^2$, then $f(-2) =$" : 35% for Intermediate.

3. Solving equations:

   "If $2a - 3 = 19$, then $a =$ " : 53% for General;

   "If $10x - 2 = 6x$, then $x =$ " : 45% for Intermediate.

4. Translating verbal statements into algebra:

   "Chris saved \$$M$ each week for W weeks and then spent \$$R$.
   The amount (in dollars) Chris had left was: ..." : 77% for Intermediate;

   "Three consecutive integers have a sum of x. The smallest of the integers is .... " : 47% for Advanced.

5. Manipulation of symbolic expressions including fractions:

   "$\frac{4a}{5} - \frac{a}{3} = ...$ " : 51% for Intermediate;

   "$x + \frac{1}{x} = ...$ " : 29% for Advanced.

(Memorandum Number 150/89 from the N.S.W. Board of Secondary Education)

These results show that less than 50% of the 1989 Year 10 students in N.S.W. managed to achieve success on many of the algebra questions set for their respective tests. The form of algebra tested in all questions was generalized arithmetic and the alphabetic symbols in every case represented numbers.

Implications for research. Overall, these poor performance figures on algebra topics taught in many countries and the evaluators' comments quoted above both support the need for research of the type undertaken in this study, in which the cognitive processes involved in early algebra are examined to shed some light on the reasons for the poor understanding shown by many as they begin their study of algebra. The study set out to examine the cognitive steps required by beginning algebra students to enable them to develop their understanding of the concept of algebraic symbols as representing numerical variables. Knowledge gleaned from this examination could help future efforts to increase students' success rates with tasks that involve the use of letters as unknown numbers and as variables.

Shaping the Research Agenda

Difficulties with early algebra are not just a recent occurrence. Both researchers and teachers have attempted over the years to elucidate the reasons for the difficulties and to find ways to handle them. Previous research studies on the difficulties students have with basic concepts in algebra fall into two groups, namely, those that investigate cognitive levels involved in understanding early algebra, and those that are more concerned with teaching initiatives. Now, with a larger pool of students doing algebra, there are more people needing the skills required to address algebraic problems and it becomes more urgent to take up where the earlier studies left off.

As Chapter 2 is devoted to a discussion of relevant psychological learning theories found helpful in pursuing the objectives of the research and to an account of recent research about early algebra, only a brief mention of these is made here.

Psychological Learning Theories

The theories of learning that are applied in the study are those of Biggs and Collis (1982, 1991), Halford (1987), and Fischer (1980). These were found applicable to mathematics and they gave consideration to learning environments.

Previous Research

The foundation for the research is the work of Collis (1972, 1975a, 1975b), Harper (1979), Küchemann (1980, 1981), Booth (1983, 1984a), Rosnick and Clement (1980), and MacGregor (1989, 1991). This study proposes to continue on from Booth's last stage, namely that of teaching, by monitoring concept development during the early algebra teaching and learning phase.

Possible Scope for the Research Study

Following the Research Agenda Conference on Algebra, held in Athens, Georgia, in 1987, Kieran and Wagner (1987) reported that "in generating issues for the research agenda, four broad categories emerged - - content, learning, instruction, and representation" (p. 433). Reflection on this statement helps place the present project in perspective relative to the wider research effort. Of the points they mentioned with respect to "content", some attention is given to "interconnections between symbolic manipulation and conceptual understanding in algebra" (p. 433). A key focus is the students' level of understanding for symbols. The need expressed under the heading "instruction" for "improved theories of learning in order to have a possible effect on algebra instruction" (p. 433) is kept in mind. The development of basic concepts in algebra is studied in the context of the instructional stage for introducing algebra to secondary school students, and references are made to theories of learning which took account of the learning environment. A response is made to the expressed need for study of "the extent to which dynamically-linked representations enhance or inhibit metacognitive processes" (p. 433). The inclusion of classes taught with the aid of concrete manipulatives allowed observations to be made about whether or not these aids helped develop and/or clarify cognitive and metacognitive processes.

The main area addressed in this thesis, however, is that of learning, which Kieran and Wagner (1987) described as including

> issues such as the characterization and development of algebraic thinking, levels of understanding in algebraic thinking with respect to specific concepts and processes, and the identification of difficulties inherent in the learning of algebra. (p. 433)

Wagner and Kieran (1989) also published a statement headed "An Agenda for Research on the Learning and Teaching of Algebra", as "an edited and elaborated version of the working group's thinking up to 1988" (p. 220). The working group referred to had generated suggestions for future research questions when they were at the Athens conference.

Two of the aspects suggested with regard to defining what students learn when they *learn algebra* are at least partly examined, namely:

"a. What are students' intuitive, pre-instructional ideas about various algebraic concepts?

b. What characterizes students' post-instructional attainment of conceptual and procedural knowledge?" (Wagner & Kieran, 1989, p. 226)

Much of the research dwells on the suggested topic of levels of understanding:

> What are the *levels of understanding* in algebra with respect to specific concepts/processes? (e.g., variable, ...)
> a. How can we characterize/define these levels?
> b. Are there cognitive hierarchies ... ?
> ...
> e. Are there constraints on the rate of development through the identified levels of understanding due to general stages of cognitive development?"
> (Wagner & Kieran, 1989, p. 227)

Some attention is given to the suggestion that misconceptions be researched:

"What are *common misconceptions* that students acquire in algebra and how do these misconceptions develop?" (Wagner & Kieran, 1989, p. 227).

A moderated response is given to the following challenge, which was also recommended for research:

"How can general *theories of learning* be elaborated so that they are more applicable to algebra?" (Wagner & Kieran, 1989, p. 229).

The study does not revise or extend the theories chosen but applies them to the learning of algebra. There is elaboration in the sense of demonstrating how they could be applied to algebra.

Intervention Efforts to Overcome Student Difficulties

Many people have worked at overcoming the difficulties that beginning students have with understanding the meaning and use of algebraic symbols and with developing the concept of a numerical variable. These include classroom teachers, text-book writers, curriculum organisers, and researchers. Approaches differ widely, as the following examples show.

1. Sawyer's bags of stones (Sawyer, 1964). Diagrams help students visualize a concrete referent. As Sawyer points out, his "bag of stones" matches the ancient Egyptians' practice of picturing an unknown number as a heap of stones.

2. Introducing Algebra with Math-Tiles (Mason & Broom, 1979). Letters are used as labels for geometric shapes.

3. Herscovics' arithmetic approach (Herscovics & Kieran, 1980). Formal arithmetic supplies the basis from which to build algebra.

4. The approach in which letters stand for objects such as apples and bananas, as in the textbooks Maths 8 by Lynch, Parr, Picking, & Keating (1980), and Progress Maths Year 8 by Singleton (1989).

5. Booth's computer metaphor (Booth, 1983). Memory locations are used for housing values given to variables and students are asked to write instructions to tell the computer how to perform mathematical functions. A discussion of Booth's work is presented in Chapter 2.

6. A Concrete Approach to Algebra (1982) prepared by The Mathematics and Computing Department of Brisbane College of Advanced Education. Geometric shapes are used to represent letters with certain "values". Apart from the suggestion that "values" be associated with the geometric manipulatives, the approach is similar to that presented in "Introducing Algebra with Math-Tiles".

7. The Hawaii Algebra Learning Project (Rachlin, 1987). This concentrates on developing problem-solving processes through the teaching of algebra.

8. The South Notts Algebra materials (Wigley, Rooke, Hart, & Bell, 1984). These are based on the assumption that "an improvement in algebra skill is likely to come only with increased understanding" (p. 6). The activities start from "number situations" such as Piles of Stones, Number Routes, Arithmagons and Number Chain Puzzles.

9. The N.S.W. Algebra Research Group's A Concrete Approach to Algebra (Quinlan, Low, Sawyer, & White, 1989). Physical manipulatives are used for several representational models.

Limiting the Scope

Before stating the objectives for this study, the restrictions placed on its scope are considered.

The form of algebra to be considered is that which fits the description of generalized arithmetic in which alphabetic symbols are used as numerical variables for treating arithmetical relationships in a general way. An example of an algebraic expression within this form of algebra is '$p + r$', when it is interpreted to have the meaning of the sum of any two numbers, 'p' and 'r', regardless of what values the two symbols might have.

The title of this thesis, "Developing an Understanding of Algebraic Symbols", implies the possibility of investigating the two dimensions of studying the cognitive development as a subjective process and of examining ways in which teaching activities might influence this development. The first dimension is the major component, with only passing comment on the second.

Research focused on comparing teaching approaches or assessing the influence of intervention teaching does not always produce a significant result. Brophy and Good (1986, p. 329) said of projects prior to the 1970's that "there has been remarkably little systematic research linking teacher behavior to student achievement." In the 1970's, the well-resourced projects Developing Mathematical Processes and Individually Guided Education (Romberg, 1977) produced the outcome that "little evidence is available to substantiate the importance of teacher actions", according to Romberg and Collis (1987, p. 17). These researchers identified the importance of including observations of teacher actions, pupil actions, and teacher-pupil interactions for productive research. Investigators are challenged not only by these aspects but also by the need to balance characteristics of schools, teachers, classroom groups, and individual students when comparing teaching approaches. Despite these difficulties, successful research in the area has been documented.

Brophy and Good (1986) summarized examples of progress made since 1970 in Process-Product Research as well as Correlational and Experimental Studies. Newmann and Thompson (1987) reviewed high quality studies of the effects of cooperative learning on achievement in secondary schools and found that two-thirds of 37 comparisons of cooperative versus control methods "favored a cooperative learning method at the .05 level of significance" (p. 11). Peterson (1988, p. 17) reported that "recent research on training cognitive strategies has demonstrated that cognitive strategy teaching can significantly enhance students' learning." Day, Webb, Nabate and Romberg (1987) showed that the use of certain teaching materials for statistics "had a positive effect on student outcomes that was not apparent in the control group" (p. 26).

Considering the time and resources put into several large-scale projects and the uncertainty of conclusive results about the influence of teaching practice on learning, it was clear that this doctoral research project could not reasonably attempt to resolve the still open question of what teaching approach would best introduce students to the algebra of generalized arithmetic. The focus in this project was kept to finding out more about how the students learn their algebra. A better knowledge of learning pathways and the cognitive processes involved should logically precede and contribute to the development of better teaching strategies at a later stage.

Shayer (1987) discussed major problems that arise when one attempts to identify causation involved in accelerated learning and when one tries to specify the parameters for any connection between an intervention teaching effect and students' learning. He maintains that the minimum requirement in the field of interventionist theory would seem to be that there should be some model underlying the intervention, and that one should show that "the effect of the intervention was due to the model used" (p. 755), an outcome difficult to attain. To measure the amount of transfer of training, he points out that it is necessary to carry out "a post-intervention comparison of fresh learning by experimental and control groups" (p. 770). It is beyond the limitations of this research to compare teaching approaches or to evaluate precisely which changes in performance on algebraic tasks might be due to the influence of intervention teaching. Any comments included on these issues are to be regarded as subsidiary to the main theme of the thesis.

Statement of Objectives

The scope of the project was delineated in terms of a major objective and three associated objectives. The latter were considered integral to the thrust of the first objective. The four objectives were:

1. to investigate the cognitive difficulties that beginning algebra students experience in developing an understanding of the meaning and use of algebraic symbols, especially during the process of first being introduced to algebra in the secondary school;

2. to examine interrelationships between relevant measures that were used by previous investigators and were restricted to separate studies or considered as isolated measures within the one study;

3. to develop the means of measuring levels of understanding in early algebra and of identifying patterns of thought in relation to basic algebraic tasks; and

4. to measure and interpret, from a psychological viewpoint, levels of understanding for beginning algebra students and for samples of students across all the years of secondary schooling.

Three applications flowed from the objectives.

Applications. After obtaining measures for the cognitive processing of basic algebraic tasks, they will be applied in investigations such as the following:

1. to find out whether or not the level of development of one's understanding of symbols is related to one's degree of success with certain algebraic tasks,

2. to see whether or not the data support psychological reasons for the hierarchical order of difficulty for certain cognitive processes in algebra and to see whether or not such hierarchical orders of difficulty are reflected in sequential orders of learning, and

3. to distinguish between the learning paths of those who progress with algebra and those who do not progress.

Procedures Used

The procedures for achieving the objectives of the present study will be detailed in Chapter 3. Briefly, they were to follow closely the first few weeks of introducing algebra to 10 groups of Year 7 students, of ages 12 or 13 years, by observing some lessons, testing, and interviewing; to test students from Year 8 to 12 who had more experience with algebra than the beginners in Year 7; to retest the Year 7 students after a delay of about six months; to analyse the resultant data in terms of evidence for hierarchies of learning; and to relate any such findings to levels of difficulty according to theories of cognitive psychology.

Classroom factors. The focus of the study, as already explained, is the development of an understanding of algebraic symbols and their use in early algebra. Such understandings do not develop in a vacuum or from ordinary, every-day life, but are the result of deliberate teaching (cf. Swafford & Brown, 1989, p. 55, quoted on page 8 above). It was considered important to the main focus to study classes of students as they were introduced to algebra, and it was seen to be advantageous for elucidating the cognitive processes involved to incorporate classes being taught by differing approaches. The two modes selected are the traditional textbook mode and one which uses concrete manipulatives. In the main research program, of the 10 Year 7 classes studied as they began algebra, 3 classes were Textbook Classes and 7 were Manipulatives Classes. A description of classroom activities for both the Textbook approach and the Manipulatives approach will be given in Chapter 3.

Textbook classes. Textbook Classes are included in order to have a record of the progress made by students taught by "traditional" methods. The texts already provided for the students were accepted for use during the research data-collection period and teachers were free to teach in whatever way they wished, with the aid of those texts. The freedom granted to these three teachers clearly precluded the possibility of regarding this research project as a vehicle for scientifically comparing different teaching methods. These lessons were not monitored by observation.

Manipulatives classes. Of the list of special intervention efforts given above, the last approach (the ninth) was chosen to be included in the research study. The teachers of the Manipulatives Classes followed the approach developed by the N.S.W. Algebra Research Group. Reasons for this choice will be presented in Chapter 3. In contrast to the Textbook classes, the researcher did have some control over the learning environment for the Manipulatives classes and did monitor many lessons, mainly in four of these classes (and the one he taught), as specified on page 72.

Review

The justification for undertaking a study focused on early algebra has been presented as two-fold:

1. Algebra holds a place of great importance in cultural heritage, in the science of mathematics, and in school curricula; and

2. Performance on algebra by many secondary students around the world is poor, a situation calling for improvement.

The objectives for focusing on the teaching and learning of early algebra are to come to clearer insights, in terms of the understandings of cognitive psychology, of the degrees of difficulty of the learning challenges which are faced by beginning algebra students as they develop an understanding of the meaning and use of algebraic symbols and start to be able to work with the concept of a numerical variable; and to identify those very basic understandings, if there are any, which beginners need to open the way for their possible progress with higher concepts of algebra.

Content of Remaining Chapters

Chapter 2. After consideration is given to several theories of learning which appear to be relevant to the challenge of inducting young secondary school students into an understanding of basic aspects of algebra in the form of generalized arithmetic, a review of recent studies on early algebra is presented.

Chapter 3. A description is given of the pilot testing and the designing of a test instrument deemed suitable for achieving the objectives mentioned above. The methods used for collecting data for the main project by lesson observations, interviews and repeated testing are outlined, with details regarding schools, samples, classes and differing methods of teaching intervention for ten classes of beginning algebra students. Data were collected in four testings of 208 Year 7 students and one testing of another 309 students from Years 7 to 12.

Chapter 4. A global view of student responses is presented, using data obtained from all 517 students. Details are given of coding procedures used to allow management of data and to preserve information regarding the variety of cognitive response types. Comparisons are reported, where available, between outcomes from this study and those of earlier researchers.

Chapter 5. The process of reducing the number of variables by means of principal component and factor analyses and scaling procedures is described. Scales are established first for only the correct answers and then for coded responses which included incorrect answers. Underlying patterns of thought were identified in these processes, and the scale scores provided ordinal measures of the degree to which students used these cognitive pathways.

Chapter 6. Levels of understanding of symbols as given in previous research papers are synthesized into a hierarchy of difficulty. Correlational analyses are reported for examining whether or not one's degree of success with the algebraic tasks included in the test was related to the levels of development of one's understanding of symbols. Statistical support is reported for the classifications of scale measures as Correct Responses, Progress Indicators, or Hindrance Indicators. Relationships between scale measures are discussed. Data are examined throughout this chapter by combining responses from all 517 students.

Chapter 7. The investigation of a Numbers View and an Objects View for symbols provides the focus of this chapter. Subgroups of responses are considered here for the first time, such as those from the different testings of the Year 7 beginners and from classes which differed in terms of the students' mathematical ability or experience with algebra.

Chapter 8. This chapter studies changes in patterns of thinking and reports the outcomes of investigations designed to identify hierarchies of learning. Theories of learning are applied to a selected four-part test question to produce a task analysis of the cognitive steps required to answer the question at various levels of expertise. Empirical evidence is then presented in support of the ranking of certain algebraic tasks and concepts in hierarchical orders of difficulty in accordance with psychological considerations.

Chapter 9. Responses are subdivided still further so that a study could be made of differential rates of development of an understanding of algebraic symbols. By focusing on a few test items as a means of identifying students' views of symbols, hierarchies of difficulty are further elucidated and evidence for sequential learning paths is discussed. Particular attention is given to the vital question of why some students progress rapidly in algebra while others make little, if any, advance.

Chapter 10. After presenting a summary, the final chapter comments on the limitations of the study and its implications for classroom practice and future research.

CHAPTER 2

THEORIES OF COGNITION AND PREVIOUS RESEARCH

Overview

Since the major objective of the research project is to investigate the process of developing an understanding of the meaning and use of algebraic symbols, it was considered logical to turn to established theories of cognition which could profitably be applied to such an investigation. Neo-Piagetian (or modified Piagetian) theories were found most suitable for supplying a framework and foundation to support the investigation of the variety of cognitive demands faced by students when they commence their study of algebra. Of those available, the Biggs and Collis, Halford, and Fischer theories were chosen. These theories are described in the first part of this chapter and are shown to be relevant to the research objectives.

This research project does not set out to evaluate, assess, or amend established theories of cognition. Others, such as the writers of the 1987 papers assembled by Demetriou (1988), have already reviewed the spectrum of neo-Piagetian theories. Rather, the selected theories are applied in the context of learning algebra in an effort to elucidate the cognitive difficulties identified in this investigation.

A second decision, deemed logical within the planning for the research project, was to build upon the work and experience of previous researchers who had found methods of studying some of the difficulties experienced by algebra students. The test instrument used in this project was largely constructed from ideas developed during the last two decades of research in the area. In the latter part of this chapter, an account is given of relevant research projects conducted by Collis, Harper, Küchemann, Booth, Rosnick and Clement, and MacGregor. These projects were clearly related to the theories of cognition chosen and appeared to be particularly appropriate as they formed a progression and provided the foundation for a broad-based project that had the potential to address the objectives defined in Chapter 1. Measures of various aspects of the development of an understanding of algebraic symbols were derived from these studies. Whereas such measures had previously been restricted to the separate studies, they could now be considered side by side: Interrelationships could be examined and combinations could be used to form new measures.

Theories of Cognition

This study concentrates on the cognitive challenges faced by students in their study of algebra in the secondary school, especially during the early stages. To achieve some degree of understanding of the levels of difficulty within these challenges, it was necessary to consider appropriate theories of cognition.

As was explained in Chapter 1, a strict experimental comparison of differing and different teaching methods and a detailed study of the impact of intervention teaching on concept development are beyond the scope of this research project. The fact that a large proportion of the data analysed in the study was associated with a period of teaching intervention gave reason, nevertheless, for applying theoretical stances which considered environmental influences alongside maturational influences. Clearly in such a context, theories which deny a place for environment as an influence on the development of cognitive skills raise questions about the activities of teachers, as Bidell and Fischer (in press) say:

> The traditional separation of cognitive structure from context-embedded activity creates a readiness dilemma, placing educators in a helpless position -- either waiting for cognitive structures to develop and then leaping in to fill them with knowledge, or ignoring the day-to-day business of schooling in an attempt to stimulate cognitive development.

The three theories chosen to provide a framework for investigating the cognitive difficulties experienced by students in their study of the algebra of generalized arithmetic were those of Biggs and Collis, Halford, and Fischer. Other neo-Piagetian theories were considered, such as the six discussed by Sternberg (1987) alongside Fischer's and Halford's, namely, those of Case, Demetriou and Efklides, Farrar, Pascual-Leone, Shayer, and Siegler. The three selected theories were deemed suitable in the context of mathematics and they acknowledge the importance of the environment in the learning process. They have much in common with Piaget's theory. All include a number of general structural levels of development associated with approximate age ranges, and all have higher order structures subsuming the lower orders. Moreover, these neo-Piagetian theories give more emphasis to context-specific development than does Piaget, as was pointed out by Case (1987b, p. 777):

> Although Piaget himself devoted considerable energy to defending the primacy of logical structures, most of the present theorists are content to acknowledge that different domains have their own structures, and that the experience on which acquisition of these structures depends has a strong domain-specific component.

All of the chosen theorists have acknowledged the importance of the learning environment on cognitive development:

1. Biggs and Collis (1991, p. 68) recognized that "appropriate social support from parent or teacher may enable the child to operate at a higher level than without such support", while also recognizing the place of maturation in the course of the development of higher order thinking;

2. Halford (1987, p. 625) saw that "performance will be a function of maturation and experiential variables", especially for children performing below their theoretical limit on certain tasks; and

3. Fischer was well aware of "the potent effect of environmental support on developmental level" (Fischer & Silvern, 1985, p. 638), while keeping in mind the maturational constraints on performance.

All three theories accept that development occurs at different rates for different people and within different contexts and, except for Halford, attempt to explain this in terms of various substages of development. They also consider applications that go beyond the logico-mathematical thought that drew much of Piaget's attention and they allow for the possibility of extending a student's cognitive development within a particular context by providing appropriate educational experiences. Let us examine them in turn in more detail.

Biggs and Collis: Modes of Functioning and Learning Cycles

Biggs and Collis (1982) realized the limitation of a cognitive development theory that held almost rigidly to a framework in which learners had to pass through one stage before they could cope with the next. It was the learners who carried labels such as "sensorimotor" or "concrete-operational" until they moved to the next higher category. Biggs and Collis saw that by endorsing a stage theory, then "it would follow that it is pointless to instruct children in material that requires thinking at a higher stage than that at which they are currently capable of thinking" (p. 20). Biggs and Collis thought of a clear-cut way around this impasse: They changed the frame of reference so that their point of departure was learning quality and not the developmental stage of the child. Children's performances on learning tasks could then take over the role of carrying the labels for different stages so that on one task they might perform at the "concrete operational" level but on another the same children might show that they can work at the "formal" level. Their 1982 theory has since become associated with "the SOLO Taxonomy", a method of analysing the Structure of the Observed Learning Outcome. Qualities of the observed outcomes were categorized according to a neo-Piagetian sequence of stages which had been previously identified by Collis (1972, pp. 311 - 318; 1978, pp. 245 - 247) as:

early concrete-operational (7 to 9 years) - the later stages of Piaget's Stage IIA,
middle concrete-operational (10 to 12 years) - Piaget's Stage IIB,
concrete generalizations (13 to 15 years) - Piaget's Stage IIIA, and
formal operational (15 or more years) - Piaget's Stage IIIB.

For the SOLO Taxonomy, Biggs and Collis (1982, p. 19) also included a pre-operational stage (4 to 6 years) which corresponded approximately with the early stages of Piaget's Subperiod IIa, according to Battro's description of Piaget's Subperiods (Battro, 1973, p. 166).

In the 1982 exposition of their theory (Biggs & Collis, 1982), the following descriptors of the quality of outcomes were introduced: prestructural, unistructural, multistructural, relational and extended abstract. These terms were first used as sequential descriptors (p. 25) of the five developmental stages but were later used (p. 216) to refer to learning cycles associated with each stage. Characteristic modes of functioning were also given for each of the stages. To avoid confusion about the SOLO Taxonomy theory, more recent publications which are clearer (Collis & Campbell, 1987, Biggs & Collis, 1991, Collis & Biggs, 1991) have been used to assemble the following summary.

First, the modes of functioning are defined and then the learning cycle associated with each mode is explained.

The modes of functioning. The five modes of functioning identified by Biggs and Collis (1991, pp. 62 - 64) are as follows.

1. *Sensorimotor* (from birth). The elements operated upon are nearby objects and the operations involve the management and coordination of motor responses with respect to these objects. This mode of functioning copes with more and more complex situations as one gets older and leads to *tacit* knowledge as exemplified by people of whatever age who can skilfully carry out quite complicated motor activities such as those involved in dancing, golfing or sailing, even though they may not be able to explain how the individual sensorimotor skills involved in these activities were performed.

2. *Ikonic* (from 18 months). The elements which are manipulated in this mode are signifiers which stand for objects and events. Actions are internalized by imagining them, forming an internal picture or ikon. This mode is a prerequisite for language development and is the basic mode for *intuitive* thinking, a mode in which intellectual adults often come to see solutions to challenging problems. Ikonic thought can influence emotions, as is the case when it is associated with enjoyment of the arts in which this mode plays an important part.

3. *Concrete-symbolic* (from around 5 or 6 years). Here there is a significant shift towards abstraction. "The elements develop from mere signifiers to concepts and

operations which are manipulated using a logic of classes and equivalences; both elements and manipulations being directly related to the real world" (Collis & Campbell, 1987, p. 4). This mode opens up the possibility of acting on our environment through the medium of symbolic systems such as written language, mathematics, maps, and musical notation. A major task in primary and secondary schooling is mastery of these systems and of their application to real-world problems. This mode of learning leads to *declarative* knowledge which allows the concrete world to be described in symbolic terms.

4. *Formal* (from around 14 to 16 years). In this mode, theoretical constructs are able to be manipulated. Neither the elements nor the operations need a real-world referent. Abstract elements are used in thinking processes which involve the formation of hypotheses and propositional reasoning. The mode enables a person to attain *theoretical* knowledge such as would be required by undergraduates if they are to have "a workable grasp of an abstract academic discipline" (Collis & Biggs, 1991, p. 189).

5. *Postformal* (from about 20 years). Performances in this mode demonstrate development beyond the formal one by, for instance, having such a clear overview of a subject discipline that its basic tenets can be challenged or research can be carried out to extend understanding in the area.

<u>Learning cycles</u>. Biggs and Collis (1982, 1991) propose that associated with each of the above modes of functioning is a learning cycle consisting of five basic levels. These levels are observed through responses that are, in increasing order of complexity: prestructural, unistructural, multistructural, relational and extended abstract.

1. *Prestructural* responses imply that the learner is not operating at the level of abstraction required for the mode being considered and they belong to a prior mode of functioning. The response is not logically related to the data presented in the question.

2. *Unistructural* responses utilize just one relevant aspect of the data in the mode being considered.

3. *Multistructural* responses are identified by the way they process several disjoint aspects of the relevant data. The learners have taken notice of more and more relevant or correct features but have not integrated them.

4. *Relational* responses characteristically reveal an integrated understanding of the relationships between the different aspects being considered in the particular mode of functioning.

5. *Extended Abstract* responses represent a level of abstraction which makes use of a higher order principle and thus extends into the next mode and becomes the first, or unistructural, level of that next mode.

The following example illustrates the levels of responses within the concrete-symbolic mode. It was analysed by the authors of the SOLO Theory (Biggs & Collis, 1982, pp. 68 - 70; Collis & Biggs, 1979, pp. 91 - 93) and parts of it were used in the test instrument devised for the present project.

> You are to decide whether the following statements are true always, sometimes, or never. Put a circle around the right answer. If you put a circle around 'sometimes' explain when the statement is true. All letters stand for whole numbers or zero (e.g., 0, 1, 2, 3, etc.)
>
> 1. $a + b = b + a$ Always / Never / Sometimes, that is when .....
> 2. $m + n + q = m + p + q$ Always / Never / Sometimes, that is when .....
> 3. $a + 2b + 2c = a + 2b + 4c$ Always / Never / Sometimes, that is when .....
>
> (Biggs & Collis, 1982, p. 68)

Unistructural responses. One trial with numbers was used for making a decision. Each letter was considered as representing one and only one number. Students working at this level generally succeeded with the first part but were unable to manage the next two parts.

Multistructural responses. Two or more trials were used as bases for making a decision. The numbers were selected at random and the trials were treated as isolated pieces of information. The students at this level of thinking relied on the use of several specific numbers to replace the pronumerals and so succeeded with only the first item.

Relational responses. Students operating at this level

> had not developed the concept of pronumeral well enough to consider it as a variable but thought instead of letters in the terms as representing 'all the numbers that one could readily think of'. ... They confined their attention to interrelationships within the concrete data as they saw it, not thinking beyond it to the whole cardinal number system.
> (Biggs & Collis, 1982, p. 69)

Students giving relational responses worked with the concept of the letters as standing for a "generalized" number, so that '*a*', for instance, was regarded as an entity in its own right but restricted to having properties the same as any number with which they had previous experience. They were unable, even so, to proceed to the correct final step in Items 2 and 3. In Item 2 they did not imagine that '*n*' could equal '*p*' even though they realised that '*n*' and '*p*' could have any values from a large range of possibilities. Moreover, with Item 3 they were not able to deduce that '*c*' equalled zero even if they were able to conclude that '$2c = 4c$'.

Extended abstract responses produced correct answers to all items because the pronumerals were regarded as variables. Students operating at this level were able to take all possibilities into account because they were using high-level logical skills. As the previous responses were carried through at the concrete-symbolic level, here the extended abstract form of response fell into the category of the first level in the next mode, that of formal, and could be classified as unistructural in that mode, at least in Items 1 and 2. Item 3 could be considered as multistructural in the formal mode as several steps need to be completed even when using the variable idea. This meant, in fact, that those students who had attained the level of thinking in terms of variables were able to complete the three given items successfully because they did not need to revert to the numeral system through a series of trial numbers but could make comparisons and decisions while working with the alphabetic symbols.

Multimodal learning. In 1982, Biggs and Collis wrote that their theory had produced a marriage between the cyclical nature of learning and the hierarchical nature of cognitive development. Each level had its own integrity but served to supply the building blocks for the developmental transition to the next higher mode which subsumed the earlier mode. At that time, they pointed out that "such subsumption is not entire, however, as the learner had the option of operating at lower levels" (Biggs & Collis, 1982, p. 219). The idea that the learners could opt to work at lower levels than the highest of which they were capable has since been crystallized into a concept which they described more recently as "multimodal learning" (Biggs & Collis, 1991, p. 67).

Figure 2·1 summarizes the four possibilities which emerged when a multimodal outlook was accepted. This incorporation of the multimodal outlook has made the Biggs and Collis theory even more relevant to the study of the development of concepts for basic algebra in the context of different teaching activities, as attention was now drawn to the following possibilities: steady progress upwards through the developmental modes of functioning (Line A), progress limited to the learning levels within one mode of functioning (Line B), "top-down" facilitation of lower order learning through multimodal learning whereby higher modes assist learning in lower modes (Line C), and "bottom-up" facilitation of higher order learning in which lower modes are invoked in a multimodal progress to higher modes (Line D). For example, the use of concrete manipulatives as an aid to developing an understanding of simple algebraic expressions could well be an example of the latter possibility.

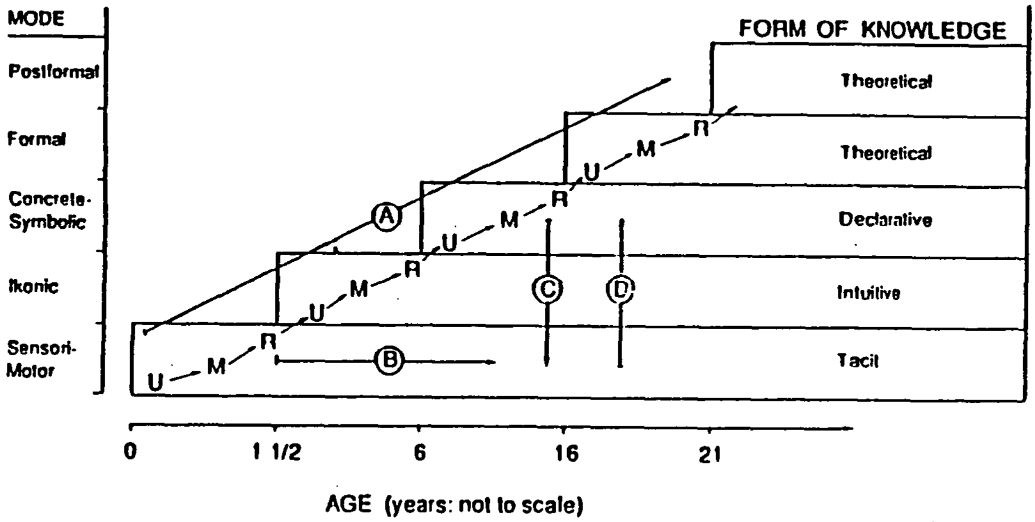


Figure 2·1. Modes, learning cycles and forms of knowledge (from Biggs & Collis, 1991, p. 66)

Halford's Structure Mapping Theory

The structural complexity of thought processes was given a central role in Halford's theory of cognitive development. Halford (1982) pointed out that thought itself, being essentially symbolic, is directly dependent on internal representation and that "cognitive structures, like concepts, ... consist of nothing more than an internal representation of some set of elements and a set of relationships between the elements" (p. 27). His theory provided a way of analysing the structure of cognitive processes in terms of representations and their interrelationships. He maintained that four major levels of cognitive development could be defined in terms of different levels of structure mappings which progressively require greater information-processing capacity.

Description of Structure Mapping. Halford (1987) explained that a structure consisted of a set of elements associated with each other by a set of relations. He defined a structure mapping as:

> a rule for assigning elements of one structure to elements of another, in such a way that any functions or relations between elements of the first structure will also be assigned to corresponding functions or relations in the second structure. (p. 611)

His four levels for structure mapping were defined as follows:

1. Element mappings are mappings in which single elements are mapped from one system to another. Criteria for element mappings are either similarity or convention. Similarity justifies an element representation when an object or event is represented by an image. Convention validates other representations such as using a word to represent an object or event. In element mappings, only one element needs to be considered at a time. (Example: categorizing a pet as a dog.)

2. Relational mappings are those involving the mapping from one structure to another of pairs of elements and a relationship between them. It is the similarity of the relationships between the elements that gives validity in this case. The mapping is independent of element similarity and convention, giving relational mappings more flexibility and generality than element mappings. To establish relational similarity, two elements must be considered in each mapping. (Examples: mare is to foal as mother is to child, and "big stick is larger than small stick" maps to "adult is larger than infant".)

3. System mappings are mappings in which three elements with relationships between two of the pairs are mapped from one structure to another. "Each element in one structure must be mapped into one and only one element in the other structure (i.e., uniquely) and there must be a consistent correspondence between the relations in the two structures" (Halford, 1987, p. 614). The relationships in one structure need not be the same as those in the other structure nor need they be similar, but they need to correspond consistently from one system to the other. System mappings are independent of both element and relational similarity and convention, giving them yet more generality and flexibility than relational mappings. Three elements must be considered in each mapping decision. (Examples: ordering and transitivity, such as when using heights of liquid in three equal-diameter cylinders to compare the volumes of liquid in each.)

4. Multiple system mappings involve the mapping of four elements in each system with corresponding relationships. The processing load is higher than it was in the previous mappings because now at least four elements must be considered in every mapping decision. (Example: "2 times 3 times 4 is the same as 4 times 3 times 2" maps onto, say, "3 times 5 times 9 is the same as 9 times 5 times 3", and both of these map onto the algebraic rule that "*a* times *b* times *c* is the same as *c* times *b* times *a*".)

The levels and their corresponding structure diagrams are set out in Table 2·1, with further examples.

Table 2-1
Halford's Structure Mapping Levels

LEVEL	STRUCTURE DIAGRAM	No.of ele-ments	EXAMPLES
ELEMENT MAPPINGS	a ↕ x	1	simple categories (e.g., my pet is a dog)
RELATIONAL MAPPINGS	a __R__ b ↕ ↕ x __R'__ y	2	simple binary relations (e.g., is less than, equals); simple analogies (e.g., mare to foal as mother to child)
SYSTEM MAPPINGS	a __R__ b __R__ c ↕ ↕ ↕ x __R'__y __R'__z	3	simple binary operations (e.g., ordering); interpretation of algebraic expressions containing single arithmetic operations (e.g., $3y$, or $y+3$)
MULTIPLE-SYSTEM MAPPINGS	a _R_ b _R_ c _R_ d ↕ ↕ ↕ ↕ w _R'_x _R'_y _R'_z	4	interpretation of algebraic expressions containing compositions of arithmetic operations [e.g., $3y+2$, or $a(b+c)$]

Note. Based on Halford, 1987, pp. 616 & 628.
↕ indicates a correspondence between elements in the mappings.
__R__ and __R'__ indicate relationships between elements.

Halford carried out experiments in which the median ages of the subjects for mastering the levels of structure mapping were found to be 1 year for element mappings, 2 years for relational mappings, 3 years for system mappings, and 11 years for multiple system mappings. Allowing for individual differences due to variations in capacity, he suggested that "system mappings might be attained by 10 percent [*sic.*] of three-year-olds, 25 percent of four-year-olds, 50 percent of five-year-olds, 75 percent of six-year-olds, and so on" (Halford, 1987, p. 629). Moreover, he assessed algebraic rules for high school as consisting of compositions of binary operations and hence requiring multiple-system mappings, and concluded that "ability to relate algebraic rules to an underlying representation of the concept should emerge at approximately 11 years" (Halford, 1987, p. 633).

Importance of structural correspondences. A key issue in the analysis of teaching and learning mathematics was identified by Halford and Boulton-Lewis (1989) when they wrote that "the recognition of correspondences between structures ... is central to mathematics learning at all levels" (p. 40). They admitted that the problem of how abstractions develop out of experience was far from solved and categorized the explanation of how people progress from representing constants to representing variables as a major problem. Nevertheless, they promoted the proposition that generalizations such as algebraic laws are developed by the learner's recognition of the structural correspondence between the abstraction and one or more specific examples of its application. They pointed out that both analogy theory and representation theory depend on structure mapping theory, as both depend on mapping one structure into another: in the case of analogies, from one mental structure to another, and from a cognitive structure to an environmental structure in the case of cognitive representations.

There is a strong emphasis on the role of structure mapping in their theory for analysing the value and limitations of analogues in mathematics teaching, indicating that structure mapping could be a suitable tool for analysing the value of any concrete model for helping to develop some concept(s) in mathematics. As an example of this, let us look at making use of arithmetic to develop a generalization in algebra, using the following sequence of steps: (a) the recognition of the similarity in structure of several arithmetical examples of the generalization, and (b) the recognition of the similarity of the structure of the algebraic generalization with the common structure of the numerical examples.

Halford and Boulton-Lewis (1989) proposed a hierarchy of mappings which could lead to the understanding of the algebraic generalization given by:

$$a(b+c) = (a \times b) + (a \times c).$$

To simplify the presentation, the following specific case of this generalization will be used in Figure 2·2 as an illustration of the hypothesized sequence:

$$2(y+3) = 2 \times y + 2 \times 3.$$

The sequence requires the recognition (in steps E and F) of the similarity of structure between examples employing different numerical values for 'y', such as '5', '1', or '4.5'. This leads to the mapping (in step G) of the recognized common structure onto the structure of the algebraic generalization.

In terms of Halford's categories, steps A and B of Figure 2·2 are examples of system mappings and the remaining steps are examples of multiple-system mappings.

$$\begin{array}{lccccccc}
\text{A} & 2 & (5+3) & = & 16 & & & \\
 & \updownarrow & \updownarrow & & \updownarrow & & & \\
 & 2 & \times\ 8 & = & 16 & & & \\[1em]
\text{B} & & (2 \times 5) & + & (2 \times 3) & = & 16 & \\
 & & \updownarrow & & \updownarrow & & \updownarrow & \\
 & & 10 & + & 6 & = & 16 & \\[1em]
\text{C} & 2 & (5+3) & = & (2 \times 5) & + & (2 \times 3) & \\
 & \updownarrow & \updownarrow & & \updownarrow & & \updownarrow & \\
 & 2 & \times\ 8 & = & 10 & + & 6 & \\[1em]
\text{D} & 2 & (1+3) & = & (2 \times 1) & + & (2 \times 3) & \\
 & \updownarrow & \updownarrow & & \updownarrow & & \updownarrow & \\
 & 2 & \times\ 4 & = & 2 & + & 6 & \\[1em]
\text{E} & 2 & (5+3) & = & (2 \times 5) & + & (2 \times 3) & \\
 & \updownarrow & \updownarrow & & \updownarrow & & \updownarrow & \\
 & 2 & (1+3) & = & (2 \times 1) & + & (2 \times 3) & \\[1em]
\text{F} & 2 & (5+3) & = & (2 \times 5) & + & (2 \times 3) & \\
 & \updownarrow & \updownarrow & & \updownarrow & & \updownarrow & \\
 & 2 & (4.5+3) & = & (2 \times 4.5) & + & (2 \times 3) & \\[1em]
\text{G} & 2 & (5+3) & = & (2 \times 5) & + & (2 \times 3) & \\
 & \updownarrow & \updownarrow & & \updownarrow & & \updownarrow & \\
 & 2 & (y+3) & = & (2 \times y) & + & (2 \times 3) &
\end{array}$$

Figure 2·2. Structure mappings from arithmetic to algebra (Adaptation of Figure 7 in Halford & Boulton-Lewis, 1989. $\updownarrow$ indicates a correspondence between elements in the mappings)

In using this paradigm, Halford and Boulton-Lewis (1989) suggest that each correspondence should be learned so well that retrieval is automatic before progressing to the next . "The load imposed by one structure mapping must be reduced to zero before the next structure mapping is undertaken, otherwise the cumulative load will become excessive" (p. 31). Thus, in the case outlined in Figure 2·2, the correspondence in part A must be available by immediate recall before proceeding to the correspondence in part B, and the latter must be learnt before progressing to the correspondence in part C, and so on. Approaching the algebraic generalization

$$2(y+3) = 2 \times y + 2 \times 3$$

from arithmetic is here shown to be quite complex for students.

Halford's (1987) method of assessing the complexity of a cognitive task is in terms of the number of independent dimensions in the structure. Halford argues that this method is mathematically based and so is not as heavily dependent upon intuition as is a processing model.

Fischer's Skill Theory

As Case (1987b) points out, "In Halford's work, the basic units are the symbol, and the conceptual framework that results from relating symbols to each other ... [but] ... in Fischer's system the basic units are skills" (p. 782). Bidell and Fischer (in press) define a skill as "a control structure governing a specific class of actions that a person can perform in a specific context" and they emphasize that a skill is not an attribute of the person alone nor the environment alone but of a person-in-a-context. The example they give is the counting skill that a young child can construct for the express purpose of setting knives and forks around a table.

The Fischer skill theory was considered relevant to analyses of cognitive growth that could take into account the activities which students experience in a classroom: "Because skill theory connects organism and environment instead of separating them, it transforms much traditional wisdom about the application of cognitive developmental theory to education" (Bidell & Fischer, in press).

Table 2·2 summarizes the main features of the theory with examples for each of the levels he distinguished.

Three tiers and four levels. Fischer identified three tiers in cognitive development, namely, sensorimotor, representational and abstract, and he specified four levels within each tier, namely, single set, mapping, system, and system of systems, as detailed in Table 2·2. In a similar fashion to Biggs and Collis (1991, p. 67), he saw the fourth level of a tier as the first level for the next highest tier, e.g., level 7, the system of representational systems is the level of single abstract sets which "subsume the representational and sensory-motor [*sic*] sets from earlier tiers" (Fischer, 1980, p. 487). He located levels relative to stages in schooling, e.g., level 6, representational systems, at grade (or primary) school; level 7, single abstract sets, in early high school; and level 8, abstract mappings, in late high school.

Table 2·2

Fischer's Levels of Cognitive Development

LEVEL[a]	DESCRIPTION	AGE [a,e] first seen	EXAMPLE
1 Sensori-motor 1	Single sensorimotor action	few months	look at toy OR grasp toy [b]
2 Sensori-motor 2	Mappings of sensorimotor actions	middle of first year	look at toy in order to grasp it [b]
3 Sensori-motor 3	Systems of sensorimotor actions	11-13 mths	attention across experiments, e.g., vary the way pieces of bread are dropped and watch outcomes [b]
4 Sensori-motor 1/ Represent-ations 1	System of sensorimotor systems OR single representational set	early pre-school years 18-24 mths	pretending that a doll is walking [c]
5 Represent-ations 2	Representational mappings	late pre-school years 4-6 yrs	pretending that two dolls are Mum and Dad interacting [c]; ordering on one dimension, e.g., weight [b]
6 Represent-ations 3	Representational systems (also called concrete operations) [c]	primary school years 6-8 yrs	pretending that two dolls are Mum and Dad as well as a doctor and a teacher simultaneously [c]; calculation and explanation of concrete arithmetic problems, e.g., 9 + 7 = 16 [d]
7 Represent-ations 4/ Abstrac-tions 1	Systems of representational systems OR single abstractions (also called formal operations) [c]	early high school 10-12 yrs	general definitions of arithmetic operations [d]
8 Abstrac-tions 2	Abstract mappings	late high school 15-16 yrs	general relations of two similar arithmetic operations, e.g., + - [d]
9 Abstrac-tions 3	Abstract systems	adulthood 19-21 yrs	general relations of two dissimilar arithmetic operations, e.g., + ÷ [d]
10 Abstrac-tions 4/ Principles	Principles OR systems of abstract systems	adulthood 24-26 yrs	principles underlying the four arithmetic operations [d]

[a] Rose & Fischer (in press), p. 12. The "previous" numbering of levels from 1 to 10 is given together with the "current" labels for levels. [b] Rose & Fischer (in press), pp. 75 - 83. [c] Fischer & Lamborn, 1989, p. 40. [d] Fischer, Hand & Russell, 1984, pp. 48 - 50. [e] Fischer, 1980, p. 522.

Fischer symbols. Fischer developed a system of symbols for designating different aspects of his skills theory. For instance, he used bold-faced capital letters to designate sensorimotor sets and italic capital letters to designate representational sets. The symbols used for the first seven levels are summarized in Table 2·3, which appears as Table 3 in Fischer (1980, p. 490).

Table 2·3
Symbols for Fischer's Sensorimotor & Representational Levels of Skills

Level	Name of structure	Sensorimotor sets [a]	Representational sets	Abstract sets [b]
1	Single sensorimotor set	$[^{1}\mathbf{A}]$ or $[^{1}\mathbf{B}]$		
2	Sensorimotor mapping	$[^{2}\mathbf{A} - {}^{2}\mathbf{B}]$		
3	Sensorimotor system	$[^{3}\mathbf{A}_{G,H} \leftrightarrow {}^{3}\mathbf{B}_{G,H}]$		
4	System of sensorimotor systems, which is as single representational set	$\begin{pmatrix} ^{4}\mathbf{A}^{R} \leftrightarrow {}^{4}\mathbf{B}^{R} \\ \updownarrow \\ ^{4}\mathbf{C}^{R} \leftrightarrow {}^{4}\mathbf{D}^{R} \end{pmatrix}$	$\equiv [^{4}R]$	
5	Representational mapping		$[^{5}R - {}^{5}T]$	
6	Representational system		$[^{6}R_{J,K} \leftrightarrow {}^{6}T_{J,K}]$	
7	System of representational systems, which is a single abstract set		$\begin{pmatrix} ^{7}R^{E} \leftrightarrow {}^{7}T^{E} \\ \updownarrow \\ ^{7}V^{E} \leftrightarrow {}^{7}X^{E} \end{pmatrix}$	$\equiv [^{7}\mathbf{E}]$

Note. After Fischer, 1980, p. 490.
[a] Sensorimotor sets continue after Level 4, but formulas become so complex that they have been omitted. To fill them in, simply replace each representational set with the sensorimotor formula for Level 4. [b] Development through the abstract tier shows the same cycle as development through the sensorimotor and representational tiers. Abstractions are built from representational and sensorimotor sets in the same way that representations are built from sensorimotor sets.

Whereas in some recent approaches to cognitive development (such as Case, 1987b, and Pascual-Leone, 1987) large-scale changes in development are accounted for in terms of the number of items in short-term store or working memory, "for skill theory they entail a fundamental change in the organization of behavior" (Fischer & Pipp, 1984, p. 53). Skill theory does not assume that once a person reaches a certain stage or level then that person has most of the skills for that level. Rather, skill theory refers to an "optimal level" which "specifies the upper limit on the complexity of skill that an individual can control" (Fischer & Pipp, 1984, p. 47). Below that limit, behaviour varies widely across levels and, for a person to function at the upper limit, environmental factors were considered to play an important role in either inducing or supporting the higher mode. In one experiment (Fischer & Pipp, 1984, pp. 55 - 57), the importance of environmental factors was highlighted. It was found that a practice-and-support condition provided the greatest opportunity for optimal performance, in contrast to the spontaneous condition in which subjects were tested without any associated help. Optimal performance could improve and the emergence of a new developmental level, characterized by a new and qualitatively different type of skill structure, was indicated by a cluster of spurts in optimal performance evident across a wide range of domains. The upper bound affected only modestly most of the systematic changes in skills but, taking this into account, skill theory offers five transformation rules in a method of analysing the processes of skill acquisition below the upper limit.

Five transformation rules. Fischer (1980) detailed five transformation rules which specified how less-complex skills are transformed into the more complex. These rules were particularly helpful in teasing out the reasons for fine changes in difficulty levels of some of the algebra tasks used in the present study. In order of complexity they are: substitution, differentiation, focusing, compounding and intercoordination. Only the last was classified as describing a combination of skills that produce macrodevelopment, that is, development from one level to the next. If any skill combination is to occur, a co-occurrence of at least two skills must be experienced. Furthermore, as Fischer and Lamborn (1989) pointed out, "the likelihood of co-occurrence is affected by many factors, including task, optimal level, environmental support, and emotion" (p. 45).

Substitution means putting a new element into some previously established structural relation. An example would be transforming the skill of driving one model of car to driving a similar but different model. The Fischer (1980, p. 501) diagram for substitution at a specific level (Level 5 in the example) is of the type:

Sub $[{}^5T — {}^5P\,] = [{}^5T — {}^5P_1]$, where ${}^5P_1$ is the substitute set.

Differentiation is a by-product of one of the other transformation rules and involves taking a previously single element and separating it into two aspects which are then controlled by the subject. Depending on which other transformation is involved, differentiation can be either microdevelopmental (within a level) or macrodevelopmental (across levels). Consider the case of the child working with a "gadget" consisting of a weight hanging on the end of a cord which goes over a pulley at the edge of a table and is connected by means of a spring to a fixed point on the table. An important step in understanding relationships within the gadget is to differentiate between the two partial lengths of the cord, namely, that part which is horizontal and the other part which is vertical. Differentiation of a specific set is designated:

Diff $C = C_H, C_V$. (Fischer, 1980, p. 502)

Focusing is concerned with a moment-to-moment shift of attention between skills without involving relations between skills (Rose & Fischer, in press). A teacher or researcher can change the focus engaging a student by directing attention to one aspect of a problem or another, but it is when shifts in focus are controlled by the student that a new skill is learnt, and this step is transitional to forming a higher compounded skill. Fischer uses the "greater than" symbol to represent shifts in focus, so that a shift in focus from skill '*e*' to skill '*f*' which can be consistently controlled by the child is represented as:

Foc $(e, f) = [e > f]$. (Fischer, 1980, p. 500)

Compounding produces relatively large developmental steps within a developmental level and is the most important form of transformation within a level whereby two or more skills are combined to form a more complex skill. For example, a learner-driver advances from a shift-of-focus use of the brake pedal and the accelerator pedal to the compounded skill whereby the driver focuses on the motion of the car and uses either the brakes or accelerator as needed. Fischer (1980, p. 499) uses the addition symbol to signify compounding, as in:

$[{}^5T — {}^5P\,] + [{}^5P — {}^5R\,] = [{}^5T — {}^5P — {}^5R\,]$.

Intercoordination is the mechanism of transformation between levels and is the process of combining skills at one level to form a skill at a developmentally more advanced level. Fischer maintains that the process "is gradual and continuous" (1980, p. 498). Some object or event in the environment induces the person to relate two skills to each other after they had been functioning independently. The person gradually intercoordinates the skills after working out the relationship between the two skills with that object or event. Rose and Fischer (in press) give the example that a complex compounding of the skill of driving a car would be learning to drive an 18 - wheel truck. Fischer uses a multiplication symbol to designate intercoordination. For instance, the child using "the gadget" may understand how the vertical and horizontal lengths of the cord change and even how the changes relate to each other, but it is not until the constancy of the total length is recognized that intercoordination is achieved. This was depicted by Fischer (1980, p. 498) as follows:

$[{}^5C_H — {}^5C_V\,] \cdot [{}^5C_V — {}^5C_H\,] = [{}^6_1C_{H,V} \leftrightarrow {}^6_2C_{H,V}]$.

Collis (1986) explained that "Fischer makes the point himself that these [transformation] rules are probably not exhaustive and that future research will show whether (and what) additional rules are required" (p. 7). Moreover, Fischer (1980) warned that it is not an easy matter to apply the theory to a new skill domain, because it requires a careful descriptive analysis of the specific skills that are to be developed in that domain. Nevertheless, these transformation rules provide a system for examining and coding the mechanisms of changes in cognitive development. Applications to aspects of the process of developing an understanding of early algebra are included in Chapter 8 of this thesis.

Previous Research on Algebra

Much of the groundwork for this research project was laid by researchers over the past two decades. The following pages describe and discuss research projects which produced findings that were useful in determining which aspects of cognitive difficulty would be included in the study and that were a source of useful ideas for measuring students' levels of development of an understanding of algebraic symbols.

Collis

Collis' 1972 project. Collis (1972) reported findings from a series of research projects carried out in the Hunter Valley District of N.S.W. The main objective was to try to define the meaning of concrete and formal operational thinking in the context of elementary mathematics. Items were devised which controlled the level of abstraction in the operations required and in the elements to be operated on, as well as the number of operations to be used. The level of abstraction in the elements was raised by using large numbers or algebraic symbols.

The preliminary experiment involved 101 girls aged from 8 to 17 years and provided evidence that the level of abstraction in the operations or relationships had more influence on the degree of difficulty of items than did the level of abstraction of the elements involved. It was found, too, that it was not until the later years of secondary school that students were able to work within unfamiliar mathematical systems.

In a major follow-up study in the same series, the subjects were 30 children from each of eleven age groups, from 7 to 17 years, with approximately equal numbers of boys and girls. It was confirmed that operations were the prime cause of difficulty and that there was an interaction with the types of elements. Students at the concrete operational level could appear to be thinking at a formal level until they were asked to work with unusual operations or elements. Collis (p. 159) pointed out that it was possible that students might regard a teacher as a "good" mathematics teacher when what the teacher was doing was using techniques which effectively allowed the students to avoid working at the formal, abstract level.

Another study of similar magnitude in the 1972 series probed the effect of setting items which specifically required the subjects to work within an unfamiliar system, such as working with an operation, '*', defined by the statement

$a * b = a + 2 \times b$.

It was found that only those students who were able to work formally with abstraction achieved satisfactorily on items which required them to work with just one defined operation and found items involving two operations more difficult. Many students showed that they were not working at the formal level by reverting to the properties of familiar operations such as addition and were unable to work with the defined operations.

Yet another major study within the 1972 project verified that students younger than about 10 years are unable to cope with using more than one operation even when closure was readily available. Between about 10 and 14 years, students were able to work with two operations "unless a doubt is raised concerning the possible uniqueness of part or all of the expression by the introduction of either a non-numerical element or an unfamiliar operation" (p. 213). It seemed that true formal operational ability generally appeared only after about 14 years.

<u>Collis' 1974 project</u>. A later study by Collis (1975a) in England in 1974 identified age-related differences in the meanings children give to alphabetic symbols in algebra. The research strategy involved testing 180 children spread across the ages 10 years to 15 years, and interviewing one-third of them. The test was in three parts, each with six questions. The first twelve of these gave the participants two or three equations and asked them to derive the relationship between some of the letters.

e.g., find the relationship between x and y

given $k = x + a$, and $y + a = k$, (40% correct)

or given $3x = a$ and $a + 3y = 180$ (15% correct).

The final six questions asked participants to decide whether given statements were true always, never or sometimes, and if they chose "sometimes" they had to say when they were true. Sample statements were:

$a + b = b + a$, $m + n + q = m + p + q$ and $a + 2b + 2c = a + 2b + 4c$.

Some of these were used earlier to illustrate the categories of response types according to the SOLO taxonomy of Biggs and Collis (1982). In this section, the focus is on the meanings that the students gave to the symbols in the algebra presented to them, and, incidentally, indicates why this style of question was included in the present research project.

The younger students generally regarded each letter as standing for only one number and tended to use just one trial value in a problem situation. The middle age groups tended to substitute a variety of numbers as possible values for letters, in the hope that they might find the "right" value appropriate to the problem in question. As they were using this trial and error approach, it seemed that each value stood alone as a specific meaning for the letter under scrutiny. Only the older students showed an understanding of a letter as a generalised number, that is, as an entity in its own right representing a class of numbers without requiring a specific value to be stated. A small proportion of participants worked with letters at a yet higher level, viewing them as variables. For example, 25% were able to conclude that '$n = p$' was the condition for making '$m + n + p = m + p + q$'. Here, one needed to be aware that each of the letters stood for a great range of numbers and that it would be possible for both 'n' and 'p' simultaneously to have the same values.

The lessening of the student's reliance upon empirical reality to support cognitions was also registered. This showed up as a development of the degree of tolerance for lack of closure with mathematical items. At the lowest operational level, the student needed to replace two elements connected by an operation with a third element. Later, the outcome of an operation could be regarded as necessarily unique but the student no longer needed to make the actual replacement with a third element. The stage of concrete generalizations was reached when the ability to refrain from closure became general, on the condition that a unique result was obtainable at any time, if required. When closure was accepted in the formal sense, the student showed the ability to work on operations themselves without the need to relate the elements or the operations to a physical reality, allowing the person to deal with variables as such and work on relationships rather than unique results (Collis, 1975a, pp. 5 - 6).

Collis was able to gather evidence that the children had difficulty with questions that required them to coordinate two skills, even when closure was a possibility at each step. Test questions were used which asked students to identify arithmetical operations and to then use that information to evaluate some unknown. In the example question below, part (b) showed up the two-stage difficulty with the students succeeding more readily on the other two parts. Part (c) was adapted for use in the present study. The success rates were 63.3% for part (a), 50.0% for part (b), and 78.3% for part (c).

2. (a) $8 * 2 = 11 * 5$ $\quad * = ...$
(b) If $15 * 5 = 9 * 3$ and $8 * 2 = a * 3$ find a
(c) Look at this statement $\quad 3 * 4 = 6 * a$

If this statement is true, then (There may be more than one correct answer)
* could be x yes ... no ... can't tell ... If 'yes' then a must be ...
* could be + yes ... no ... can't tell ... If 'yes' then a must be ...
* could be + yes ... no ... can't tell ... If 'yes' then a must be ...
* could be - yes ... no ... can't tell ... If 'yes' then a must be ...
(Collis, 1975a, p. 121).

Harper

The research completed by Harper in 1979 contributed greatly to preparing the way for the present study. He analysed in depth the range of levels of meaning given to algebraic symbols by students across the ages of 11 years to 18 years. His study deserves close consideration because it is so relevant to this thesis. Several of his research tasks were used in the present study because of the potential they possess for contributing data to further the understanding of the cognitive challenges faced by beginning algebra students.

Harper (1979) was concerned with the question of why so many students fail to learn mathematics. He proposed that many of them had difficulties due to the nature of mathematics as a system of language systems, in which each system has its own reality, harbouring its own meaning for key concepts. He concentrated on the language of algebra and declared that the main aim of his study was to attempt to explain the statement quoted earlier on page 4 that "the frontier between arithmetic and algebra is considered to have been crossed when the learner adopts the usage of the letter as a 'numerical variable' " (p. 3).

He carried out a total of 144 interviews of students in two schools in England so that he could assess the views they had of the meaning(s) for algebraic symbols in the form of letters. They were spread across the range of secondary school years with average ages from 11 years 9 months in Year 1 to 17 years 3 months in Year 6 (Harper, 1981, p. 173). He restricted his sample to those who could correctly answer an introductory task which tested whether the students had the necessary arithmetical ability to deal with equations involving more than one letter and who understood the meaning of terms such as "the value of x" in that context.

Four types of tasks were devised, with subtasks in each of three of them. The range of meanings given to algebraic symbols in the responses was classified into three levels as follows (Harper, 1979, p. 244):

1. *fictitious measures* - the letter as an object with a unique, unknown content, thus allowing "null" or "concrete" variation;

2. *discovered content* - the letter as a pigeon hole or box for numerals, allowing some potential for variation; and

3. *species* - the letter regarded as a symbolic variable, or variation in itself.

These classifications become clearer when examples are considered from the student responses to the tasks he set.

Harper's Parallel Lines Task. In three subtasks, pairs of parallel lines, one red and the other green, were labelled with lengths in centimetres given as numbers or as algebraic symbols (e.g., as '*b* cm' and '*a* cm' for Subtask 3). Participants were asked questions about which line was longer, and when one could be longer than the other or equal in length to it.

Those responding at the fictitious measures level regarded the letters as temporary substitutes for the numerals which would be obtained if the lines were measured, and so did not allow the values of the letters to change in any real sense. At the next level, that of discovered content, there was a generally inconsistent use of the letters as organizers of perception, whereas, at the level of species, the students were able to transcend any ordering suggested by the geometry of the figure and did not see the content for the letter as given by the outcome of applying a measuring process to the line. In relation to this task, Harper (1979, p. 186) highlighted the importance of being able to consider algebraic symbols as naming a variety of numerals and also as useable as a unique entity, thus showing an ability to transcend geometric orderings.

He reported that performance improved with the year levels. Taking Subtask 3, which was used as an item in the test instrument devised for the present study, the percentages who used the species concept at each year level were:

8.3 (Yr.1), 20.8 (Yr.2), 29.2 (Yr.3), 29.2 (Yr.4), 33.3 (Yr.5), 83.3 (Yr.6).

The outcomes indicated that the strong geometric features of the tasks effectively distracted many students from the implications of the algebraic descriptions of the lengths of the lines. Many found it difficult to consider that the lines were merely sketches of two lines as examples taken from an unlimited assortment of possible lines. The variability of length built into the algebraic symbols eluded them. They attempted to introduce variability in terms of perspective within the sketches, or by imagining that the lines might be leaning towards the viewer at one end. The tasks were successful in fulfilling the purpose Harper intended by the Parallel Lines Task, namely, to decide which pupils were allowing variation in a geometrical setting by using the species notion to organize perception.

The other tasks he set will now be considered, keeping in mind their purpose in Harper's view, which was to study pupils' interpretation of letters in a non-geometrical setting, particularly with regard to their readiness to allow variation.

<u>Harper's Equations Task</u>. Given different equations relating 'x' and 'y', participants were asked questions about the relative sizes of 'x' and 'y' (e.g., in equation '$2x + y = 9$', which was used in Subtask 2).

At the lowest level, the level of fictitious meaning, students displayed the following tendencies: (a) They assumed that each letter had an indeterminate content; (b) they thought that the letters had an ordering dictated by the coefficients in an equation; (c) they did not allow the value of the letter to change in any real sense and so manufactured variation by using operations on the letters or changing the actual context, such as changing the given coefficients; (d) they sometimes imagined that they had found the true content of the letters; (e) they were restricted in their substitution strategies; (f) they wavered between different response modes for different equation types; (g) they appeared to see indeterminate equations as collections of individual units, each with a fixed content.

At the level of discovered content, students treated the equations either as statements including boxes into which numerals could be posted or as covarying systems. Those who suggested a number of integer solutions gave no indication that any strategy is involved other than a direct replacement of the letter by a number, much as one might place objects in a box one by one, removing each one prior to replacing it with a second. This revealed sporadic rather than fluid thinking about the variables in the equation. Whole numbers were each welcomed as a possible replacement, however, without any false ordering or preference. The implications for covariance given by the equations were taken into consideration, and so these students avoided number pairs which were unable to make the particular equation true.

The species level was reflected in an ability to work with a system of covarying numeral pairs with a natural usage of negative numbers and the inclusion of the possibility of fractional values. The proportion of students who operated at this species level increased according to the year level, the response percentages being:

8.3 (Yr.1), 15.3 (Yr.2), 20.8 (Yr.3), 29.2 (Yr.4), 47.2 (Yr.5), and 72.2 (Yr.6) for all three Subtasks taken together. For Subtask 2, which was incorporated into the assessment test for the present research project, the percentages were:

8.3 (Yr.1), 16.7 (Yr.2), 25.0 (Yr.3), 33.3 (Yr.4), 50.0 (Yr.5), and 83.3 (Yr.6) The data indicated that these questions were successful in discriminating between students according to their view of the letters in the given equations. The tasks were sufficiently difficult to enable progressive improvement in performance across the years of secondary school to be registered.

Harper's Literal Numbers Task. In this task, participants were asked to compare two algebraic expressions (such as '$t + t$' and '$t + 4$', in Subtask 1) and to define conditions for them to be equal or for one to be greater than the other. The purpose of the task was to find the extent to which students were prepared to accept a letter in itself as a non-ordered entity.

It was found that some of those at the fictitious level of meaning for letters used some type of false ordering which was only sometimes corrected during the interview. Students at this level often exhibited a desire for ordered content, such as an ordering determined by an alphabetic code; appeared to see literal numbers as objects with fixed (undetermined) content; did not regard repeats of the same letter as necessarily having the same value within a particular question; had little respect for the formal syntax of the algebraic language; and were prone to errors such as claiming that '$t + 4 = 4t$'. They appeared to find security by dealing with an ordered world of static entities with fixed measures and found it very difficult to accept that more than one ordering could exist at any particular instant (Harper, 1979, p. 146). They did not use a matching strategy to decide the comparative literal or numerical content for a letter, as this seemed to be inconceivable or of no use in mathematics given that the letters had unknown values. Thus, in Subtask 1, they did not compare 't' with the value '4' and in Subtask 2 in which they were asked about relative sizes of '$m + m$' and '$m + k$', they could not see the sense in comparing 'm' and 'k'. Harper denigrated the stance that allows letters in algebra to be regarded as objects, a viewpoint which is examined later in this thesis:

> Such responses suggest that the pupil continues to look upon 'letters' as objects (such as oranges and pears) which have a fixed content (analogously a fixed mass) which cannot be known. Whatever 'variation' exists here would appear to be a 'hypothetical variation'.
> (Harper, 1979, p. 150)

Beyond this lowest level, participants were able to match strategies across literal numbers to obtain a literal or numerical content for a letter. Such responses were either at the true level of species where a relationship applied for all numerals, or at the level of discovered content where the relationship was true for some (unspecified) member of a collection of numerals, in the sense of the classical, Diophantine unknown. Diophantus considered unknowns as exclusively conventional numerals (Harper, 1987, p. 82) and, while aware that some problems had general solutions, "he cannot *express* a general solution" (Harper, 1987, p. 87). The Diophantine view could be shown by the strategy of giving multiple substitutions. Harper found that with this literal numbers task alone he could not identify students using the species notion. For instance, the claim that '$k = m$' could be based on meanings for the letters that are either the classical Diophantine unknown or the species concept, and it is difficult to distinguish these here. Students at the discovered content level, if not the species level, did not show false ordering or false content and were prepared to impose relationships between letters and numerals.

For all three subtasks, there were 24.8% of responses at the fictitious measure level, 17.8% at the discovered content level, and 57.4% at a more algebraic level. The corresponding percentages for Subtask 1, which was used in the test designed by the writer, were: 27.1%, 12.5% and 60.4%. All Year 6 students used the higher algebraic approach to each of the three subtasks but the same was not true about any other Year group. These outcomes suggest that the subtasks were successful in distinguishing those who regarded the letters as non-ordered entities from those who did not. Harper noted that all students who transcended ordering in the Parallel Lines Task gave algebraic responses to the first two subtasks in the Literal Number Task.

Harper's Zetetic Task. In this task, the problem presented to the participants was "a reformulation of a problem first posed and solved by Diophantus, and later by Vièta using his more sophisticated language system" (Harper, 1981, p. 172). The title, no doubt, was derived from the Zetetica of Vièta "in which the author treats in his own way a large number of Diophantus' problems" (Heath, 1910, p. 27). The problem was:

"If you are given the sum and the difference of any two numbers, show that you can always find out what the numbers are. Make your answer as general as possible" (Harper, 1979, p. 84).

The data confirmed that there is an age-related transition in approaches from rhetorical to Diophantine to Viètan (Harper, 1981, p. 173). As explained in Appendix 3F, a question of this type was found, after trialling, to be not efficient enough for inclusion in the final test instrument.

Relevant outcomes. Harper's study highlighted the importance of the student's understanding of the symbols of algebra, which is the focus of the writer's research project. Harper found that the nature of the tasks he used caused pupils to switch between different interpretations of the letter and that, in each Year-group, there were students at each of the levels. His data showed that the majority of pupils completed secondary school studies without appreciating the symbolic conception of number, and that "many pupils complete their mathematical studies at the fifth-year level devoid of any clear understanding of the language of algebra" (1979, p. 380).

He made the point that just as pre-Viètan mathematicians had difficulty making headway with algebra, so students would find the same difficulty until they accommodated the symbolic number concept. In so emphasizing the importance of the interpretation of the letter, he questioned Collis' view (1972) that the degree of success with an algebra problem depends mainly on the number of operations involved, and the view of Brown and Küchemann (1976, 1977) that it depends, in an arithmetic setting without letter-symbols, on the type of operation involved. For example, Harper's Equations Subtask 3 involved one operation (in '$5x = y$') yet was found harder than Subtask 2 which involved two operations (in '$2x + y = 9$'). It should be noted, however, that the concept of negative numbers needs to be applied in Subtask 3 to identify when 'x' is more than 'y'. This could be the additional, and telling, factor not allowed for by concentrating only on the number and type of operations involved in a question.

Harper's expectations regarding the acceptance of lack of closure extended those envisaged previously by Collis (1975a). Firstly, according to Harper, such acceptance opened the way to the view that the letter had a potential variation across a range of numerical values. Secondly, in order to acquire the concept of a variable, this acceptance needed to extend to regarding the letter as a non-ordered entity which could identify all possible numerals simultaneously and as useful to avoid mentioning all these numerals.

He recommended that classroom models should be devised for developing the species concept, especially that the letter is introduced intentionally to save the labour of mentioning numerals and does not just stand in place of a numeral or a measured outcome.

Küchemann

Küchemann (1980) analysed the data on algebra obtained in the CSMS research program described on page 9 above. As statistics regarding success rates on questions from this research were given in Chapter 1, the focus here is on clarifying the levels of thinking revealed by students who responded either correctly or incorrectly to test items. Küchemann's (p. 49) classifications of children's differing interpretations of letters are used again here. These six categories were developed from Collis' findings in his 1974 project (Küchemann, 1984, pp. 114, 115). and Collis' ideas "formed an important basis for constructing the [CSMS] Algebra test" (Küchemann, 1978, p. 23).

1. Letter evaluated. This category applied to responses where the letter was assigned a numerical value, for instance, by using the 'alphabet code' (whereby an answer '35' was given for the area of a rectangle of dimensions 'e + 2' by '5' by assigning the value '5' for 'e'). Other items included in this category asked for a numerical value but it was not necessary to manipulate any letters first.

e.g., "What can you say about a if $a + 5 = 8$?"

"What can you say about m if $m = 3n + 1$ and $n = 4$?"

2. Letter not used. Here the letter was ignored or kept without giving it a meaning, as when children wrote '$7n$' or just '7', instead of '$3n + 4$' when they are asked to "Add 4 to $3n$". They could succeed at this level with items such as "Add 4 onto $n + 5$" by leaving 'n' untouched and simply adding the '4' to the '5'.

3. Letter as object. This label covered misunderstandings of the use of alphabetic symbols in algebra whereby they were regarded as a shorthand for the names of objects or as objects in their own right, such as thinking of '$2a + 5a$' as "2 apples and 5 apples" giving a total of "7 apples" or simply as "2 a's and 5 a's, which made 7 a's altogether". Thinking of letters in this way enabled some expressions to be successfully simplified but at other times it was quite inappropriate, for instance, when the letter was meant to represent a number of objects as in "a apples", and not the

object itself.

4. Letters as specific unknowns. Here the letters were regarded as unique but unknown numbers, and students showed that they could operate on them without using specific numerical values.

e.g., "Multiply $n + 5$ by 4."

"Part of the figure [a sketch was given] is not drawn. There are n sides altogether all of length 2. Write an expression for the perimeter."

5. Letter as generalized number. In this type of question, the possibility that a letter could take more than one value needed to be considered. These were found to be very difficult.

e.g., "What can you say about c if $c + d = 10$ and c is less than d ?"

"$L + M + N = L + P + N$ is Always Sometimes (say when) Never true."

6. Letter as variable. Students who had attained the level of viewing algebraic symbols as variables realised that they represented a range of unspecified values, and these students were able to identify and work with systematic relationships between sets of possible values. The variable notion was used when answering the item "$a = b + 3$. What happens to a when b is increased by 2 ?" if the relationship between 'a' and 'b' is interpreted as "a is always 3 bigger than b", rather than "this particular a is 3 bigger than this particular b". Küchemann argued that when a second (or higher) - order relationship was established between expressions, the respondent showed ability with the concept of variable. The only CSMS item he admitted to this category was:

"Which is larger $2n$ or $n+ 2$? Explain."

Levels of understanding. Küchemann (1980, 1984) selected 30 of the 51 test items and spent a lot of time and effort in sorting them into four levels of understanding. A condition for being grouped in any one level was high correlation between items. Items at Levels 1 and 2 were those that could be solved by evaluating the letters or regarding them as objects without having to operate on letters as unknowns. The degree of complexity involved in answering items distinguished Level 2 from Level 1. For items at Levels 3 and 4, the letters had to be treated as specific unknowns, generalized numbers or variables. Again, the complexity associated with the items distinguished between these two levels. The levels were compared with Piagetian stages on the basis of general descriptions of Piaget's stages, Collis' research findings in the area of mathematics, and empirical evidence from the comparison of performance on the algebra items with performance on one of the Piagetian tasks as developed and administered by the science wing of the CSMS study. The correspondences were given as:

Level 1	Below late concrete
Level 2	Late-concrete
Level 3	Early-formal
Level 4	Late-formal
	(Küchemann, 1981, p. 117)

Examples of items in each level are given in Table 2·4, the levels being appropriate for responding correctly to the items.

Table 2·4

Levels of Cognitive Difficulty for Samples of CSMS Algebra Items

LEVEL	ITEM	Küchemann's COMMENTS
1	What can you say about a if $a + 5 = 8$?	Letter evaluated
1	$2a + 5a =$	Letter as object
2	What can you say about m if $m = 3n + 1$ and $n = 4$?	Letter evaluated but need to cope with (temporary) ambiguity of $m = 3n + 1$
3	4 added to n can be written as $n + 4$. Add 4 onto $3n$.	Letter as specific unknown. n has to be operated upon, not avoided (as in response $7n$) or ignored (as in response 7)
3	What can you say about c if $c + d = 10$ and c is less than d ?	Letter as generalized number
4	Is the following always, never or sometimes true? $L + M + N = L + P + N$	Letter as generalized number. M and P can represent a range of values, which may coincide
4	Multiply $n + 5$ by 4	Letter as specific unknown. Here it is necessary to coordinate two operations and to recognise the ambiguity of an answer like $4 \times n + 5$
4	Which is larger, $2n$ or $n + 2$? Explain	Letter as variable (2nd order relationship). As n changes the difference between $2n$ and $n+2$ changes so for some value of n, $2n$ may be less than $n+2$.

Note. From Küchemann, 1980, pp. 64 - 69.

Of the items listed in Table 2·4, all but the first three were adapted and incorporated into the test instrument constructed by the writer for this project.

It may be noted that Küchemann classified the item about the relationship

$$L + M + N = L + P + N$$

as involving an understanding of algebraic symbols as "generalized numbers". Collis was inclined to go further when he used the case of

$$m + n + q = m + p + q$$

and stated that, to succeed on this type of question, students needed a level of understanding that was at least that of "generalized number" but he suspected that the question was "beginning to tap the level of abstraction where the term 'variable' might be used rather than 'generalized number' " (Collis, 1975a, p. 47).

Incorrect answers to CSMS items were also classified by Küchemann into the four levels and it was found that responses to some items spread across several levels. An example of an item yielding responses spread across all four levels was the item "Multiply $n + 5$ by 4". The incorrect response '20', indicating that the letter was not used, was classified as Level 1, while the similar non-use of the letter to give the response '$n + 20$' was classified as a Level 2 response. A Level 3 response was '$4 \times 5 + n$', which was incorrect but treated the letter as a specific unknown. The correct answer, '$4n + 20$', was regarded as a Level 4 response.

Students were described as being at a particular level if they correctly answered about two-thirds of items at that level but not at a higher level. Only 34 children out of the 2 854 participants were unable to be classified into levels according to this criterion. Table 2·5 gives the percentages of students at each level, excluding the unclassified 34 children, and includes in brackets the cumulative percentages which tally the percentages of students at a given level or above that level.

Table 2·5

Percentage Frequencies for Levels Based on CSMS Algebra Items

LEVEL →	0	1	2	3	4
CLASS (AGE) ↓	PERCENTAGES				
2nd Yr. (13 yrs.)	10 (99[a])	50 (89)	23 (40)	15 (17)	2 (2)
3rd Yr.(14 yrs.)	6 (99[a])	35 (93)	24 (58)	29 (34)	6 (6)
4th Yr. (15 yrs.)	5 (99[a])	30 (93)	23 (63)	31 (40)	9 (9)

Note. From Küchemann , 1980, p. 75. Cumulative Frequencies in brackets.
[a] It is unclear why 100% was not given here.

In the summary section, under the heading "Implications for Teaching", Küchemann (1981, p. 118) stated that the majority of the students in the study were not able to cope consistently with items that can properly be called algebra, meaning items where the use of letters as unknown numbers cannot be avoided. They seemed to be at the Piagetian stage of concrete operations. To alleviate the problems, at least for students at Levels 1 and 2, ways of making algebra more plausible were considered, such as using concrete embodiments and inducing conflict. However, after studying some specific examples of such tactics and admitting that he had a scarcity of information, Küchemann (1980, p. 191) expressed the view that these approaches were unlikely to eliminate entirely the mismatch between the students' level of understanding of algebra and the cognitive demands of what they were being taught.

Booth

Booth (1983) conducted an investigation into the reasons underlying particular errors in elementary algebra which had been shown by the CSMS study to be prevalent among 13 to 15 year-old children. The research proceeded in three phases: the conceptual bases for errors were explored in 72 interviews, interaction with a remedial program was monitored for three small groups, and the effects of a teaching program were analysed for eleven whole classes.

The following difficulties were identified in the interviews:

1. Letters were treated as objects rather than as representing numbers.

2. Regarding letters as generalized numbers seemed to go against a natural tendency to interpret letters as standing for specific numbers.

3. Different letters meant different values (e.g., '$y = p$' was often disallowed in relation to the Collis-type question about the equality of '$x + y + z$' and '$x + p + z$').

4. Informal methods were applied which did not lend themselves easily to symbolization, sometimes because of the arithmetical methods used.

5. Algebraic symbolization was not readily accepted as appropriate.

6. Conjoining was used in cases such as '$a + b$' to produce 'ab' as an "answer", maybe because '$a + b$' was viewed as a description of an operation rather than an acceptable answer.

7. The need for brackets was not appreciated, perhaps because the context was considered sufficient to define the order of operations for a particular case.

The small groups teaching phase lasted the equivalent of six 35-minute lessons and there were two follow-up tests, one given immediately after the teaching and the other given four months later. A computer metaphor was used in which there were memory locations for housing values given to variables and students were required to write instructions to tell the computer how to use mathematical functions and to record the machine's output. This approach was chosen so that attention was focused on precision in making explicit the procedure by which a problem was to be solved, and to provide a rationale and a mechanism for using generalized numbers.

Sample questions from Worksheet 4 in this small-scale teaching phase were:

1. Add any number I give to itself.
8. Multiply together any two numbers I give.
15. Multiply any number I give by itself, and then take away another number.

(Booth, 1983, p. 353)

Figure 2·3 shows how the Machine Model was presented on the students' worksheets. The worksheets also provided pairs of rectangles, the left-hand one labelled "Instruction Pad" and the right-hand one "Print-out Pad".

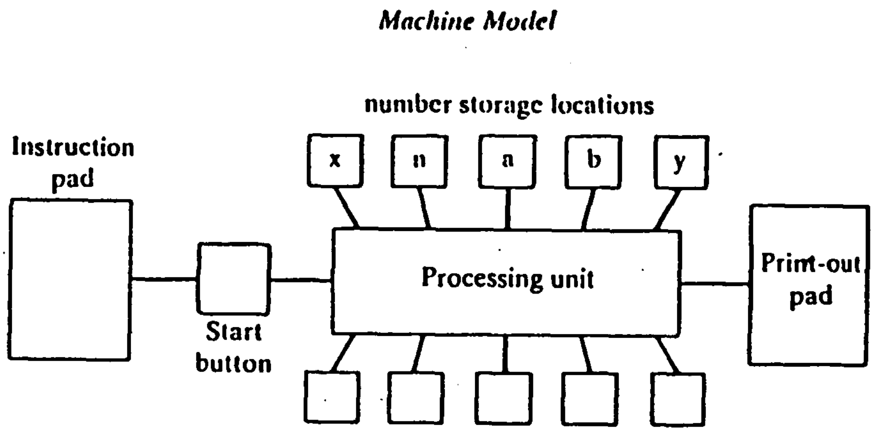


Figure 2·3. Booth's machine model
(from Booth, 1984a, p. 105)

The concept presented for alphabetic symbols was that they were labels for number storage locations and were such that, when used in a computer instruction, they called up the numbers stored in those locations. As Booth (1983) put it, the machine model included "a set of number storage locations labelled by letter call-signs, e.g. '*x*' in an instruction calls up the number currently stored in that particular location" (p. 361). It would seem that from this model students could have developed an understanding of algebraic symbols at Harper's "discovered content" level in which letters were regarded as pigeon holes for numerals. Thus the machine model could have hindered the development of the wider notion of algebraic symbols as numerical variables able to represent simultaneously a variety of numbers.

The small-scale teaching phase proved to be motivating for the students and they could see the sense in using letters to write general "rules" for the machine so that whole classes of problems could be solved. They were surprised that sometimes the print-out "answers" were the same as the instructions. They were prepared to avoid conjoining for addition and also to use brackets within the context of the machine. There was an improvement in their understanding of letters as representing a range of possible values, although they tended to keep to integers and to regard different letters as standing for different numbers. Booth (1983, p. 200) pointed out that students seemed to take different meanings for some expressions she used as the teacher. When told that '*a*' and '*b*' could stand for numbers that were "the same or different", some took the meaning of "or" as exclusive rather than inclusive. Similarly, when it was explained that a letter could stand for "any number", some interpreted this as meaning "any particular one you choose".

Booth conducted whole-class teaching experiments for which she provided detailed teachers' notes and student worksheets which she had trialled with four small classes. The approach included amendments to that used with small groups, in response to the previous findings. These amendments included emphasis on the use of non-integers, the importance of brackets, the meaning of "any" number, the possibility that two different letters could have the same value, and the meaning of notation such as '$3a$'. Class discussion was used to clarify the conventions for writing inputs and outputs and for interpreting any letters used. Seven volunteer teachers participated in the study by teaching the research module, and classes of four other teachers were used as control classes. The main gains were in accepting unclosed, algebraic answers to instructions; avoiding conjoining in algebraic addition; and, for more able groups only, an improved level of understanding of the meaning of letters in algebra. The program did not make much impact on the students' use or non-use of brackets.

Booth found that further difficulties such as the following were substantiated by the research:

1. Some believed there was a "pattern" in the relationship between letters and the numbers they represented, e.g.,

   a) 'x', 'y', 'z' related to 3, 4, 5 or 10, 20 ,30, etc.

   b) 'y' is "higher" than 'p', and

   c) a fixed alphabetical substitution applied.

2. The meaning of the letter was sometimes ignored, e.g.,

   a) Problems such as "Simplify $2a + 5b + a$" were treated as mere manipulation of symbols and "rules" invented to govern manipulation, such as adding up all the numbers, then writing down letters (according to various rituals); and

   b) For several who gave '$8y$' as their answer to the question "Add 3 to $5y$". Booth, through interviews, found that their level of interpretation of the meaning for the letter 'y' had little influence on this outcome as some of them regarded 'y' as an object, others as just a letter, and others as a number.

3. Some did not understand that an "answer" could be an algebraic expression, e.g., "I can't do it because I don't know what the numbers are". This revealed the problem with acceptance of lack of closure. When pushed, they fell back on

   a) measuring (if a possibility) to get a value,

   b) assuming a particular value and hence getting an "answer", or

   c) using alphabetic substitution.

4. There was confusion over the distinction between letters representing numbers of objects or relating to the value of a measure and letters thought of as representing the object or measure itself.

5. Notational convention caused confusion, such as

   a) taking '$4y$' as '$4 + y$', or as '$40 + y$', or '4 y's' instead of "4 times the value of y",

b) ignoring brackets, and

c) making use of conjoining of terms in order to obtain an "answer".

The results of Booth's research suggested "a possible association between acceptance of the notion of generalized number and the attainment of particular levels of reasoning" (Booth, 1983, p. 274). However, maturation of itself did not assure progress with this notion. These findings called for further research. It was thought important in this thesis to include several items to sample performance requiring the notions of generalized numbers and variables in algebra. Booth had used only two such items in her tests. She found that young teenagers were capable of responding to teaching aimed at moving them from informal methods to more formal mathematical methods. The research underlined the nature of many of the challenges for beginning algebra students, such as the differences between arithmetic and algebra, the notation itself and the conventional uses of notation, and the cognitive stance needed for acceptance of lack of closure and for working with generalizations rather than specific numerical cases (cf. Booth, 1984b).

Rosnick and Clement

Research on students' ability to manage the interface between mathematical symbols and verbal descriptions of real-world problems was conducted at the University of Massachusetts in 1980. A common error pattern was identified by Rosnick and Clement (1980) and described as "the reversal misconception" (p. 3). One of their research problems became well-known in the literature and will be referred to as the "professors-and-students problem". It was adapted later for the research instrument used in the present study. The 1980 version was:

> Write an equation using the variables S and P to represent the following statement: "There are six times as many students as professors at this university." Use S for the number of students and P for the number of professors. (p. 4)

Rosnick and Clement found that only 63% of a group of first-year engineering majors correctly answered this question, and that two-thirds of those wrong wrote an equation in reversed form as '$6S = P$'. Further, only 43% of the social science students tested were correct. They interviewed many students, some under a video-camera, to identify the cause of the errors, and found that it was not the specific wording used in a word problem that was the primary cause of the reversal error since the same error was quite common also in translations from pictures and data tables to equations. The interviews revealed that the students had difficulties with the conceptualization of the basic ideas of equation and variable. Some, for instance, thought that the larger coefficient should be associated with the larger variable. The researchers found that some students "directly identify the letter S as a label standing for 'students' rather than making the proper interpretation that *S* means the 'number of students' " (p. 6). Another confounding influence was the misconception that the equals sign meant things like "for every" or "is associated with" rather than "is numerically equal to". They tried to remedy the reversal problem for nine students by pilot tutoring interviews, reporting that at least seven of these students maintained the reversal misconception, which indicated that this misconception was not a superficial one.

The researchers followed up these initial findings by designing a more systematic teaching strategy which focused on the idea that letters in equations are variables that are meant to be replaced with appropriate numbers. Six students enrolled in the first year of a rigorous calculus course were taught by this new strategy and their performance improved. However, the interviewers concluded that their understanding of the basic concepts of equation and variable remained, for the most part, unchanged and made the point that written answers that are correct may be poor indicators of understanding. This comment highlights the role of interviews in research, and gave impetus to the decision to conduct interviews in this present study.

In their conclusion, they stated their beliefs that teachers should not assume that students will develop correct concepts of equations and variables by osmosis, that students should be encouraged to view equations in an operative way, representing active operations on variables that create an equality, and that "it is essential that students be able to view variables as standing for number" (p. 23). They regarded this last conception as "a fairly abstract one and, for that reason, a very difficult one to teach" (p. 23). This aspect of understanding the symbols used in the algebra of generalized arithmetic was a conscious concern in this thesis. Rosnick and Clement were critical of educational programs that concentrated on manipulative skills and urged that more attention be given to conceptual development in mathematics education.

Clement (1982) presented further analyses of the interview protocols collected during the Rosnick and Clement (1980) research project. He clarified the conclusions that incorrect answers generally resulted from word order matching or static comparisons (for example, the symbol with the larger coefficient represented the variable with the larger value), while success was achieved when some hypothetical active operation was invented, such as increasing the number of professors sixfold. He stressed that despite being successfully taught a standard method for a mathematical skill, students may still possess intuitive, non-standard methods that can compete for control.

<u>Follow-up studies</u>. There has been a number of follow-up studies since the Rosnick and Clement 1980 study, each of them adding to the store of knowledge

about the complexities of relating algebraic equations to real-world situations and the difficulties people have with algebraic symbols. Chapter 7 of this thesis reports further insights gained in this area from responses to the writer's 1990 version of the professors-and-students problem.

Lockhead (1980) collected data from 200 academics and 150 teachers and, finding that apart from those in the physical sciences there was an overall success rate on the professors-and-students problem of only about 50%, he urged that serious attention be given to the fact that the interpretation of mathematical statements is fundamentally a confusing process. Rosnick (1981) stressed the need for students to distinguish between letters used as labels and letters used as variables which stood for some number or numbers of things after he had given a multiple-choice form of the professors-and-students problem to about 150 undergraduates and had found that only 60% made correct choices. Rosnick (1982) analysed clinical interviews of ten college students and suggested that students who use a letter inconsistently in a problem may be shifting unconsciously from one interpretation to another. He found that qualitative and multiple quantitative attributes were associated with letters, such as a label for items purchased, the cost of one item, or the total cost of all items. He urged teachers to pay more attention to the way variables are generated and interpreted in problem-solving contexts. Kaput and Sims-Knight (1983) reported that just over 50% of secondary students responded to the professors-and-students problem with the reversal error and deduced from their data that students had difficulty in translating a multiplicative quantitative relationship into an algebraic equation containing two variables, whether the translation was from a representation in words or in images. Dubé (1990) reported that about one-quarter of 240 Grade 12 students in a national high school in Papua New Guinea used a "holistic" (p. 9) approach in which the entire problem was considered as a global entity. The rest used an analytical approach, making use of semantic and mathematical reasoning after breaking the problem into known and unknown aspects. About 40% were correct on the professors-and-students problem regardless of the approach used.

However, attention is drawn especially to the study by MacGregor (1989) for this resumé of the related research.

MacGregor

MacGregor (1989) examined nine reports of studies on the professors-and-students problem and extracted the following five causes of students' difficulties in writing simple equations: (a) use of algebraic letters as abbreviated words, (b) literal substitution from words to symbols, (c) selection of the wrong cognitive frame, (d) conflict between the syntax of ordinary language and the syntax of algebra, and (e) the misleading influence of a mental image of two sets of objects.

After trialling several questions dealing with the formation of equations or their interpretation, a different list of factors was proposed to account for the errors recorded, namely: (a) meaning of words, (b) understanding written text, (c) understanding of the concepts of "more than" and "times", (d) meaning of algebraic letters, (e) conventions of mathematics notation, and (f) appropriate use of the "equals" sign. These factors were critically examined in the main research project in which a 12-item written test was answered by Year 9 students and the responses obtained from 235 of these students were analysed. Some of these students were also interviewed. Sample test questions were:

> 3. At a meeting there were five more women than men. There are 25 women. How many men are there?
> 6. If 6y = d, which is the bigger number, y or d ?
> 10. 'The number y is eight times the number z.' Write this information in mathematical symbols. (p. 90)

Using the responses on the first eight items of the test and the factors derived from earlier trialling, predictions were made regarding the error patterns expected for the remaining four questions. Discrepancies noted between the expected and the actual error patterns were of such a scale that it was concluded the proposed factors were insufficient to explain the difficulties of some items. This led to the rather audacious claim that "the difficulty of these items can not be explained by existing theories in the mathematics education literature" (p. 151). Question 10, quoted above, was designed to eliminate errors due to some of the previously-nominated causes and yet half of the responses from the Year 9 students were reversed equations or expressions, such as '$8y = z$', '$z = 8y$', and '$8y \times z$'. Some errors were accounted for by factors unrelated to algebraic skills, such as inadequate reading skills, poorly-developed concepts of the operations of arithmetic, and careless, informal, or imprecise use of mathematical notation. Other errors were caused by misconceptions about the meaning of algebraic letters and misunderstanding of what an equation signifies. Evidence was presented for the use by students of intuitive models simulating the semantic features or surface structure of a problem rather than its logical form. Semantic and syntactical analyses were found to be independent, so that meaning was often constructed without, or prior to, syntactic analysis. Most students disregarded the syntactic form of the stated problem and errors such as their reversed equations were written attempts to represent cognitive models of compared unequal quantities. It was concluded that the learning of algebra and the development of formal reasoning were obstructed by intuitive strategies used in processing natural language and in everyday reasoning. Joint research by psycholinguists and mathematics educators was recommended.

An adaptation of Question 6 above was included in the test used by the writer.

The last of a series of propositions treated in Chapter 8 considers the responses to this question in relation to responses by the same students to the professors-and-students problem.

Review and Forecast

This chapter has given reasons for choosing three neo-Piagetian theories of cognition as structural bases for analysing the process of developing an understanding of the meaning and use of algebraic symbols. These theories have been outlined, preparing the way for their application to the data obtained in this study.

Acknowledgment is given to the importance to the present project of the work done by other researchers over the past two decades in identifying difficulties commonly experienced by students in their efforts to understand algebra, and in developing techniques for measuring relevant aspects of students' levels of understanding.

While the next chapter describes the methodology adopted for this project, it will be seen to have developed from the work of the earlier researchers, especially with respect to the test items used for data collection.

CHAPTER 3

RESEARCH METHODOLOGY FOR TRIALLING PROGRAM AND MAIN RESEARCH PROGRAM

Overview

There are five parts to this chapter:

1. The first section describes the pilot program of trialling test items, and interacting with secondary students in order to become familiar with difficulties they might have with early algebra. Essential to the methodology was the continuing analysis of the outcomes from these trials so that items could be evaluated and retained, deleted, or adapted, and that new items could be considered to serve the purpose of measuring appropriate aspects of learning. Relevant documentation is included in the appendices for this section.

2. In the second section, a resumé of the final test items is presented.

3. The third section discusses reasons for omitting some of the trialled test items.

4. Section Four outlines the research methodology, showing how its five main aspects were relevant to the research objectives.

5. The final section describes the process of carrying out the main research program. A description is given of the samples of subjects who participated by being members of the classes monitored during their first three weeks of algebra, and/or by completing the test and, for some, by being interviewed. The classroom activities used during the teaching intervention period are described, and details of the testing program and the interview program are then presented.

Section One: Preliminary Investigations and Trialling of Test Items

The preliminary investigations were designed to find out the ways that young secondary school students thought about some of the basic elements of the algebra of generalized arithmetic and to formulate test questions suitable for measuring levels of understanding of some of these elements. The investigations consisted of trialling a variety of testing procedures, short sessions of teaching intervention by the researcher, and several student interviews. The interviews were directed towards clarifying students' answers to the test questions as a means of checking if the format of the questions could produce written answers that accurately reflected students' thoughts.

Likewise, the short teaching interventions conducted by the researcher were aligned with the principal objective as they provided the opportunity to identify some of the difficulties experienced by students in the early stages of learning algebra and to assemble some evidence as to whether or not the trial test questions would be sensitive to any changes in understanding during the period of the teaching sessions. The teaching made use of concrete manipulatives to clarify the meaning of algebraic symbols as used in first degree algebraic expressions.

Detailed summaries of the timetables for the 1989 and 1990 preliminary investigations may be found in Appendix 3A.

The groups of subjects who responded to a variety of test questions during the process of developing the research test items are described in Table 3·1. Manipulatives classes were those taught by a concrete approach to algebra and the Textbook classes were those taught by a more traditional textbook approach.

Table 3·1

<u>Description of Student Groups Involved in Trialling</u>

GROUP	DESCRIPTION	SAMPLE SIZE	Av.AGE in years
I	Manipulatives Year 7 Class School X 1989	20	13.12
II	Textbook Class Year 7 School X 1989	24	13.14
III	Manipulatives Class Year 7 School Y 1989	28	13.10
IV	Textbook Class Year 7 School Y 1989	18	12.89
V	Two Year 8 classes School X 1989	42	14.07
VI	Australian Catholic University students 1990	61	20.50
VII	University of Tasmania students 1990	36	25.63
VIII	Year 7 Class in School X 1990	19	11.86

Table 3·1 presents a summary of the sample sizes and average ages of the various groups of students involved in the trialling stage. All 136 students involved in the 1989 activities had been introduced to algebra whereas, in 1990, some trial research data were collected from 19 students (Group VIII) before they started their classroom work on algebra. The participating students were mainly from two Hobart secondary schools, one a boys' school ("School X") and the other a girls' school ("School Y"). A few questions were also trialled with groups of tertiary undergraduate students, some in Sydney and others in Hobart. Some of the investigations were conducted in the last few weeks of the 1989 school year and others in the early months of the 1990 school year.

A major objective in these preliminary investigations was to develop test items which might measure and discriminate between students' levels of understanding of the meanings for, and use of, alphabetic symbols in algebra. Consequently, the structure of the test often changed from trial test to trial test. Copies of the various test papers used over these months are assembled in Appendices 3B to 3I.

Table 3·2 summarizes the testing program, giving the test titles, references to the appendices, the groups of students who responded to each test, and the dates on which the tests were used. The appendices listed in the table include not only copies of the tests but also data on performance outcomes and comments on the findings.

Table 3·2
Summary of Trialling of Test Questions

TEST TITLES	APPENDICES	GROUPS of SUBJECTS	DATES
Brain-Box Quiz No.1	3B & 3D	I, II	24 Nov.89 and 8 Dec.89
Brain-Box Quiz No.2	3C & 3D	III IV V	8 Dec.89 12 Dec.89 7 Dec.89
Algebra Project 1990	3E	Part of Group V, VI	2 Mar.90, 12 or 14 or 15 Mar.90
1990 Algebra Project	3F	Part of Group V	16 Mar.90
Algebra Project 1990 + Q.3 from 1990 Algebra Project	3E 3F	VII	29 Mar.90 or 5 Apr.90
Yr.9 Test 1990	3G	Part of Group V	5 Apr.90
New Test 2 1990	3H	III, IV	5 Apr.90
Algebra Project New Test 1990	3I	VIII	5 Apr.90

The trialling of test items was spread across several months. Each stage of trialling led to decisions about the final test instrument in terms of which items to omit or amend and which new items were needed to cover aspects not yet tested.

<u>Trialling in School X</u>. The author taught a mixed ability class of Year 7 students (Group I, Table 3·1) in School X for three lessons towards the end of the 1989 school year. The activities used were taken from parts of the first two units of "A Concrete Approach to Algebra" (Quinlan et al., 1989) and included the modelling of expressions such as '$C + 3$', '$2C$' and '$2(C + 3)$' in terms of a number of square centimetres of area (selections from Unit One Worksheet Two), and expressions such as '$2y + 6$', '$2(y + 3)$', '$(y + 4) + (y + 2)$', '$x + 2y + 4$' and '$2(y + 2x)$' in terms of an objects-and-containers model (selections from Unit One Worksheet Three and Unit Two Worksheet One). Integer values were used most of the time, but the models were applied briefly for non-integral values and zero cases, as well as for subtraction.

This class and another Year 7 class (Group II, Table 3·1) from School X completed the "Brain Box Quiz No.1" test twice: once before the Group I teaching intervention and once after it, as detailed in Table 3·2. No special input was given to Group II between the tests.

For most of the 16 questions in the test used, there were no statistically significant differences from pretest to posttest. Such an outcome was understandable as the students had studied algebra during the year and were not likely to alter their perceptions of algebra over a time-interval of a few days. As Appendix 3K records, significant improvements were recorded by Group I on parts of four of the questions and by Group II on parts of five of the questions, and both groups improved significantly on the test total. It seemed that the practice effect entailed in responding to the same test twice was more influential than the intervention lessons.

<u>Trialling in School Y</u>. The researcher took two lessons with the Year 7 students in Group III (Table 3·1) in School Y. As these students had fewer distracting end-of-year activities than those in School X, the lessons were used with increased efficiency. The teaching approach was similar to that used in School X. The area model (from Unit One Worksheet Two) was used in the first of these lessons, and time was available for the students in groups to represent cases in which the numerical variable was not only integral but also fractional, to consider the zero case, and to manipulate areas to represent subtraction, as in Question 7 of Appendix 3N, one of several activities written during the research program to supplement those published in Quinlan et al. (1989). In the second lesson, the objects-and-containers model (from Unit One Worksheet Three and Unit Two Worksheet One) was used by the student groups to model first degree functions, such as '$2(y + 3)$' and '$2(y + 2x)$', and to model cases in which the variable was zero or fractional, as well as model and discuss the case in which 'x' equalled 'y' when working on Question 4 of Unit Two Worksheet One. The worksheets were the same as those used in School X over three

lessons. Only a selection of activities from each worksheet was covered in the time available.

Group III and another Year 7 class from School Y (Group IV, Table 3·1) completed the "Brain Box Quiz No.2" both before and after the teaching intervention for Group III, as reported in Table 3·2. Between the tests, Group IV continued with their regular mathematics program.

Group III improved significantly from pretest to posttest on substantially more measures than did Group IV, according to the *t*-tests reported in Appendix 3K, Table 3K·1: Group III showed significant improvement on overall test scores and on 12 of the 17 items tests, whereas Group IV improved significantly on just 3 test questions. These differing outcomes could not be readily explained as a consequence of practice from repeated testing since both groups had this practice. It was noted that only Group III improved on Item 6, one of the more difficult items from Küchemann's 1989 study, asking students to compare the values of the two functions '$n + 2$' and '$2n$'. The intervention teaching for Group III seemed to be a telling factor but it was not clear whether it was dependent upon the activities used or simply upon the fact that this group had been given some special algebra lessons. In any case, the test items were seen to be able to detect changes in response patterns.

Implications from Year 7 trialling. The impression was gained that many of the 90 Year 7 students who participated in the trialling were still at the formative stage in their understanding of the meaning and use of algebraic symbols. This made it difficult, for instance, to categorize them according to whether or not they viewed letters in algebra as standing for numbers or as standing for objects. It was found, instead, that many changed their mind from one context to another. For example, 29 of the students who responded to Brain-Box Quiz No. 1 accepted an object as a meaning for a letter in only one of the options given in the first two questions, choosing either "a sheep" or "an apple". Similarly, 48 accepted just one of the options "an apple" (Question 2) or "a cabbage" (Question 14), while 34 accepted both "an apple" and "a number of apples in a box" in Question 2. There were also changes from test to test, such as, in Question 2(c)(i), 14 changed their minds from "No" to "Yes", while 14 changed from "Yes" to "No" in their options about whether or not '*a*' could equal "an apple". Similarly, in Question 2 (c) (ii) which asked whether or not '*a*' could equal "the number of apples in a box", 14 changed from "Yes" to "No" and 17 changed in the opposite direction.

Indications such as these led to a broadening of the scope of the test questions about the ways in which students viewed the letters of algebra, as it did not appear wise in the main study to concentrate too much on comparing an "objects view" with a "numbers view".

Year 8 trialling. The test entitled "Brain Box Quiz No.2" was completed once by two Year 8 classes (Group V, Table 3·1) in School X. The test outcomes obtained from Year 7 students the first time they did the test were compared with those obtained from Year 8 students who were tested just once.

The 42 Year 8 students from School X were significantly better than the 44 School X Year 7 students on parts of eight of the test questions, as well as on total test scores. Details of the appropriate *t*-tests are summarized in Table 3K·2 of Appendix 3K. Here was some tentative support for the possibility that, if a developmental sequence exists for learning algebra, perhaps some of the test items could provide data for investigating such a sequence. It was noted, however, that the Year 7 students were significantly better than the Year 8 students on parts of two test items.

The Year 8 students from School X scored significantly better than the 46 Year 7 students from School Y on parts of 10 of the test questions. The latter had better scores on parts of two test items, as reported in Appendix 3K, Table 3K·2. Again, there is at least tentative evidence that the items in the research test instrument were able to detect differences in performance which could be related to age and/or experience with algebra, and so they were suitable for investigating the possibility of the existence of developmental stages in the learning of algebra.

The fact that the Year 8 students scored significantly better than the Year 7 students on several questions gave encouragement for using at least some of those test questions for measuring development in understanding of basic algebra.

For Brain Box Quiz Nos. 1 and 2 percentage frequencies of responses, together with comments on the outcomes, have been assembled in Appendix 3D.

Sessions with Year 9 volunteers. Twice a week for six weeks in early 1990, some volunteer Year 9 students in School X (from the 1989 classes of Group V, Table 3.1) spent half an hour or more on algebra with the writer. The numbers varied from five to eight. These sessions were particularly helpful not only for the students but also for the writer as they provided the opportunity to follow at close quarters how students thought about basic issues in algebra and to trial new activities in algebra.

The sessions began by building geometric patterns out of matches, making use of the approach detailed in association with Unit One Worksheet One of Quinlan et al. (1989), which led the students to express generalizations in everyday language about the number of matches needed for any number of units in these patterns. Although the students were in Year 9 and were about 14 years of age, some had great difficulty forming generalizations based on patterns built out of concrete materials. A worksheet exercise (like that in Appendix 3M) was trialled in which the students were asked to take the generalizations they had derived from the patterns and translate them from everyday language so as to include the use of a letter to stand for an unknown number of units in the pattern. The outcome was unexpected. Some of these Year 9 students had great difficulty managing the concept that a letter could stand for an unknown number and they revealed a strong tendency to think of letters as abbreviations for objects or as representing the objects themselves. This experience led to the decision to include such an exercise in the main teaching intervention program.

Some weeks later, the use of the objects-and-containers model for the expression '2(3*y*)' led to an extended discussion. One student was convinced that there was a rule which dictated that the meaning for the expression was '2 x 3 x 2 x *y*'', so that it came to '24' if '*y*' was '2'. He refused to accept the result of '12', obtained by building '3*y*' with the model and then building it again to get 6 containers each holding 2 objects. He was arguing from his conviction that the "rule" was to multiply everything inside the brackets by '2', a rule which worked when there was an addition operation within the brackets as in '2(*y* + 3)', but not if the brackets enclosed a multiplication operations as in '2(3*y*)'. A transcript from a tape-recording of the discussion is given in the last few pages of Appendix 3G. The incident led to the inclusion of '2(5*y*)' as part of the Item 3 substitution question in the final test (Appendix 3I), and it strengthened the researcher's resolve to incorporate the use of models for Manipulatives classes in the main research. The models provided a talking-point for students and so helped them clarify their thinking, at the same time revealing their thinking more clearly to the researcher. The manipulatives approach appeared to assist students in their metacognition, a process of reflecting on their thinking, as illustrated in the discussion in Appendix 3G. This was an outcome relevant to one of the research objectives stated in Chapter 1.

<u>Trialling with university students</u>. As was reported in Tables 3·1 and 3·2, some assistance in determining the items to be used in the final test instrument was obtained from university students, namely, Group VI from the Australian Catholic University and Group VII from the University of Tasmania.

In Group VI,

28 were 2nd Year Primary B.Ed. students,

18 were 2nd Year Secondary B.Ed. students, and

15 were 3rd Year Secondary B.Ed. students.

In Group VII,

19 were 2nd Year Primary B.Ed. students, and

17 were Primary Dip. Ed. students.

Only about 40% of these teacher education students correctly responded to the professors-and-students problem. Associated reversal errors appeared to be linked with one's view of the meaning of the symbols used. Details are available in Appendix

3E or in relevant comments on page 111. About one-fifth of Group VII succeeded with the Harper's Parallel Lines task, the trialling of which is discussed as "Question 3" in Appendix 3F. The two questions became Items 7 and 11 respectively in the final version of the test.

Trialling of complete test format. The final version of the test instrument (Appendix 3I) was trialled with Group VIII who were Year 7 students and had not started their classroom study of algebra. It was expected that some questions would be completely foreign to students who had not been introduced to algebra. However, the trialling group gave correct answers to at least some parts of all questions except Items 4 and 12. The range of marks was from 15 to 29 with an average of 19.5 and a standard deviation of 4.40. Although the maximum possible mark was 65, these outcomes were acceptable. The test was considered to provide a suitable data collection instrument for the research project.

The test provided insights into views held by the uninitiated. Thus, the path was clear for examining, in the main study, whether or not preconceptions were resistant to the effects of teaching and experience or were readily modifiable. As instances of misconceptions that arose from this initial trialling with beginning algebra students, we have the following incorrect sets of answers which were often the result of correctly applying logical but uninformed thinking to unfamiliar problems:

1. In Item 2(i), using an alphabetic code led one student to the decision that '*d*' is greater than '*a*' "because it's the fourth letter", and then, using 'y = 25' in Item 3, he wrote

   $2y = 50$ (=2x25), $2(y+5) = 32$ (=2+25+5), $2(5y) = 127$ (=2+5x25).

2. In Item 3, conjoining was interpreted as addition to give the results

   2y = 5 (=2+3), 2y+5 = 2(y+5) = 2(5y) = 10 (=2+3+5), 3y-y = 3 (=3+3-3).

3. In Item 3, coefficients were regarded an indices to produce

   $2y = 9$ (= 3^2), $2y+5 = 14$ (=9+5), $3y - y = 24$ (= 3^3 - 3).

4. In Item 5, instead of writing '$p + r$', over 25% gave a numerical answer, and over 20% wrote 'pr'.

5. In Item 8(a), slightly more than half the students concluded that '*a*' represents "apples" given that "3*a* represented 3 apples".

6. In Item 8(b), more than 40% gave the place-value result that '*a*' equalled '6' if '3*a* = 36'.

7. In Item 9, over 30% wrote answers such as '*n*9', '7*n*', or '*n*20', instead of '*n*+ 9', '3*n* + 4', and '4*n* + 20' respectively, and a few more than 25% gave a numerical answer.

Section Two: The Final Test Items

A summary is presented in Table 3·3 of the objectives of the items which were eventually included in the final test instrument, a copy of which is in Appendix 3I.

Table 3·3
Summary of Objectives for Items Included in the Final Test Instrument

ITEMS	OBJECTIVES
1	arithmetic processes and finding values for symbols
2	symbols as generalized numbers; relationships
3	conventions used for writing algebraic expressions; substitution
4	interpretation of symbols as numbers or objects; real-life context
5	operation on symbols in real-life context
6	symbols as objects, or numbers which are not only positive integers; relationships; abstract context
7	interpretation of symbols as people or numbers; relationships; real-life context
8	symbols as numbers or objects; use of conjoining for multiplication
9	operations on symbols in abstract context
10	symbols as numerical variables; relationships; one operation
11	symbols as numerical variables; relationships; geometric context
12	symbols as numerical variables; relationships; two operations
13	symbols as numerical variables; relationships; equation context
14	interpretation of, and operations on, symbols in real-life context
15	symbols as numerical variables; relationships; multiple operations

The main criterion for the inclusion of items in the final test instrument was that they sought responses which were relevant to the objectives of the research project by providing information about the way students viewed algebraic symbols and/or about their ability to use symbols. Other more general criteria were also considered, such as wording clarity and item formatting. Open-ended items were considered with caution to ensure that the types of responses elicited from the students were relevant to the objectives and were not so diverse as to become unmanageable.

Within this general context, the search was for items which might measure the degrees of ability students had for:

1. distinguishing between symbols being thought of as representing numbers of objects or people and as representing non-numerical objects or people;

2. recognizing that algebraic symbols may take fractional, zero, and negative values, not just positive integral values;

3. interpreting algebraic conventions for describing first degree functions;

4. carrying out operations with symbols without knowing their numerical values and, in this way, accepting lack of closure;

5. interpreting the meanings of symbols used with real-life referents;

6. managing relationships between two variables or two functions;

7. forming mathematical generalizations and expressing them in algebraic terms; and

8. making use of the concept of variable.

To ascertain the history of the trialling for each of the items which became components of the final test, a summary is presented in Table 3·4.

Table 3·4

Summary of Trialling of Final Test Items

Final Test ITEM	B.B. Quiz No.1	B.B. Quiz No.2	Alg. Proj. 1990	1990 Alg. Proj.	Yr.9 Test 1990	New Test 2 1990	Algebra Project New Test 1990
1				6		1	1
2	10 iii,iv	12 iii,iv				(ii) 9	2
3	6[a]	8[a]			5[a]	7[a]	3
4	15 i[a]	13			3	10	4
5	3[a]	3[a]					5
6	14[a]	16[a]					6
7			1		1	6	7
8		7[a]			(d) 4		8
9	9	11					9
10				2		3	10
11				3		5	11
12	13[a]	6[a]				4	12
13				1[a]	2	2	13
14	8[a]	10[a]				11	14
15	16	17			6	8	15

Note. The numbers give the item numbers in the respective trial tests.
[a] denotes that the format of the trialled item was later changed.

The first two tests listed, namely Brain Box Quiz No.1 and No.2, were administered twice to Year 7 students late in the school year of 1989, and the second of these was given to Year 8 students once at about the same time. The other tests were administered to different groups between early March and early April in 1990. By the time the test entitled "1990 Algebra Project" was used on 16 March, 1990, some three weeks before the main research program actually began, each item had been trialled at least once. From the 1989 trials, only four items were preserved unchanged and six more found their way into the final test after some modifications. Five other items emerged in 1990. Table 3·4 gives the cross-referencing of the numbering of the final items against the numbering of corresponding trialled items. When the items selected for the final test are discussed in Chapter 4, the outcomes from the trialling are shown to have provided the background for their inclusion.

Section Three: Comments on Some Omitted Items

Several of the items trialled during the pilot studies were omitted from the final test instrument. Three groups of such items, labelled A, B and C in what follows, merit some comment. Performance statistics obtained when these items were trialled may be found in Appendix 3D.

Group A: Choose Meanings for Letters from Given Options. Items 1, 2, 7, 14 (a) from Quiz 1, which were numbered respectively 1, 2, 9 and 16 (a) in Quiz 2 formed Group A. Table 3·5 gives correlations between items in Group A.

The questions in Group A unnecessarily duplicated information. As Table 3·5 shows, responses to the parts of Item 14 (a) that required an acceptance of small positive integers, zero, negatives, or fractions as possible values for an algebraic variable were significantly and positively correlated with corresponding parts of Items 2 and 7. The option "the number of apples in a box" in Item 2 also correlated positively and significantly with responses involving three of the number choices in Item 14 (a). The fact that these correlations were significant supported the recognition that there was little need for so many questions on these issues. Item 1 did not correlate with the other items as well as did Item 14 (a). Moreover, Item 1 was shown by the trialling to be rather limited in usefulness because the context of the number of sheep in Quiz 1 and the number of cows in Quiz 2 restricted the range of valid number types to zero or positive integers.

Consequently, only Item 14 (a) was retained from Group A, this item becoming part of the final item, Item 6, which also included the option "the number of apples in a box" from Item 2.

Table 3·5
Correlations for Responses on Four Possibilities for Numerical Variables

Letter represents Small Integers > 0	2 (a) (i)	7 (i)
14 (a) (i)	.1559 *	.2375 **
Letter represents Zero	2 (b) (i)	7 (ii)
14 (a) (ii)	.1435 *	.4255 ***
Letter represents Negative	2 (b) (iii)	7 (v)
14 (a) (iii)	.1640 *	.3005 ***
Letter represents Fraction	2 (b) (ii)	7 (iii)
14 (a) (iv)	.3702 ***	.3570 ***

Note. N=132: Yr.7 School X (44),Yr.7 School Y (46), Yr.8 School X (42). Correlations are for responses when students did the test for the first time. Items numbered as in Quiz 1.
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Group B: Choose Equivalent Expressions from Given List. Items 4 and 11 from Quiz 1 with Item 14 from Quiz 2 comprise Group B.

In each of the questions of this group, the students were given a set of expressions and were asked to select those that correctly described some aspect of a particular environment. For the first question, Item 4 Brain-Box Quiz 1, each expression was to be judged on whether or not it represented the total number of milk cartons that could fit into 4 crates, each of which held '*n*' cartons, and so the letter symbol stood for some unknown number of milk cartons. In contrast, Item 11 of the same test used letters to represent geometric objects, with '*s*' representing a square and '*t*' a triangle. These two items were used in Quiz 1 in an effort to find out if the "numbers view" or the "objects view" of letters might influence performance in this type of exercise. The objects view item was replaced in Quiz 2 by another numbers view question (Question 14), using the numbers of sweets in sets of boxes as the referent.

It was found that there was not a significant difference whether the referent used encouraged the numbers view of letters or the objects view. This claim was established by *t*-tests between the average scores on Item 11 for Quiz 1 (objects view) and Item 14 for Quiz 2 (numbers view). The *t*-values were too small for any difference to be statistically significant.

It was similarly established that the inclusion of diagrams did not significantly assist students and could, in fact, be a hindrance. The average scores fell when diagrams that led to a numbers view were included, but rose when the diagrams led to an objects view. However, the changes were significant in only one case: On the posttest scores for Groups III and IV (Year 7 from School Y), the average score on the question without diagrams (Item 4) was significantly better ($p < 0.050$) than the average for the question with diagrams (Item 14).

These questions were not considered very useful in the study. They were concerned with the degree of knowledge which students demonstrated about certain conventional uses for algebraic symbols, a form of knowledge assessed by other test items, such as Items 3 and 9. They did not give clear distinctions between the effects of diagrams or of referents of different types.

Group C: Simplify Algebraic Expressions. Group C consisted of the parts of Item 5 (a) in both Quiz 1 and Quiz 2.

This question asked students to simplify expressions written in terms of two variables, 'a' and 'b'. They were trialled as examples of questions which were found to be common to several of the assessment programs discussed in Chapter 1. The success rate was about 60% for the first three parts of the item trialled and a little less than 40% for the last part, which required a simplification for '$5a + 3b + 2a - 4b$'. A variety of errors was noted, including the conjoining of terms. As mentioned in Chapter 2, both Küchemann and Booth pointed out that students can achieve well in such exercises even if they regard letters as objects to be manipulated. Thus these exercises were not considered very useful in this study of students' views of algebraic symbols.

Symbol manipulation. The art of manipulating symbols to convert one form of an expression into an equivalent form was not included in the main study, although the possibility of including this aspect was considered in the Groups B and C questions just discussed. It was thought that data on the levels of skill with the manipulation of symbols would not contribute much, if anything, to a study of the understanding of the meaning of the symbols. As Kieran (1989b, p. 164) stated,

> There is ample empirical evidence to show that students are able to manipulate symbolic expressions and equations with a great deal of control and success, but still not be able to do much else in algebra.

Section Four: Methodology for Main Study

To attain the research objectives, the methodology was determined so as to include:

1. written testing;
2. interviews;
3. intervention teaching for about three weeks as algebra was introduced to Year 7 students in their first year of secondary school;
4. data collection from the beginning Year 7 students at four stages: before starting classroom algebra, after about a week and a half of lessons, after about three weeks of lessons, and again about six months later; and
5. data collection from students in classes across Years 7 to 12, covering a range of mathematical ratings (such as "Advanced" or "Slow Learners").

These five aspects of the methodology are now examined in turn.

1. Written Testing

In developing the written test instrument, a selection of questions used in previous research projects was incorporated because they probed students' understanding of the meanings and use of algebraic symbols. Some of these questions were adapted to clarify the information sought. Questions used by Harper (1979) in interviews were trialled in written form to ensure that the range of responses which Harper obtained in interviews could also be obtained by means of written responses. They were then incorporated in the test. Other questions were designed by the researcher, trialled and, if necessary, re-written and included in the test instrument if they were judged useful in pursuing the objectives. A global view of the response patterns to the test is presented in Chapter 4, and evidence to support the combination of clusters of items into scales is detailed in Chapter 5.

Special features. The test instrument provided data across a base which was broader than that of previous research in similar areas of interest. For the first time, comparisons could be made (in Chapters 4 and 8, for example) between responses by the same subjects to questions based on work by Collis (1975a), Harper (1979), Küchemann (1980), Rosnick (1981), Booth (1983) and MacGregor (1989). Furthermore, original items were included which employed a more direct approach than had been used previously for obtaining data on such issues as the readiness of students to allow variables to assume values other than positive integers, and to view symbols, inappropriately, as representing non-numerical objects.

2. Interviews

Written tests were followed up by interviewing samples of students. This was an important part of the methodology because it was thought that it would provide insights into the ways students were thinking that either clarified what they had written in their test responses or extended and interpreted the information they had recorded in writing. Year 7 students were targeted for interviews, particularly those who had changed their minds on certain answers from one test to the next. However, to elucidate various aspects of students' thinking, some older students were also interviewed. Almost all interviews were recorded either on audio-tape or video-tape, the exceptions being for a few students who preferred not to be taped. Interview extracts are interspersed throughout the thesis, particularly in Chapters 6, 7, and 8.

3. Intervention Teaching for Year 7 Students

Intervention teaching in the main study was monitored over a period of approximately three weeks while Year 7 students in 10 classes were introduced to secondary school algebra. By following closely the lesson sequences for some classes, useful information was gathered for elucidating the cognitive processes by which students might move towards developing an understanding of the meaning and use of algebraic symbols.

Classroom observations. Schools were selected to ensure the inclusion of classes using a manipulatives approach. Except for Class 3 (Table 3·6), Year 7 mathematics lessons in each school were time-tabled for the same periods in the day. Consequently, observations were restricted mainly to Classes 3, 5, 6, and 8 - as well as Class 1 (taught by the researcher). Textbook classes were not monitored by observation. If the project had included the comparison of teaching methods as an objective, a very different structure would have been necessary.

Textbook classes. Three classes used a traditional textbook approach and are referred to as Textbook Classes. Information about the teaching methods used in these classes was obtained simply in the form of the textbook references. Details of the textbook pages used during this period have been assembled in Table 3·8. These classes were, in a sense, "Uncontrolled Classes" as the researcher did not have control over the teaching style used. For instance, one of these classes (in School C) used concrete manipulatives for building geometric patterns in the first few lessons whereas the other two classes did not. Generally, the lessons were based on textbook exercises in the traditional way: Introductory examples would be considered by the class as a whole and then similar examples would be attempted by the students as they worked individually. Teachers were asked simply to include exercises which dealt with the interpretation of algebraic expressions before the end of the three weeks.

Manipulatives classes. Seven of the participating classes were classified as Manipulative Classes because they were introduced to algebra with the aid of concrete manipulatives. The teaching approach used in the Manipulative Classes was controlled in such a way that the researcher knew the sequence of lessons and the method of presenting the lessons, and so these seven classes could be regarded "Controlled Classes". The lessons in these classes followed an approach developed by the N.S.W. Algebra Research Group, for whom the writer was the Project Coordinator during the four years of the action research development phase which led to the publication of four booklets for teachers, entitled "A Concrete Approach to Algebra" (Quinlan et al., 1989). Details of the approach are given as the chapter unfolds. During the three weeks of liaison, the teachers followed the teaching methods given in parts of the first two booklets. Information about how the teachers used the worksheets was obtained by working with the teachers when they were preparing some of their lessons, observing lessons, and tape-recording segments of lessons. Supplementary activities and worksheets were written by the researcher in response to perceived learning difficulties and a number of these were used by some teachers. These as-yet-unpublished activities and worksheets have been assembled in Appendices 3M, 3N, 3P and 3Q. Details about which parts of the worksheets were used with which classes during the intervention stage are presented in Table 3·8.

Choice of manipulatives approach. Among reasons for this choice of teaching approach were the following:

1. Research was scarce in the area of introducing algebra either with or without the aid of concrete manipulatives and this study was a chance to incorporate this aspect of research within the embrace of a project which was planned to investigate student difficulties in early algebra, using a broad-based methodology.

2. The approach made use of generalizations from geometric patterns and three models, namely, an area model, an objects-and-containers model, and a length model, each of which stressed that algebraic symbols stood for numbers and not objects. An approach using models in which algebraic symbols represented numbers was in contrast to approaches numbered 2, 4, and 6 in the section of Chapter 1 which dealt with types of intervention efforts. These latter three approaches respectively used letters to represent labels for geometric shapes, objects such as apples and bananas, and geometric shapes with certain "values". Examples showing that the models used by the Manipulatives Classes were adaptable for modelling fractional and/or zero values for variables, although not negatives, are given in Appendix 3N for the area model, and Quinlan & Collis (1990, p. 445) for the objects-and-containers model.

3. The models used seemed to rate well on the following criteria for scientific analogues as set out by Gentner (1982): clarity, richness, systematicity, abstractness, validity, and scope.

4. Mappings between the concrete system which the models provided and the

algebraic system for first degree expressions adequately passed Halford's commutativity test (Halford & Wilson, 1980, p. 372), and the two systems were sufficiently isomorphic (Coombs, Dawes & Tversky, 1970, p. 11) to allow mappings from algebra to concrete and vice versa. A sample commutativity test and some examples of mappings are given in Part I of Appendix 3P.

5. The models provided a vehicle for making explicit the students' thought patterns. Correct and incorrect ways of viewing algebraic ideas were made visibly clear by the ways students manipulated the models. Many beginning students, for instance, were observed to show their misunderstanding of the algebraic expression '$2(y + 3)$' by building it incorrectly with the objects-and-containers model which used a student-selected number of small objects in a container to represent 'y'. They built the expression by placing 3 objects beside two containers holding the same number of objects instead of putting 6 objects beside the two containers. Students within groups were able to argue about which representation was correct. Teachers could readily observe how the students were thinking, not only by following their discussions but also by observing the ways they used the models. Metacognition was assisted by the presence of the manipulatives. Thus, the incorporation of Manipulatives Classes into the main research program provided another avenue for obtaining clues about the elusive pathways followed as cognitive processes led students to the development of better understandings of basic algebra.

6. Some control, even if it were minimal, could be exercised over the teaching activities used in the Manipulatives Classes by having teachers follow a similar approach during their introduction of algebra to their students. It was planned to work with teachers who were familiar with the concrete approach rather than work with teachers new to the method, in which case they would have had to be trained in the approach. There would then be the prospect of working with such teachers as they tried the method for the first time. The author did the teaching for the one group of Hobart students who participated in the main study, and the other six Manipulatives Classes were taught by teachers in New South Wales who were familiar with the approach. There was one exception: In School D, Class 10 was taught by a teacher who had transferred from teaching Primary classes and was teaching Year 7 mathematics for the first time. The influence of the troublesome variable of how a subject is taught was partly reduced by exercising some control over the teaching method employed in seven of the classes.

Two approaches. The project did not set out to compare contrasting teaching methods, for reasons spelt out in Chapter 1. Knowledge about the controlled teaching approach in selected classes was intended to help elucidate the cognitive processes which the students experienced in developing their early concepts of what algebra was all about. The inclusion of some classes being taught by the traditional textbook method did, however, allow some passing comment, mainly in Appendix 7B, on whether or not different teaching and learning activities may have had different effects on the way the students developed their understandings of the meaning and use of algebraic symbols.

4. Repeated testing for Year 7 students

The methodology of using repeated tests was planned to provide measures of the development of students' understanding of the meaning and use of algebraic symbols over a period of time. The procedure made this research project somewhat distinctive. Testing was carried out before any classroom lessons on algebra in order to provide a record of the students' views prior to being taught algebra. Hence it was possible to examine (as in Chapters 7 and 9) whether or not these prior views were retained or changed during the teaching phase. Therefore, there was a mechanism for focusing on the persistence or otherwise of viewpoints and then investigating reasons for the variations. Furthermore, testing after about one and a half weeks and again after three weeks provided data which allowed the investigator to study rates of development for different students and to ponder over explanations for the differing rates and for the sequences of change in students' viewpoints about algebraic symbols. Chapters 8 and 9 report on investigations of these issues. Delayed posttests after about six months registered the effects of time and, perhaps, further experience with algebra, on the responses given by students to the test items. It was decided to use the same test repeatedly to make comparisons of responses easier and to avoid bringing extraneous variables into the arena, such as influences that could result from changes in the wording of questions.

5. Testing across Years 7 to 12

Giving the test to classes of different abilities across Years 7 to 12 in 1990 was seen as a means of contributing to the objectives in several ways. The data gave a broad picture of development from the "beginners" in Year 7 to the "experts" in the top-level classes of Year 12. Such data were valuable in assessing the degree of challenge in certain concepts and, consequently, in contributing to an understanding of the hierarchies of learning in relation to the tasks involved in the test items. They also provided a context which placed the achievement levels of the beginners in perspective. The collection of test responses from classes of different mathematical abilities provided information allowing comparisons of performance across different age groups and different ability groups. Chapter 8 reports findings obtained from a study of responses given by the different class groups. Some students from Advanced level classes were re-tested on selected questions in July 1991 in order to elucidate earlier findings which appeared paradoxical.

Section Five: The Main Research Program

The Schools

During 1990, data were collected from 517 students in four schools from student samples across the secondary grades Years 7 to 12 by means of the written test instrument (Appendix 3I), which has been described above in detail, and by means of observing lessons and interviewing a selection of students. For a subgroup of 115 students, a follow-up, which consisted of a short test and interviews, was conducted in the middle of 1991.

The four schools were:

School A, a girls' school in Hobart;

School B, a boys' school in the western suburbs of Sydney;

School C, a girls' school situated beside School B; and

School D, a coeducational school on the north-west fringes of Sydney.

The Participating Classes

Table 3·6 summarizes general information about the ten participating Year 7 classes which were tested four times in 1990. The respective schools are listed and the average ages for each class at the time of their first test are tabulated. The teaching intervention sessions started with the lessons following the first testings, the dates of which are listed. The type of teaching approach used is described simply as either a Manipulatives approach or a Textbook approach. The mathematics ability ratings for the classes were determined by the mathematics department of the schools. In Schools C and D, Year 7 mathematics classes were graded according to ability, the term "Advanced" being applied to the highest ability classes. In School D there were three Advanced classes of similar ability ratings but in School C, of the two Advanced classes who participated, Class 7 was rated as of higher mathematical ability than Class 6. In the other schools, A and B, the policy was not to grade the Year 7 mathematics classes except that, in each case, one slow group was taught separately. Class 3 was one such slow group and the other classes from these two schools were classified as "Mixed" ability. These classifications were useful for comparing the progress of students of different abilities and for allowing comparisons to be made of test responses from classes of similar rating across the Year levels, as in Chapter 8.

Table 3·6

Summary of Information About Year 7 Classes Tested Four Times

CLASS (Code)	SCHOOL	Average AGE[a] in years	DATE of first test (1990)	TYPE[b]	No.[c]	ABILITY rating	GENDER
1	A	12.79	6 Apr.	M	8 (8)	mixed	girls
2	A	12.43	6 Apr.	T	15(13)	mixed	girls
3	B	12.75	30 Apr.	M	8 (7)	slow	boys
4	B	12.50	30 Apr.	T	21(20)	mixed	boys
5	B	12.56	30 Apr.	M	23(21)	mixed	boys
6	C	12.40	3 May	M	28(21)	advanced	girls
7	C	12.43	3 May	T	29(26)	advanced	girls
8	D	12.60	4 June	M	25(25)	advanced	coed.
9	D	12.56	4 June	M	25(23)	advanced	coed.
10	D	12.62	5 June	M	26(22)	advanced	coed.

[a] AGE when given their first test. [b] TYPE: M = Manipulatives, T = Textbook
[c] The numbers of students who responded to each of the first three tests are shown, followed by the numbers (in brackets) of those same students who also completed the fourth test. The absentees for the fourth test were spread across the ability levels.

Particular attention was given to beginning algebra students in ten Year 7 classes by the following four-fold strategy:

1. monitoring their first three weeks of classroom work on algebra;

2. administering the test instrument three times during this period: before any lessons on algebra, after about a week and a half, and again after the three weeks;

3. interviewing a selection of the students with regard to their test responses; and

4. administering a delayed posttest and conducting associated interviews.

Table 3·7 describes all the participating classes which were not included in Table 3·6. The latter classes were tested just once in 1990, apart from Classes 33, 34, 35, 54 and 55, who were given a short second test in July 1991. Two Year 7 classes (Classes 14 and 15), both groups of slow learners, are included in Table 3·7 as they were tested just once, unlike the other Year 7 classes listed in Table 3·6.

A timetable of events for the main study is given in Appendix 3L.

Table 3·7
Summary of Information About Classes Tested Only Once or Twice

CLASS (code)	SCHOOL	YEAR	Average AGE[a] in years	DATE of 1990 test	No.	ABILITY rating	GENDER
14	C	7	13.08	3 Dec.	12	slow	girls
15	D	7	13.20	3 Dec.	12	slow	coed.
24	C	8	13.48	15 June.	26	advanced	girls
25	D	8	14.00	4 Dec.	18	slow	coed.
26	D	8	14.63	4 Dec.	10	slow	coed.
31	A	9	14.90	19 Oct.	11	average	girls
32	A	9	14.70	19 Oct.	10	slow	girls
33	B	9	14.54	12 June	26	advanced	boys
34	C	9	14.50	14 June	28	advanced	girls
35	D	9	14.63	21 Jun.	21	advanced	coed.
44	C	10	15.81	13 June.	15	average	girls
45	D	10	15.52	21 June.	11	advanced	coed.
46	D	10	15.73	21 June.	11	mixed	coed.
51	C	11 2U	16.50	14 June.	17	average	girls
54	C	11 3U	16.42	14 June.	32	advanced	girls
55	D	11 3U	16.49	21 June.	16	advanced	coed.
61	B	12 MIS	17.35	13 June.	11	slow/ average	boys
62	B	12 3U	17.67	12 June.	7	advanced	boys
63	B	12 4U	17.43	12 June.	13	advanced	boys

Note. Year 11 and 12 classes followed the N.S.W. Courses: Mathematics in Society (MIS) (lowest level), or 2 Unit (2U), or 3 Unit (3U), or 4 Unit (4U) (highest level).
[a] AGE when given their test in 1990

Descriptions of the Teaching Interventions

The program of teaching interventions is summarized in Tables 3·8 and 3·9. The first of these tables gives a brief description of the teaching activities used between Test 1 and Test 2, with references to the appropriate publications and appendices (for unpublished material), and the second does likewise for the activities used between Test 2 and Test 3.

Table 3·8

Classroom Activities Between Test 1 and Test 2

Class	Type	No.of lessons	References	Description of Activities
1	M	5	QLSW U1W1 App.3M QLSW U1W2 Qq.1-4 App.3N Q.5	From geometric patterns to generalizations in everyday language; in symbols. Area Model Qq.1 - 5
2	T	5	MGCT	Using symbols as pronumerals Ex.4A - Ex.4C
3	M	5	As for Class 1	As for Class 1
4	T	5	MCW	Geometric patterns; Start on pronumerals Ex.5:01 - Ex.5:03 Q.1
5	M	5	As for Class 1	As for Class 1
6	M	5	As for Class 1 + App.3N Qq.7-12	As for Class 1 + Qq.7 - 12
7	T	5	MCW	Geometric patterns (used matches & cubes); Pronumerals Ex.5:01 - Ex.5:03 Q.7
8	M	7	As for Class 6 + App.3N Q.6	As for Class 6 + Q.6
9	M	7	QLSW U1W1 App.3M QLSW U1W4 Qq.1-5	From geometric patterns to generalizations in everyday language; in symbols. Length Model Qq.1 - 5
10	M	5	QLSW U1W1 App.3M	From geometric patterns to generalizations in everyday language; in symbols.

Note. M = Manipulatives; T = Textbook; Q = Question; App. = Appendix; MGCT = McLeod, Ganderson, Creeley, & Tanti (1988); MCW = McSeveny, Conway, & Wilkes (1989); QLSW = Quinlan, Low, Sawyer, & White (1989); U = Unit; W = Worksheet.

Table 3·9

Classroom Activities Between Test 2 and Test 3

Class	Type	No.of lessons	References	Description of Activities
1	M	4	QLSW U1W3 Qq.1-9 QLSW U2W1 Qq.1-4	Objects-and-Containers Model One variable: Qq. 1 - 9 Two variables: Qq.1 - 4
2	T	4	MGCT	Manipulating symbols; Substitution Ex.4D - 4F
3	M	4	QLSW U1W3 Qq.1-10 App.3P Q.11	Objects-and-Containers Model One variable: Qq. 1 - 11
4	T	4	MCW	Using pronumerals; Start on expressions Ex.5:03 - Ex.10:01
5	M	5	As for Class 3	As for Class 3
6	M	6	QLSW U1W3 Qq.1-10 App.3P Q.11 QLSW U2W1 Qq.1-9 App.3Q Q.10	Objects-and-Containers Model One variable: Qq. 1 - 11 Two variables: Qq.1 - 10
7	T	6	MCW	Pronumerals for general rules; Substitution; Start on graphs Ex.5:03 - Ex.10:03 Q.4
8	M	3	QLSW U1W3 Qq.1-3,7,8 App.3P Q.12 QLSW U2W1 Qq.1-4	Objects-and-Containers Model One variable: Qq. 1-3,7,8,12 (first use of Q.12) Two variables: Qq.1 - 4
9	M	3	As for Class 8	As for Class 8
10	M	3	QLSW U1W2 Qq.1-4 App.3N Qq.5, 6	Area Model Qq. 1 - 6

Note. M = Manipulatives; T = Textbook; Q = Question; App. = Appendix; MGCT = McLeod et al. (1988); MCW = McSeveny et al. (1989); QLSW = Quinlan et al. (1989); U = Unit; W = Worksheet.

Comments on Classroom Activities

The teachers introduced students to the idea of a numerical variable and the use of algebraic symbols in first degree expressions, without directly coaching them for the test items. An exception is discussed on pages 305 to 307.

As there were two main approaches utilized in the intervention teaching, comments on classroom activities are presented in two parts.

<u>Textbook classes</u>. In School A, Class 2 used a textbook which asked the students from the very first exercise to use alphabetic symbols "to represent numbers" (McLeod et al., 1988, p. 71), for instance, to express the amount by which 'x' is larger than 'y'. They were given practice in translating written statements into algebraic symbols to form expressions or equations. By Test 3, they were starting to evaluate expressions for given values of variables.

In Schools B and C, both Classes 4 and 7 used another textbook. In the lessons between the first and second tests, they covered a section of the text which gave them a somewhat similar start to the Manipulatives classes in that they spent time working on generalizations without using symbols. Class 4 worked from diagrams and tables rather than from geometric patterns actually built from concrete materials, whereas Class 7 used matches and small cubes to build some of the patterns. By about the fifth lesson, pronumerals were introduced by the statement "A pronumeral takes the place of a numeral" (McSeveny et al., 1989, p. 106), in the context of simplifying the writing of rules for generalizations. Both classes had practice in writing rules using symbols and calculating values from rules, and then started on another chapter which introduced several conventions for the use of symbols in algebra. By the time the third test was administered, Class 7 also had the experience of substituting values into a variety of algebraic expressions and had started graphing relationships by using sets of points.

<u>Manipulatives classes</u>. The activities used when introducing algebra with the aid of manipulatives were taken from the student worksheets given in Quinlan et al. (1989) or from the additional activities given in Appendices 3M to 3Q. For these students, the first few lessons were based on using everyday language to describe generalizations about the number of matches needed for constructing various geometric patterns. Algebraic symbols were then introduced to stand for indefinite numbers of units in the patterns, and were written into the previously-derived generalization statements which had been assembled and presented in a form similar to the worksheet in Appendix 3M. Thus when these students first met letters in algebra, they were used to represent numbers for which no particular value was nominated and they were met in a context of abstractions which described generalizations about familiar concrete geometric patterns. After this introduction, Class 9 used a length model (Unit One Worksheet Four) as a way of representing algebraic expressions, while the other six classes used an area model (Unit One Worksheet Two and Appendix 3N). All except Class 10 then used an objects-and-containers model for algebraic expressions in either one variable (Unit One Worksheet Three and Appendix 3P) and/or two variables (Unit Two Worksheet One and Appendix 3Q). Teachers

stressed that the algebraic symbols stood for numbers rather than objects when introducing each model - the number of square centimetres of area, the number of small objects inside a container, or the number of centimetres of length.

The numerical values for the symbols were chosen by the students so that, within the class, any symbol took various values simultaneously, enabling the concrete approach to avoid the pitfall of "inherent particularity ... which runs entirely opposite to the inherent generality and abstractness of algebraic statements" (Kaput, 1987, p. 352). However, the researcher became aware that most of the exercises being used involved mappings, in the Halford (1987) sense, from algebra to model rather than from model to algebra. Mappings in either direction were logically possible if the structure provided by the models was isomorphic with the structure of the type of algebraic expressions being modelled. This meant that the students were being drilled in choosing some arbitrary value for the variable so that they could then model expressions containing that variable. Teachers in School D were alerted to the lack of exercises in which the mapping went in the reverse direction and so were requested to direct students' attention to the fact that if different groups had modelled, say, '$2C + 3$' using various values for 'C', it was still true that each group had modelled '$2C + 3$'. In other words, the algebraic expression '$2C + 3$' expressed a generalization which described what each group had built. A new exercise (Appendix 3N, Question 6) was written in which students were instructed that "All answers are to be given in terms of k". It was used by Class 8 only. This was an effort to concentrate on the acceptance of generalizations expressed in algebraic form as "answers" to questions.

The Testing Program

Testing in 1990. Table 3·10 details the dates of completion of the written tests by those Year 7 students who responded to the test four times during 1990. There were 208 who completed it the first three times. On the final testing, 22 of these were absent. The absentees included students from each ability level, as detailed in Table 3·6.

The test was also given once in June 1990 to 13 classes from Schools B, C and D, spread across Years 8 to 12. In October 1990, two Year 9 classes of different ability rankings in School A completed the same test and, in December 1990, four slow-learner classes responded to it: one Year 7 class from School C, one Year 7 and two Year 8 classes from School D. In total, the test was administered once to 19 classes. Details were given in Table 3·7.

Table 3-10
Dates of Testing for Year 7 Classes in 1990

CLASS	1st Test	2nd Test	3rd Test	4th Test
1	6 Apr.	20 Apr.	27 Apr.	19 Nov.
2	6 Apr.	20 Apr.	27 Apr.	19 Nov.
3	30 Apr.	14 May	23 May	26 Nov.
4	30 Apr.	14 May	23 May	26 Nov.
5	30 Apr.	15 May	22 May	26 Nov.
6	3 May	15 May	24 May	3 Dec.
7	3 May	15 May	24 May	3 Dec.
8	4 June.	19 June.	21 June.	3 Dec.
9	4 June.	19 June.	21 June.	3 Dec.
10	5 June.	19 June.	21 June.	3 Dec.

Throughout the 1990 testing program, the writer marked the test papers to give each student a test score, coded the variety of responses to each part of each test item, and entered the coded records into a computer file, ready for analysis. The school principals, heads of the mathematics departments and teachers of the participating classes were given copies of the test scores of the students from their school, and some outcomes from the computer analyses of the responses.

Testing in 1991. A short test was given to a selection of five Advanced classes in July 1991. This follow-up testing was undertaken in response to the finding that in 1990 about 30% of Years 9 to 12 Advanced students selected "an object like a cabbage" and "an object like a pear" as options for the meaning of '*c*' in '$c + d = 10$', when responding to Item 6 (a). The paradox that these data appeared to enshrine was that these students, by giving such responses, apparently accepted that algebraic symbols could stand for physical objects, and yet worked with the notion that algebraic symbols represented numerical variables when responding to other items. Interviews were the most direct way of clarifying the thought processes which led to these responses.

Retesting on items relevant to the perplexing responses was judged necessary prior to interviewing a sample of students and so a short test, headed "New Test 1991", was assembled which consisted of Items 6 (a), 7 and 15 from the test instrument, as shown in Appendix 3R. The first item was the one which actually gave rise to the paradox when the options "an object like a cabbage" and "an object like a pear" were chosen as meanings for '*c*' by Advanced level students. Item 7 provided further data on the same basic issue of whether or not students regarded the symbols as standing for numbers or objects (in this case, for numbers of people or simply people - professors or students). The third question measured students' ability to work with the concept of a numerical variable and, as it consisted of four parts of differing difficulty, it provided measurements on a scale from zero to four.

This test was administered to five of the Advanced classes who had completed the full test in 1990, namely, Year 12 Three Unit Mathematics classes in two of the schools (Classes 54 and 55) and Year 10 Advanced Mathematics classes, one in each of three schools (Classes 33, 34, and 35). A total of 115 students were tested and 52 of these were interviewed. The outcomes are reported in Chapter 7.

The Interview Program

Interviews were integral to the research methodology in pursuing the objective of finding out more about the difficulties students experienced in understanding the meaning and use of algebraic symbols.

Interviews in 1990. The 1990 interviews concentrated on Year 7 students as beginning algebra students were the particular focus of the research. In all, 170 Year 7 interviews were conducted by the researcher with 94 different students, some of whom were interviewed two or three times.

Criteria used for selecting students for interviews were:

1. That the student sample provided a reasonable cross-section of the range of responses;

2. Important issues were embedded in student responses;

3. Students had changed their minds about some responses from test to test (for interviews conducted from Test 2 onwards) - although some students who were not making any apparent progress were also interviewed;

4. When students were grouped for an interview before the third test, they had answered most of the questions in a similar way (to avoid additional "teaching"); and

5. When interview groups were selected following the third test, some groupings included students who had contrasting views on certain issues, as opening up new ideas at this time was not a concern.

Students were generally interviewed in pairs, with just a few interviews for a group of three, and some students were interviewed individually. The advantage of having more than one student at a time was that they could interact with each other during the interview, rather than solely with the interviewer. Discussions between students revealed the way they thought and added to the information obtained by reacting to the interviewer's promptings. Grouping the students also made for more efficient use of the limited time available. Taking some students individually allowed for deeper probing of how students thought about various issues. An example of an interview in which two students interacted is as follows:

Interview extract. (Student 'Ka', Year 7, School D, after Test 1, 5 June, 1990, aged 12 years 6 months, with another student, 'Ke'. The interviewer was the writer, denoted as 'E' for Experimenter.)

This extract illustrates the occurrence of interaction between students during an interview session. Student 'Ka', at a very early stage in beginning algebra, was operating mainly at a level for understanding the meaning of algebraic symbols such that she consistently used the technique of giving sample replacement values for symbols when trying to answer general questions such as the Harper-style Question 12. She had started to show some signs of developing the concept of variable but her test responses also recorded that on seven occasions she denied symbols their true freedom as variables. The idea of a generalized number was also beginning to grow. This extract shows how she was dependent on giving '*n*' a value in order to compare the two expressions given in Question 12. Student 'Ke' had not yet grasped the notion that the algebraic symbols stood for numbers that can vary, and was really at a prestructural level, meaning that she did not understand the question under discussion.

E	All right. What did you think about Q 12? Just turn over another page.
Ke	I didn't get these ones.
E	Don't worry about it.
Ke	I just didn't get why is '$2n$' larger?
Ka	Because '$2n$' stands for 2 times '*n*', and '$n + 2$' is just '*n* ' plus '2'. So say '*n*' stands for 3, then 2 times '*n*' is 6, and 3 + 2 is five, so '$2n$' is larger than '$n + 2$'. Then say the '*n*' stands for 1, 2 times 1 is 2, and then 1 + 2 is 3, so '$n + 2$' is larger.
E	Yes, you put those down here. So you are thinking about using different numbers?
Ka	Yes, but it depends what the '*n*' stands for, before you can say which is the larger.
Ke	Yes, we need a value for '*n*'.

Two tables are presented below, one to summarize the timing of the interviews and the other, the number of interviews taken per student.

Table 3·11 shows that 28 students were interviewed soon after they had completed the test for the first time. These interviews were to ensure that the researcher was clear about certain misconceptions these students had expressed in their responses to the test items before they started classroom algebra. Following Test 2, 48 interviews were conducted; following Test 3, another 44; and, after Test 4, another 50.

Table 3-11

Frequencies of 1990 Interviews After Each Year 7 Test

CLASS	AFTER TEST NUMBER				TOTAL INTERVIEWS
	1	2	3	4	
1		8	8	8	24
2				8	8
3		4	4	6	14
4		6	4	3	13
5		7	4	3	14
6		4	7	4	15
7		4	6	4	14
8	11	4	6	6	27
9	9	5	5	5	24
10	8	6		3	17
TOTAL INTERVIEWS	28	48	44	50	170

Interviews prior to the fourth test were recorded on audio-tape, except for two cases when students preferred not to be taped, and interviews following the delayed posttest were video-taped, after obtaining parental permission. Transcripts of all interviews were organized.

Mapping exercises similar to those in Part I of Appendix 3P were used in delayed posttest interviews. Students were presented with several modelled examples of a particular algebraic expression or identity and were asked to express in algebraic terms the generalizations represented by the models. These exercises involved mappings from models to algebra and their importance was recognized as a result of insights gained while carrying out the research. They highlighted the power of algebra to express generalizations succinctly. Additions to Unit One Worksheet Three in the form of Questions 13 and 14, as presented in Appendix 3P, Part II, were written after the final 1990 interviews and thus were not used with whole classes during the research period.

Based on the 50 video-taped interviews, a pilot investigation of long-term effects of manipulatives indicated that students from Manipulatives classes were not dependent on the manipulatives they had used when being introduced to algebra some six months previously but could still make good use of them for self-correction. As this was a minor project, only a brief comment is made here and further details are in Appendix 3P, Part III.

As Table 3·12 shows, of the 94 students interviewed, 32 were interviewed twice and another 22 were interviewed three times. Analyses of the findings from these interviews contributed valuable information to the research objectives by clarifying the meanings behind some of the written responses given in the testing process and by expanding on ideas not fully expressed in written form.

Table 3·12
Frequencies of 1990 Interviews Per Year 7 Student

CLASS	NO.of INTERVIEWS per STUDENT			TOTAL STUDENTS
	1	2	3	
1			8	8
2	8			8
3	4	2	2	8
4	5	4		9
5	4	5		9
6	5	2	2	9
7		4	2	6
8	6	6	3	15
9	5	5	3	13
10	3	4	2	9
TOTAL STUDENTS	40	32	22	94

Interviews in 1991. The final round of interviews was conducted in July 1991 as a means of elucidating the paradox mentioned above, namely, why Advanced students who managed the concept of a numerical variable would claim that an algebraic symbol, set in an arithmetic context, could represent an object like a cabbage or a pear. Table 3·13 sets out the distribution of the students interviewed in terms of their classes and whether or not they had chosen the options "an object like a cabbage" and "an object like a pear" from Item 6 (a) of their test (Appendix 3R). As with all the other analyses of data, the discussion of the findings about the paradox occurs in a later chapter, in this case, Chapter 7.

Table 3·13
Frequency Distribution of Students Interviewed in 1991

SCHOOL	B	C	D	C	D	TOTAL STUDENTS
CLASS	33	34	35	54	55	
YEAR LEVEL 1991	10	10	10	12	12	
Interviewees chose objects	3	7	8	8	2	28
Interviewees did not choose objects	2	7	3	5	7	24
Total Interviewed	5	14	11	13	9	52
Total chose objects	3	7	11	9	2	32
Total did not choose objects	20	24	14	16	9	83
Total Tested	23	31	25	25	11	115

Review and Forecast

This chapter contains a description of preliminary investigations, undertaken in late 1989 and early 1990, for determining an appropriate methodology and for devising a research test instrument capable of providing data relevant to the stated objectives and the intended investigations. The methodology has been indicated and the rationale for this methodology was explained in terms of the research objectives and the planned investigations. The rationale for choosing the final test items and the history of their development have been described. The benefits accruing from the trial teaching interventions and their influences on the main research program have been delineated. Using a descriptive rather than an analytic mode, the major factual aspects of the main research program have been supplied in terms of the schedule of events, the student samples, the teaching interventions, the testing, and the interviewing.

Chapter 4 assembles descriptions of the data collected by the operation of the research methodology. The chapters which follow report on various levels of analysis to which the data were subject and the interpretation of the outcomes of these analyses.

CHAPTER 4

STUDENT RESPONSES: A GLOBAL VIEW

Overview

The major objective of this chapter is to present a global view of the vast array of empirical data obtained in the form of students' responses to the newly-designed test instrument. Response types are organized and summarized in considerable detail for two main reasons: firstly, to describe the essential ingredients for the analyses discussed in later chapters and, second, to record with clarity this expansion of the data base for research in the area. Types of responses to each test item are discussed and tables for the participants as a whole give not only item facility levels (proportions correct) but also percentage frequencies of each class of response.

The view presented now does not comment on the characteristics of responses by particular subgroups within the total body of 517 students in the main research program. Treatment of the various research investigations in later chapters will identify and discuss variations between subgroups and comparisons with data provided by the Year 7 students when they completed the test for the first, second, third, and fourth times. These variations in performance are summarized in Chapter 8 and its appendices, not in the form of tables of percentage frequencies (as given in this chapter for all 517 students), but as graphs of average scale scores. The discussion of the formation of scales from the responses is given in Chapter 5, and tables showing the allocation of response types to scales are presented in Appendix 5F.

For each test item, background to its inclusion in the test is given in terms of outcomes from trialling, frequency distributions are presented in category and ordinal forms of data from the main research project, and comparisons are made, if possible, with data obtained by other researchers.

The chapter begins by describing the coding procedures used for the management of the data. Frequency outcomes from this coding are then discussed.

Student Responses

Responses by 208 Year 7 students on their third testing were used in assembling the frequency distributions reported in this chapter. These were the responses they gave following their first few weeks of classroom algebra when they knew at least something about the ideas being tested, although they were far from being experts in the field. Table 3·6 (p. 77) describes the classes of these Year 7 students and Table 3·10 (p. 83) lists the dates of their testing program. The responses of 309 other participants have also been included in the frequency distributions. These were the students from Years 7 to 12 who completed the test only once, as reported in Chapter 3. Table 3·7 (p. 78) provides descriptive details of their classes.

A copy of the test appears in Appendix 3I.

Coding the Responses

Three coding approaches were used for data management: one for calculating an overall test score, one for recording the categories of answer types, and the third for condensing the categories into ordinal form.

Coding for test score. Firstly, responses were classified as either correct or incorrect in order to allocate a total test score to each student. The name "Score" was given to this variable, the maximum value for which was 65. Responses were treated as ordinal variables in tallying the scores. For most items, the dichotomy between correct and incorrect answers resulted in values of either "1" for correct or "0" for incorrect. There were, however, the following three cases in which the ordinal scoring was a little different:

1. The four parts of Item 15 were considered to be useful indicators of the degree of students' ability for working with the important concept of a numerical variable. Therefore, in this case, "2" points were allocated for answers which correctly stated when the given expressions were equal and "0" points otherwise.

2. For part (c) of Item 6, the ordinal range was extended by recording "3" for the best answers, namely, '*c*' was described as "less than 5" and specific mention was made of at least one of the facts that it could be zero or negative or fractional; "2" for answers which gave '*c*' as "0, 1, 2, 3 or 4", "0, 1, 2, 3, 4, 5", or "less than 5" without mention of the zero, negative or fractional possibilities; "1" for writing that '*c*' was "1, 2, 3 or 4" or "1, 2, 3, 4 or 5"; and "0" for other answers.

3. In Items 10, 11, 12, and 13, a score out of 4 was decided by assessing the answers to the four parts of each question as a unit.

Data coded in this ordinal form were considered appropriate for use in statistical analyses such as factor analyses.

It was necessary to distinguish between "omit" and "missing" categories of responses. There were several items (e.g., Item 15) in which incorrect responses to one part of the question implied a lack of a written response to some other part or parts. Care was taken not to exclude students from analyses on this account. On each follow-up part they were given a score of "0" in the ordinal scale to register the fact that they had really given an incorrect answer to that part. The score of "0" was the "omit" category. Even if students had correctly answered one part but had left blank an associated part, again a score of "0" was allocated for the omitted response. Such procedures in scoring were important steps in retaining subjects in later analyses which involved such variables alongside other variables. In this way, they were not excluded from further analyses involving that part of the question by being classified as "missing", that is, as candidates who had not expressed an opinion.

Category coding. Category data were assembled, coded and entered into a computer file to describe the variety of answers obtained to each of the questions.

Some questions, such as each part of Item 6 (a), required only three categories to cover all response types, namely, "Blank" for omit, "1" if incorrect and "2" if correct. Most questions required more categories to preserve the wealth of information supplied by the student responses. For instance, 10 categories were used to keep a record of the range of answers to the following questions: the second and third parts of Item 4, each part of Item 7, and each part of Items 10, 11, 12, and 13. Many more categories were obtained in other questions when, as in Item 9 and the last two parts of Item 14, the actual values of numerical answers were recorded, even though all of these answers were inappropriate. In the case of Item 14, the variety of algebraic answers which resulted was also classified by even more categories.

Responses to each part of Item 11, the Parallel Lines Task, serve as an example of 10 categories derived from just one question:

1 repeats question, e.g., in part (b), "When the red line is shorter";
2 no idea, e.g., talk about the thickness of the lines or their colour;
3 focus on geometry, e.g., perspective, leaning lines;
4 mixture of algebra and geometry;
5 incorrect algebra, e.g., 'a' and 'b' must stand for different numbers;
6 correct algebra, e.g., in part (b), "If a > b";
7 "always" or "now";
8 "never"
9 "I don't know how long the lines are";
10 "blank" if the part was omitted.

Only some statistical procedures were appropriate for data in the category form of coding. Frequency counts of the responses in each category were helpful in probing the ways that students thought about the basic ideas of algebra. Cross-tabulations were also helpful, for example, in determining the degree of persistence of certain viewpoints from one test to the next.

Ordinal coding. The third mode of coding used in managing the data was achieved by judging the relative merits of different categories of answers and then allocating numerical values to the categories for producing a set of ordinal variables. For most items, there were only three scale scores, namely:

"2" for correct answers,

"1" for incorrect answers, and

"0" for "omit", as distinct from "missing".

In some items, however, the range of scores was extended. In such cases, the allocation of scores will be explained below as each item is treated in turn. For the categories listed above for Item 11, an ordinal scale was applied as follows:

"4" points for category 6;

"3" points for categories 4, 5 and 9;

"2" points for categories 3, 7 or 8;

"1" point for categories 2 and 1;

"0" for omitting a certain part, provided that at least one other part had been answered; and

"Blank" or "Missing" if none of the item parts had been answered.

Frequency Distribution of Test Scores

Table 4·1 summarizes the overall test scores in a grouped frequency table and Figure 4·1 presents the resultant distribution in the form of a histogram.

Table 4·1

Grouped Frequency Data for Total Test Scores

FREQUENCY RANGE	PERCENTAGE
1 - 10	8.5
11 - 20	22.3
21 - 30	20.7
31 - 40	20.3
41 - 50	21.4
51 - 60	6.8

Note. $N = 517$. Year 7 scores are from Test 3 for those who did the test more than once. Mean = 30.06, standard deviation = 14.16, 1st quartile = 18, 2nd quartile (median) = 30, 3rd quartile = 42.

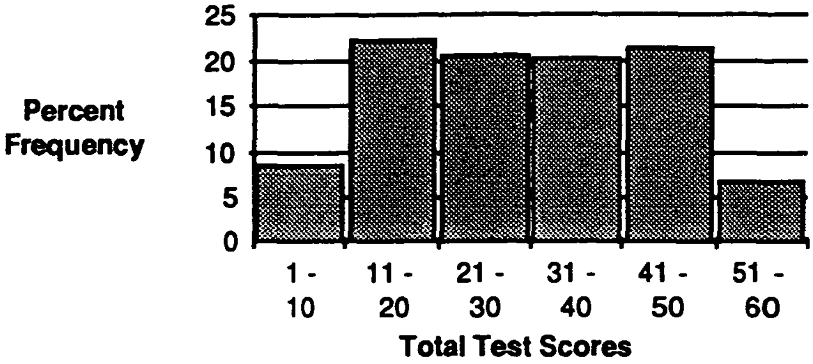


Figure 4·1. Percentage frequency histogram for total test scores (N = 517, using Test 3 for Year 7 students)

Attention is drawn at this stage simply to the large spread of scores. Closer investigations of the composition of the scores and distributions for various subsets of students are presented in later chapters. The maximum possible score was 65 and the actual range of scores was from 1 to 60.

Individual Items

Frequencies of the response types for each question are now presented by means of frequency tables based on category coding and graphs based on ordinal coding. As mentioned earlier, the frequency distributions are based on responses of all 517 students, taking Test 3 responses of Year 7 students who did the test more than once. The tables define the allocation of ordinal scale scores to the various categories and explanatory comments accompany them. The background leading to the inclusion of the items is given in terms of the trialling outcomes. General comparisons are made between the new frequency data and corresponding data from earlier research projects for those test items which were based on the earlier research.

Item 1

1. Look at this: 3 * 4 = 6 * y
Tick the correct answers below.

(a) * could be ADD (+): .... yes ...... no .... can't tell

If you ticked "yes", then y must be ....

(b) * could be TIMES (x): ...... yes ...... no .... can't tell

If you ticked "yes", then y must be ....

(Collis 1975a, adapted)

This item was taken from a four-part question used in the study by Collis (1975a) and discussed in Chapter 2, pages 40 and 41. Responses from trialling students are summarized and discussed in Appendices 3F (Table 3F·6), 3H (Figure 3H·1) and 3J (Table 3J·1). Although only slightly less than one-third of the Year 7 beginners were successful, about three-quarters of the Years 8 and 9 students had both parts correct. The question was found to be useful in the present context for measuring the ability of students to work with symbols, namely, '*' for an arithmetic process and 'y' for a number-to-be-found, and to be flexible in their thinking by considering both addition and multiplication possibilities for the same equation. Moreover, the values given for 'y' revealed the ways they had thought about the problems.

Tables 4·2 and 4·3 summarize the response frequencies for students in the main study.

Table 4·2

Percentage Frequencies of Responses to Subparts (i) of Item 1

CATEGORY DATA			ORDINAL DATA		
Response Type	% 1a i (ADD)	% 1b i (TIMES)	Score	% 1a i (ADD)	% 1b i (TIMES)
Omit	5.4	5.2	Missing	5.4	5.2
Can't tell	6.8	5.2	1	see below	see below
No	12.4	14.5	1 (total)	19.1	19.7
Yes	75.4	75.0	2	75.4	75.0

Note. $N = 517$. "see below" indicates where to find % tally for given score on ordinal scale. Subpart (i) asks about meanings for '*'.

As indicated in Table 4·2, just over 5% of students completely omitted the question. Three-quarters of the students correctly chose "Yes" from the options given in subpart (i) of both parts, and were allocated an ordinal score of "2": 75.4% for (a) and 75.0% for (b). Responses in the two erroneous categories, "Can't tell" and "No", were grouped in the ordinal scale under the common score of "1", a score which accounted for 19.1% of cases in part (a) and 19.7% in part (b). The 19.7% for 1(b)(i) was the sum of the tabulated 5.2% ("Can't tell") and 14.5% ("No"). The 19.1% for 1(a)(i) was obtained in a similar way, as 6.8% plus 12.4%, allowing for the rounding off of the smaller percentages.

Almost one-quarter of the students omitted subpart (ii) after answering subpart (i), as shown in Table 4·3. In the case of those students who chose "No" or "Can't tell" in subpart (i), logically they were not required to answer the follow-up part asking for the value of 'y'. These students were regarded as having given an incorrect answer to the follow-up part of the question and were thus allocated the "omit" score of "0". The score of "0" was also allocated to those who correctly chose "Yes" as their first response but then did not write a value for 'y', since such an omission indicated an inability to work out the 'y' value after having made an attempt to answer the question. On the other hand, those who did not answer any aspect of, say, part (a) of the item were relegated to the "missing" category and were excluded from any further analyses involving that part of the item. A similar approach was used for the other items composed of interrelated parts.

Table 4·3

Percentage Frequencies of Responses to Subparts (ii) of Item 1

CATEGORY DATA			ORDINAL DATA		
Response Type	% 1a ii (ADD)	% 1b ii (TIMES)	Score	% 1a ii (ADD)	% 1b ii (TIMES)
Omit i & ii	5.4	5.2	Missing	5.4	5.2
Omit ii, not i	24.6	24.4	0	24.6	24.4
Algebraic	0.2	0.2	1	see below	see below
Incorrect nos.: 2,4,6,7,8,9,12	4.6	-	1 (total)	4.8	-
Incorrect nos.: 1,6,7,8,10,12,16	-	8.5	1 (total)	-	8.7
Correct No. i: 1; ii: 2	65.2	61.7	2	65.2	61.7

Note. $N = 517$. ' - ' denotes "Not Applicable". "see below" indicates where to find % tally for given score on ordinal scale. Subpart (ii) asks about meanings for 'y'.

About 5% did not answer subpart (ii) after choosing "Yes" in subpart (i), as can be deduced by subtracting the total percentage who gave some answer to (ii) from the percentage who answered "Yes" in part (i), specifically, in the case of part (a),

the percentage who answered part (ii) = 0.2 + 4.6 + 65.2 = 70.0, and

the percentage who answered "Yes" to subpart (i) = 75.4 (Table 4·2).

Hence, the percentage of those who omitted (ii) after choosing "Yes" for (i) is 5.4% (= 75.4 - 70.0), assuming that those who chose "Can't tell" or "No" logically did not answer (ii).

In part (a), 4.6% gave incorrect values for 'y' and 8.5% did the same in part (b). One student (0.2%) gave algebraic responses to each part. Those who gave an incorrect value for 'y' or an algebraic response merited a score of "1" on the ordinal scale and those who wrote the correct value were scored at "2".

The most common incorrect 'y' value in part (b) was "6", given by 4.8% who probably thought along the lines

3 x 4 = 12 and 6 + 6 = 12,

with the '*' representing multiplication for the left-hand side of the equation and addition for the right-hand side.

Figure 4·2 reports the response rates on Item 1 in graphic form using the ordinal scale, as set out in Tables 4·2 and 4·3, whereby "2" was the score for correct answers, "1" for incorrect answers, and "0" for omitted answers. The item outcomes confirmed that over 60% of the students were able to understand the arithmetical processes involved and to apply appropriate procedures to evaluate 'y' successfully.

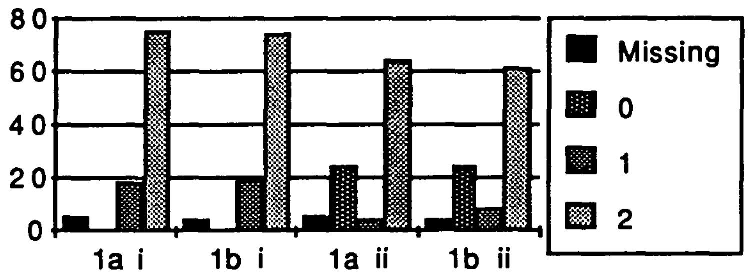


Figure 4·2. Percentage frequencies of responses to Item 1 when grouped to form an ordinal scale (N = 517)
[a: re. '+'; b: re. 'x'; i: re. '*'; ii: re. 'y']

Comparison with Collis (1975a). Consideration was limited to the possibilities of addition and multiplication whereas the Collis (1975a) question, on which it was based, took all four arithmetical operations in turn. Collis reported overall performance statistics for his version of the question without a breakdown for each operation. Of 60 students spread across ages 10 to 15 years, 47 (nearly 80%) were classified as successful (a deduction from Table 3.2, Collis, 1975a, p. 58). The average success rate was lower in the present study, with 517 students aged from 12 to 17 years, as only 53.8% (a figure not reported in Table 4·3) had the value of 'y' correct for both parts of Item 1.

Item 2

2. (i) If a and d are any two numbers, which, if either, is the bigger? Give a reason for your answer.

....

(ii) If y and d are two positive numbers and 6y = d, which is the bigger number, y or d? ....

(MacGregor, 1989, adapted)

The trialling outcomes for the two parts of this item, which are discussed in Appendices 3D (Figure 3D·10), 3H (Figure 3H·9) and 3J (Table 3J·1), showed that this item was able to provide data which could be useful in following up MacGregor's work (1989) by investigating the possible influence of an abstract or a real-life context on judgements about relationships between variables. Reversal of the relative sizes of variables was found, in trialling, to be less common in this item, set in an abstract context, than in Question 7, a corresponding item presented in a real-life setting.

Table 4·4 shows how ordinal scores were allocated to response categories and reports the frequency distributions of responses obtained in the main study.

Table 4·4

Percentage Frequencies of Responses to Item 2

CATEGORY DATA		ORDINAL DATA	
Response Type	**% 2i a**	**Score**	**% 2i a**
Omit	21.3	Missing	21.3
a	5.8	1	see below
d	10.8	1	see below
other error	1.2	1 (total)	17.8
same	2.3	2	see below
not sure	2.5	2 (total)	4.8
neither	56.1	3	56.1
Response Type	**% 2i b**	**Score**	**% 2i b**
Omit a & b	21.1	Missing	21.1
Omit b, not a	4.8	0	4.8
irrelevant	5.0	1	see below
alphabetic	13.0	1 (total)	18.0
neither	56.1	2	56.1
Response Type	**% 2ii**	**Score**	**% 2ii**
Omit	10.1	Missing	10.1
6*y*	0.8	1	see below
"none" or "both"	2.4	1	see below
y	12.8	1 (total)	15.9
d	74.1	2	74.1

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.

Both parts of Item 2 asked students to relate the sizes of two variables, testing their ability to regard algebraic symbols as representing more than one possible number. A specific relationship was not given in part (i) but a multiplicative relationship was specified in part (ii). Part (i) resulted in useful information about student views regarding letters in algebra, such as deciding the order of size by reference to the alphabet. The question produced information about students' acceptance of letters as numerical variables and, in part (ii), their ability to work with a covarying pair of unknowns.

In part (i), students were first asked, in what is now called subpart (a), to indicate the letter standing for the bigger number, and then, in what is here referred to as subpart (b), were asked to give a reason for their first answer. For the subpart (a) response, the range of the ordinal scale was extended. Those who clearly stated that neither '*a*' nor '*d*' was the larger (as their values were not given) were allocated a score of "3", whereas a score of "2" was allocated to those whose answers were border-line to being correct. The latter group's responses included "Unsure" or "They are the same". Those with clearly incorrect answers were given a score of "1". In subpart (b) of part (i) and in part (ii), the standard three ordinal scores were used, namely, "0" for omit, "1" for an incorrect response, and "2" for the correct response.

Slightly more than half (56.1%) were able to reason correctly about the variables in the open-ended case given in part (i) of the item. The 16.6% favouring either the '*a*' or the '*d*' in subpart (a) included 13.0% who did so by quoting reference to the order of the letters in the alphabet in subpart (b). Two other students argued for '*a*' as being the better "grade".

Figure 4·3 depicts the frequencies of the scores on the ordinal scale which was derived from the categories of responses as set out in Table 4·4.

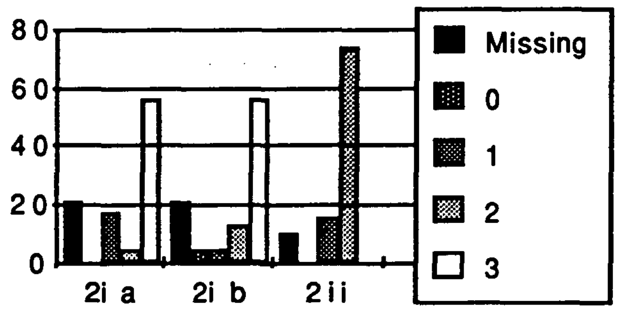


Figure 4·3. Percentage frequencies of responses to Item 2
when grouped to form an ordinal scale ($N = 517$)

Comparison with MacGregor (1989). Part (ii) was a revision of a question, referred to in Chapter 2 and used by MacGregor (p. 90). The inclusion of the words " If y and d are two positive numbers" eliminated the uncertainty noted by MacGregor (p. 121) about whether or not students should be expected to consider negative or zero possibilities. Nearly three-quarters (74.1%) of the students gave correct answers, while 12.8% made the reversal error by claiming that '*y*' was larger than '*d*'.

MacGregor had reported that, despite the disparity between the age groups of the participating subjects, a comparable 13.2% of her 235 Year 9 subjects made the same reversal error while 77.9% were correct (p. 122). The last of the propositions treated in Chapter 8 discusses the success rate on this item compared with the rate on Question 7, the professors-and-students problem, and a comment comparing the frequencies of the reversal error for the same two questions is made at the end of Chapter 7.

Item 3

3.	If y = 3, what is the value of	(i)	2y ?	....
		(ii)	2y + 5 ?	....
		(iii)	2(y + 5) ?	....
		(iv)	2y + y ?	....
		(v)	3y - y ?	....
		(vi)	2(5y) ?	....

Outcomes from trialling are presented in Appendices 3D (Figure 3D·6), 3G (Comments on Question "5"), 3H (Figure 3H·7) and 3J (Table 3J·1). Except for the Year 7 beginners, the majority of students succeeded with each part. The two parts featuring brackets were generally found harder than the other parts. The last part was included after an extended discussion with one of the Year 9 1990 students of School X, as reported in Appendix 3G.

The numerical answers given to the six parts of the question were suitable for measuring the students' understanding of basic conventions for the use of algebraic symbols such as conjoining for multiplication and determining the order of operations by means of brackets. The questions took the students from algebra to arithmetic and did not require an understanding of the concept of variable.

All numerical answers were recorded to help identify categories of responses. The numbers in answers revealed the methods used by the students when interpreting the given expressions and evaluating them. For instance, the most common incorrect numerical answer to part (i) was '5' for '2*y*', indicating that the 2.7% of students who gave this answer had added '2' and '3' instead of multiplying them.

Table 4·5 summarizes the major features of the response types in the main study for the first three parts of the item, with the last three parts appearing in Table 4·6.

Only those who omitted all six parts of the item were classified as "Missing", and those who omitted some parts were given a score of "0" on the Ordinal Scale for whatever part(s) they omitted. A score of "1" was allocated to answers retaining the symbol 'y', "2" to incorrect numerical answers, and "3" to the correct numerical answers. For both parts (i) and (ii), the success rate was high, being above 85%. The brackets in part (iii) made it the hardest part of the item for some students yet more than three-quarters of them succeeded in writing '16'. About 7% of students left 'y' as part of the answer.

Table 4·5

Percentage Frequencies of Responses to Item 3 parts (i) to (iii)

CATEGORY DATA				ORDINAL DATA			
Response Type	% 3i	% 3ii	% 3iii	Score	% 3i	% 3ii	% 3iii
Omit i to vi	1.0	1.0	1.0	Missing	1.0	1.0	1.0
Omit this part	-	-	1.3	0	-	-	1.3
Algebraic	7.4	6.8	7.0	1	7.4	6.8	7.0
Add not mult. (i 5,ii 10,iii 10[a])	2.7	2.5	3.9	2	see below	see below	see below
Part use of distrib.law (iii 11 or 13)	-	-	3.5	2	-	-	see below
Place value[#] $2y = 20 + y$ (i 23,ii 28,iii 28)	1.5	1.6	0.4	2	see below	see below	see below
Other incorrect	1.2	2.7	5.8	2 (total)	5.4	6.8	13.5
Correct	86.3	85.5	77.2	3	86.3	85.5	77.2

Note. $N = 517$. ' - ' denotes "Not Applicable". "see below" indicates where to find % tally for given score on ordinal scale. [a] the result '10' in part iii could have been obtained by ignoring 'y'. [#] place value errors are discussed on page 169 in Chapter 5.

Table 4·6 records the percentage frequencies for the last three parts of Item 3. The last part caused nearly as much difficulty as did the earlier part with brackets, with just under 80% being successful. Finding such problems with brackets corresponded with the results obtained otherwise by Booth (1983). Parts (iv) and (v) were found to be a little more difficult than the first two parts of the item.

Overall, achievement was high, possibly because the question took the students away from algebra back into the more familiar scene of arithmetic. To be successful, however, they needed to understand the conventions for writing first degree expressions in algebra. The question tested this aspect of the use of symbols. Figure 4·4 displays the high success rates achieved by the students in managing the skills necessary for this item. This outcome suggested that a closer look at the rapid progress in the skill of substitution by the beginning students could be profitable. Responses to

Item 3 were used in analyses of hierarchies of cognitive difficulty, as reported in Chapter 8.

Table 4·6

Percentage Frequencies of Responses to Item 3 parts (iv) to (vi)

CATEGORY DATA				ORDINAL DATA			
Response Type	% 3iv	% 3v	% 3vi	Score	% 3iv	% 3v	% 3vi
Omit i to vi	1.0	1.0	1.0	Missing	1.0	1.0	1.0
Omit this part	0.2	1.0	2.5	0	0.2	1.0	2.5
Algebraic	7.5	6.6	7.0	1	7.5	6.6	7.0
Add not mult. (iv 8,v 3[a];vi 10[b])	2.3	3.7	1.9	2	see below	see below	see below
Misuse of distrib.law (vi 60)	-	-	1.4	2	-	-	see below
Place value[#] (iv 26,v 30, vi 28 or 106 or 253 or 630)	1.5	1.2	1.4	2	see below	see below	see below
Incorrect operation(s): (iv 12 or 18,v 12, vi 16 or 17)	2.1	2.7	2.7	2	see below	see below	see below
Other incorrect	2.1	3.3	2.7	2 (total)	8.1	10.8	10.1
Correct	83.2	80.7	79.5	3	83.2	80.7	79.5

Note. $N = 517$. ' - ' denotes "Not Applicable". "see below" indicates where to find % tally for given score on ordinal scale. [a] the result '3' in part v could have been obtained by "cancelling" 'y'. [b] the result '10' in part vi could have been obtained by ignoring 'y'. [#] place value errors are discussed on page 169 in Chapter 5.

Figure 4·4 portrays the frequencies of ordinal scores on each part of Question 3.

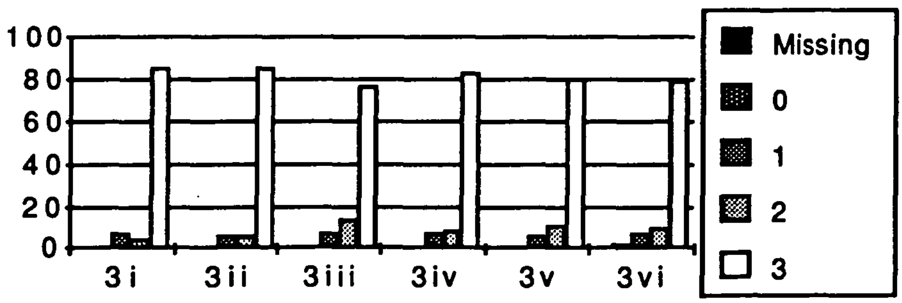


Figure 4·4. Percentage frequencies of responses to Item 3 when grouped to form an ordinal scale (N = 517)

Item 4

4. **If the expression 4g + 8 represents a number of flowers, could 4g represent the number of flowers in 4 same-sized bunches of flowers ? YES / NO.**

If YES, the **g** represents ....

If YES, the **8** represents ....

The wording of this item was changed from "4 equal bunches" (in the test "Brain-Box Quiz No.1") to "4 same-sized bunches" so as to avoid the possibility of the word "equal" being a source of misinterpretation of the question. The revised version was trialled in "Brain-Box Quiz No.2", "New Test 2 1990" and "Algebra Project New Test 1990". In the 1989 trialling, reported in Appendix 3D (Figure 3D·15a), only about 5% of Groups I to V were correct in their interpretation of '*g*', regardless of which form of wording was used. When some of the same students were re-tested in 1990, only about 10% gave a correct interpretation of '*g*', as Appendices 3G and 3H record. The Year 7 beginners (Appendix 3J, Table 3J·1) could not make any sense of the question. This question was kept even after the trialling had indicated that students found it difficult. It was one way of inviting students to record the variety of ways in which they interpreted algebraic symbols in a real-life context.

In the main study, apart from the first response, the success rate was low, with 10.1% correct on the second response and 20.5% correct on the last response. The one-in-ten students who correctly answered the middle part gave the meaning of '*g*' as the number of flowers in a bunch, thus showing a clear understanding of the meaning and use of symbols in the given real-life setting. Table 4·7 summarizes the response frequencies.

Almost half the students omitted answers to the second and third parts of Item 4, although only one-quarter had excluded themselves by choosing the "No" option in the first part. The meaning of '*g*' was described in terms of objects rather than numbers of objects by about one-fifth of the students, showing that they did not clearly recognize that, in the given context, the algebraic symbols had a numerical connotation. Others gave an incorrect number meaning such as "the number of bunches" or "any number of flowers", rather than the desired response, "the number of flowers in a bunch". Ten students claimed that '*g*' stood for 'grams', two wrote "the number of grams", and two more used the alphabetic code which gave '*g*' a value of '7'. As regards the meaning of the '8', fewer students wrote in terms of objects, while 16.2% gave an incorrect number meaning such as "the number in a bunch" or simply "any number", rather than relate it to the given situation as, say, "8 flowers not in bunches", or "8 extra flowers".

Table 4·7

Percentage Frequencies of Responses to Item 4

CATEGORY DATA				ORDINAL DATA			
Response Type	% 4a	% 4b	% 4c	Score	% 4a	% 4b	% 4c
Omit a to c	13.3	13.3	13.3	Missing	13.3	13.3	13.3
Omit this part	1.5	33.1	36.6	0	1.5	33.1	36.6
'g' or '8' = flowers or bunches	-	20.1	7.7	1	see below	see below	see below
Other incorrect (a No; b, c not a number)	26.9	9.5	5.6	1 (total)	26.9	29.6	13.3
Number, but incorrect	-	13.9	16.2	2	-	13.9	16.2
Correct part a (Yes)	58.2	-	-	2	58.2	-	-
Correct parts b, c	-	10.1	20.5	3	-	10.1	20.5

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.
' - ' denotes "Not Applicable".

Figure 4·5 displays the ordinal scores. These scores were graded in parts (b) and (c) so that "3" was allocated for correct answers, "2" for answers that were incorrect but gave a numerical interpretation for 'g' or '8', and "1" for incorrect answers which gave interpretations in terms of objects.

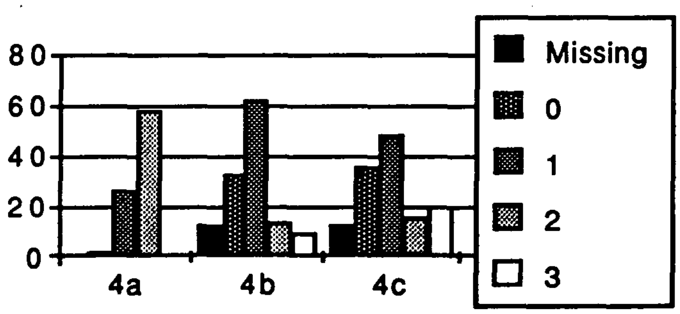


Figure 4·5. Percentage frequencies of responses to Item 4
when grouped to form an ordinal scale ($N = 517$)

Item 5

5. In a football match, one team scored **p** points and the other scored **r** points

How many points altogether were scored in the match? ....

(Booth 1983, adapted)

Item 5 was an adaptation of Booth's (1983) question "West Ham scored x goals and Manchester United scored y goals. What can you write for the number of goals scored altogether?" (p. 350). In early trialling, the wording was adjusted according to current sporting events but was generalized for the final version. As Figure 3D·3 in Appendix 3D records, about 70% of the 1989 trial students were successful, about 20% gave a conjoined answer, and the rest gave some number. Only 3 of the 19 beginners from the 1990 Year 7 trial group succeeded (Appendix 3J, Table 3J·1).

The item gave a measure of students' readiness to accept lack of closure by writing symbols in an answer and performing the addition operation on them even though their values were unknown. Hence, the item measured aspects of the ways students viewed the symbols. It also measured their degree of skill in understanding that the convention '$p + r$' expresses the sum of 'p' and 'r', whereas the conjoined form, 'pr', is the conventional way of writing the product of 'p' and 'r'. Those who wrote 'pr' perhaps indicated a degree of lack of closure, wanting to express an "answer" for the sum of 'p' and 'r'. Students showing the strongest tendency to seek closure gave arbitrary numerical answers to the question.

Table 4·8 summarizes the outcomes and defines the ordinal scale derived from the responses to Question 5.

Table 4·8

Percentage Frequencies of Responses to Item 5

CATEGORY DATA		ORDINAL DATA	
Response Type	% 5	Score	% 5
Omit	8.9	Missing	8.9
Uncertainty or other error	6.4	1	see below
alphabetic: $p+r = 34$	2.1	1	see below
Other number	7.5	1 (total)	16.1
pr	11.8	2	11.8
$p + r$	63.2	3	63.2

Note. $N = 517$. "see below" indicates where to find % tally on ordinal scale.

The ordinal scale scores are presented in Figure 4·6, showing clearly that over 60% of the students succeeded in stating that '$p + r$' (score of "3") was the acceptable way to describe the total number of points in the football match described in the question. Incorrect answers which were algebraic (viz., 'pr') merited a score of "2", and other incorrect answers were given a score of "1".

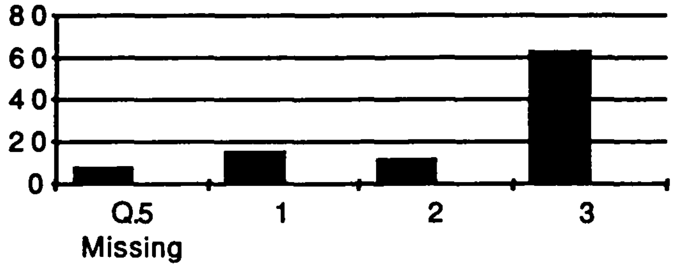


Figure 4·6. Percentage frequencies of responses to Item 5 when grouped to form an ordinal scale ($N = 517$)

Comparison with Booth (1983). Nearly 10% of the students tested in the main research study sought closure by expressing the total score as a number rather than in symbols. Some of these said that '$p = 16$' and '$r = 18$', using the position of the given letter-symbols in the alphabet, and so gave the score as '34'. Over 10% made the error of conjoining to write 'pr' rather than '$p + r$'. Booth (1983, p. 137) also reported similar errors in responses to the question from which Item 5 was derived when she interviewed 28 students who were experiencing difficulty with algebra.

Item 6

6. (a) If **c + d = 10**, tick ALL the meanings that **c** could have:

3 10 12 7.4 the number of apples in a box

an object like a cabbage an object like a pear

(b) If **c + d = 10**, what happens to **d** as **c** gets bigger?

....

(c) If **c + d = 10**, and **c** is always less than **d**, what values may **c** have?

....

(Harper, 1979, adapted, & Küchemann, 1980, adapted)

The origin of this item is two-fold. It is based on the following question from the CSMS study:

"What can you say about c if $c + d = 10$ and c is less than d ?"
(Küchemann, 1980, p. 67)

In addition, the question was an adaptation of one of Harper's (1979) Equations Tasks, namely, the task in which students were asked to compare the values of 'x' and

'y' in the equation '$x + y = 10$'. It paralleled the Harper question "When is the value of 'x' less than the value of 'y'?"

Parts (a) and (b) were added to direct students' attention to the fact that their values could include zero, fractional and negative possibilities, and to the covarying relationship between the two variables in the given equation.

The format and the selection of options for part (a) were the outcome of trialling several similar questions in 1989, as recorded in Appendices 3B, 3C and 3D. Part (a) proved useful as a measure of the types of possible meanings students were prepared to accept for alphabetic symbols in algebra. As Figure 3D·14a in Appendix 3D reports, the given number options for 'c' were accepted in the following order of popularity: '3', '7.4', '10', and '12', with '12' being by far the least popular. Investigating reasons for such variations was considered to be possible material for the main study. This item proved valuable also in producing another outcome noted as worthy of further study, namely, that about 20% chose as meanings for 'c' objects like a cabbage or a pear. This aspect is investigated in Chapter 7.

A total of nine responses were requested in Item 6. For data from the main study, these are reported in four stages, starting with the first five answers to part (a). Students who selected at least one of the seven options in part (a) were considered to have made a judgement about each option so that the only students recorded as "Missing" for any sub-part of part (a) were those who did not select any of the options.

Table 4·9 reports the frequencies of the responses to parts (i) to (v) of Item 6 (a) and the frequency distribution of the ordinal scores is displayed in Figure 4·7. A score of "2" was allocated on the ordinal scale whenever a student selected one of the first five options, these being the options in which numerical interpretations were given as possible meanings for 'c' in the equation '$c + d = 10$'. Those who rejected any of these options were given a score of "1".

Table 4·9

Percentage Frequencies of Responses to Item 6 (a) parts (i) to (v) ($N = 517$)

Response Type	Ordinal Score	6a i c = 3	6a ii c = 10	6a iii c = 12	6a iv c = 7.4	6a v c = no. in box
Omit	Missing	6.8	6.8	6.8	6.8	6.8
Reject	1	6.6	29.6	70.4	19.3	52.4
Accept	2	86.7	63.6	22.8	73.9	43.8

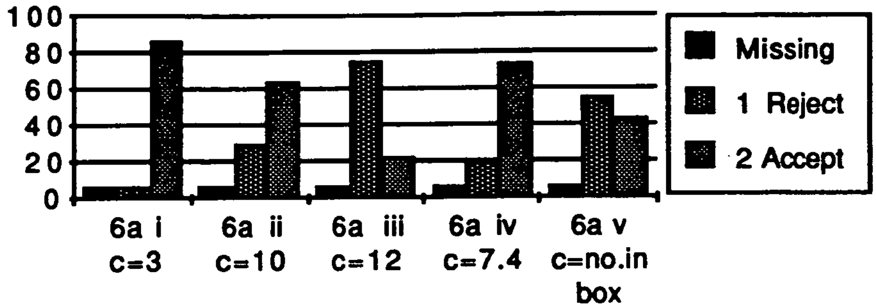


Figure 4·7. Percentage frequencies of responses to Item 6(a) parts (i) to (v) as an ordinal scale ($N = 517$)

The rates of acceptance and rejection for the last two options in part (a) of Item 6 are presented in Table 4·10 and Figure 4·8.

Table 4·10
Percentage Frequencies of Responses to Item 6 (a) parts (vi) and (vii) ($N = 517$)

Response Type	Ordinal Score	6a vi c = cabbage	6a vii c = pear
Omit	Missing	6.8	6.8
Accept	1	21.3	20.9
Reject	2	72.0	72.3

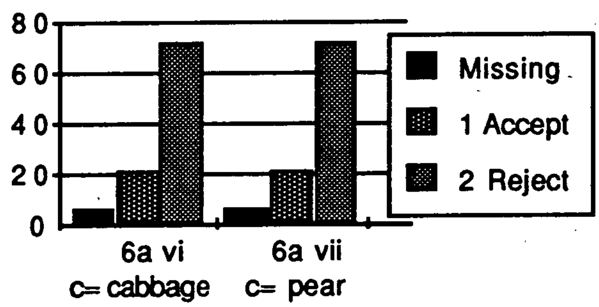


Figure 4·8. Percentage frequencies of responses to Item 6(a) parts (vi) and (vii) as an ordinal scale ($N = 517$)

These last two choices tested whether or not students would accept objects like a cabbage or a pear as possible meanings for '*c*'. To merit a score of "2", they were expected to reject such options. A score of "1" was allocated to approximately one-fifth of the students who accepted them. Of particular relevance to the study were follow-up interviews in 1991 as part of the investigation of the students who accepted objects as meanings for algebraic symbols and still managed the variety of cognitive tasks in the other test items. Investigations of the "cabbage" and "pear" responses are reported in Chapter 7.

In part (b), students expressed various opinions about what would happen to '*d*' if '*c*' became larger, given that '$c + d = 10$'. However, nearly 85% reasoned correctly that '*d*" would get smaller. Table 4·11 details the frequencies of the types of responses and Figure 4·9 displays the percentage frequencies.

Table 4·11

Percentage Frequencies of Responses to Item 6 (b) ($N = 517$)

CATEGORY DATA		ORDINAL DATA	
Response Type	% 6b	Score	% 6b
Omit	6.4	Missing	6.4
No change or other error	5.6	1	see below
Bigger	3.7	1 (total)	9.3
Smaller	84.3	2	84.3

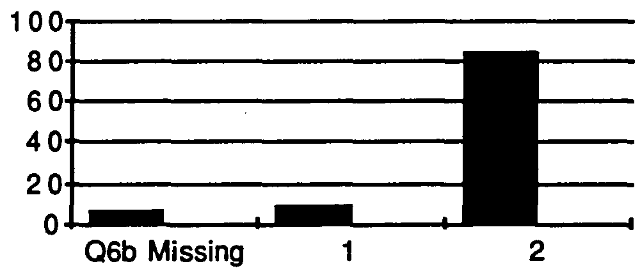


Figure 4·9. Percentage frequencies of responses to Item 6(b) when grouped to form an ordinal scale ($N = 517$)

In part (c), the highest ordinal value of "5" was awarded to those who indicated that '*c*' had to be less than 5 and specifically mentioned at least one of the following possibilities: that it could be zero, negative, or fractional. Simply specifying that '*c*' was "less than 5" merited a score of "4". Those who wrote "less than 4" were grouped with those who listed integer values down to one and were scored at "2", while students who listed zero among other valid integers were given a score of "3". Those who gave only one number were given a score of "1". It was decided to disregard the erroneous inclusion of '5' in integer answers.

The variety of response types in part (c) is listed in Table 4·12 and the frequencies of the derived ordinal scores are presented in Figure 4·10. As the figure displays, this item succeeded in distinguishing students according to several levels of correct responses in terms of the mathematical variability (cf. Dienes, 1963) applicable to the symbols in the problem.

Table 4·12

Percentage Frequencies of Responses to Item 6 (c)

CATEGORY DATA		ORDINAL DATA	
Response Type	% 6c	Score	% 6c
Omit	12.4	Missing	12.4
One value e.g., 3 or 4	3.4	1	see below
other error	18.8	1 (total)	22.2
1, 2, 3, 4 (5)	37.3	2	37.3
0,1, 2, 3, 4 (5)	12.4	3	12.4
< 5	8.3	4	8.3
< 5 + zero or negative or fractional	7.4	5	7.4

Note. $N = 517$. "see below" indicates where to find % tally on ordinal scale.

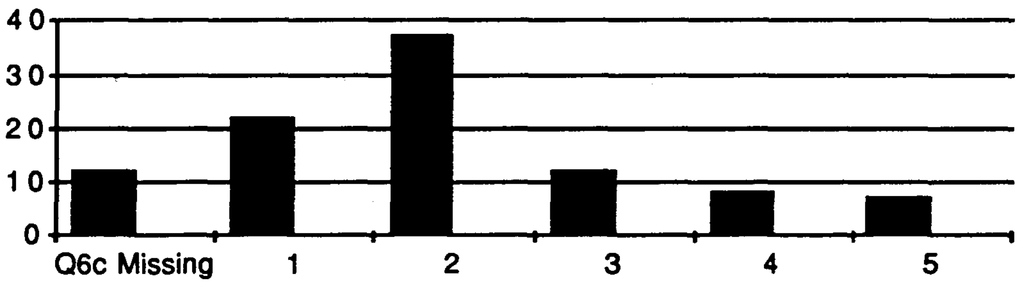


Figure 4·10. Percentage frequencies of responses to Item 6(c) when grouped to form an ordinal scale ($N = 517$)

Comparisons with previous research projects. Küchemann (1980) regarded his question as one which required students to view letters as generalized numbers but found that nearly 40 percent of responses gave just one value for '*c*'. In the 1989 trialling, over half (56.8%) of students gave answers to part (c) which included at least four values for '*c*' (Appendix 3D, Figure 3D·14c). Nearly half of the Group VIII beginners tested in early 1990 also gave at least four values for '*c*' (Appendix 3J). This was an improvement on the 30% (Küchemann, 1980, p. 57) who did likewise in the CSMS study and it indicated that the inclusion of the introductory parts had possibly alerted students to using the notion of generalized numbers. Küchemann (1981, p. 108) found that nearly 40% of his large sample (about one thousand) of 14-year-old students gave only one value for '*c*' and about 30% gave four or more values. Only 3.5% of the 517 students in the present study, which included many younger than 14, gave simply one value for '*c*' and 65.4% gave at least four possibilities for '*c*'. It seems probable that rewording the question slightly and using parts (a) and (b) of Item 6 as a means of leading the students into careful consideration of the question had at least some influence on the attainment of such improved outcomes.

Harper (1979) used four questions in interviews about the equation '$x + y = 10$', one of which was equivalent to the problem posed in Item 6 part (c). He reported (pp. 339 - 340) that 45.8% of the students in his sample responded at an algebraic level. This meant that they showed an understanding of the symbols as covarying over sets of numerals (p. 175). Only 15.7% of the students in the main project gave answers rated at ordinal scores of "4" or "5" by providing responses which could be considered truly algebraic. However, another 49.7% gave answers scored at "2" or "3" which showed that they had an understanding of a covarying pair of variables, even though they mentioned only integers.

Item 7

7. At a certain university there are six times as many students as there are professors.
This fact is represented by the equation **S = 6 P.**

CIRCLE YOUR CHOICES IN THE FOLLOWING QUESTIONS:

(a) In **this equation**, what does the letter **P** stand for?
- (i) Professors
- (ii) Professor
- (iii) Number of professors
- (iv) Students
- (v) Student
- (vi) Number of students
- (vii) None of the above
- (viii) More than one of the above (if so, indicate which ones)
- (ix) Don't know.

(b) In **this equation**, what does the letter **S** stand for?
- (i) Professors
- (ii) Professor
- (iii) Number of professors
- (iv) Students
- (v) Student
- (vi) Number of students
- (vii) None of the above
- (viii) More than one of the above (if so, indicate which ones)
- (ix) Don't know.

(Rosnick 1981, adapted)

Item 7 was a redrafting of the multiple-choice version of the professors-and-students problem, given in Rosnick (1981, p. 419) as:

> At this university there are six times as many students as there are professors. This fact is represented by the equation **S = 6 P.**
>
> A) **In this equation, what does the letter P stand for?**
> i) Professors
> ii) Professor
> iii) Number of professors
> iv) None of the above
> v) More than one of the above (if so, indicate which ones)
> vi) Don't know.
>
> B) What does the letter S stand for?
> i) Professor
> ii) Student
> iii) Students
> iv) Number of students
> v) None of the above
> vi) More than one of the above (if so, indicate which ones)
> vii) Don't know.

Rosnick's discussion of responses to his version made it clear that more options were required. He found that over 22% chose option (i) ('*S*' for "professor") in part (B) and that all of these chose option (vi) ("Don't know") in part (A), which indicated that they sought the option " '*P*' for student". The Rosnick version did not fully allow students to express such a reversal misconception, which seems strange following the evidence for such a tendency already accumulated by Rosnick and Clement (1980), as discussed in Chapter 2. The revised multiple-choice item was used in all the 1990 trial tests except the "1990 Algebra Project".

The question was found difficult by the Years 8 and 9 students in the trialling. As the statistics in Appendices 3E (Figure 3E·1), 3G (Figure 3G·1) and 3H (Figure 3H·6) record, about 10% of them were correct on both parts of the item. Another 60% of these secondary students chose options which indicated that they thought the symbols represented people rather than numbers of people, and nearly 20% chose a mixture of these views. The question was further tested by obtaining responses from samples of university students, some in Sydney and others in Hobart, to test whether or not older students could manage the thinking required for success. Of the university students in Group VI who were preparing to teach mathematics in secondary schools, there were 63.6% who had both parts correct (Appendix 3E, Figure 3E·1). By coincidence, one group of 18 of these students were given the question immediately following a lecture which included an explanation of the very same professors-and-students problem and yet seven of them gave incorrect responses. About one-quarter of the tertiary students studying for primary teaching answered the two parts of the item correctly.

The item was kept as valuable for contributing to a knowledge of how students interpreted algebraic symbols within an equation which stated the relationship between numbers of students and numbers of professors in a real-life context. Analysis of the trialling data confirmed that the item also provided information regarding an aspect of the reversal misconception which proved worthy of further investigation: This error was in evidence mainly from those who thought the symbols represented people rather than numbers of people. Perhaps holding the view that algebraic symbols may represent non-numerical objects (or people) could inhibit the ability to work with numerical variables in items such as Item 7. (See Chapter 7 for follow-up analyses based on responses in the main study.)

Table 4·13 details the percentage frequencies in the main research data of different response types and defines the ordinal scale derived from these response types. The order of merit allocated to the category data resulted in ordinal scale scores as follows:

"5" Correctly chose "number of ..." without reversal (e.g., "number of students" for '*S*');

"4" Chose "number of ..." with reversal (e.g., "number of students" for '*P*');

"3" Chose people rather than numbers of people or a mixture of these possibilities, but without reversal (e.g., "number of students" and "students" for '*S*');

"2" Chose people rather than numbers of people or a mixture of these possibilities, but with reversal (e.g., "student" and "students" for '*P*');

"1" Chose the option "None of the above"; and

"0" Omitted one part although attempted the other (e.g., omitted part (b) after part (a) was answered).

A little more than half the students either related the symbols to people or to a combination of people and numbers of people without making the reversal error (Ordinal Scale score of "3"), showing that they did not appreciate that the symbols in the algebraic equation, '$S = 6P$', necessarily stood for numerical variables. By combining the total frequencies of the scale scores "2" (includes reversal error) and "3" (excludes reversal error), the total proportion of students who failed to appreciate the numerical meaning for the symbols was 64.4% in part (a) and 66.7% in part (b).

The tally of those who opted only for numbers of people, obtained by adding the frequencies of score "4" (with reversal) and score "5" (no reversal), was 32.0% in part (a) and 29.6% in part (b). These students recorded their belief that the symbols represented numbers rather than people.

The success rate was just over 25% (Ordinal Scale score of "5") in each part of the item, with actually 22.6% having both parts correct. Figure 4·11 presents graphs of the rates of responses on each of the ordinal scale scores for both parts of Item 7.

Table 4·13
Percentage Frequencies of Responses to Item 7

CATEGORY DATA			ORDINAL DATA			
Response Type	% 7a	% 7b	Score 7a	% 7a	Score 7b	% 7b
Omit a & b	1.9	1.9	Missing	1.9	Missing	1.9
Omit this part	0.6	1.0	0	0.6	0	1.0
vii None	1.2	0.8	1	1.2	1	0.8
iv Students	9.5	33.8	2	see below	3	see below
v Student	1.7	6.0	2	see below	3	see below
2 or more of iv, v, vi	2.7	12.0	2 (total)	13.9	3 (total)	51.8
i Professors	29.8	5.0	3	see below	2	see below
ii Professor	7.5	7.4	3	see below	2	see below
2 or more of i, ii, iii	13.2	2.5	3 (total)	50.5	2 (total)	14.9
vi No.of students	4.1	25.3	4	4.1	5	25.3
iii No.of professors	27.9	4.3	5	27.9	4	4.3

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.

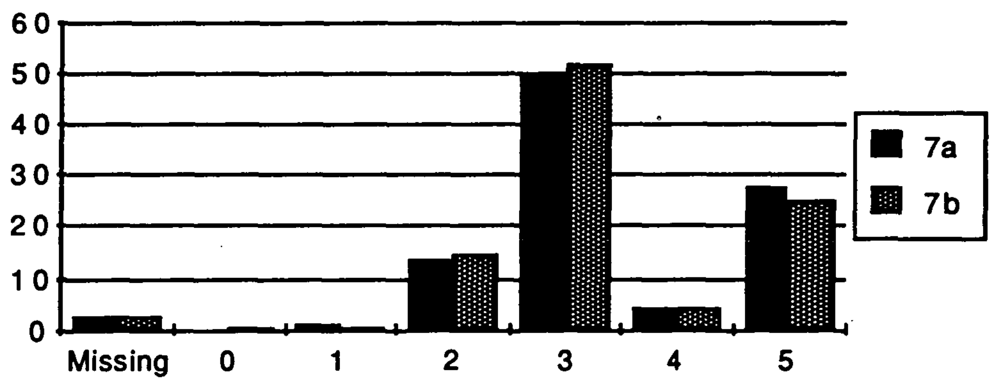


Figure 4·11. Percentage frequencies of responses to Item 7 when grouped to form an ordinal scale ($N = 517$)

Comparison with Rosnick (1981). The question was an extended version of the multiple-choice question by Rosnick (1981) who reported that less than 60% of the 152 first- and second- year university students he tested successfully chose "number of professors" for the meaning of '*P*'. He expressed surprise that over 22% of his subjects chose "professor" as the meaning for '*S*'. Figures in Table 4·13 indicate that the frequency of the reversal error was about the same with the 1990 subjects. In part (a), the proportion of students who made the reversal error of relating the symbol '*P*' to student(s) with one or both of the options "student" or "students" (some of these also chose "number of students") was 13.9% (Ordinal Scale score of "2"), with another 4.1% making the same error by choosing only the "number of students" option (Ordinal Scale score of "4"). The corresponding figures for relating the symbol '*S*' to professor(s) in part (b) were 14.9% (score of "2") and 4.3% (score of "4").

Because of the objective to focus on meanings given to algebraic symbols, outcomes from this well-known item were of particular relevance. Further investigations based on responses to Item 7 are discussed in Chapter 7.

Item 8

8. (a) If **3a** represented 3 apples, what would **a** represent? ....

(b) If **3a** = **36**, what would be the value of **a** ? ....

(c) Kay and Ray say that **3a** + **2b** could represent the total number of people seated in a restaurant, some at **3** large tables (the same number at each) and some at **2** smaller tables (the same number at each).
Tick ONE of the following to show how strongly you agree or disagree with Kay and Ray:

I strongly agree.... I agree.... I disagree.... I strongly disagree....

(d) Jack and Jill say you must not add **3a** and **2b**.
CIRCLE ONE of the following as what you consider to be the BETTER reason:
(i) because it would be like trying to add 3 apples to 2 bananas.
(ii) because **a** and **b** stand for numbers but you do not know what the numbers are.

The strategy behind part (a) of the question was that of asking students to consider the implication of regarding '3*a*' as '3 apples', namely that the algebraic symbol would be representing an object rather than a number. In the 1989 trial, about one-third of students correctly pointed out that '*a*' would stand for simply 'an apple' or 'one apple' under the given conditions (Appendix 3D, Figure 3D·17a). None of the Year 7 beginners in the trialling saw this (Appendix 3J, Table 3J·2).

Part (b) incisively tested the degree of commitment students had to the belief that conjoining in algebra meant multiplication, and so it tested their degree of understanding of the use of this convention for writing algebraic expressions. Like Item 3, it took the students to a specific numerical value for the letter and nearly 80% of Groups I to V were able to do this (Appendix 3D, Figure 3D·17b). Part (b) also revealed that a minority of these students held misconceptions about '3*a*' such as a place value interpretation (resulting in '6' as an answer) and a conjoining-for-addition notion (giving '33' as an answer). The high frequency of such misconceptions, especially the first one (42.1%), amongst the Group VIII beginners supported the methodology plan to test Year 7 students in the main study before they started their classroom lessons on algebra. Only one of these beginners had part (b) correct.

About 70% of Groups I to V students agreed with Kay and Ray's statement in part (c), showing a willingness to view algebraic symbols as representing numbers with unstated values in a real-life setting (Appendix 3D, Figure 3D·17c). Just over 30% of the Group VIII beginners agreed with the statement (Appendix 3J, Table 3J·2). This part of the item was considered useful in contributing information on students' views about symbols.

Part (d) was a combination of two parts of the version of the item used in "Brain-Box Quiz No.2" (Appendix 3C), namely:

(d) Jack and Jill say you must not add **3a** and **2b** because it would be like trying to add 3 apples to 2 bananas. **Tick ONE** of the following to show how strongly you agree or disagree with Jack and Jill:

I strongly agree.... I agree.... I disagree.... I strongly disagree....

(e) Joanna and Joshua say you must not add **3a** and **2b** because **a** and **b** stand for numbers but you do not know what the numbers are. **Tick ONE** of the following to show how strongly you agree or disagree with Joanna and Joshua:

I strongly agree.... I agree.... I disagree.... I strongly disagree....

Trialling with these two parts separately showed that some students tended to change their point of view from one context to another. As Figures 3D·17d and 3D·17e in Appendix 3D record, nearly 70% agreed with the objects argument of Jack and Jill and nearly 60% agreed with the numbers argument of Joanna and Joshua. These two parts were combined in an effort to get a commitment by students to either the "numbers view" or the "objects view" of the letters in this case. The new version was first trialled with Year 9 students (Appendix 3G - see Question "4" comments) who showed a strong tendency to think of symbols as representing objects rather than numbers. The enforced dichotomy placed 61.1% in the objects category and 33.3% in the numbers category. When this new version was trialled with Group VIII using "Algebra Project New Test 1990", it was found that nearly 80% of the beginners (Appendix 3J, Table 3J·2) chose the numbers view, prompting interest in investigating whether algebra teaching displaced this correct view with one which saw algebraic symbols as representing objects.

In reporting on the responses in the main study, the four parts of Item 8 are taken in turn as independent questions.

In Part (a), nearly 20% of the students realized that the implication in the statement was that '*a*' represented an object, which was one apple (Ordinal Scale score of "3"). Others may have thought that '*a*' equalled the number of apples but, in the given case, the value of the number was '1'. These students were scaled at "2". The most common error, by nearly 40%, was to take '*a*' as meaning "apples", as if the '3*a*' in algebra was simply shorthand for the wording "3 apples".

Table 4·14 details the frequencies of responses and the allocated scale scores, which are presented in Figure 4·12.

Table 4·14
Percentage Frequencies of Responses to Item 8 part (a)

CATEGORY DATA		ORDINAL DATA	
Response Type	% 8a	Score	% 8a
Omit	6.4	Missing	6.4
"apples"	37.9	1	see below
other error	15.3	1 (total)	53.2
"1" or "apple"	17.0	2	see below
"no.of apples"	4.6	2 (total)	21.7
"an apple" or "one apple"	18.8	3	18.8

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.

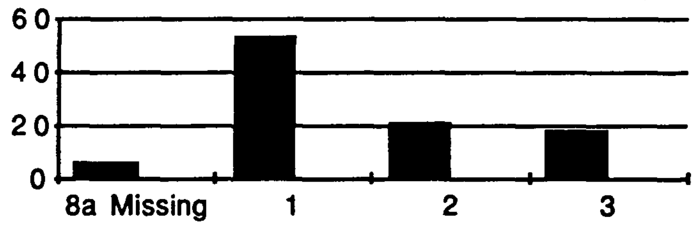


Figure 4·12. Percentage frequencies of responses to Item 8 (a) when grouped to form an ordinal scale ($N = 517$)

In part (b), almost three-quarters of the subjects succeeded in solving the equation '$3a = 36$'. However, nearly 15% wrote either '6' (taking a place value interpretation of '$3a$') or '33' (taking '$3a$' to mean '$3 + a$'). The outcomes are presented in Table 4·15 and Figure 4·13.

Table 4·15

Percentage Frequencies of Responses to Item 8 part (b)

CATEGORY DATA		ORDINAL DATA	
Response Type	% 8b	Score	% 8b
Omit	4.6	Missing	4.6
not a number	1.9	1	see below
wrong no. (not 6 or 33)	5.8	1	see below
6	11.0	1	see below
33	3.5	1 (total)	22.2
12	73.1	2	73.1

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.

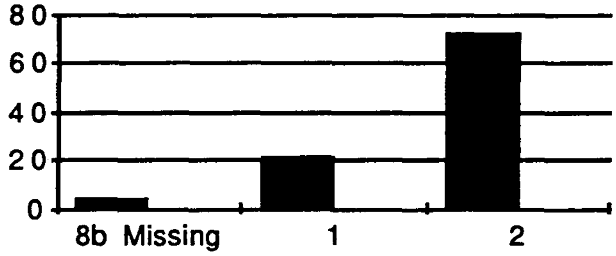


Figure 4·13. Percentage frequencies of responses to Item 8 (b) when grouped to form an ordinal scale ($N = 517$)

Part (c) provided an in-built ordinal scale from "4" for "strongly agree" to "1" for "strongly disagree". Students were expected to agree with the use of the expression '$3a + 3b$' to represent the number of people seated at 5 tables as described in the question. One third strongly agreed with the given statement and almost another 40% simply agreed. The frequencies are reported in Table 4·16 and Figure 4·14.

Table 4·16

Percentage Frequencies of Responses to Item 8 part (c) ($N = 517$)

CATEGORY DATA		ORDINAL DATA	
Response Type	% 8c	Score	% 8c
Omit	3.9	Missing	3.9
strongly disagree	10.4	1	10.4
disagree	13.5	2	13.5
agree	39.1	3	39.1
strongly agree	33.1	4	33.1

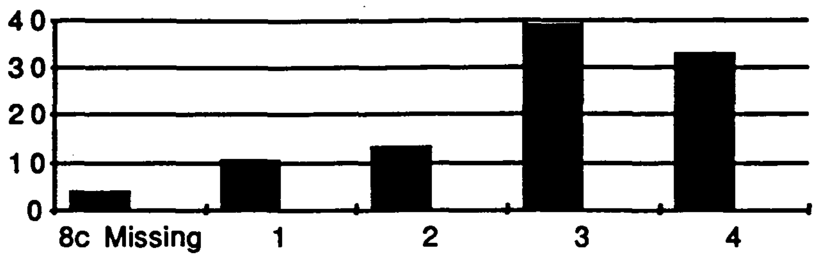


Figure 4·14. Percentage frequencies of responses to Item 8 (c) as an ordinal scale ($N = 517$)

The last part of Item 8 was designed to place students in either the category who thought of algebraic symbols in terms of objects (ordinal score of "1") or the category of those who thought of them in terms of numbers (ordinal score of "2"). The percentage choosing the numbers views was only 15 percentage points above the 40% who chose the objects view, as recorded in Table 4·17 and Figure 4·15.

Table 4·17

Percentage Frequencies of Responses to Item 8 part (d) ($N = 517$)

CATEGORY DATA		ORDINAL DATA	
Response Type	% 8d	Score	% 8d
Omit	4.4	Missing	4.4
i symbols as fruit	40.2	1	40.2
ii symbols as numbers	55.3	2	55.3

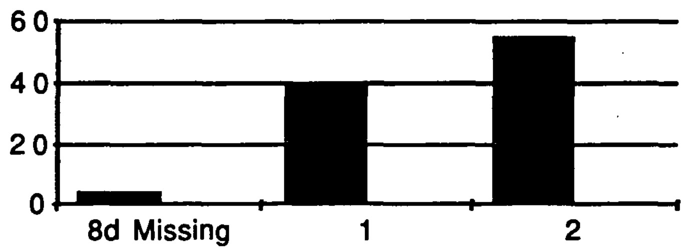


Figure 4·15. Percentage frequencies of responses to Item 8 (d) as an ordinal scale ($N = 517$)

Item 9

9. (i) Add **4** onto **n + 5.** ....

(ii) Add **4** onto **3n.** ....

(iii) Multiply **n + 5** by **4.** ....

(Küchemann, 1980)

These questions have been discussed in Chapter 2. They were taken from the CSMS study except for the omission of the introduction to part (ii) in the original question, which read:

"4 added to *n* can be written as *n* + 4. Add 4 onto 3*n*." (Küchemann, 1980, p. 67). This change did not appear to affect the success rate, which was about 35% in the CSMS study and also in the trialling with Groups I to V (Appendix 3D).

The item was found to be valuable in the research as it measured ability with operations on unknowns at various levels of difficulty. The success rates for the 1989 trial students dropped from about 70% for part (i) to about 35% for part (ii) and 20% for part (iii), as Figure 3D·9 in Appendix 3D records. There was only one correct answer from Group VIII (Appendix 3J, Table 3J·2). Such outcomes provided encouragement for investigating hierarchies of difficulty in the main study.

The frequencies for the variety of outcomes obtained from students in the main study are listed in Table 4·18.

Table 4·18

Percentage Frequencies of Responses to Item 9

CATEGORY DATA				ORDINAL DATA			
Response Type	% 9i	% 9ii	% 9iii	Score	% 9i	% 9ii	% 9iii
Omit i to iii	5.4	5.4	5.4	Missing	5.4	5.4	5.4
Omit this part	0	0.2	2.1	0	0	0.2	2.1
Ignore '*n*' i 9, ii 7, iii 20	6.8	3.9	4.1	1	see below	see below	see below
Alphabetic: $n = 14$ i 19, or 23, ii 46, iii 76	1.2	0.6	0.6	1	see below	see below	see below
Other Nos.	7.4	11.0	10.8	1 (total)	15.3	15.5	15.5
Other wrong algebra	8.3	4.4	4.4	2	see below	see below	see below
Conjoined i 9*n* (or *n*9), ii 7*n*, iii 20*n*	11.8	23.2	12.8	2	see below	see below	see below
Faulty algebra i 4*n*+5, iii 4*n*+5 or *n*+20	1.9	-	10.4	2 (total)	22.1	27.7	27.7
Correct i *n*+9, ii 3*n*+4, iii 4*n*+20 or 4(*n*+5)	57.3	51.3	49.5	3	57.3	51.3	49.5

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.
' - ' denotes "Not Applicable".

The highest score of "3" on the ordinal scale was for the correct algebraic answers, "2" was the score attained by those who gave incorrect algebraic answers, and "1" by those who gave numerical answers. Figure 4·16 summarizes the frequencies of the ordinal scale scores on Item 9, as defined in Table 4·18.

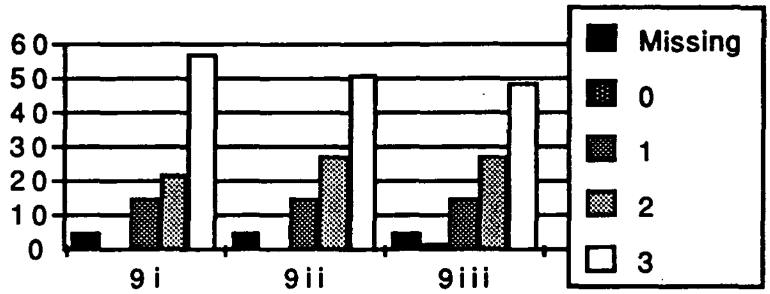


Figure 4·16. Percentage frequencies of responses to Item 9 when grouped to form an ordinal scale ($N = 517$)

Comparison with Küchemann (1981). Küchemann (1981, p. 108) reports not only the percentage correct but also percentage frequencies for several of the errors found when these questions were used with about one thousand students at the end of their third year of secondary school. Table 4·19 compares figures from the two studies.

Table 4·19

Comparison of Item 9 Percentage Outcomes from Quinlan & Küchemann Studies

ITEM	RESPONSE	QUINLAN 1990	KÜCHEMANN 1981
9 i	$n + 9$ #	57.3	68
9 i	9	6.8	20
9 ii	3n + 4 #	51.3	36
9 ii	$7n$	23.2	31
9 ii	7	3.9	16
9 iii	$4n + 20$ # or $4(n + 5)$ #	49.5	17
9 iii	$4n$+5 or 4 x n+5 or $n + 20$	23.2	50
9 iii	20	4.1	15

Note. 1981 percentages from Küchemann, 1981, p. 108.
# denotes correct answer.

The students in the Quinlan study ranged from the first to the sixth year of secondary school. As a group, they achieved better than those in the Küchemann study, except for the first part, which, as Küchemann pointed out (1981, p. 109) could be correctly answered by those who did not know how to work with the symbol 'n' and simply operated on the more familiar numbers. The success rates for all three parts were similar in the Quinlan study, in contrast to the results obtained from earlier trialling and those obtained by Küchemann. Perhaps it could be expected that many of the students in the later years of their schooling would have developed the skills needed for these questions and would have achieved better results overall.

Items 10, 11, 12, and 13. These four items were modelled on items used by Harper (1979) in interview format. In order to find out whether the range of responses which he found by means of interviews could be replicated by using the items as part of a written test, Items 10 and 11 and another similar to Item 13 were trialled in the test "1990 Algebra Project". The written form did produce the range of responses (Appendix 3F, Tables 3F·1 to 3F·3) and so it was decided to include these questions in the research test. The questions have been discussed in Chapter 2. An equivalent of Harper's Zetetic Task (Harper, 1979, p. 84) was trialled as Question 4 (about piles of stones) in the same test but was deleted from the research instrument because only a few of the students who had it correct used algebra in their method (Appendix 3F, Table 3F·4) and it seemed not to be helpful in casting any new light on the research problem at the level of this study.

Item 10

10. This question is about **t + t** and **t + 4**.

(a) Which is larger, t + t or t + 4 ? WHY?

- -

(b) When is t + t larger?

- -

(c) When is t + 4 larger?

- -

(d) When are they equal?

- -

(Harper, 1979)

This was Subtask 1 in Harper's Literal Numbers Task, which was discussed in Chapter 2. It was less complicated than the other two subtasks, which respectively dealt with the pairs '$m + m$, $m + k$' and '$a + b + 3$, $a + c + 4$'. Trialling with a sample of Group V students when they were in Year 9 produced a spread of response types which matched those obtained by Harper in an interview mode. Less than 60% of the answers were correct (Appendix 3F, Table 3F·2). It was, therefore, thought that the item was sufficiently difficult as well as useful in measuring student readiness to compare abstract entities in the form of algebraic expressions whose values were left dependent on whatever value 't' happened to have.

Harper grouped responses into categories as follows:

"A" False ordering without correction,

"B" False ordering with correction,

"C" Numerical replacements, and

"D" Algebraic.

The interview format gave his students the opportunity to correct themselves as they discussed the various aspects of the item, thus giving rise to the "A" and "B" categories. These were inseparable in the written format and were grouped under a score of "1" on the ordinal scale applied to the written responses, as defined in Tables 4·20 and 4·21. Harper's "C" category corresponded with a scale score of "2". From the written answers, the algebraic category "D" was split into two: algebraic correct (with a score of "4") and algebraic not quite correct (with a score of "3"). Responses from the students in the main study are set out in Tables 4·20 and 4·21.

Table 4·20

Percentage Frequencies of Responses to Item 10 part (a)

CATEGORY DATA		ORDINAL DATA	
Response Type	% 10a	Score	% 10a
Omit a to d	6.0	Missing	6.0
Omit a	1.7	0	1.7
$t + 4$	10.4	1	see below
$t + t$	6.2	1	see below
t not known	7.4	1	see below
wrong idea	8.7	1 (total)	32.7
variable notion but incorrect	9.3	3	9.3
variable notion and correct	50.3	4	50.3

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.

Table 4·21

Percentage Frequencies of Responses to Item 10 parts (b) to (d)

CATEGORY DATA				ORDINAL DATA			
Response Type	% 10b	% 10c	% 10d	Score	% 10b	% 10c	% 10d
Omit a to d	6.0	6.0	6.0	Missing	6.0	6.0	6.0
Omit this part	7.2	10.3	10.3	0	7.2	10.3	10.3
always	1.7	3.5	0.8	1	see below	see below	see below
never	1.7	2.9	4.6	1	see below	see below	see below
wrong idea	28.4	25.7	25.0	1(total)	31.9	32.1	30.4
one number as replacement	2.3	1.7	-	2	see below	see below	-
more than one number as replacements	0.2	1.2	-	2 (total)	2.5	2.9	-
algebra not quite correct e.g., ii $t \geq 5$	9.1	7.0	-	3	9.1	7.1	-
algebra correct	43.3	41.8	53.4	4	43.3	41.8	53.4

Note. N = 517. "see below" indicates where to find % tallies on ordinal scale. ' - ' denotes "Not Applicable".

The percentage frequencies of the ordinal scores obtained by those who answered the written test are depicted in Figure 4·17.

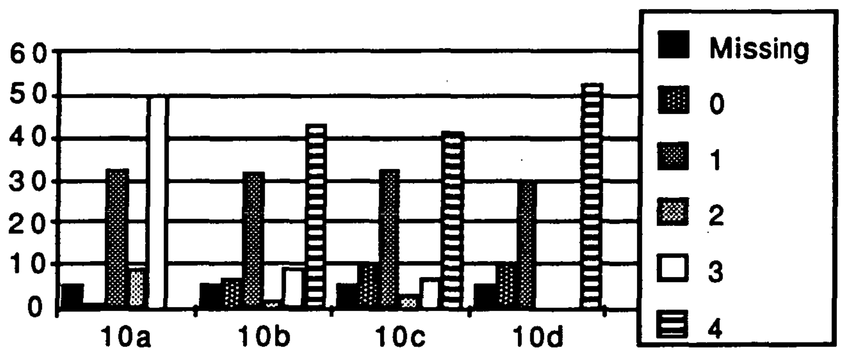


Figure 4·17. Percentage frequencies of responses to Item 10 when grouped to form an ordinal scale (N = 517)

Comparison with Harper (1979). The ordinal scores were applied to each part of the question in the 1990 main testing program whereas Harper's categories were applied globally to each subject's handling of the four parts of the item. To compare the outcomes from the two studies, as presented in Table 4·22, the percentage frequencies based on ordinal scores for each part were considered in two ways. Firstly, a statistical estimate of the frequencies of outcomes, had the parts of the item been assessed globally, was calculated as the average frequency for each ordinal category, taking into account only those parts of the question to which a particular score was applicable. Secondly, the maximum frequency was examined, since Harper could well have been guided by the highest level of performance exhibited by any subject when he classified that subject according to the categories. Because those who were interviewed were obliged to at least attempt this item, while those answering the written test had the option of omitting it, all percentages based on the written test were corrected, for the Quinlan data in Table 4·22, to include only those who attempted at least one part of the item.

Table 4·22
Comparison of Item 10 Outcomes from Quinlan & Harper Studies

QUINLAN 1990			HARPER 1979		
Ordinal scale score	Average %	Maximum %	Harper category	%	Description
0	7.9	10.9	-	-	Omit some part(s)
1	33.8	34.8	A & B	27.1	"False ordering"
2	2.9	3.1	C	12.5	Numerical replacements
3 & 4	57.0	63.4	D	60.4	Algebraic

Note. Percentages for Quinlan calculated after deleting all missing values. Harper percentages deduced from Tables 20 (a) and 20 (b), Harper, 1979, pp. 351, 352.
' - ' denotes "Not Applicable".

The 144 subjects in the Harper study were spread across the six years of secondary school as were the 517 in the present study. The distributions per grade and per ability level were not, however, identical: Harper's subjects were evenly spread across the years and were selected from those who were rated as having high academic ability relative to the total school population. Subjects in the present study were spread across the full range of academic ability and 45% of them were from the first year of secondary schooling. As regards the highest category of achievement, the algebraic level, the outcomes were similar at about 60% success rate. The strategy of using numerical replacements was more common in the Harper study (at 12.5%) than the Quinlan study (at about 3%). In Chapter 9, further comment is made on this point when sequences of learning are addressed. Other investigations based in part on responses to Item 10 are also reported in the same chapter.

Item 11

11. This question is about the two lines shown in the sketch.

(a) Is the red line longer than the green line,
the green line longer than the red line,
are they equal in length, or
could any of these be possible? WHY?

(b) When is the green line longer than the red line?

(c) When is the red line longer than the green line?

(d) When are they equal in length?

(Harper,1979)

As was discussed in Chapter 2, this Harper (1979) question was designed to find out the degree to which students allowed variation in a geometrical setting by using the variable notion to organize perception. It was the third subtask in Harper's Parallel Lines Task. His other parallel lines subtasks showed line pairs labelled respectively as "2 cm, 4 cm" and "*p* cm, *p* cm". The chosen subtask appeared to allow the most scope for writing about algebraic relationships and was thus selected for trialling. What was sought in the trialling was mainly a variety of responses of the types Harper obtained by an interview approach. Such a variety was found, covering truly algebraic responses as well as various ways of reasoning from a geometric perspective.

The university students tested in the trialling found the task difficult. Only 19.4% of Group VII, who were preparing to teach in primary schools, gave a correct algebraic answer (Appendix 3F, Table 3F·3). However, success was attained by half of the 30 students from Group V who were tested when they were in Year 9 (Appendix 3F, Table 3F·3). About 20% of a group of Year 8 students gave correct algebraic responses (Appendix 3H, Figure 3H·5) and only one student from the beginners in Year 7 did likewise (Appendix 3J, Table 3J·3). The question was considered relevant and useful, especially in conjunction with the other Harper questions.

The response rates for the 517 students in the main study are recorded in Table 4·23 in which the ordinal scale for this item is defined. In the highest category, with a score of "4", were correct algebraic answers to the parts of the item while the category for answers which unsuccessfully attempted an algebraic approach was allocated a score of "3". Those who succumbed to the strong geometric distraction written into the item (the red line was always sketched as longer than the green line) were given a score of "2" and those who made other errors scored "1".

Table 4·23

Percentage Frequencies of Responses to Item 11 parts (a) to (d)

CATEGORY DATA					ORDINAL DATA				
Response Type	% 11a	% 11b	% 11c	% 11d	Score	% 11a	% 11b	% 11c	% 11d
Omit a to d	5.0	5.0	5.0	5.0	Missing	5.0	5.0	5.0	5.0
Omit this part	2.1	9.1	10.4	10.8	0	2.1	9.1	10.4	10.8
Repeats question	-	6.6	7.2	12.0	1	-	see below	see below	see below
No idea	2.1	3.3	3.1	2.1	1 (total)	2.1	9.9	10.3	14.1
geometric argument	51.3	23.4	14.7	15.5	2	see below	see below	see below	see below
always	-	0.8	18.6	1.0	2	-	see below	see below	see below
never	-	11.0	1.9	0.6	2 (total)	51.3	35.2	35.2	27.7
geometry & algebra	4.1	1.2	0.6	1.7	3	see below	see below	see below	see below
incorrect algebra e.g., $a \neq b$ ever	1.9	4.1	2.3	4.8	3	see below	see below	see below	see below
uncertain about lengths	1.2	1.0	0.8	0.6	3 (total)	7.2	6.2	3.7	7.2
correct algebra	32.3	34.6	35.4	35.2	4	32.3	34.6	35.4	35.2

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.
' - ' denotes "Not Applicable".

The ordinal frequency data are graphed in Figure 4·18. The success rate (with score "4") is seen to be fairly consistent across the four parts of the item, intimating that once students were thinking along the correct path for one part of the item they generally managed to think correctly for each of the other parts, whereas those who did not use an algebraic approach for some part were very likely not to use that approach in each of the other parts. In parts (b) and (c), about 35% of students used a geometric approach (score "2") and a similar number gave a correct algebraic answer (score "4").

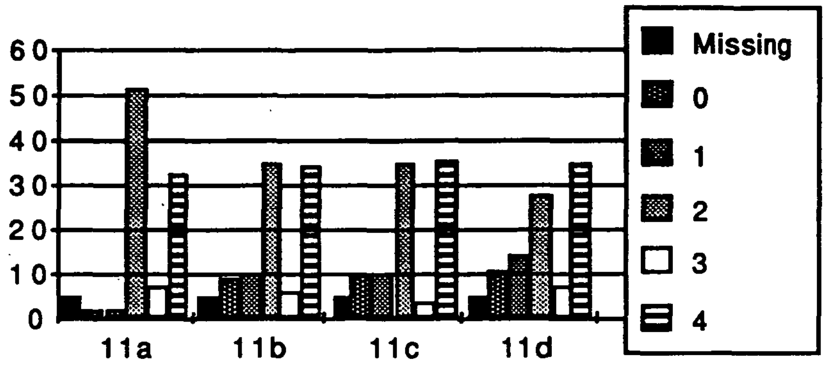


Figure 4·18. Percentage frequencies of responses to Item 11 when grouped to form an ordinal scale ($N = 517$)

Comparison with Harper (1979). Harper (1979) used simply two categories: algebraic correct as one category, and non-algebraic together with incorrect algebraic as the other. The procedures explained in the discussion of Item 10 were followed when comparing the outcomes from Item 11 in the two studies, and the percentage figures are given in Table 4·24. The 1990 students were slightly more successful than the 1979 students.

Table 4·24
Comparison of Item 11 Outcomes from Quinlan & Harper Studies

QUINLAN 1990			HARPER 1979		
Ordinal scale score	Average %	Maximum %	Harper category	%	Description
0	7.9	10.9	-	-	Omit some part(s)
1 & 2 & 3	55.0	63.7	A	66.9	Incorrect
4	36.2	44.6	B	33.1	Algebraic

Note. Percentages for Quinlan calculated after deleting all missing values. Harper percentages deduced from Tables 16 (a) and 16 (b), Harper, 1979, p. 326.
' - ' denotes "Not Applicable".

Item 12

12. This question is about **2n** and **n + 2**.

(a) Which is larger, **2n** or **n + 2** ? WHY?

- -

- -

(b) When is **2n** larger?

(c) When is **n + 2** larger?

(d) When are they equal?

(Küchemann, 1980, adapted to Harper 1979 style)

This item originated in a CSMS question which appeared to be too succinct if it were to test students on their ability to work with the concept of a variable. The original version was:

"Which is larger, $2n$ or $n + 2$? Explain.".

Küchemann put the question in the category of testing the concept of a variable, as was discussed in Chapter 2. He reported that only 6% of the 14-year-olds in the CSMS study gave acceptable answers (Küchemann, 1980, p. 69). Students from Groups I to V were even less successful with the item in the Küchemann form (Appendix 3D, Figure 3D·13).

Use of a Harper style format, as shown in Item 12, focused the students' attention on several possibilities in turn, with the expectation that they would consider seriously the fact that 'n' could take any of a large range of values. When the expanded version was trialled using test "New Test 2" (Appendix 3H) with a group aged about 14 years, it was found that on part (a), which was equivalent to the CSMS question, 13.7% succeeded in communicating the fact that the outcome depended upon the value of 'n'. Nearly one-quarter were correct on part (d) and over 20% gave at least one correct instance for the other two parts. The new version clearly improved the chances that the question would enable students to reveal their thoughts about the issues inherent in the original question. It produced responses which could be categorized in line with classifications used by Harper for similar questions. As Figure 3H·4 in Appendix 3H recounts, judged on answers to all parts of the item, nearly 10% attained the "algebraic" category, about 15% "border-line algebraic", about 5% "placeholder", and the majority (about 60%) were in the "fictitious" category because they assumed that multiplication gave a greater result than addition. None of the Year 7 beginners in the trialling succeeded with any part of the item (Appendix 3J, Table 3J·3). The item was preserved as useful, especially in association with the other Harper questions used, and as a parallel to part (iii) of Item 15 which asked about the relationship '$2a = a + 2$'.

Tables 4·25 and 4·26 report the percentage response rates based on the main 1990 data and define the ordinal scale derived from the response categories.

Table 4·25

Percentage Frequencies of Responses to Item 12 part (a)

CATEGORY DATA		ORDINAL DATA	
Response Type	% 12a	Score	% 12a
Omit a to d	6.2	Missing	6.2
Omit a	0.4	0	0.4
$2n$ or n+2	41.8	1	see below
same	6.0	1	see below
other incorrect	5.4	1 (total)	53.2
one number as replacement	2.5	2	2.5
correct	37.7	4	37.7

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.

Table 4·26

Percentage Frequencies of Responses to Item 12 parts (b) to (d)

CATEGORY DATA				ORDINAL DATA			
Response Type	% 12b	% 12c	% 12d	Score	% 12b	% 12c	% 12d
Omit a to d	6.2	6.2	6.2	Missing	6.2	6.2	6.2
Omit this part	8.7	8.9	10.4	0	8.7	8.9	10.4
always	12.6	1.2	3.7	1	see below	see below	see below
never	5.0	16.2	11.0	1	see below	see below	see below
wrong idea	23.0	25.1	25.3	1(total)	40.6	42.6	40.0
one number as replacement	1.7	6.2	-	2	see below	see below	-
more than one number as replacements	0.2	1.0	-	2 (total)	1.9	7.2	-
algebra not quite correct e.g., (b) $n \geq 3$	10.1	5.6	-	3	10.1	5.6	-
algebra correct	32.5	29.6	43.3	4	32.5	29.6	43.3

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.
' - ' denotes "Not Applicable".

Taking average percentage rates across all four parts of the item and deleting the "missing" cases, the 1990 students exhibited the following distribution in terms of the Harper (1979) categories:

"Algebraic" category (scale score "4"): 38.1%,
"Border-line algebraic" category (scale score "3"): 8.4%,
"Placeholder" category (scale score "2"): 4.1%, and
"Fictitious" category (scale score "1"): 47.0%.

Figure 4·19 displays the frequencies for the ordinal scale scores and, in so doing, graphs the rates of the Harper categories.

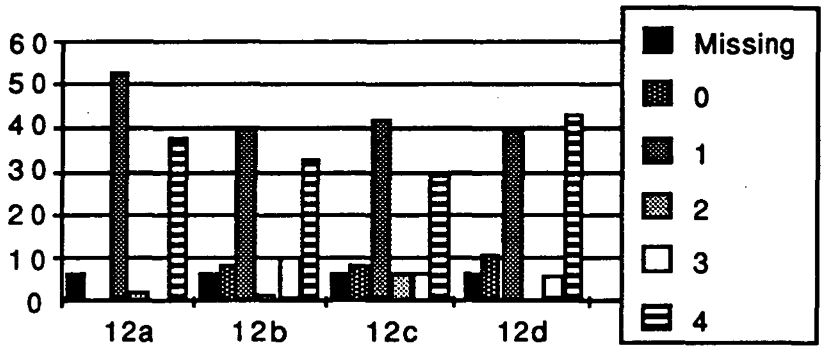


Figure 4·19. % frequencies of responses to ordinal scale for Item 12 ($N = 517$)

Comparison with Küchemann (1981). Item 12 was based on a question in the 1981 Küchemann study but was redrafted along the lines of the Harper (1979) items used in the test instrument as Items 10, 11, and 13. The success rate with the Harper-style format was from about 30% to about 40% for the various parts of the item. Küchemann found that only about 6% of the one thousand 14-year-olds in his study succeeded with the question in the form used in the CSMS study.

The new version of the question enabled more students to reveal their thoughts about the underlying problem. Chapter 9 reports the outcomes of further investigations involving responses to Question 12.

Item 13

13. This question is about x and y in **the equation 2x + y = 9.**

(a) If the equation is true,
is the value of x always, sometimes or never
greater than the value of y? WHY?

(b) When is the value of x greater than the value of y?

(c) When is the value of x equal to the value of y?

(d) When is the value of x less than the value of y?

(Harper, 1979)

This question, discussed in Chapter 2, was Subtask 2 of Harper's Equations Task, the other two involving equations '$x + y = 10$' and '$5x = y$'. On trialling the '$x + y = 10$' subtask in test "1990 Algebra Project" (Appendix 3F, Table 3F·1), it was found that 13.3% gave algebraic responses and 30.0% gave border-line algebraic responses. The subtask using the equation '$2x + y = 9$' was considered one which would provide greater insights into students' thinking. It included two arithmetic operations on a covarying pair of numerical variables, and one term of the equation tested the students regarding the significance of having '2' as a coefficient. Harper found that students being interviewed gave "fictitious" level answers by referring to the coefficient in distorted ways, such as saying that 'x' would always be bigger than 'y' because of the '2' in front of it. In the 1990 trialling when the task was set in written form, similar references to the coefficient were obtained from two Year 9 students (Appendix 3G - comments on Question "2") and from two Year 8 students (Appendix 3H - comments on Question "2"). Thus, this additional ingredient in the question was able to reveal whether or not students understood the use of coefficients and, thus, to contribute further information in line with the objectives of the study.

The chosen item was first tested in "Yr. 9 Test 1990" and the responses were spread across a variety of levels, with about 30% giving correct algebraic answers for parts (b) and (d) and nearly 40% answering part (c) correctly. For the item as a whole, slightly more than 20% were categorized as at the "algebraic" level, and the same proportion were at the "border-line algebraic" level. Over a quarter gave "fictitious" level responses and a little more than 10% simply listed one or more numerical examples (Appendix 3G, Table 3G·1). Testing with younger students from Groups III and IV when they were in Year 8 showed that they had great difficulty with the question, a few of them managing to reach the "border-line algebraic" level, with nearly 40% resorting to a list of numerical examples (Appendix 3H, Figure 3H·2). Group VIII students found the item too difficult (Appendix 3J, Table 3J·3), as could be expected for beginners. The item was considered valuable in assembling information relevant to the objective of investigating students' difficulties with the meaning and use of symbols.

The categories of responses in the main data revealed a lot about how the students thought about algebraic symbols in the context of the equation '$2x + y = 9$'.

Table 4·27 presents the frequencies of different response types and clarifies how they were matched to an ordinal scale from "0" to "4". The main features of the distribution are evident in the graphs of Figure 4·20.

Table 4·27

Percentage Frequencies of Responses to Item 13 parts (a) to (d)

CATEGORY DATA					ORDINAL DATA				
Response Type	% 13a	% 13b	% 13c	% 13d	Score	% 13a	% 13b	% 13c	% 13d
Omit a to d	14.7	14.7	14.7	14.7	Missing	14.7	14.7	14.7	14.7
Omit this part	4.3	10.3	12.8	13.7	0	4.3	10.3	12.8	13.7
always or never	5.4	5.2	8.5	5.2	1	see below	see below	see below	see below
depends on coefficient	6.2	4.4	3.9	2.9	1	see below	see below	see below	see below
x, y not related	23.8	6.0	5.4	6.6	1	see below	see below	see below	see below
no idea	8.1	27.1	25.1	25.7	1 (total)	43.5	42.7	42.9	40.4
one number pair as replacements	4.1	10.4	-	4.6	2	see below	see below	-	see below
more than one number pair as replacements	0.6	1.9	-	1.9	2 (total)	4.6	12.4	-	6.6
algebra not quite correct e.g., (b): $x \geq 4$	0.6	6.2	-	10.1	3	0.6	6.2	-	10.1
algebra correct	32.3	13.7	29.6	14.5	4	32.3	13.7	29.6	14.5

Note. $N = 517$. "see below" indicates where to find % tallies on ordinal scale.
' - ' denotes "Not Applicable".

Figure 4·20 indicates that the modal score for each part of the item was "1", illustrating that the most common errors were misunderstandings of the problem. Over 30% either had no idea what to do or failed to recognize the covariant relationship between 'x' and 'y' which is imposed by the equation. Approximately another 10% appeared to distort the significance of the coefficient of 'x' by responding with "always" or "never", or by explicitly expressing the misconception that the size of a variable was determined by the size of its coefficient. Some of these students suggested that the coefficient be changed to achieve the result that, say, 'y' could be greater than 'x'. Student success in parts (a) and (c) was approximately double that in the other two parts.

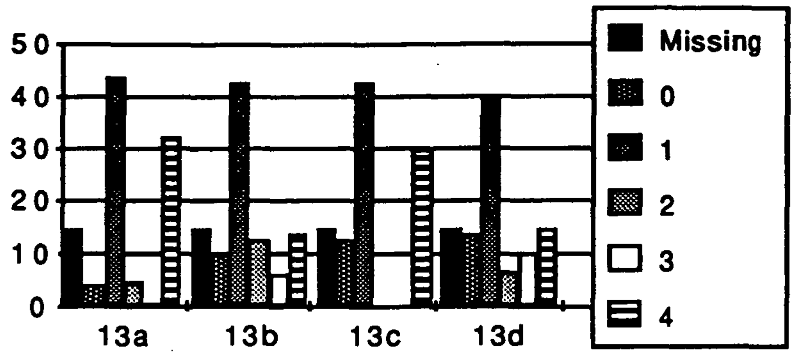


Figure 4·20. Percentage frequencies of responses to Item 13 when grouped to form an ordinal scale ($N = 517$)

Comparison with Harper (1979). To compare the outcomes of the written test with those obtained by Harper in 1979 by means of interviews, the procedures explained in the discussion of Item 10 were applied here also. Table 4·28 summarizes the comparisons.

Table 4·28
Comparison of Item 13 Outcomes from Quinlan & Harper Studies

QUINLAN 1990			HARPER 1979		
Ordinal scale score	Average %	Maximum %	Harper category	%	Description
0	12.0	16.1	-	-	Omit some part(s)
1	49.8	51.0	A	18.1	"Fictitious measure" - wrong idea
2	9.2	14.5	B	30.6	"Placeholder" - lists one or more examples
3	6.6	11.8	C	15.3	"Border-line algebraic" e.g., (b) $x \geq 4$
4	26.4	37.9	D	36.1	"Algebraic" e.g., (b) $x > 3$

Note. Percentages for Quinlan calculated after deleting all missing values. Harper percentages deduced from Tables 18 (a) and 18 (b), Harper, 1979, pp. 339 - 340. ' - ' denotes "Not Applicable".

The proportion of students who succeeded with the item was similar in both studies (score "4" and category "D"). However, about half of the 1990 students were unable to make much progress with the item (score "1"), compared with less than one-fifth of the 1979 students (category "A") who had a similar lack of success. As in

Item 10 (See Table 4·22.), those tested by interview responded more with numerical examples (category "B") than did those who were tested by writing (score "2").

Item 14

14. For a school excursion, 3 buses take **f** students each and
4 cars take **g** students each.

(i) CIRCLE the ONE which best says what the value of **3f** tells us:

(a) 3 buses x f students

(b) How many students took buses

(c) That there are the same number of students on each bus

(d) Three buses, f students

(e) The number of buses which take the children

(ii) Give the total number of students taken by these buses and cars. ....

(iii) One car leaves early with **g** students. How many students remain? ....

This question was created to obtain information on students' ability to interpret the meanings of letters when they referred to a real-life context and to carry out operations on numerical variables without knowing their values.

Part (i) was originally an open-ended question expressed as "What does **3f** tell us?" in Brain-Box Quiz No.1, and as "What does the value of **3f** tell us?" in Brain-Box Quiz No.2. Less than 20% gave correct responses when the question was asked in either of these forms (Appendix 3D, Figure 3D·8). The options in the final multiple-choice form were simply selections from the answers obtained using the earlier format. The new format was first trialled in "New Test 2 1990". The order of popularity of the choices (Appendix 3H, Figure 3H·11a) was (a), (c), (d), (b) and (e). Nearly half of the Year 8 students in this trialling chose option (a) which was incorrect on the fine technicality that to be correct it would have had to be written as "3 buses x **f** students *per bus*". The possibility that the question might have been improved by deleting this option was not put to the test.

Parts (ii) and (iii) identified those not ready to accept the lack of closure required for carrying out operations on variables before knowing their numerical values. Approximately 5% of students in the 1989 trials simply wrote an arbitrary number (Appendix 3D, Figure 3D·8). These parts also identified students who ignored the conventions appropriate for writing the algebraic expressions required or who incorporated only some of the given data into their thinking. For part (ii), nearly 40% were correct in the 1989 trials and, when students in Groups III and IV were re-tested in 1990, the success rate was nearly 50%. Part (iii) was found more difficult with the corresponding rates of success being about 20% in 1989 and 30% in 1990 (Appendix 3D, Figure 3D·8 and Appendix 3H, Figure 3H·11b). Only two Year 7 beginners made any headway at all with the item (Appendix 3J, Table 3J·2).

In the trialling, the item achieved the objectives for which it was written.

For the main data, Table 4·29 summarizes the response rates on the multiple-choice question in part (i) of the item. The most popular choice (by 41.4%) was option (a), for which students were allocated an ordinal scale score of "2". By selecting option (b), "how many students took buses", 20.3% were correct and merited a scale score of "3". The other choices were each given a score of "1".

Table 4·29

Percentage Frequencies of Responses to Item 14 part (i)

CATEGORY DATA		ORDINAL DATA	
Response Type	% 14i	Score	% 14i
Omit i, ii, iii	4.8	Missing	4.8
Omit i	1.9	0	1.9
Choice e	3.1	1	see below
Choice d	19.1	1	see below
Choice c	9.3	1 (total)	31.5
Choice a	41.4	2	41.4
Choice b	20.3	3	20.3

Note. $N = 517$. "see below" indicates where to find % tally on ordinal scale.

The next two parts of Item 14 required students to operate on symbols without knowing their numerical value. As in Items 5 and 9, some students avoided the use of algebraic symbols in their answers by simply replacing the letters by arbitrarily-chosen numbers. Approximately one-eighth of the students did this in Item 14 and were given a scale score of "1". Those who could not operate on symbols but resisted reverting to an arithmetic substitute were given scores of "2", while those who wrote answers in algebra were given a score of "4" if they were correct and "3" if incorrect. The rates of occurrence of these types of responses are given in Table 4·30 which also summarizes the system for the ordinal scaling.

Table 4·30

Percentage Frequencies of Responses to Item 14 parts (ii) and (iii)

CATEGORY DATA			ORDINAL DATA		
Response Type	% 14ii	% 14iii	Score	% 14ii	% 14iii
Omit i, ii, iii	4.8	4.8	Missing	4.8	4.8
Omit i	12.4	13.3	0	12.4	13.3
arbitrary no.	12.0	11.2	1	12.0	11.2
f, *g* unknown	2.7	2.1	2	2.7	2.1
incorrect algebra e.g., iii: 3*g*	16.6	27.3	3	16.6	27.3
correct	51.5	41.2	4	51.5	41.2

Note. $N = 517$. "see below" indicates where to find % tally on ordinal scale.

The frequency distribution of the ordinal scores is graphed in Figure 4·21, showing that a little more than half the students were correct on part (ii) and a few more than 40% had part (iii) correct.

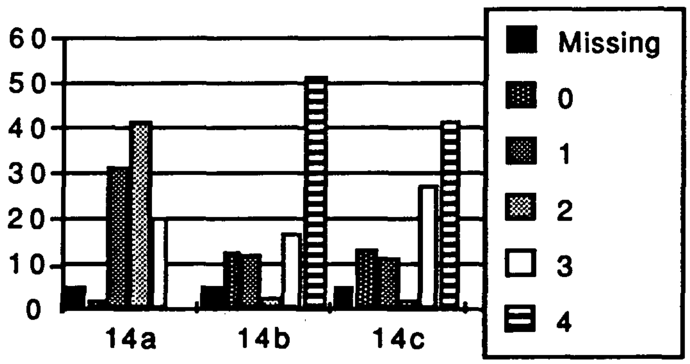


Figure 4·21. Percentage frequencies of responses to Item 14 when grouped to form an ordinal scale ($N = 517$)

Item 15

15. Decide whether the following statements are TRUE always, never or sometimes. Tick the correct answer. If you tick "**true only when** ..", write **when it is true**. All the letters stand for whole numbers or zero (0, 1, 2, 3, 4, ...)

(i) **a + b + c = a + x + c**

O true always
O never true
O true only when ....

(ii) **2a + 3b + 7 = 5a + 7**

O true always
O never true
O true only when ....

(iii)	2a = a + 2	O true always O never true O true only when ....
(iv)	a + 2b + 2c = a + 2b + 4c	O true always O never true O true only when ....

(Collis, 1975a, adapted)

The value of the information obtainable from items of this type has been described in Chapter 2. They were considered to provide valid measures of the students' ability to work with the concept of numerical variables. Little comment is needed here except, perhaps, to point out that part (iii) was not in the Collis (1975a) study but follows the same pattern of presentation as the other three parts which were in his study and is complementary to Item 12 which asked students to compare 'n + 2' with '2n'. It was noted that 78% of those who did "New Test 2 1990" responded in consistent ways to both these complementary questions.

The item proved most suitable for providing data for investigating variations in the levels of difficulty across different algebraic tasks. Figure 4·22 summarizes the success rates recorded in different stages of trialling. Part (i) was found to be the least difficult for all the groups tested. Part (iii) was the next in line, except for the 1990 Year 9 students who found part (ii) easier. Part (iv) was found the most challenging by all groups. Explanations for such outcomes are examined in Chapter 8.

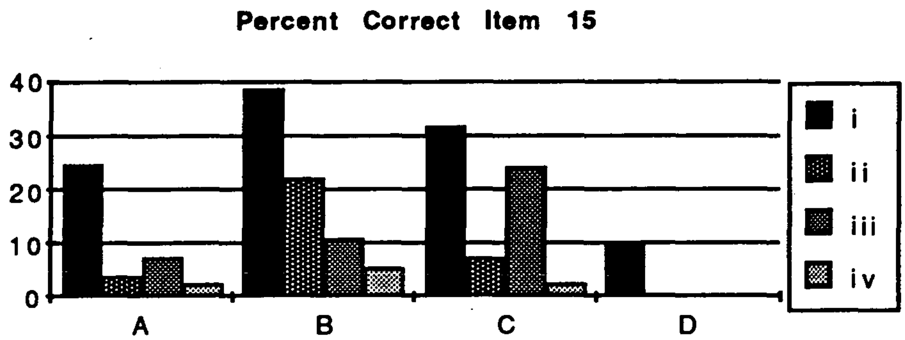


Figure 4·22. Comparison of success rates on the four parts of Item 15 (A - Groups I to V in 1989; $N = 132$ [Appendix 3D, Figure 3D·16]; B - Part of Group V in 1990; $n = 18$ [Appendix 3G - Question "6"]; C - Groups III and IV in 1990; $N = 41$ [Appendix 3H, Figure 3H·8]; D - Group VIII in 1990; $N = 19$ [Appendix 3J, Table 3J·2])

Table 4·31 presents the main study outcomes from Item 15 in two stages. In the first subpart, students were asked to choose whether the given statements were "true always", "never true", or "true only when ...". In the second subpart they were asked to write when the statement was true, if they had chosen the last option. To retain student responses in computer analyses, a code of "0" was given for one subpart whenever it was left blank after answering the other corresponding subpart. A score of "1" was allocated, however, if a wrong response was given. Those who had any subpart correct were given a score of "2".

Table 4·31

Percentage Frequencies of Responses to Item 15 parts (a) to (d)

CATEGORY DATA					ORDINAL DATA				
1st subpart Response Type	**% 15i**	**% 15ii**	**% 15iii**	**% 15iv**	**1st subpart Score**	**% 15i**	**% 15ii**	**% 15iii**	**% 15iv**
Omit 1st subpart	7.2	8.9	7.9	7.9	Missing	7.2	8.9	7.9	7.9
always	10.6	15.1	15.3	9.3	1	see below	see below	see below	see below
never	32.7	42.4	34.0	70.6	1 (total)	43.3	57.4	49.3	79.9
when	49.5	33.7	42.7	12.2	2	49.5	33.7	42.7	12.2
2nd subpart Response Type	**% 15i**	**% 15ii**	**% 15iii**	**% 15iv**	**2nd subpart Score**	**% 15i**	**% 15ii**	**% 15iii**	**% 15iv**
Omit both subparts	7.2	8.9	7.9	7.9	Missing	7.2	8.9	7.9	7.9
Omit "when" subpart	44.7	59.0	50.1	80.7	0	44.7	59.0	50.1	80.7
"when" incorrect	7.7	12.2	6.2	7.2	1	7.7	12.2	6.2	7.2
"when" correct	40.4	19.9	35.8	4.3	2	40.4	19.9	35.8	4.3

Note. $N = 517$. "see below" indicates where to find % tally on ordinal scale.

Figures 4·23 and 4·24 display the ordinal score frequencies. For each of the four parts of the item, it can be seen that those who scored "1" for having made incorrect choices in the first subpart (recorded as part 'a') were amongst those who scored "0" on the second subpart (part 'b'). In terms of those who scored "2" on the second subpart, the order of difficulty was, from easiest to hardest, part (i), part (iii), part (ii), and part (iv). An analysis of these difficulty levels is pursued in Chapter 8.

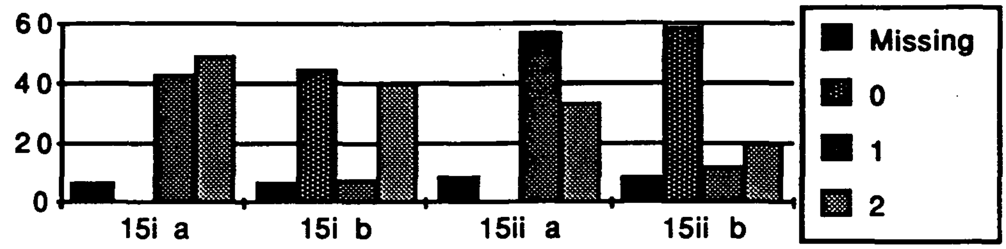


Figure 4·23. Percentage frequencies of responses to Item 15 parts (i) and (ii) when grouped to form an ordinal scale ($N = 517$)

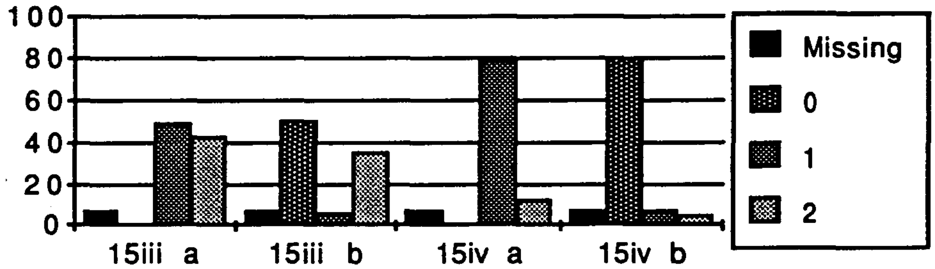


Figure 4·24. Percentage frequencies of responses to Item 15 parts (iii) and (iv) when grouped to form an ordinal scale ($N = 517$)

Comparison with Collis (1975a). The format of Item 15 was based on test items used by Collis with students spread evenly across ages 10 to 15 years. Table 4·32 presents the success rates on three parts of Item 15 in both the Collis study and this study.

Table 4·32
Comparison of Item 15 Outcomes from Quinlan & Collis Studies

Q.15	QUINLAN 1990		COLLIS 1975a	
part	Equation	% correct	Equation	% correct
i	$a+b+c = a+x+c$	40.4	$m+n+q = m+p+q$	25.0
ii	$2a+3b+7 = 5a+7$	19.9	$a+2b = 2a+b$	23.3
iv	$a+2b+2c = a+2b+4c$	4.3	$a+2b+2c = a+2b+4c$	5.0

Note. 1975 frequencies derived from Tables 2.3 & 2.4, Collis, 1975a, pp. 26 & 28.

Part (iv) was identical with one of those items (Item 5, Collis, 1975a, p. 118), with part (i) similar to Collis' Item 3(a) except that the letter-symbols used were changed. Part (ii) was based on the same principle as Item 4 from the Collis source, but it was changed from a consideration of '$a + 2b = 2a + b$' to a consideration of '$2a + 3b + 7 = 5a + 7$', and so was not as comparable as were the other two parts. Part (iii) was a new question using the same format and, by focusing on the case '$2a = a + 2$', it was a way of re-testing part (d) of Item 12, which has been discussed above. It was found that 93.5% of those who had Item 15 part (iii) correct also had Item 12 part (d) correct, and 77.9% of those who had Item 12 part (d) correct also had Item 15 part (iii) correct.

The 1975 data were collected from 60 students who did the questions individually in the presence of Collis as researcher. Collis' subjects returned a slightly higher success rate than did the 1990 subjects for parts (ii) and (iv), while the latter group were more successful on part (i). The fact that different letter symbols were used in the Collis study was discounted by both Booth and Collis as a likely influence on the differences in outcomes: Booth reported that results were similar when she used, in Collis style, considerations of both '$x + y + z = x + p + z$' and '$L + M + N = L + P + N$' (1983, p. 350), and Collis obtained similar results using '$a + b + c = a + b + d$' and '$m + n + q = m + p + q$' (1975a, p. 26).

Comparisons of Item 15 (i) results in four studies. Part (i) was used, sometimes with different letters, in the CSMS study (Küchemann, 1980), and in the Mathematical Understanding of Taiwan students project (MUT). Lin (1988) compared outcomes from the CSMS and MUT studies. Table 4·33 adds data from the present project and from Collis (1975a) to Lin's comparisons.

Table 4·33
Comparison of % Response Rates on Item 15 (i) in Four Studies

Responses →	"Never"			"Sometimes" with correct explanation		
Age-group	13	14	15	13	14	15
CSMS[a]	56	51	50	11	25	27
MUT[b]	27	17	14	24	46	52
QUINLAN[c]	38	39	61	29	51	27
COLLIS[d]	N.A.	N.A.	N.A.	20	60	50

[a] Data from Lin (1988, p. 481); about 1 000 students in each age group. [b] Data from Lin (1988, p. 481); nearly 200 in each age group. [c] $n = 69$ for 13 year-olds, 97 for 14 year-olds, and 33 for 15 year-olds. [d] Data from Collis (1975a, pp. 26 and 28); 10 in each age group. 'N.A.' denotes "not available".

The Taiwan study was investigating whether or not differences in language and society between England and Taiwan influenced children's mathematical understanding. Lin (1988) noted that more than half of the English students thought that different letters represented different numbers, and "compared with English students, there are more Taipei students who view word symbol [*sic*] as a generalized number or as a variable" (p. 480). When percentages were assembled according to age and regardless of class groupings for Australian students in the Quinlan study, it was found that these students had a slightly higher success rate than the Taipei students, except for the poorer results from the small subgroup of 33 students in the 15-years age-group. The latter group's rate of success was the same as that for the CSMS group of 15-year olds. The other students in the Quinlan study recorded more than twice the success rate of the corresponding groups of the CSMS students. Data obtained by Collis (1975a) showed that another small sample of Australian students produced success rates which were higher than those for the CSMS students and similar to or higher than those for the MUT students. As student samples were not matched on ability, background, or sample size, there is little justification for assuming that these comparisons would apply generally.

Review and Forecast

The chapter has described data which not only contributed to the store of research information regarding mathematics education but formed the ingredients for analyses to be discussed in remaining pages of this thesis. The presentation of overall results in percentage frequencies format was global in that it dealt with test responses for the total population of subjects.

The management of data has been explained in terms of recording information in categories which led to the formation of ordinal scales. Procedures for dealing with blank answers spaces have also been described. The influences of trialling on the determination of the composition of the final test have been discussed. For test items which were based on previous research, a general comparison of frequency outcomes has been recorded.

To help identify the major constructs embedded in the mass of data, procedures were undertaken for factor analyses and the formation of scales. Scales then made further data management more efficient by reducing the number of variables. The scores of each scale variable were attained by summing scores for the individual variables making up the scale. Chapter 5 is devoted to the methods used to determine which variables could justifiably be included together in such scales. The overriding importance of cognitive support for grouping variables is stressed.

We turn first to consider principal component analysis and factor analysis, two related statistical procedures used in establishing the scales.

CHAPTER 5

ESTABLISHMENT OF SCALES

Overview

This chapter is devoted to the statistical processes used to identify suitable measures for relevant cognitive constructs while reducing the number of variables for ease of data management.

Outcomes from a variety of factor analyses and principle component analyses are presented at the beginning of the chapter. Clusters of variables indicated by such analyses were evaluated for their suitability as members of linear scales. Other groupings of student responses were also tested for suitability as scale measures of common thought processes which sometimes included incorrect answers. The statistical criteria for acceptance of groups of items into scales are detailed.

The report on the establishment of scales is presented in three parts:

1. The first concerns the scales which were built on correct responses and which purported to measure the degree to which students had developed correct understandings and skills in early algebra.

2. The second part concentrates on scales which were called *Progress Indicators*. These were groupings of responses which included the correct answers as well as some responses which, while not fully correct, indicated that the students were progressing towards the correct way of thinking about the algebra in the question.

3. The third type of scales measured misconceptions and incorrect procedures, and these scales were called *Hindrance Indicators*.

The Progress and Hindrance Indicators were of critical importance for many of the analyses discussed in the chapters which follow.

Factor Analyses

One class of analyses was directed towards the establishment of scales, a procedure which focused on searching for possible commonalities of cognitive processes reflected in the student responses across test items. A unifying purpose of multidimensional scaling techniques is to identify "whatever pattern or structure may otherwise lie hidden in a matrix of empirical data" (Shepard, 1972, p. 1). This search was assisted by using the computer for factor analyses and for testing the reliability of scales formed by grouping certain responses together. Before any groupings were accepted, the researcher imposed the criterion of an intelligible rationale for considering the responses as belonging to the same scale.

A major benefit from the establishment of factors and scales and is that, by concentrating on relationships within sets of variables, a reduction in the dimensionality of the data is achieved (cf. Cureton & D'Agostino, 1983, p. 2). The number of variables is reduced "with as little loss of information as possible" (R. J Harris, 1975, p. 23), making data management considerably more efficient when interrelationships between dimensions are examined.

Ordinal Variables. Chapter 4 has described the formation of Ordinal Variables from the categorization of the types of answers students gave to the test questions. Data expressed in this form were used for input to the factor analysis program available in the SPSS^X Version 2.1 computer package, while excluding responses in the missing values category.

Dichotomous Variables. Factor analysis outcomes were also examined using the data in a Dichotomous Variable form. This was the form in which a score of "1" was allocated for a correct answer and a score of "0" for an incorrect answer or one that fell into the omit category, while leaving the missing values category out of consideration. (The distinction observed between missing values and omit categories was explained in Chapter 4.) The Dichotomous Variable form was ordinal in that a correct response merited a higher score than an incorrect one, but there were only the two possible scores, in contrast to the Ordinal Variable form in which a greater variety of ordered scores was allocated. R. J. Harris (1975, pp. 226 - 231) argued convincingly for the acceptance of the application of multivariate procedures to data presented in dichotomous form, and the results of factor analyses described later in this chapter support his contention.

Missing values excluded. One of the difficulties faced was the loss of student inclusions in analyses on account of "missing values" somewhere in their test responses. From the 517 students who completed the research test, 241 (i.e., 46.6%) responded to the questions in such a way as to avoid having any blank spaces classified as "missing values". Only these 241 students were included in the first batch (Table 5·2) of factor analyses since all the test responses were analysed together. Two methods were applied to check on how representative these students were of the whole sample population. Firstly, to check that there was a spread of ability levels, the number from each Year group was examined and, secondly, the number in each quartile subdivision based on the overall test score was investigated. It was found that the representation was satisfactory. There were at least some students from each of the Year groups from Year 7 to Year 12, with more than 40% of the beginning Year 7 students included. Although the representation in the lower quartile divisions was somewhat small, there was representation from each of the four groups bounded by the test quartiles. Tables 5A·1 and 5A·2 of Appendix 5A provide the details.

Missing variables included. As a way of checking the robustness of the findings from the analyses carried out on the subgroup of 241 students, factor analyses were also carried out on data from all 517 students. The method used to ensure the inclusion of all students was to allocate a zero score to all blank answer spaces. The logic in doing this was based on the assumption that if a student left an answer out then the student did not know how to do it and could be classified as incorrect on that question, thus meriting a zero score. As the following details confirm, general agreement was found between factor analyses which included all students and those which included only those who answered all the questions in such a way as to avoid having any missing value classifications.

The inclusion of missing values was a technique used only in the second batch (Table 5·3) of principal components analysis and factor analysis methods of exploration for groupings of variables likely to be useful in the formation of scales. In all other analyses throughout the rest of the thesis, missing values were excluded.

Variety of factor methods. C. W. Harris (1967) compared a variety of methods for factor analysis that were available to him at the time he was applying the procedure. His recommendations within a strategy for factor studies were to use several computing methods on the same data, compare the outcomes, and "regard as the important substantive findings those factors that are robust over methods - but only those" (p. 369). C. W. Harris' advice was followed by submitting the test data to the seven factor analysis methods available in the SPSS^X Version 2.1 package. The seven methods are referenced below by the abbreviations given in SPSS^X User's Guide (SPSS Inc., 1986, p. 718), which are as follows:

PC	Principal components analysis
PAF	Principal axis factoring
ALPHA	Alpha factoring
IMAGE	Image factoring
ULS	Unweighted least squares
GLS	Generalized least squares
ML	Maximum likelihood.

On the question of distinguishing between types of analysis, R. J. Harris (1975) points out that "authors differ in whether they consider principal components analysis to be a type of factor analysis or a distinct technique" (p. 23). In this account of the outcomes attained by the application of both approaches to the same research data, the terms "factors" (for factors or principal components) and "factor weightings" (for factor weightings or for principal component weightings) will be applied indiscriminately for the sake of simplicity of presentation. As will be seen, the outcomes were closely similar regardless of the method applied. Both approaches aim to reduce the dimensionality of a variable set by forming "factors" which are linear combinations of the original set of variables and are fewer in number. The basic difference is that principal components analysis operates on the total variance in the set, whereas all factor analysis models explicitly separate the unique variance from the common variance (R. J. Harris, 1975, p. 25; Dunteman, 1984, p. 183). The unique variance is that part of the variance of responses to any item which is independent of responses to all other items.

When all seven available methods were applied to the data in Ordinal Variable form with the exclusion of missing values, the analyses were completed in four cases, namely PC, ALPHA, IMAGE and ULS. The same four methods also attained completion when the input data was in the form of Dichotomous Variables with the exclusion of missing values. In both cases, the PAF analysis was arrested with the finding of a "communality greater than one" and each of the other two methods led to a "no local minimum" finding. When the data were admitted in Dichotomous form with the inclusion of missing values, the five methods that produced completed factor analyses with the whole student population were PC, ALPHA, IMAGE, ULS and GLS. The completed analyses included Varimax rotation of the factor loadings to assist interpreting the outcomes by "maximizing the variance of the squared loadings in each column" (Nie, Bent, Jenkins, Steinbrenner, & Hull, 1975, p. 485), thus simplifying the columns. All loadings of factors on variables reported below are those in the factor pattern matrix (cf. Cattell, 1978, p. 235) obtained after convergence using Varimax rotation was attained.

The number of variables in each of the analyses was 62 and, consequently, there were 62 variables in each of the factors produced in the analyses. In each of the four successfully completed methods of analysis in which missing values were excluded, the number of factors extracted was 16 for Ordinal Variable input and 17 for Dichotomous Variable input. When missing values were included so that all the student population was considered in the factor analyses, the number of factors derived in all five successful methods was 14.

To allow comparisons, factor loadings for the three PC analyses of the data and one ALPHA analysis are tabulated in appendices, as follows:

Appendix 5B: using PC with ordinal data, rejecting missing values;

Appendix 5C: using PC with dichotomous data, rejecting missing values;

Appendix 5D: using PC with dichotomous data, keeping missing values; and

Appendix 5E: using ALPHA with ordinal data, rejecting missing values.

Clusters

Timm pointed out that a principal component analysis could be used to assist in the "understanding of the dependencies existing among variables of a set and also determining whether subsets of variables cluster, or go with one another" (1975, p. 528). The same held true for factor analyses. Within each "factor" (factor or principal component), clusters of variables were identified by the relative sizes of the factor loadings. Variables in the cluster had higher factor loadings within the given factor, which will be called the *Cluster Factor*, than they did in any other factor, and their loadings were higher than loadings for other variables within the Cluster Factor.

As an example of the way the cluster variables were identified within a rotated factor pattern matrix, Table 5·1 presents the crucial information about factor loadings for a cluster which appeared on all of the analyses referred to above. It was the cluster called *Cluster E* which reflected the strong affinity between responses to each of the four parts of Question 11 about the lengths of two parallel lines. The cluster variables were a subset of the 62 variables which formed a certain factor, referred to as the Cluster Factor.

Table 5·1
Factor Loadings for Cluster E From Item 11

Question	Factor Loadings WITHIN Cluster Factor	Highest Factor Loadings OUTSIDE Cluster Factor
11 (b)	.92	.18
11 (c)	.91	.21
11 (d)	.83	.25
11 (a)	.77	.11
12 (a) [outside Cluster E and with next highest loading in Cluster Factor]	(.22)	(.67)

Note. $n = 241$. PC method of factor analysis with Ordinal Variables as input and rejecting all missing values. Question 12 (a) used as referent from outside the cluster.

It can be seen that Table 5·1 lists the loadings within the Cluster Factor for the test items which formed the cluster, and also loadings surrounding these within the matrix. It includes the next highest loading to be found within the Cluster Factor but loaded on some variable, namely scores on Question 12 (a), which did not belong to the cluster. This provides some idea of how distinct the cluster is within that Cluster Factor. The next highest loading outside the Cluster Factor for each constituent question is located in some other factor of the matrix. These next highest loadings check that the cluster variables have their highest loadings within the Cluster Factor, as expected, and they also give some indication of the degree of "ownership" which the Cluster Factor has of the cluster variables. For the example given, the loadings for the cluster variables within the Cluster Factor are noticeably larger (.77 or more) than their loadings on some other factor (.25 or less). For Question 12 (a), the loading on the Cluster Factor is small (.22) but the loading on some other factor is large (.67), indicating that it "belongs" to the latter factor and not the factor within which the Item 11 variables were clustered. These figures show that there is strong statistical support for considering the formation of Cluster E from the four parts of Item 11.

Factor Analysis Outcomes

The data bank used as input for the principle components and factor analyses was the same as that described and discussed throughout Chapter 4. It was based on responses from all 517 students, selecting Test 3 responses from the Year 7 students who completed the test more than once. Percentage frequencies for all item scores were presented in Chapter 4 for the total student group.

Table 5·2 summarizes the clusters identified by the eight factor analyses completed when missing values were excluded. This is followed by Table 5·3 which summarizes five analyses on dichotomous variables while including missing values. The clusters are listed in the order in which they occurred in the PC method of analysis when the data were presented as Ordinal Variables.

The test items were such that, even when the data were coded to form dichotomous variables, the factor analysis outcomes were very similar to those obtained by using the test responses in ordinal data form, thus endorsing R. J. Harris' (1975) argument for using dichotomous variables in multivariate analyses. The similarity between the findings of all eight analyses in Table 5·2 showed that, while responding to C. W. Harris' (1967) recommendation to seek factors that are robust, a number of clusters appeared likely to be suitable for the formation of bona fide scales. Clusters E, G, H, I, K, and L were identical in all eight analyses. With the other clusters, there were variations from one analysis to the next but there were questions in common across these variations. For instance, in Cluster A, the test questions which were common were all parts of Questions 10, 12 and 15 (iii) and, in some analyses, Cluster A also included parts of Item 6 or 8 or 14.

Table 5-2
Item Clusters From Factor Analyses Which Excluded Missing Values

CLUSTER		METHOD OF ANALYSIS ("ALL" means cluster of all parts of questions listed)							
		ORDINAL DATA				DICHOTOMOUS DATA			
Cluster	Questions	PC	ALPHA	IMAGE	ULS	PC	ALPHA	IMAGE	ULS
A	6c,10, 12,15iii	ALL + 6a iii, 8a - 6c	ALL + 6a iii, 8a	ALL + 6a ii, iii,iv, 8a,c	ALL + 6a ii,iii, 8a,14i	ALL - 6c	ALL	ALL + 6a ii,iii, 14i	ALL
B	5, 9, 14ii,iii	ALL	ALL	ALL - 14ii,iii	ALL - 14ii, - 14iii + 6b	ALL + 2ii	ALL - 14ii + 6b,8a	ALL + 6b,8a	ALL - 14ii, - 14iii, +6b,8a
C	3, 8b	ALL	ALL	ALL + 2ii	ALL + 2ii	ALL	ALL	ALL + 2ii	ALL
D	13	ALL	ALL	ALL + 14i	ALL	ALL	ALL	ALL	ALL
E	11	ALL	ALL	ALL	ALL	ALL	ALL	ALL	ALL
F	1	ALL	ALL	ALL - 1b	ALL	ALL	ALL	ALL	ALL +2ii
G	4	ALL	ALL	ALL	ALL	ALL	ALL	ALL	ALL
H	15ii	ALL	ALL	ALL	ALL	ALL	ALL	ALL	ALL
I	6a v,vi,vii	ALL	ALL	ALL	ALL	ALL	ALL	ALL	ALL
J	7	ALL +8d	ALL +8d,14i	ALL +8d	ALL +8d	ALL	ALL	ALL	ALL
K	15iv	ALL	ALL	ALL	ALL	ALL	ALL	ALL	ALL
L	2i	ALL	ALL	ALL	ALL	ALL	ALL	ALL	ALL
M	15i	ALL + 6c	ALL	ALL	ALL	ALL	ALL	ALL	ALL
N	6a i,iv,2ii	ALL	ALL	-	ALL - 2ii	ALL	ALL	-	ALL - 2ii
O	8c,d	-	-	-	-	ALL + 8a	ALL	ALL	ALL
% Variance Explained		74.5	66.4	64.3	67.7	72.7	63.3	58.1	63.9

Note. Years 7 to 12, with Test 3 responses for Year 7 students. $n = 241$.

Table 5·3

Item Clusters From Factor Analyses Which Included Missing Values

CLUSTER		METHOD OF ANALYSIS ("ALL" means cluster of all parts of questions listed)				
Cluster	Questions	PC	ALPHA	IMAGE	ULS	GLS
A	6c,10, 12,15iii	ALL	ALL + 2i, 8a	ALL + 8c - 15iii - 6c	ALL + 8c - 15iii - 6c	only 10 + 8d
B'	5, 9, 14	ALL	ALL	ALL + 2ii,6c,8a, 15iiib	ALL + 2ii,8a	ALL + 2ii,8a
C	3, 8b	ALL	ALL	ALL + 6b	ALL	ALL + 6b
D	13	ALL	ALL	ALL	ALL	ALL
E	11	ALL	ALL	ALL	ALL	ALL
F	1	ALL +6b	ALL +6b	ALL	ALL +6b, c	ALL
G	4	ALL	ALL	ALL	ALL	ALL-4a + 12a,d, 15iii
H & M	15i,ii	ALL	ALL	15i only	ALL	ALL
I'	6a vi,vii	-	-	ALL	ALL	ALL
J	7	ALL	ALL	ALL + 8d	ALL	ALL + 8d
K	15iv	ALL	ALL	ALL +15ii, 15iiia	ALL	ALL
L	2i	ALL + 2ii	ALL	ALL	ALL	ALL + 4a
P	6a i to v	ALL + 6a vi,vii	ALL + 6a vi,vii	ALL	ALL	ALL + 6c
% Variance Explained		69.7	61.6	57.5	62.7	63.0

Note. Dichotomous Variables. Years 7 to 12, with Test 3 responses for Year 7 students who did the test more than once. $N = 517$.

Several clusters were so robust that they appeared also in factor analyses which included responses from all 517 students by scoring "missing values" as "0" as if they represented errors. Table 5·3 is presented for comparison with Table 5·2 to confirm the strength of the clustering of certain variables and, hence, the likelihood that valid scales could be formed from at least some of these clusters. In both tables, the following clusters included the same group of questions: A, C, D, E, F, G, J, K, L. When all students were included, Questions 15 (i) and 15 (ii) were grouped together, whereas they were grouped separately (as Clusters H and M) when students with missing values were excluded. With the full cohort of students, Cluster B became Cluster B', a cluster which included one additional part, namely, part (i) of Question 14, while Cluster I became Cluster I' by dropping one part, namely, part (v) of Question 6(a). A new cluster, Cluster P, which incorporated the first five parts of Question 6(a), was identified with the larger group but not the subgroup. Other clusters were revealed only in the analyses with the subgroup: Cluster N (Questions 2 ii, 6a i, iv) and Cluster O (Questions 8c, d).

Conclusion From Factor Analyses

The factor analysis outcomes summarized in Tables 5·2 and 5·3 gave strong indications that scales could be formed by grouping responses to test questions which measured similar cognitive constructs or ability with similar algebraic skills. This high probability was confirmed by trialling various combinations of responses and subjecting them to statistical scrutiny to establish linear scales, as is explained below.

Scale Formation and Analysis

The next major stage in establishing scales was to study the outcomes from statistical analyses of scales formed by clustering together those test variables which seemed likely to meet the requirements for satisfactory scales. The SPSS^X Version 2.1 program "Reliability" (SPSS Inc., 1986, pp. 856 - 872) was used for statistical analyses of possible scales.

The outcomes from the thirteen completed factor analyses were used as a guide in determining combinations of variables that would be trialled for scaling. While there were some variations between clusters from one factor analysis to the next, certain groupings of variables appeared consistently. These were the groupings tested for the possibility of forming scales, provided that there was some underlying cognitive association or similarity which could justify treating the variables within the one scale score. The underlying cohesive aspect was identified in the naming of the scale, if this could be done efficiently (e.g., the VBL Variable Scale grouped items which measured the degree to which students had developed the notion of algebraic symbols as numerical variables). The values of the factor loadings in the rotated factor pattern matrices signposted the most likely groupings of variables for successfully forming scales. If all the cluster members had high positive loadings, this was taken as a sign that they could "belong" to the same scale. Cluster E was used above as an example of a cluster with high factor loadings on each variable in the cluster. However, if loadings within a cluster varied in sign or if some of them were small compared with others, they indicated that the members of that cluster were not likely to scale equitably together.

Four Requirements for Scaling

Four requirements were imposed when accepting any grouping of variables for the formation of a scale. The first requirement had a psychological base and the other requirements were of a statistical nature. They are explained in turn.

1. Cognitive unity was evident.

The first requirement was that the items to be included should be related to each other in a cognitive sense. A rationale needed to be clear for grouping them together and this requirement was assessed by studying the cognitive processes and demands associated with responding to the actual test items.

2. Cronbach's *alpha* reliability coefficient, α, was satisfactory for the particular cohort of students.

Amongst the meanings that Cronbach (1951, p. 331) claimed for his *alpha* reliability coefficient, the following were the most relevant to the procedures for establishing groups of items for measuring a common cognitive process:

> (a) α is the mean of all possible split-half coefficients; ...
> (d) α estimates, and is the lower bound to, the proportion of test variance attributable to common factors among items. That is, it is an index of common-factor concentration ... ;
> (e) α is an upper bound to the concentration in the test of the first factor among the items.

For the present discussion, the word "test" will be replaced by "scale". Interpretation (a) means that α indicates the degree of reliability of measurement provided by the scale, or its "accuracy or dependability" (Cronbach, 1951, p. 297). Aspects (d) and (e) were of particular interest in the process of establishing scales: Higher α values indicated higher concentration of common purpose in the measurements made by the scale items. As Cronbach points out, "What is required [for a test to be interpretable] is that a large proportion of the test variance be attributable to the principal factor running through the test" (1951, p. 320).

3. Item-item correlations were sufficiently high.

Correlations provide a measure of the degree of similarity between response patterns to pairs of test items. For items to belong to a scale grouping they should be measuring the same factor of information. Thus, the expectation would be that they should be reasonably well correlated. This would indicate substantial internal consistency and make the scale interpretable psychologically.

After examining patterns across data concerning many scaling trials, it was noted that the average item-item correlation was a very good indicator of the likelihood of a successful outcome. Groupings which fared better as scales tended to have the better mean item-item correlations.

Groups of items which produced any negative correlations were disallowed.

4. Corrected item-scale correlations were sufficiently high.

A corrected item-scale correlation is defined by the statement "This is the correlation between that item's score and the scale scores computed from the other items in the set" (SPSS Inc., 1986, p. 861). For an item to merit a place in a set to form a scale it needs to have at least a reasonable correlation with the total scale score calculated from the other items, otherwise it would be questionable that it could make a justifiable contribution to the scale. As Dunn-Rankin (1983, p. 92) pointed out, this correlation acts as a discrimination index for each item and a high correlation here indicates internal consistency and support for the inclusion of that item. Agreeing with his advice, items which were negatively correlated with a scale total were eliminated from that scale.

Example of Test for Satisfying Scale Requirements

To illustrate the process of scale-formation, requirements for forming a scale from the four items in Cluster E are now considered. The factor loadings detailed in Table 5·1 indicate that responses to the parts of Question 11 are interrelated and could possibly match the four requirements listed above for establishing a scale. These are applied in turn.

Cognitive unity. Cluster E was composed of scores on the four parts of the same test item, Question 11, which used one of Harper's (1979) Parallel Lines Subtasks. Each part contributed to the measurement of only one cognitive aspect or skill. In this case, the key requirement for success was to overcome the strong and deliberate geometric distracter in the question so that judgements could be made based on the algebraic significance of defining the lengths of the lines as "*a* cm" and "*b* cm" respectively. Those who successfully applied algebra to the four parts of Item 11 and who thus showed their ability to work with the variable or species interpretation of algebraic symbols, were allocated higher scores than the other students in both the ordinal and the dichotomous forms of the data. Support for some underlying unity was also given by the clustering of the four parts of Question 11 by the factor and principle component analyses, as reported in Tables 5·2 and 5·3.

The *alpha* reliability coefficient, α. The *alpha* value was very high, at .92 for the data in ordinal form and .95 in dichotomous form, as recorded in Table 5·4. (The *alpha* values for all scales established for this thesis were at least .50.)

Item-item correlations. The inter-correlations between the scores on the four parts of Item 11, as displayed in Table 5·4, were sufficiently high to indicate that students attained similar response patterns to each part. The average item-item correlations were .72 for ordinal data and .82 or .81 for dichotomous data. The minimum correlations were .55 or higher. The maximum correlations, between scores on parts (b) and (c), were so close to 1, at .94 or .98, that they indicated that these two parts of the item measured approximately the same cognitive process, which was the thinking required to apply algebra to the consideration of when one or other of the two sketched lines was the longer. It was decided to include scores on both of these parts in the scale, giving credit to those students who successfully applied the required thinking in both parts.

The Pearson Product-Moment Correlation formula produced similar correlation coefficients for both forms of the data, ordinal (scored as specified on page 126) and dichotomous (scored as explained on page 143). Table 5·5 and the ensuing comments compare similar paired correlations for Correct Response Scales. As the correlations were generally of comparable value for the two scoring systems, Pearson Correlation Coefficients are used throughout this thesis. The alternative Phi Coefficient is not invoked for dichotomous data.

Table 5·4

Statistical Analysis of Scale Formed From Item 11

SCALE PL	ORDINAL DATA excluding missing values $n = 241$	DICHOTOMOUS DATA excluding missing values $n = 241$	DICHOTOMOUS DATA including missing values $N = 517$
Cronbach's Alpha	.92	.95	.95
Range of item-item correlations	.55 - .94	.68 - .98	.67 - .98
Average of item-item correlations	.72	.82	.81
ITEM	CORRECTED ITEM-SCALE CORRELATIONS		
11a	.59	.75	.73
11b	.91	.94	.94
11c	.91	.95	.95
11d	.85	.86	.86

Corrected Item-Scale Correlations. For each of parts (b), (c) and (d), the correlation with the score attained by summing scores on the other parts of Item 11 was .85 or higher. In the case of part (a), which was a general introductory question, the corresponding correlation was lower but yet sufficiently high, from .59 to .75, to be acceptable as an indicator that this part, with the others, contributed positively to the formation of the scale total.

Conclusion. This scale was assessed as satisfying the four requirements and it was called the PL *Parallel Lines Scale*, giving students scores from 0 to 4 by summing their scores on each part of Item 11 without using any weighting. Thus, a student who had, say, two parts of Item 11 correct was allocated a score of "2" on the PL Scale.

Similar analyses were carried out on all the scales which were accepted for use in investigations reported later. The establishment of all these scales is described next.

The Scales

As mentioned in the overview of this chapter, three classes of scales were established, namely, those built on correct responses only; those constructed from near-correct as well as correct responses, giving what are called Progress Indicator Scales; and those built upon responses which indicated incorrect understandings of basic algebra, resulting in the Hindrance Indicator Scales . These scales provided the ordinal measures used in the investigations reported later.

Summary of allocation of response types to scales. Appendix 5F summarizes the allocation of various response categories to the scales discussed below. The summary takes each test question in turn and matches the different response types with their respective scales.

Statistics for Scales Based on Correct Responses

Fourteen scales were accepted as measures of correct thinking. For the component variables in these scales, the highest scores were reserved for correct answers to the corresponding questions. In the case of the ordinal data, errors were allocated graded scores according to the types of errors but, with the dichotomous data, all errors were scored at zero. The same scale groupings were supported for both the ordinal and dichotomous forms of the data. Table 5·5 summarizes the relevant features of the analyses.

Table 5·5
Summary of Analyses of Scales for Correct Responses

Questions	Scale (Cluster)	No.of Items	No.of Cases	Cognitive Focus	*Alpha*	Item-item Correlations		Corrected Item-Scale Correlations	
						Mean	Range	Mean	Range
6c,10, 12, 15iii	VBL (A)	11	407	Variable	.94 (.93)	.66 (.53)	.42 - .95 (.16 - .89)	.79 (.70)	.68 - .87 (.27 - .78)
5,9, 14ii,iii	SYM (B)	6	440	Symbols in answers	.87 (.86)	.58 (.50)	.43 - .84 (.34 - .79)	.70 (.65)	.63 - .77 (.56 - .74)
3, 8b	SUBS (C)	7	489	Substitution & Solving	.93 (.90)	.66 (.59)	.35 - .94 (.33 - .89)	.79 (.72)	.42 - .88 (.46 - .84)
13	EQN (D)	4	441	Covariance in Equation	.92 (.86)	.74 (.62)	.64 - .89 (.50 - .87)	.81 (.71)	.74 - .90 (.69 - .75)
11	PL (E)	4	491	Variable in Geometry	.92 (.95)	.72 (.81)	.55 - .94 (.69 - .98)	.81 (.87)	.59 - .91 (.75 - .95)
1	AR (F)	4	471	Arithmetic Operations	.82 (.84)	.61 (.57)	.42 - .82 (.39 - .74)	.71 (.68)	.67 - .75 (.62 - .75)
4	FL (G)	3	448	Alg. for no. of flowers	.78 (.62)	.61 (.36)	.57 - .64 (.25 - .45)	.68 (.44)	.63 - .71 (.38 - .51)
15i,ii	BXBA (H&M)	4	461	Variable $b = x$; $b = a$	.85 (.84)	.65 (.57)	.48 - .94 (.43 - .83)	.74 (.68)	.71 - .80 (.62 - .73)
7	PS (J)	2	507	Covariance Prof-Student	.88 (.89)	.78 (.80)	.78 - .78 (.80 - .80)	.78 (.80)	.78 - .78 (.80 - .80)
15iv	CZ (K)	2	476	Variable $c =$ zero	.92 (.67)	.90 (.56)	.90 - .90 (.56 - .56)	.90 (.56)	.90 - .90 (.56 - .56)
2i	AD (L)	2	407	Independent Variables a,d	.84 (.94)	.77 (.89)	.77 - .77 (.89 - .89)	.77 (.89)	.77 - .77 (.89 - .89)
6a vi,vii	NJCP (I')	2	482	Reject $c =$ Objects	.95 (.95)	.91 (.91)	.91 - .91 (.91 - .91)	.91 (.91)	.91 - .91 (.91 - .91)
6a i,iv	C2 (N - 2ii)	2	437	Generalized Number	.58 (.58)	.46 (.46)	.46 - .46 (.46 - .46)	.46 (.46)	.46 - .46 (.46 - .46)
8b,10d, 11d,12d, 13c,15	EQL#	9	385	Equality	.82	.32	.08 - .72	.35	.12 - .62

Note. Figures not in brackets: Ordinal Variables; Figures in brackets: Dichotomous Variables. Missing values rejected. Years 7 to 12, with Test 3 responses for Year 7.
EQL Scale: Dichotomous Variables only and not forecast by factor analysis.

The assessments are summarized in Table 5·5 in terms of the four scaling requirements listed above, namely, the cognitive focus, the Cronbach *alpha*, the item-item correlations, and the corrected item-scale correlations. The number of cases varies from one scale to the next, depending upon the pattern of missing values. The scales are described and discussed in a sequence determined by the cluster order in Table 5·2, the order of factor allocation using the PC method of analysis as presented in Appendix 5B. This decision resulted in sequencing the scales according to the number of items grouped within them.

Individual scales. The statistical support was considered sufficient to justify grouping the items into scales as listed in Table 5·2. Brief comments are appropriate about the cognitive focus of each scale and about aspects of component items.

1. The VBL Scale, the Variables Scale, was based on Cluster A which was embedded in the factor which explained more of the overall variance in the test responses than did any other factor in each of the 13 analyses summarized in Tables 5·2 and 5·3. This claim may be verified for each of the four examples of factor loadings detailed in Appendices 5B to 5E by comparing the eigenvalues for the factor F1 (which included Cluster A) with eigenvalues for the other factors. Cluster A consisted of Question 6 (c) ($c + d = 10$, $c < d$, c = ?), all four parts of Question.10 (compare '$t + t$' and '$t + 4$'), all four parts of Question 12 (compare '$2n$' and '$n + 2$'), and the two parts of Question 15 (iii) (when does '$2a$' equal '$a + 2$'?), a total of 11 test responses. In Questions 10 and 12, students needed to consider the letters as standing for a variety of possible numbers for identifying conditions for pairs of expressions to be equal or different in value. Question 15 (iii) was actually a repeat of part of Question 12 in a different form, which explains why the factor analysis program had clustered Question 15 (iii) with Question 12. Question 6 (c) extended students to think of symbols 'c' and 'd' as representing numerical variables which had a covariate attribute imposed by the equation '$c + d = 10$'. To merit a score of "1" for the VBL Scale, students were required to give responses to Question 6 (c) which allowed 'c' to take on any value less than 5 and not simply integer values such as 4, 3, 2, or 1. The cognitive focus common to the eleven questions was the concept of an algebraic variable, so the scale was called the VBL Variables Scale.

The two statistical analyses detailed in Table 5·5 both gave reasons for reducing these eleven variables of the VBL Cluster into one scale variable. The values for *alpha* and the correlations were supportive. The ordinal data resulted in higher correlations as it included the error pattern of the students and these were, apparently, persistent from one question to the next. There were some very high item-item correlations such as .95 between the two parts of Question 15 (iii) using ordinal data and .89 using dichotomous data. Although these figures indicated that the pairs of responses were practically measuring the same skills or understandings, it was decided to include them in the scale tallies.

2. The SYM Scale, the Symbols Scale, was highlighted by Cluster B and

comprises responses to all three parts of Question 9 (e.g., "Add 4 onto $3n$"), Question 5 (points in football match), parts (ii) and (iii) of Question 14 (algebra for numbers on a school excursion). To succeed on each of these six items, students needed to accept a symbolic rather than numeric format for their answers. They had to operate on algebraic symbols without knowing their numerical values. The items measured their willingness to accept lack of closure and to leave an unevaluated symbolic expression as an answer.

3. The SUBS Scale, the Substitute and Solve Scale, was derived from Cluster C. Six of the component items in this scale were parts of Question 3 and tested students' ability with substitution into algebraic expressions. The seventh item was Question 8 (b) about solving the equation '$3a = 36$'. The cognitive link between the first six components was clear: Each part asked students to substitute the value '3' for 'y' in a first degree algebraic expression, thus testing their understanding of the conventions for writing the expressions in algebra. The exercise in solving the equation resulted in a numerical answer, as was the case with the other items in the scale. The cognitive link was judged to be the interpretation of conventional algebraic forms followed by movement from algebraic symbols via arithmetic to produce an answer in the more familiar numerical form rather than in letters. Each of the factor analyses and principal components analyses indicated that these seven items should be clustered together. The *alpha* values were .9 or more and justified the grouping of these items even without the further support of the positive correlations, as summarized in Table 5·5. Responses to the first two parts of Question 3 recorded high correlations (.95 for ordinal data and .90 for dichotomous data), indicating that they were measuring closely-related skills: Part (ii) simply required students to add '5' to their answer in part (i). Both were, however, included in the scale.

4. The EQN Scale, the Equation Scale, was directly derived from Cluster D and consisted of the four parts of the Harper (1979) Equations Subtask about the equation '$2x + y = 9$' in Item 13 of the test. The cognitive processes required for correctly answering this question included the recognition that the value of one of the given symbols, 'x' or 'y', was determined by the value of the other. Therefore, they did not have independent arbitrary values, as some students erroneously thought.

5. The PL Scale, the Parallel Lines Scale, based on responses to Question 11, was used earlier as an illustration of the methods of analysis applied to the scales. Further comment is not considered necessary.

6. The AR Scale, the Arithmetic Scale, combined scores on the four subparts of Item 1, which had been grouped into Cluster F by the factor analysis calculations. These subparts were cognitively related in that they asked students to interpret the equation '$3 * 4 = 6 * y$' for the arithmetic processes of addition and multiplication, and then to solve for the value of 'y'.

7. The FL Scale, the Flowers Scale, was named after the content of the test item which contributed the scores for the scale components. It was Item 4 which tested students' ability to interpret algebraic symbols presented in a real-life context about a number of flowers, some in bunches and some not. Forming a scale from the three parts of the item made cognitive sense. Either the question was interpreted in the way it was intended or students made little progress with it. Only a minority of students succeeded to any great extent, as was reported in Chapter 4. Support for forming the FL Scale was given by statistics in Table 5·5, mainly those based on ordinal data. The figures were lower for dichotomous data, especially under the heading "Item-item Correlations", showing that the tendency to get all parts completely correct or wrong was not very strong.

8. The BXBA Scale, the $b=x$, $b=a$ Scale, was composed of the four subparts in Questions 15 (i) and 15 (ii). The questions required the recognition that two variables could take a range of values and that it was possible for both variables to have the same value simultaneously. In part (i) it was a case of seeing that '$b = x$' so that '$a + b + c$' could equal '$a + x + c$', and in part (ii) that '$3b$ ' equalled '$3a$', leading to '$b = a$' so that '$2a + 3b + 7$' could equal '$5a + 7$'. Thus part (ii) required one step more than part (i). These were clustered in two groups by the analyses summarized in Table 5·2, but were grouped into one cluster in four of the analyses shown in Table 5·3. Taking all this information into account and seeing all questions as measuring the ability of the students to work with the variable notion, it was decided to test the grouping of all four subparts into one scale by means of the Reliability program. The outcome was quite satisfactory, giving a strong *alpha* value of .84 or .85 and averages of about .6 for item-item correlations and .7 for corrected item-scale correlations.

9. The PS Scale, the Professors-and-Students Scale, consisted of the two parts of Question 7. To score "1" on this scale, students had to avoid the reversal error and choose descriptors of the letters 'P' and 'S' which were in terms of numbers of people rather than simply people. The assessment statistics strongly supported this scale.

10. The CZ Scale, the c=zero Scale, was built from the two subparts of question 15 (iv), the hardest item on the test. Only 22 students had both subparts of this question correct. Despite the difficulty for most students, the statistical analysis reported in Table 5·5 supported the grouping of the scores on the two subparts.

11. The AD Scale, the a, d Scale, was a grouping of answers to the two subparts of Question 2 (i) which asked students to compare the values of two unrelated algebraic symbols. The correlations between the ideas presented in these subparts were very high and the scale was strongly endorsed.

12. The NJCP Scale, the No Cabbage or Pear Scale, derived from Cluster I' in Table 5·3, took in only the last two subparts of Question 6 (a). The factor loadings for Cluster I from Table 5·2, made up of the last three parts of Question 6 (a), consistently

gave part (v) a sign for its factor loading that was different from the sign attached to the loadings for parts (vi) and (vii). This outcome forecast that there would be difficulty trying to group the three parts in the one scale. It was found better to group only the last two parts, each of which was scored in favour of those who rejected the objects cabbage and pear as options for the meaning of the numerical variable 'c'.

13. The C2 Scale, the Two 'c' Values Scale, simply grouped responses to two subparts of Question 6 (a), namely '$c = 3$' and '$c = 7.4$'. A score of "1" was allocated for each correct acceptance of these options for 'c' given that '$c + d = 10$'. The *alpha* value of .58 was the lowest in the list given in Table 5·5 but it was still considered supportive of the scale formation, as were the correlations of .46.

14. The EQL, Equality Scale, consisted of nine questions which tested whether or not students understood the meaning of "equals" (or '='). Dichotomous variables were created by allocating a score of "1" if a student correctly answered any question in the group. The questions used in forming the EQL Equality Scale were not clustered by the factor analyses reported above because they had strong affiliations with other questions. Some tabulated correlations were a little low but all were at least positive, and the *alpha* value was a satisfactory .82. This scale variable was useful, in Chapters 6 and 8, for investigations related to Kieran's work (1981a) on the difficulties students have with the concept of equality.

Advantages of Correct Response Scales. The formation of scales which concentrated responses to groups of questions produced variables which measured students' level of ability in operating with certain modes of thinking. These scale measures were more representative of students' cognitive processes than were measures taken by single items. The claim that they were more representative is based on the fact that test items were accepted as members of the same scale only if the students' pattern of response to those items was sufficiently consistent, as assessed by the statistics presented. By concentrating a set of measures taken from groups of items into one scale score for the one construct, a gain in reliability is attained:

> To the extent that fewer parameters are estimated from the same data, each is generally based upon a larger subset of the data and, so, will have greater statistical reliability.
> (Shepard, 1972, p. 2)

In the algebra research projects referred to in earlier chapters, data were generally treated in terms of individual test or interview items. Much of the 1990 data, on the other hand, was organized into scales. The scale scores provided ordinal measures of different aspects of understanding the basic concepts of early algebra, ready for the investigations reported in the chapters which follow.

Three of these scales incorporated scores on items which were previously treated in separate studies. Other scales combined scores on two or more items within particular studies. Thus, the use of scaling was one way of attaining the second research objective, as set out on page 17 above, of incorporating interrelationships between measures formerly considered only in separate studies or treated in isolation within the one study. Details of the sources for the scale items incorporated in Correct Responses Scales are summarized in Table 5·11 at the end of this chapter.

The number of variables encompassed by the 14 scales for correct responses was 53. This left responses to only 9 test questions as isolated measures. Adding these 9 to the 14 scale measures, the original 62 test variables were now reduced to 23. Hence, the objective expressed at the start of the chapter, to reduce the number of variables, was achieved. The list of test questions not included in the scales reported in Table 5·5 is: 2(ii), 6 (a) parts (ii), (iii) and (v), 6 (b), 8 parts (a), (c) and (d), and 14 (i). Several of these questions were used in other scales which included some incorrect responses so that an assessment could be kept of error patterns and near-correct judgements. Such scales are discussed below.

Scales for Progress Indicators

Responses classified as Progress Indicators were either correct or in error categories which could be considered as on-the-way-to-being-correct. Evidence for the fact that scores on these scales indicated some signs of progress is presented in Chapter 6. In terms of the ordinal scale scores, to merit a score of "1" on a Progress Indicator Scale a student score on a particular item either equalled the maximum score for that item's ordinal scale or a score near the maximum. The allocation of ordinal scale scores to each test item was detailed in Chapter 4.

To establish these Progress Indicator Scales, the four criteria listed earlier were applied and the computer program Reliability was again used. However, it was found that attempts to use factor and principle component analyses were not helpful because, as Appendix 5F indicates, error categories and near-correct categories for responses from some questions were allocated to different scales as they reflected different ways of thinking about a common problem. Thus responses to common items were grouped into more than one scale and this fact distorted the computer findings from factor or principle component analyses. In contrast, for scales based on correct responses, each correct test response was a contributor to simply one scale, but for the Equality Scale.

Two examples clarify the process of forming Progress Indicator Scales.

1. For the NRPS Scale, No Reversal for Professors-and-Students Scale, which measured the ability to avoid the reversal error in Question 7, a recoded "1" score was allocated for ordinal scores of "5" (for correct) and "3" (for a choice which avoided the reversal error but regarded the letters as standing for people). All other responses were scaled at "0".

2. Question 7 responses were scaled differently when placing them in the NBR Scale, the Numbers View Scale, to measure the tendency to choose options which included the words "number" or "how many" to describe the meanings for symbols. Responses which were scored at "5" (for correct) or "4" (for a number choice which indicated a reversal error) merited inclusion as a Progress Indicator with a score of "1", while all other scores were rescaled to "0". There were two other items included in the NBR Scale, namely, Question 6 (a) part (v) and Question 14 (i). In the first of these, students who correctly chose "the number of apples in a box" as a possible option for '*c*' were scored at "1" and incorrect responses were allocated "0". In the second item, both the correct choice ("How many students took buses") and the incorrect number choice ("The number of buses which take the children") were scored at "1", and other responses were given a score of "0".

The only Progress Indicator Scales which did not include some responses at the highest end of the ordinal scale were the three scales concerned with the frequency with which students used the technique of giving numerical examples or replacement values, instead of the expected more general answer. These are the last three scales in Table 5·7, namely, the 1REP, 2REP and 12REP Scales.

Assessment statistics for Progress Indicator Scales are assembled in Table 5·6.

Advantages of Progress Indicator Scales. The 11 scales based on Progress Indicators enriched the possibilities for investigating certain aspects of the significance of the data obtained from testing. The previously-described dichotomous variables gave measures only of whether students were correct or not in their test responses. Now, scales were developed which included not only the correct responses but also those which were not correct yet reflected an understanding of some important issue in early algebra, such as the fact that the symbols stood for numbers and not people. These scales also had an advantage over the scales based on the data in ordinal form. The latter incorporated measures of the spread of response types, from correct to partially incorrect to seriously incorrect. The new scales identified which aspect of that spread was under consideration. As with the Scales for Correct Responses, the progress Scales provided measures which could be considered as more reliable than those based on responses to single items. They grew out of identified interrelationships between items derived from previous studies or created specifically for this study. Details of sources of scale items for these scales are in Table 5·13.

Table 5·6

Summary of Analyses of Scales for Progress Indicators

Questions	Scale	No. of Items	No. of Cases	Cognitive Focus	*Alpha*	Item-item Correlations		Corrected Item-Scale Correlations	
						Mean	Range	Mean	Range
2i,6a,c, 10a,b,c, 11a,b,c,d 12a,b,c, 13a,b,d	GNV	17	325	Generalized Number or Variable	.90	.33	.10 - .99	.55	.28 - .72
6a,10b,c 12b,c, 13b,d	GN	7	411	Generalized Number & not Variable	.70	.26	.06 - .70	.42	.15 - .53
6aii,iii, iv,6c, 15ivb	FZN	5	411	Include Fractions, Zero, and/or Negatives	.55	.19	.11 - .33	.32	.22 - .38
10b,c 12b,c 13b,d	INT	6	429	Positive integers only	.79	.38	.16 - .75	.54	.49 - .60
5,9, 14ii,iii	ALC	6	440	Accept Lack of Closure	.87	.55	.36 - .97	.68	.57 - .79
4b,c	NFL	2	448	Number Notion re Flowers	.60	.44	.44 - .44	.44	.44 - .44
6a v,7, 14a	NBR	4	461	Number Notion	.55	.23	.10 - .79	.35	.16 - .56
7	NRPS	2	507	No Reversal Prof-Student	.92	.85	.85 - .85	.85	.85 - .85
10b,c, 12b,c, 13a,b,d	12REP	7	429	One or more Replacement Values Given	.55	.16	.01 - .53	.29	.07 - .42
10b,c, 12b,c, 13b,d	1REP	6	429	One only Replacement Value Given	.55	.21	.03 - .56	.32	.19 - .40
13a,b,d	2REP	3	164#	Two or more Replacement Values Given	.66	.39	.27 - .57	.47	.33 - .57

Note. Dichotomous Variables. Missing values rejected. Years 7 to 12, with Test 3 responses for Year 7. For # 2REP Scale, only Year 7 Test 2 responses were used.

Individual Scales. Brief comments are needed to explain the structure and importance of each of the Progress Indicator Scales.

1. The GNV Scale, the Generalized Number or Variable Scale, measured the extent to which students regarded symbols as standing for generalized numbers (in the

sense given by Collis, 1975a) and/or variables. The scale sampled responses across 17 items. Recognition was given for correct algebraic answers which signified the use of the variable notion and also for the less developed acceptance of the possibility that algebraic symbols could simply have more than one value. Students qualified for a score of "1" on this scale by having Question 2 (i) correct; choosing at least two values for '*c*' from the numbers given as the first four options in Question 6 (a); allowing at least four values for '*c*' in Question 6 (c); indicating that the variables in Questions 10, 12, and 13 could have two or more values; or allowing the lengths of the sketched lines, given as '*a*' and '*b*' cm in Question 13, to vary. Correct algebraic answers to Questions 10, 11, 12, and 13 were accepted into the tally for this scale as these answers insisted that the relevant symbols could take more than two values. Two subparts of Question 11 showed a correlation of .99, indicating that students responded in closely similar fashion to each. To succeed, they had to overcome the geometry of the sketches so that the algebra in the given situation could dominate and allow the relative lengths of the given lines to change. It was decided, however, to leave both subparts in the scale. The corrected item-scale correlations were quite supportive of this scale and the *alpha* value was .90

2. The GN Scale, the Generalized Number Scale, excluded responses which were algebraic and implied an understanding of symbols as true numerical variables, and included those responses which showed that the student saw that the symbols could take more than one value. The latter responses were in the form of two replacement values or of almost correct algebraic, general answers which did not make explicit the possibility of non-integral values. The only correct answers accepted into this scale record were the choices of more than one numerical value for '*c*' in Question 6 (a). The *alpha* value gave good statistical support for grouping these responses into a scale.

3. The third scale, the FZN Scale, the Fractions-Zero-Negatives Scale, kept a record of students' willingness to accept the mathematical variability (cf. Dienes, 1963) which allowed algebraic symbols to take values that were possibly fractional, zero or negative, in appropriate circumstances. To merit a score of "1" for any item in this scale, candidates were required to choose numerical options given in Question 6 (a) parts (ii) to (iv); to include at least zero in responses to Question 6 (c); to return correct algebraic answers to part (b) of Questions 10, 12, or 13; or to reason correctly that '*c*' equalled zero in Question 15 (iii). The assessment statistics were a little lower than for the GN Scale but were considered to be satisfactory.

4. The INT Scale, the Integers Scale, recorded the number of times students indicated that they were thinking in terms of integers only in parts of Questions 10, 12, and 13, as when they used one or two integer replacement values for their responses. Alternatively, they wrote answers which were algebraically correct but for the fact that

they had overlooked the need for allowing non-integer possibilities for the variables under consideration, as when '$t \geq 5$' was given in Question 10 (b) instead of '$t > 4$'. These responses were allocated scores of "2" or "3" on the ordinal scales defined in Tables 4·21, 4·26 and 4·27. An *alpha* value of .79 was recorded for this scale.

5. The ALC Scale, the Accept lack of Closure Scale, was similar to the SYM Symbols Scale, detailed in Table 5·5, and was based on responses to the same set of questions. It was intermediate in expectation between the ordinal version of the SYM Scale and the dichotomous version. The latter recorded "1" each time a student gave the correct algebraic response to any of the component questions, whereas the new scale, ALC, allocated a score of "1" whenever a student gave some answer which was algebraic, whether correct or not. In the ordinal version of the SYM Scale, graded scores were allocated depending on whether the student was correct, or gave some incorrect algebraic answer, or gave a non-algebraic answer. The values of the various statistics for the new scale are, as could be expected, intermediate between the values given in Table 5·5 for the two versions of the SYM Scale. There were two exceptions, namely, the maximum correlations both for item-item and item-scale were higher in the case of the ALC Scale.

6. Responses for the NFL Scale, the Numbers of Flowers Scale, merited a "1" score if the students wrote that the symbols had some numerical meaning rather than representing objects. If, for instance, they wrote that 'g' stood for "the number of bunches", which was incorrect, or that it stood for "the number of flowers in a bunch", which was correct, they scored "1". Other responses were scored as "0". The statistical support was sufficiently strong to accept the NFL Scale.

7, 8. The NBR and NRPS Scales were explained above as examples of Progress Indicator Scales.

9. The 12REP Scale, the Replacement Value(s) Scale brought together seven measures of student tendency to respond by giving one or two examples of replacement values for variables rather than solving the given problem in a more general way. Students qualifying for a score of "1" had responded in what Harper (1979) had called the Numerical Replacements category. This 12REP Scale provided a measure of the intensity of students' use of this form of response across subparts of questions based on the Harper (1979) research. The component questions for these scales were similar to the questions used by Harper (1979), being parts from test items 10, 12, and 13.

10. The Scale 1REP, the One Replacement Scale, used a subset of responses from the 12REP Scale by restricting scores to cases in which students had given only one numerical example to a general Harper-style problem.

11. The 2REP Scale, the Two Replacements Scale, was analysed using responses from Year 7 students on their second test. These responses rated well on the statistical analyses given in Table 5·8, even though very few students used the technique of presenting two or more examples as an answer to a general problem. Those who did were generally consistent in applying the method. The only questions found to produce reasonable statistical support were the listed three parts of Harper's Equation Task, as used in Question 13 of the test. Insufficient support was obtained by trialling questions of this type with the larger cohort of 517 students while using Test 3 results for Year 7.

Scales for Hindrance Indicators

Another set of scales was derived from various error patterns and was described as Scales for Hindrance Indicators. The term Hindrance was applied on the presumption that achievement in algebra would be hindered by the maintenance of the errors identified. Whether or not this presumption was correct was investigated and the outcomes are reported in Chapter 6.

The responses coded as scores of "1" on these Hindrance Indicator Scales were in the categories which were low on the ordinal scales detailed in Chapter 4. For example, students were allocated a score of "1" in the SC2 Scale, the second Seek Closure Scale, if, by using arbitrary numbers to replace algebraic symbols, they had scored "1" on the ordinal scales (which went up to "3" or "4") for Questions 5, 9, or 14.

As was the case with the Progress Indicator Scales, the use of factor and principle component analyses was not appropriate as variables based on different categories of responses to common questions were allocated to different scales. This gave rise to variables that were not independent and that would distort the outcomes of factor and principle component analyses. The four criteria designated earlier were applied in the establishment of these scales.

As some of the errors occurred infrequently after a few of weeks of classroom algebra, data were taken from the second test administered to beginning Year 7 students to examine the possibility of forming error scales. Other errors were more persistent and the third test data for Year 7 were used. As Tables 5·7 and 5·8 indicate, there were several suitable groupings of errors to provide scales which were used for further investigations. Sources for items used are given in Table 5·15.

Table 5·7
Summary of Analyses of Scales for Hindrance Indicators

Questions	Scale	No.of Items	No.of Cases	Cognitive Focus	*Alpha*	Item-item Correlations		Corrected Item-Scale Correlations	
						Mean	Range	Mean	Range
2i,10b,c, 11b,c,d, 12b,c 13b,c,d	NV	12	422	No Variability Allowed	.78	.24	.02 - .89	.43	.24 - .57
12, 13	SC1	8	422#	Seek Closure not work with symbols	.81	.40	.08 - .77	.56	.32 - ..67
3i,ii,v, vi, 9	CON	7	485	Conjoin for Addition	.72	.32	.10 - .87	.46	.30 - .59
5, 9, 14ii,iii	SC2	6	428#	Seek Closure not write symbols	.88	.54	.33 - .89	.68	.57 - .78
13b,c, d	CF	3	429#	Coefficient determines larger variable	.86	.68	.60 - .78	.74	.66 - .80
4b,c	JFL	2	448	Objects Notion for Flowers	.54	.38	.38 - .38	.38	.38 - .38
7, 8a	OBJ	3	476	Objects Notion for Prof-Student & '*a*'=apples	.63	.37	.14 - .79	.46	.16 - .62
7	RPS	2	507	Reversal re Prof-Student	.91	.83	.83 - .83	.83	.83 - .83
2i,6c,10 11,12,13	PRE	19	339	Prestructural errors	.90	.31	.05 - .93	.53	.31 - ..73

Note. Dichotomous Variables. Missing values rejected. Years 7 to 12, with responses for Year 7 students from Test 3 except # from Test 2.

Advantages of Hindrance Scales. These scales were valuable as they measured the tendency of students to make similar errors consistently in their thinking about the test problems. Scale scores registered the extent of repeated errors and reflected students' modes of thinking more thoroughly than did scores on individual items. In the SC2 Scale, for example, those qualifying for a score of "1" had responded to questions which required answers in terms of algebraic symbols by replacing the symbols with arbitrarily-chosen numbers. In this way, they avoided the use of algebra and reverted to the more familiar arithmetic so that they could achieve closure in the form of a numerical answer.

Scales reported in Table 5·7 are discussed first, and this is followed by those in

Table 5·8. Although some of the assessment statistics were low for some scales, support was considered sufficient for accepting these groupings of items into scales.

Individual scales.

1. The NV Scale, the Non-Variable Scale, consisted of twelve items, one being from Question 2 (i) and the rest being subparts of Questions 10 to 13, which were Harper-style questions. The scale scores measured students' tendency to respond with a rigid point of view in excluding variation for the symbols in the problem. They had answered by giving only one numerical example or writing "never" or "always" when these were inappropriate. An understanding of the symbols as true numerical variables would have made it clear that many other options were actually available.

2. The SC1 Scale, the first Seek Closure Scale, measured the strength of the trend to avoid working with symbols whose values were unknown. This was shown by students who, for their responses to parts of Items 12 and 13, wrote such things as "What does '*n*' equal?" or "I don't know the value of '*x*' or '*y*' ". As very few Year 7 students responded in these ways after more than two weeks of algebra, the data in the analysis for this scale was taken from their second test and placed with data from the other classes.

3. The CON Scale, the Conjoin Scale, measured the frequency of the error of conjoining for addition, an error noted by Küchemann (1980) and Booth (1983). Four parts of Question 3 identified this error whenever students used addition instead of multiplication to evaluate expressions such as '2*y*'. The list of such responses was given in Tables 4·5 and 4·6 in Chapter 4, under the category score of "2". The three parts of Question 9 registered the same erroneous thinking in a different form, namely, writing algebraic expressions in conjoined form (such as '9*n*') when they should have been written to show addition (such as '*n* + 9'), as listed in the category score of "2" in Table 4·18 of Chapter 4.

4. The SC2 scale, the second Seek Closure Scale, reported the frequency with which students avoided writing symbols in answers to Questions 5, 9, and 14 by, for instance, replacing them with arbitrary numbers. This expression of seeking closure was different from that registered by the SC1 Scale and was given strong support by the assessment statistics. Instances of seeking closure in the questions covered were not usual among Year 7 students after two weeks of algebra. Hence, the analysis for this scale was based on responses as described for the SC1 Scale. Statistical assessment procedures did not support uniting the SC1 and SC2 Scales.

5. The CF Scale, the Coefficients Scale, clustered the responses within Item 13 which indicated that certain students reasoned incorrectly that the coefficients associated with variables determined the relative sizes of the values of the variables. In

the given case, 'x' was regarded as larger than 'y' because the equation '$2x + y = 9$' showed '$2x$' and simply 'y'. Some suggested changing the coefficients to accommodate the possibility that 'y' could be greater than 'x'. The statistical support indicated that the same students tended to make this type of mistake in each of the subparts of the scale. The same cohort of students were used in testing this scale as for the SC2 Scale, and for similar reasons.

6. The JFL Scale, the Objects Notion for Flowers Scale, recorded the consistency of erroneous responses which revealed that some students gave symbols a meaning more in terms of objects than numbers in both parts of Item 4.

7. The OBJ Scale, the Objects View Scale, provided another measure of student tendency to interpret symbols as representing objects or people, rather than numbers. A score of "1" was allocated whenever students chose at least one of the options "Professors", "Professor", "Student" or "Students", as given in Item 7 parts (a) and (b). Those who wrote that 'a' stood for "apples" given "$3a$ represents 3 apples" in question 8 (a) were also given a score of "1" on the Objects Scale because it appeared that they were simply using 'a' as an abbreviation for the word "apples", thus treating the symbol more as an object than as a number.

A scale to combine the two objects view scales (viz., JFL and OBJ) into one scale was not supported by the statistics produced.

8. The RPS Scale, the Reversal for Professors-and-Students Scale, recorded as a single score the tendency to make the reversal error in one or both parts of Item 7.

9. The PRE Scale, the Prestructural Scale (cf. Collis, 1988, p.71; Collis & Watson, 1989, p. 181), consisted of twelve items and kept a record of many errors, including those scored under the SC1 Seek Closure Scale. Errors which showed that students had little understanding of the set problems were included in the tally for this scale, such as translating '$t + 4$' into '$4t$' in Question 10.

More Scales for Hindrance Indicators.

There were some other error patterns which generally disappeared soon after starting classroom algebra but which were scrutinized. These error patterns were noted by other researchers (such as Booth, 1983) and included the use of place value and alphabetic codes and the ignoring of symbols completely. Table 5·8 lists the assessment statistics for the scales which concentrated on these errors and which were analysed using data from only Year 7 beginners at the time of their second test, administered after nearly two weeks of algebra.

Table 5·8
Summary of Analyses of More Scales for Hindrance Indicators

Questions	Scale	No.of Items	No.of Cases	Cognitive Focus	*Alpha*	Item-item Correlations		Item-Scale Correlations	
						Mean	Range	Mean	Range
3, 8b	PV	7	197	Place Value Code	.93	.66	.34 - .97	.78	.50 - ..93
3i,ii,iii 9i,ii	IG	5	186	Ignore letter	.60	.35	.04 - .86	.41	.35 - .46
2ib,9	AL1	4	154	Alphabetic Code Qq.2i,9	.62	.50	.39 - .66	.52	.29 - .66
14ii,iii	AL2	2	199	Alphabetic Code Q.14	.92	.86	.86 - .86	.86	.86 - .86

Note. Dichotomous Variables. Missing values rejected. Test 2 responses for Year 7.

10. The PV Scale, the Place Value Code Scale, aggregated error responses across seven instances. Typical of the type of error analysed here was the response to Question 8 (b) which gave the value of 'a' as '6' given that '$3a = 36$'. Students who made such an error regarded the component parts of an algebraic term as denoting different decimal place values, so that '$3a$' translated as '30' plus 'a'. The explanations of the variety of "Place Value" errors, which were listed in Tables 4·5 and 4·6 in Chapter 4, for the six parts of Question 3 are as follows.

(i) '23' (= 20 + 3) as the value of '$2y$' when 'y' equals '3',
(ii) '28' (= 23 + 5) as the value of '$2y + 5$' when 'y' equals '3',
(iii) '28' (= 23 + 5) as the value of '$2(y + 5)$' when 'y' equals '3',
(iv) '26' (= 23 + 3) as the value of '$2y + y$' when 'y' equals '3',
(v) '30' (= 33 - 3) as the value of '$3y - y$' when 'y' equals '3', and
(vi) '28' (= 20 + 8) or '106' (= 2 x 53) or '253' (= 200 + 50 + 3) or '630' (= 200 x 3 + 10 x 3) as the value of '$2(5y)$' when 'y' equals '3'.

11. The IG Scale, the Ignore Letter Scale, recorded the number of times students simply ignored the presence of an algebraic symbol in five selected instances, namely, the first three parts of Item 3 and the first two parts of Item 9. The responses which merited a score of "1" for this scale score were as follows:

3 (i) writing '2' as the value of '$2y$' when 'y' equalled '3',

3 (ii) and (iii) - writing '7' as the value of '$2y + 5$' or '$2(y + 5)$' when 'y' equalled '3',

9 (i) writing '9' as the result for "Add 4 onto $n + 5$",

9 (ii) writing '7' as the result for "Add 4 onto $3n$".

12. The AL1 Scale, the first Alphabetic Code scale, registered the number of times students mistakenly used an alphabetic code in the second subpart of Question 2 (i) and in the three parts of Question 9. In Question 2, they revealed their erroneous thinking by directly describing the sizes of the letters '*a*' and '*d*' in terms of their position in the alphabet. In Question 9, they betrayed their use of '14' for '*n*' by the numerical answers they gave, such as:

(i) '19' (= 14 + 5) or '23' (= 14 + 5 + 4) or '74' (= 14 x 5 + 4),

(ii) '21' (= 3 +14 + 4), or '46' (= 14 x 3 + 4) or '318' (= 314 + 4), and

(iii) '76' (= 19 x 4).

13. The AL2 Scale, the second Alphabetic Code scale, grouped only two cases of this form of error, both reflected in answers to Question 14. If students used '6' as the value of '*f*' and '7' as the value of '*g*', they gave answers such as '46' (= 3 x 6 + 4 x 7) in part (ii) and '39' (= 3 x 6 + 3 x 7) in part (iii).

Combining the AL1 and AL2 Scales was not given sufficient support by the statistical assessment procedure.

Another objects scale.

Responses in Question 6 (a) which selected the options "an object like a cabbage" and "an object like a pear" scaled together strongly and were grouped to form the JCP Scale, the Cabbage and Pear Scale. As will be explained in Chapter 7, this scale was treated separately from the other scales for measuring an "Objects View" for symbols. For the present, it will not be classified as either a Progress Indicator or a Hindrance Indicator but will be left as Unclassified. The statistical support for the formation of the JCP is reported in Table 5·9

Table 5·9

Summary of Analysis of Cabbage & Pear Scale

Questions	Scale	No.of Items	No.of Cases	Cognitive Focus	*Alpha*	Item-item Correlations		Corrected Item-Scale Correlations	
						Mean	Range	Mean	Range
6avi,vii	JCP	2	482	Chose *c* = cabbage and/or pear	.95	.91	.91 - .91	.91	.91 - .91

Note. Dichotomous Variables. Missing values rejected. Years 7 to 12, with Test 3 responses for Year 7 students who did test more than once.

The JCP Scale, the Cabbage and Pear Scale, was a further measure of the students' tendency to regard symbols as representing objects rather than numbers. A score of "1" was recorded if a student chose the option "an object like a cabbage" or "an object like a pear" as a possible meaning for 'c' in the equation '$c + d = 10$', in test Item 6 (a) parts (vi) and (vii). The very high statistical support for this scale, as given in Table 5·9, indicated that students generally chose both these options or neither of them. Adding the scores for these two questions was strongly supported as a way of reducing the number of variables to be used in other investigations.

Summary of Scale Names, Foci and Sources

Tables 5·10 to 5·17 present a summary of the titles of the scales, the aspects of student thought they were designed to measure, and the sources of the component items. They are arranged, but for the JCP Scale, according to their classification as scales for Correct Responses, as Progress Indicators, or as Hindrance Indicators.

Table 5·10

Summary of Titles and Foci for Correct Response Scales

SCALE	SCALE TITLE	COGNITIVE FOCUS
VBL	Variables	Variable
SYM	Symbols	Symbols in answers
SUBS	Substitute & Solve	Substitution & Solving
EQN	Equation	Covariance in Equation
PL	Parallel Lines	Variable in Geometry
AR	Arithmetic	Arithmetic Operations in algebraic setting
FL	Flowers	Algebra for No.of Flowers
BXBA	$b=x, b=a$	Variable: $b = x; b = a$
PS	Professors-and-Students	Covariance Professors-and-Students problem
CZ	c = zero	Variable: c = zero
AD	a, d ,	Independent Variables a,d
NJCP	No Cabbage or Pear	Reject c = Objects
C2	Two 'c' Values	Generalized Number: $c = 3$; 7.4
EQL	Equality	Equality of expressions or variables

Table 5·11

Summary of Sources for Correct Response Scales

SCALE	New	Collis	Harper	Küchemann	Rosnick	Booth	MacGregor
VBL		15iii	6c,10	6c,12,15iii			
SYM	14ii,iii			9		5	
SUBS	3, 8b						
EQN			13				
PL			11				
AR		1					
FL	4						
BXBA		15i,ii					
PS					7		
CZ		15iv					
AD	2i						
NJCP	6a vi,vii						
C2	6a i,iv						
EQL	8b	15ib,iib, ivb	10d,11d, 13c	12d,15iiib			

Table 5·12

Summary of Titles and Foci for Progress Indicator Scales

SCALE	SCALE TITLE	COGNITIVE FOCUS
GNV	Generalized Number or Variable	Generalized Number or Variable
GN	Generalized Number	Generalized Number & not Variable
FZN	Fractions-Zero-Negatives	Include Fractions, Zero, and/or Negatives
INT	Integers	Positive integers only
ALC	Accept lack of Closure	Accept Lack of Closure
NFL	Numbers of Flowers	Number Notion re Flowers Item
NBR	Numbers View	Number Notion
NRPS	No Reversal for Professors-and-Students	No Reversal Professors- and-Students
12REP	Replacement Value(s)	One or more Replacement Values Given
1REP	One Replacement	One only Replacement Value Given
2REP	Two Replacements	Two or more Replacement Values Given

Table 5·13

Summary of Sources for Progress Indicator Scales

SCALE	New	Collis	Harper	Küchemann	Rosnick	Booth	MacGregor
GNV	2i,6a		6c,11, 10a,b,c, 13a,b,d	6c,12			
GN	6a		10b,c, 13b,d	12b,c			
FZN	6a ii,iii,iv	15ivb	6c				
INT			10b,c, 13b,d	12b,c			
ALC	14ii,iii			9		5	
NFL	4b,c						
NBR	6a v,14a				7		
NRPS					7		
12REP			10b,c, 13a,b,d	12b,c	7		
1REP			10b,c, 13b,d	12b,c			
2REP			13a,b,d				

Table 5·14

Summary of Titles and Foci for Hindrance Indicator Scales

SCALE	SCALE TITLE	COGNITIVE FOCUS
NV	Non-Variable	No Variability Allowed
SC1	First Seek Closure	Seek Closure: not work with symbols
CON	Conjoin	Conjoin for Addition
SC2	Second Seek Closure	Seek Closure: not write symbols
CF	Coefficients	Coefficient determines larger variable
JFL	Objects Notion for Flowers	Objects Notion for Flowers Items
OBJ	Objects View	Objects Notion for Professor-Students & 'a'=apples
RPS	Reversal for Professors-and-Students	Reversal re Professors-Students
PRE	Professors-and-Students	Prestructural errors
PV	Place Value Code	Place Value Code
IG	Ignore Letter	Ignore letter
AL1	First Alphabetic Code	Alphabetic Code Qq.2i,9
AL2	Second Alphabetic Code	Alphabetic Code Q.14

Table 5·15
Summary of Sources for Hindrance Indicator Scales

SCALE	New	Collis	Harper	Küchemann	Rosnick	Booth	MacGregor
NV	2i		10b,c,11b,c,d,13b,c,d	12b,c			
SC1			13	12			
CON	3i,ii,v,vi			9			
SC2	14ii,iii			9		5	
CF			13b,c,d				
JFL	4b,c						
OBJ	8a				7		
RPS					7		
PRE	2i		6c,10,11,13	12			
PV	3,8b						
IG	3i,ii,iii			9i,ii			
AL1	2ib			9			
AL2	14ii,iii						

Table 5·16
Summary of Title and Focus for JCP Scale

SCALE	SCALE TITLE	COGNITIVE FOCUS
JCP	Cabbage & Pear	Chose c = cabbage and/or pear

Note. This scale is discussed in Chapter 7.

Table 5·17
Summary of Sources for JCP Scale

SCALE	New
JCP	3a vi,vii

Review and Forecast

Scale scores are central to the analyses to be described and discussed throughout the following four chapters. Chapter 5 has provided the relevant details about the formation and content of the scales. Attention has been focused on variable clusters which were clearly designated by the batteries of factor analyses reported in Chapter 4. Four requirements for suitable scales were proclaimed and applied to assess groupings of variables. Fourteen Scales for Correct Responses, 13 of which were based on factor-derived clusters, were accepted, reducing the number of variables from 62 to 23. A further 25 scales were also supported by assessment procedures. Eleven of these were for Progress Indicators and 13 for Hindrance Indicators, and another scale, the JCP Scale, was left unclassified as it produced what seemed to be paradoxical outcomes which will be discussed in Chapter 7.

The scales have simplified data management and have established measures for distinguishable aspects of the thinking used by the students when they completed the test instrument. These aspects were listed under the heading "Cognitive Focus" in Tables 5·5 to 5·10. They ranged from thinking of algebraic symbols as abstract numerical variables to thinking which showed that students had little idea of the basic elements of early algebra. Tables 5·10 to 5·17 presented summaries of the scale names, the cognitive processes they were intended to measure, and the sources of the items which contributed to the scale scores. The scales were a vehicle for integrating the study with the work of researchers over the past two decades.

The scene is set for reporting investigations of the ways that scores on these measures interrelated so as to extend our understanding of the processes of learning early algebra, and of the difficulties involved. As Shepard (1972, p. 3) explained, the purpose of multidimensional scaling is

> to enable the investigator to gain a better understanding of the total underlying pattern of inter-relations [*sic*] in his data and, hence, to decide what further observations, experiments, or modifications of theory will most advance the science as a whole.

The first field of investigation is concerned with relationships between the cognitive entities measured by the established scales. This is followed by a study of whether or not the meanings students gave to algebraic symbols were related to the ways they responded to the variety of tasks encompassed by the test items. Research findings on these issues are reported in Chapter 6.

CHAPTER 6

FIRST STUDY: A STUDY OF RELATIONSHIPS WITHIN LEVELS OF UNDERSTANDING FOR SYMBOLS AND BETWEEN LEVELS AND ACHIEVEMENT ON ALGEBRAIC TASKS

Overview

The investigations described within this chapter deal with a range of views expressed by students regarding levels of understanding of the symbols used in the test instrument. Relationships are explored within these levels and between levels of understanding of symbols and achievement on algebraic tasks. The chapter has three sections.

Firstly, a synthesis is presented of hierarchies of understanding for algebraic symbols as expressed earlier by three other researchers. This led to the identification of five hierarchical levels.

Second, one of the major information outcomes from the research project is presented, namely, the proportion of students at each of these levels. An explanation is given for the way in which the scales defined in Chapter 5 are used to determine a series of frequency distribution tables. The input data are based on responses from students across Years 7 to 12, taking responses obtained from Year 7 students when they did the test for the third time. These same responses were used in Chapter 4 for the global overview of the data. Correlations within levels of understanding are included in this section.

Third, several commonly-expressed expectations about the relationship between the level of understanding of algebraic symbols and success with algebraic tasks are put to the test by examining relevant correlations. Being a correlational study, it did not address the question of cause and effect but it supplied support for Proposition 1, namely:

Students with better levels of understanding of the meaning of algebraic symbols are more likely to have higher degrees of success with algebraic tasks.

Section 1: Hierarchies of Understandings for Algebraic Symbols

Synthesis of Views on Hierarchies of Understanding.

Collis (1975a), Harper (1979), and Küchemann (1980) described the range of students' views of algebraic symbols in hierarchical terms. They based these hierarchical levels on the way they perceived the degrees of cognitive difficulty involved in each level and on consideration of the error patterns in responses to algebraic tasks they had devised. They did not pursue the question of whether or not students followed a sequential learning path whereby they gradually worked their way up the hierarchy. This issue will be examined in this and later chapters.

A synthesis of the hierarchical levels described by these researchers will be used to investigate the relationship between the levels of understanding for symbols and the degree of success attained on the algebraic tasks encompassed by the test instrument used in the study. The synthesis, summarized in Table 6·1, uses five levels to retain the differences between the ways the researchers described possible viewpoints. A discussion of these different viewpoints follows the table.

Table 6·1

<u>A Synthesis of Hierarchies of Understanding From Three Researchers</u>

Level	Clarification	Collis (1975a)	Harper (1979)	Küchemann (1980)
5	Class of numbers: no need for use of trial numbers	Variable	Species	Variable
4	Class of numbers: readily checked by using trial numbers	Generalized Number	Discovered Content: each numeral seen as a possible replacement	Generalized Number
3	Use of trial numbers a necessary process	Several Replacement Values	Discovered Content: letters as boxes into which numeral can be posted	Generalized Number and Specific Unknown
2	Use of only one trial number seen as sufficient	One Replacement Value	Fictitious Measures: Letter as object with unique content	Letter Evaluated
1	Letter as meaningless object, or stands for object, or value from place in alphabet, etc.	Prestructural	Fictitious Measures: Unfounded Ordering	Letter as Object or Ignored

<u>Collis hierarchies</u>. As explained in Chapter 2, Collis (1975a, pp. 5 - 6; 43 - 48) associated his categories of views for letters in algebra with stages of development as regards the acceptance of lack of closure. At the highest level, students could work with operations in algebra without the need to relate the elements or operations to any

physical reality or familiar numbers and were working in or near the formal mode of functioning. Those at the next level, that of Generalized Number, were using the concrete-symbolic mode and were able to refrain from actual closure of operations while being confident that some unique and familiar result was always attainable. A little lower on the continuum, students sought closure by simply using multiple trials with numbers to replace the letters and, lower still, were content with simply one trial number for making a decision about a general problem. For those not ready to interpret algebra with any appropriate sense, the general category of "Prestructural" (Collis & Campbell, 1987, p. 5) has been applied to various classes of incorrect responses at the lowest level.

Harper hierarchies. Harper (1979) assessed the level of students' understanding of algebraic symbols by means of the four different tasks described in Chapter 2. He called the highest level that of "species" (pp. 240 - 242), a level distinguished by students' ability to fuse the roles of letters representing numeral identifiers capable of naming each appropriate numeral simultaneously and letters representing some unspecified member of the group of appropriate numbers. This level was described by Collis as the variable level, shown as Level 5 in Table 6·1. Harper's exposition of the meanings which he included in the term "discovered content" (pp. 165, 170, 181) showed that he incorporated Collis' "generalized number" concept at the higher end of this category in the sense that symbols were seen in a fluid way, allowing them to represent any numeral. At the other end of this same category, Harper included the understanding of symbols as boxes into which numerals could be placed, each replacement being considered separately. This view corresponds with Collis' classification in terms of the use of multiple replacements. The discovered content category was thus considered to span Levels 3 and 4, as shown in Table 6·1. Within his category of "fictitious measure" (pp. 119, 139, 146, 151), Harper incorporated the view of letters as having some unique and fixed value or content, a level corresponding approximately with Collis' level for making use of one replacement value. Harper also included as fictitious measures those views of symbols which allowed them to represent an object such as an orange or which gave them values determined by some extraneous influence such as the position of the letter in the alphabet. This lower end of the fictitious measures category encompassed misunderstandings at a level that Collis called Prestructural. Therefore, in Table 6·1, the Harper fictitious measures category spans Levels 1 and 2.

Küchemann hierarchies. As explained in Chapter 2, Küchemann (1980) defined the highest level of understanding for algebraic symbols, the "variable" level, in the following terms: "letters are used as variables when a second-order relation is established between them" (p. 59). The only item which he considered as a measure of the notion of a variable in the test he used was that which asked students which was bigger, '$2n$' or '$n + 2$'. He argued for a second-order relationship in this item if one considered the differences in the rates of change of the two expressions (p. 63). (The item was adapted to a Harper-style question in the test used for this study.) He regarded the use of the "generalized number" concept as one characterized by the use of several replacement values while replacement cases were isolated from each other (p. 60). He acknowledged (p. 57) that the test items he used did not distinguish between those who thought of letters as taking several values in turn and those who regarded them as representing a set of values simultaneously. To take this range of meanings into account, Küchemann's generalized number category was spread across Levels 3 and 4, as shown in Table 6·1. The "specific unknown" view of letters enabled students to operate on symbols by being aware that each symbol could have a unique value, even though it was unknown (p. 49). At this level, a statement such as '$5b + 6r = 90$' would, according to Küchemann (p. 60), be considered as one which holds for a particular pair of numbers, whereas, at the generalized number level it would be thought of as being true for several pairs of numbers, the pairs being considered in isolation. Küchemann expressed some uncertainty about the difficulty level of specific unknown usage and generalized number usage. He suggested that we "regard these two ways of interpreting letters as complementary" (p. 58). He tentatively suggested that children may attain the ability to handle specific unknowns "before they conceive of generalized numbers" (p. 58). Because of the uncertainty, Küchemann's specific unknown category was placed at the lower end of his generalized number category, at Level 3, as shown in Table 6·1.

At the lowest levels, Küchemann (1980) found that some students ignored the letters completely or reduced their meaning from something quite abstract to something more concrete. They proceeded by regarding the letters as equivalent to the names of sides of a geometric figure or as objects such as bananas, without distinguishing between whether they stood for the number of objects or the objects themselves (p. 53). Within the "letter not used" category, Küchemann included the possibilities that "children ignore the letter, or at best acknowledge its existence without giving it a meaning" (p. 49). His "letter evaluated" category included cases in which students "assigned a numerical value" (p. 49) to the letter without justification, and in which "a numerical value is asked for but it is not necessary to manipulate the letter first" (p. 49). Viewing an algebraic symbol "as shorthand for an object or as the object in its own right" (p. 49) were both included in the "letter as object" category. The relative levels of cognitive difficulty associated with these three categories are dependent upon the context set by the algebraic task at hand, as was discussed in Chapter 2. Table 2·4 gave some examples. When considering hierarchies of difficulty levels, Küchemann focused on test items rather than categories describing meanings given for symbols.

Küchemann consistently placed the "letter evaluated" category at one end of his list of categories, alongside the "letter not used" category, followed by the "letter as object" category (1978, p. 23; 1980, p. 49; 1981, p. 104; 1984, p. 115). The order has been changed in Table 6·1 in an effort to describe hierarchies of difficulty in terms of different understandings for symbols. Some of the multiple interpretations in the Küchemann classifications have been eliminated. For instance, the "letter evaluated" category is aligned with Collis' "one replacement value" category at Level 2, and both ignoring a symbol and regarding it as an object are classed as Prestructural at Level 1. Ignoring an algebraic symbol was considered a likely indicator of less understanding than replacing it with an arbitrary number. To replace it, the student at least has to acknowledge its existence and to allow it to take a numerical meaning.

Investigating the hierarchical levels. The research test instrument was designed to measure the levels of understanding that students held for the meanings of letters as algebraic symbols. All levels described above were identified in the responses obtained. This chapter reports on the frequencies of the different viewpoints amongst the student sample and the degree to which they held those views. Analyses are undertaken to assess whether or not the research data supported the hierarchical ordering of the levels given in Table 6·1. Correlations are examined to assess whether or not success in the algebraic tasks tested was related to the level of understanding of symbols. A study of changes in levels and whether or not the proposed hierarchy corresponds to a sequence of learning is reported in Chapter 9. For that study, measures of the hierarchical levels are taken from a small selection of Harper-style test items, namely, parts (b) and (c) of Questions 10 and 12,.

The following section describes how some of the scales established in Chapter 5 provided tallies of the tendencies of each student to respond according to one or other of the levels.

Section 2: Scale Measures for Meanings for Symbols

Before investigating relationships between levels of interpreting algebraic symbols and success with various algebraic tasks involved in the test items, credentials had to be established for various scales as measures of those different levels of understanding. Taking the levels in turn, the appropriate scales are discussed and the distribution patterns for student responses on these scales are summarized and interpreted.

Frequency Statistics for Scales.

The tables below summarize the frequency statistics of scales for each level, using a dichotomous form of the data. They provide the following information:

1. the scale maximum. This is equal to the number of items in the particular scale, as scores on each item were either "1" or "0";

2. the number of valid cases. The only way a student could obtain a valid score on any scale was to have a valid response to each item in the scale by avoiding the missing values category, as explained in Chapter 4;

3. "A%". The percentages listed under the heading "A%" recorded the percentage who had scored more than zero on the scale. In other words, it gave the percentage who responded to at least one item in the scale at the level of meaning being measured. Percentages were of the total number of students who had valid scores for that scale;

4. "B%". The figures under the heading "B%" recorded the percentage who responded to at least 50% of the scale items at the level being measured;

5. "C%". The figures under the heading "C%" recorded the percentage who responded to at least 80% of the scale items at the level being measured;

6. the median, the score below which half the valid cases fall; and

7. the means and standard deviations. These have been averaged across the items in each scale for ease of comparison, taking into account the fact that the number of items was not the same for each scale.

The A%, B%, and C% were included to indicate clearly the general patterns of the distributions without an overload of information.

Level 5: The Variable Level

The test items which extended at least some students to operate with the concept of a variable were the Harper or Harper-style items in Questions 6 (c), 10 to 13, and the Collis or Collis-style items in Question 15. Responses to these items could, therefore, be at any of the levels listed in Table 6·1. The following five scales were composed of linear combinations of responses to these items, as detailed in Chapter 5:

VBL, EQN, PL, BXBA, and CZ.

The full range of possibilities was preserved when the responses were scored as Ordinal Variables. However, when the same responses were scored as Dichotomous variables by allocating a score of "1" for correct and "0" for any other response, they were simply measures of the highest level for those items, giving the degree to which students operated with the variable concept. To answer some of these items (e.g., Question 10 b), students had to state conditions for inequalities to hold. When scoring such items at variable level, the requirement was imposed that each response allowed the variable in the problem to take the correct range of values, a range that included the possibility of fractional, zero and/or negative values. In other items (e.g., parts of Question 11 in the PL Parallel Lines Scale), students were given credit if they responded in a way which indicated that they allowed the symbols in the problem to take any values and these values did not need to be known. Appendix 5F describes the responses accepted for the scales.

Frequency data. Table 6·2 presents relevant statistics for the scales when they were formed from the dichotomous form of the data.

Over three-quarters of the students (79.4%) revealed some evidence of being able to operate at the variable level in their responses to at least one (A%) of the 11 items in the VBL Scale, but it was only 29.7% who maintained this level for at least 9 (C%) of the questions. Nearly half scored more than "5" (B%) on the scale. The figures under the heading "C%" for the other scales show that maintaining an 80% success rate at the variable level was difficult, as only around one-quarter managed this for the PL Scale, about one-fifth for the BXBA Scale, a little more than one-eighth for Scale EQN, and less than one-twentieth for the CZ Scale.

Table 6·2

Frequency Statistics for Scales for Concept of Variable (Level 5)

Scale	Scale max.	No.of valid cases	A%	B%	C%	Median Score	Mean per item	S.D. per item
VBL	11	407	79.4	48.9	29.7	5	0.466	0.369
PL	4	491	43.2	38.1	26.9	0	0.362	0.447
BXBA	4	468	58.5	50.4	20.3	2	0.393	0.391
EQN	4	441	43.5	31.5	13.6	0	0.264	0.360
CZ	2	476	13.2	13.2	4.6	0	0.045	0.123

Note. Years 7 to 12, using Test 3 for Year 7 students. Scales from dichotomous data. Scale maximum = no.of items in scale. A% = percentage of valid cases who scored at least 1 on the scale. B% = percentage of valid cases who responded at Level 5 to at least 50% of items. C% = percentage of valid cases who responded at Level 5 to at least 80% of items. Entries ordered by values of C%.

Correlations within Level 5. The inter-correlations between the five variables scales were found to be positive and almost all highly significant (at $p \leq .001$), as shown in Table 6·3. The highly significant positive correlations indicated that students tended to respond similarly to most of these scales. Some, for instance, consistently viewed the symbols as true numerical variables and scored well on each scale, whereas others made little progress on any scales as they did not understand the symbols as variables. The numerical values of the correlations were less than .6, indicating inconsistencies in the response levels of other students.

Table 6-3
Correlations Between Scores on Variables Scales (Level 5)

r	VBL	PL	EQN	BXBA	CZ
VBL	-	.404 ***	.556 ***	.574 ***	.213 ***
PL	397	-	.375 ***	.344 ***	N.S.
EQN	384	432	-	.358 ***	.131 **
BXBA	396	449	412	-	.283 ***
CZ	396	456	417	466	-

Note. Years 7 to 12, using Test 3 for Year 7. Scales formed from dichotomous data. Figures below diagonal give number of cases.
*** $p \leq .001$, ** $.001 < p \leq .010$. N.S. not significant.

Students' performance on the CZ Scale did not correlate significantly with their performance on questions in the PL Scale, indicating that those who succeeded on the hardest question on the test, namely Question 15 (iv), did not necessarily do well on the parallel lines task. Only 22 students had both subparts of Question 15 (iv), the CZ question, correct. Cross-tabulation showed that nine of those successful on the CZ Scale actually scored zero on the PL Scale, while another nine top-scored on both these scales. However, scores on the CZ Scale correlated positively and significantly with the other scale measures at Level 5, but not as strongly as did scores on the PL Scale or scores on the rest of the variables scales. Of the scale measures listed in Table 6-3, the CZ Scale appeared to be the weakest means of measuring facility with the concept of variable, according to the correlations listed.

Explanatory note. From Table 6-3, it can be seen that scores on the five variables scales were not uncorrelated, despite the fact that the formation of groupings of items for scales was guided by the clusters identified by principal components analyses and factor analyses, as explained in Chapter 5. Even though each scale measured a somewhat distinctive aspect of student cognition, a common bond was implied - in this case, the variable view of symbols. One of the conditions governing the principal components method of analysis is that each of the components "be uncorrelated with scores on any of the preceding principal components" (R. J. Harris, 1975, p. 24). A similar condition applies to factors in most factor analysis methods

(R. J. Harris, 1975, pp. 25 - 27). However, there are two reasons why this condition did not carry over to the scales derived from such an analysis:

1. The scale components were a subset of the original set of variables, all of which were components of every principal component (and factor); and

2. Unity (+ 1) was used as the coefficient for each variable when forming the linear combination of variables for any scale, whereas the coefficients used in the linear combination for any principal component (or factor) were the calculated loadings, which varied between -1 and +1.

Level 4: Generalized Number

The scales deemed to measure the degree to which students regarded symbols as generalized numbers, in the sense explained by Collis (1975a), were described in the treatment of "Scales for Progress Indicators" in Chapter 5. As the name implies, the GNV Generalized Number or Variable Scale recorded the number of times students gave responses which could be categorized at either the Generalized Number (Level 4) or the Variable (Level 5) stage of the development of an understanding of the meaning of algebraic symbols. The GN Generalized Number Scale, recorded the number of times students indicated that they accepted more than one numerical value for the algebraic symbols in Questions 10, 12, and 13, but their tally did not include the correct algebraic answers to these questions because such answers were considered to be indicators of attaining the notion of a true numerical variable. The FZN Scale scores were gained by indicating an acceptance of fractional, zero, or negative possibilities for algebraic symbols. The INT Integer Scale kept a tally of how many responses made use of only integer values for variables in parts of Questions 10, 12, and 13. Two other scales, namely the NBR and NFL Scales, were included within Level 4 for the reason that they measured the degree to which students regarded algebraic symbols as standing for numbers of objects or people rather than the objects or people themselves. Both scales specifically focused on the number concept for interpreting algebraic symbols which were being used in that form of algebra described as generalized arithmetic.

Frequency data. The relevant statistics for the Level 4 Scales are reported in Table 6-4. Nearly all students (98.5%) responded at least once along the lines accepted in the GNV Scale and over one-quarter of students (26.2%) gave acceptable responses to at least 80% of the items included in this scale. Some students appeared to be aware that the symbols could take on more than one value but this was just a part of their movement towards the more general understanding of the symbols as representing true numerical variables which could simultaneously represent a range of numbers. It seemed difficult to distinguish whether or not they viewed the possibility of more than one value for a symbol in a multistructural sense, taking each value as a separate consideration (cf. Biggs & Collis, 1982, p. 69), or in the more abstract sense of the symbol representing multiple values as a variable. Küchemann (1980, p. 57) also expressed difficulty in discriminating between these two levels of understanding by means of the test items he used. This difficulty is explored further in Chapter 8, where stages of development over time are examined.

Table 6·4

Frequency Statistics for Scales for Concept of Generalized Number (Level 4)

Scale	Scale max.	No.of valid cases	A%	B%	C%	Median Score	Mean per item	S.D. per item
GNV	17	325	98.5	64.0	26.2	11	0.583	0.287
NFL	2	448	46.2	46.2	20.1	0	0.332	0.396
FZN	8	411	92.2	37.5	14.1	2	0.429	0.241
NBR	4	461	68.8	38.8	6.1	1	0.334	0.306
INT	6	429	39.9	21.9	4.7	0	0.160	0.253
GN	7	411	90.8	6.1	3.4	1	0.243	0.200

Note. Years 7 to 12, using Test 3 for Year 7 students. Scales from dichotomous data. Scale maximum = no.of items in scale. A% = percentage of valid cases who scored at least 1 on the scale. B% = percentage of valid cases who responded at Level 4 to at least 50% of items. C% = percentage of valid cases who responded at Level 4 to at least 80% of items. Entries ordered by values of C%.

The data on the GN and FZN Scales showed that the vast majority of students, at least once in their test responses, allowed an algebraic symbol to take more than one value (90.8%) and to include the possibility of values being other than positive integers (92.2%). Far fewer maintained these lines of thinking across 80% or more of the items included in the scales. Only 14.1% (about 1-in-7) showed such consistency in accepting values that were not positive integers, according to measures taken by the FZN Scale, and 3.4% (about 1-in-30) showed similar consistency with a multi-valued view as scored by the GN Scale. Integers-only answers were given by 21.9% of the students for at least three of the items included in the INT Scale, and nearly 40% gave at least one integer answer.

Figures for the NBR Scale under the headings A% and C% showed that about 70% at least once chose expressions which included the words "number of" or "how many" as descriptors for algebraic symbols but only 6.1% consistently did this for all four opportunities provided by items included in the scale. The corresponding figures for the NFL Scale reported that less than half the students (46.2%) wrote about the symbols in Question 4 in terms of number, and about one-fifth (20.1%) of the students did this for both subparts of the question.

Correlations within Level 4. As Table 6·5 records, correlations were not applicable between many of the Level 4 scales because of items in common to pairs of scales. Correlations were positive and highly significant (although not large numerically) between scores on the NBR Scale and those on the GNV, GN and FZN Scales, showing that students who were more inclined to describe the meanings of symbols in terms of numbers were also more likely to work with the notion of the symbols as generalized numbers, including numbers other than positive integers, in the problems given in other test items.

Table 6·5

Correlations Between Scores on Generalized Number Scales (Level 4)

r	GNV	GN	FZN	NBR	NFL	INT
GNV	-	N.A.	N.A.	.287 ***	N.S.	N.A.
GN	411	-	N.A.	.287 ***	N.S.	N.A.
FZN	369	314	-	.347 ***	N.S.	N.S.
NBR	317	317	362	-	N.S.	N.S.
NFL	304	316	338	411	-	.085 *
INT	325	411	369	398	386	-

Note. Years 7 to 12, using Test 3 for Year 7. Scales formed from dichotomous data. N.A. not applicable (common items). Figures below diagonal give no. of cases. *** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$. N.S. not significant.

Scores on the NFL Scale did not correlate significantly with the other Level 4 scale measures, except for the very low correlation with the INT Scale scores (.085, the square of which indicated that only 0.7% of the variance was shared). This NFL Scale measurement of the tendency to think of symbols as representing numbers was taken from Question 4 which used a real-life setting in terms of bunches of flowers, and may have been subject to some influence such as the semantic processing involved, a process which MacGregor (1991, p. 123) found to proceed independently of syntactic processing. Caution was indicated regarding the use of Question 4 outcomes.

The fact that scores on the INT Integers Scale did not correlate significantly with those on the FZN Scale was probably because only one item, Question 6 (c), in the FZN Scale allowed the possibility of submitting an integer answer. All six items of the INT Scale, on the other hand, asked for a range of appropriate values for variables, and students were able to respond in ways which restricted answers to integers or included the possibility of fractional values and even zero and/or negatives as well. There was not any significant relationship between students' responses to the two sets of items. Students who responded to the INT Scale items with ranges of values that intimated that they were thinking only of integers were not automatically those who rejected non-integer options in most of the items of the FZN Scale, nor were they automatically those who accepted most of the non-integer values. It seems that students were comfortable with different number fields for different questions.

Level 3: Several Replacement Values and Specific Unknown

Scale 2REP tallied the number of times students partially answered general questions by giving two or more examples of solutions in Question 13. In Chapter 5, this scale was defined and classified as one of the "Progress Indicators", bearing in mind that success in giving correct numerical examples showed an understanding of the problem, even though the tendency to list replacement values for algebraic variables was a distraction from the wider generalization inherent in the algebra presented. This classification is in concordance with the views expressed by Collis (1975a), Harper (1979) and Küchemann (1980) on the use of replacement values.

Frequency data. Table 6.6 assembles percentage frequencies for scores on all three scales designed for recording the use of replacement values. The first is for Level 3 understanding of symbols, and the other two are for Level 2.

Table 6·6

Frequency Statistics for Scales Measuring Understanding at Level 3 or 2

Scale	Scale max.	No.of valid cases	A%	B%	C%	Median Score	Mean per item	S.D. per item
2REP	3	441	4.3	0.7	0.2	0	0.017	0.090
1REP	6	429	20.7	2.3	0.2	0	0.053	0.122
12REP	7	429	27.3	1.2	0.2	0	0.064	0.125

Note. Years 7 to 12, using Test 3 for Year 7 students. Scales from dichotomous data. Scale maximum = no.of items in scale. A% = percentage of valid cases who scored at least 1 on the scale. B% = percentage of valid cases who responded at Level 2 or 3 to at least 50% of items. C% = percentage of valid cases who responded at Level 2 or 3 to at least 80% of items. Entries ordered by values of C%.

The figures for the 2REP Scale show that only 4.3% of students at least once responded by giving two or more replacement values instead of a more general answer to the problems in Question 13. Only one student (0.2%) persisted with this approach for all three questions incorporated in the scale. The median was zero since 95.7% of students completely avoided this type of replacement approach. The infrequent use of the strategy of giving two or more examples as an answer to a general problem restricted the 2REP Scale to three items as statistical support was insufficient for the inclusion of other test items.

Level 2: Single Replacement Value

A record was kept by the 1REP Scale of the number of times students gave only one example of possible values for algebraic variables as partial answers to more general problems. Scores for the 12REP Scale, on the other hand, were allocated for responses in the form of one or more replacement example per problem.

Frequency data. As Table 6·6 shows, one-fifth of the students (20.7%) gave a single numerical example in at least one of the six questions which contributed to the 1REP Scale: these were subparts of Questions 10, 12, and 13, as detailed in Chapter 5. Only one student (0.2%) used the same approach in all six questions. Almost 80% of students avoided this form of response, accounting for a median score of zero on this scale. Table 6·6 also includes data on the 12REP Scale which recorded frequencies with which students used one or more replacement examples in the six questions that were used in the 1REP Scale or in part (a) of Question 13. More than a quarter of the students (27.3%) scored at least "1" on the 12REP Scale.

Correlations between scores on the scales at Levels 2 or 3 were not reported as the three scales at these levels had items in common.

Level 1: Prestructural Errors

Errors of different types were coded and several scales were formed to tally the frequency with which students made particular errors, as described in Chapter 5 under the heading of "Hindrance Indicators". The error scales relevant to the meanings that students gave to algebraic symbols were:

OBJ - taking symbols to stand for people (professors or students) rather than numbers of people in Question 7, or reading '*a*' as an abbreviation for the word "apples" given that "3*a* represents 3 apples" in Question 8 (a);

JFL - taking symbols to stand for objects (such as flowers or bunches) rather than numbers of objects in Question 4;

SC2 - seeking closure by replacing symbols by arbitrarily-chosen numbers rather than giving general answers in terms of symbols;

PRE - responding at a prestructural level by showing little understanding of the given problem, such as simply repeating the question in different words;

CF - allowing the coefficients in the equation '$2x + y = 9$' to determine which variable is the greater;

NV - denying variation for the values of symbols;

PV - using place-value code for coefficients, giving, for example, the meaning '23' to '$2y$' when 'y' equals '3';

CON - considering addition to be represented by conjoining, thus regarding '$2y$' as '5' when 'y' equals '3';

AL2 - using an alphabetic code to give 'f' a value of '6' and 'g' a value of '7' in Question 14;

SC1 - stating the need to know values for symbols as a reason for not attempting parts of Questions 12 and 13;

AL1 - ordering 'a' and 'd' in terms of their relative positions in the alphabet in Question 2 (i), and replacing 'n' by '14' in Question 9; and

IG - ignoring symbols altogether.

Frequency data. Table 6·7 summarizes the relevant statistics for these Level 1 scales. The most common misunderstanding about symbols was expressed in the OBJ Scale by the 28.0% who were registered in the "C%" column of Table 6·7 because they had expressed an objects view in all three items which contributed to the scale scores. In the professors-and-students problem, they chose options which indicated that they regarded both of the letters 'S' and 'P' as standing for people rather than for numbers of people and, in Question 8 (a), they wrote that 'a' meant "apples", given that "3a represented 3 apples" (cf. Booth, 1983, p. 268). Another half (51.8% = 79.8% - 28.0%) of the students made such errors once or twice, giving a total of 79.8% who had at least one such error. There were 61.7% of students who chose people instead of numbers of people in both parts of Question 7, and these were largely responsible for the fairly high mean item score of 0.581. Only one-fifth (20.2% = 100% - 79.8%) of students completely avoided these errors. The semantic processing involved in these questions could well have been a significant influence (cf. MacGregor, 1991).

In Question 4, symbols were used in relation to a real-life situation. Close to one-third (30.1%) of the students interpreted symbols in terms of objects rather than numbers at least once and nearly one-tenth (9.4%) made the same error twice, as is shown by the figures for the JFL Scale.

Table 6-7

Frequency Statistics for Scale for Basic Misunderstandings (Level 1)

Scale	Scale max.	No.of valid cases	A%	B%	C%	Median Score	Mean per item	S.D. per item
OBJ	3	476	79.8	65.3	28.0	2	0.581	0.362
JFL	2	448	30.1	30.1	9.4	0	0.198	0.327
SC2	6	440	27.3	17.7	7.0	0	0.148	0.281
PRE	19	335	86.3	29.9	6.3	5	0.338	0.273
CF	3	441	7.9	3.6	1.6	0	0.044	0.167
NV	12	422	66.2	7.3	1.1	1	0.152	0.180
PV	7	489	11.9	1.4	0.8	0	0.027	0.107
CON	7	485	27.8	2.3	0.6	0	0.084	0.164
AL2	2	492	1.0	1.0	0.6	0	0.008	0.084
SC1	8	436	6.9	1.4	0.2	0	0.018	0.084
AL1	4	392	16.6	0	0	0	0.045	0.110
IG	5	485	9.5	0.2	0	0	0.026	0.088

Note. Years 7 to 12, using Test 3 for Year 7 students. Scales from dichotomous data. Scale maximum = no.of items in scale. A% = percentage of valid cases who scored at least 1 on the scale. B% = percentage of valid cases who responded at Level 1 to at least 50% of items. C% = percentage of valid cases who responded at Level 1 to at least 80% of items. Entries ordered by values of C%.

Of the two scales which recorded the tendency to seek closure, the SC2 Scale detected that the more common form of this error was the replacement of symbols by numbers, chosen either quite arbitrarily or worked out from an alphabetic code (cf. Booth, 1984a, p. 102; Collis, 1975a, p. 43). By doing this, over a quarter of the students (27.3%) avoided writing symbols in at least one of their answers and reverted to the more familiar arithmetic to obtain numerical answers to general questions. Seven percent used the same erroneous approach in five or six of the six questions used in the SC2 Scale. Only 6.9% wrote that they could not proceed without knowing the values of letters in at least one of the eight questions in the SC1 Scale, and one student did the same for all eight responses.

The PRE Prestructural Scale registered that basic misunderstandings were expressed by almost all students (86.3% scored at least "1" on this scale). Nearly 30% showed misconceptions on at least half of the 19 items used for the scale and 6.3% recorded this tendency on at least 15 items. Attention is given in the following chapters to the strength of possible effects of such misunderstandings on student progress in algebra.

According to the scores on the CF Coefficients Scale, 7.9% of students at least once gave coefficients in Question 13 too powerful a role. About half of these (3.6%) made the error at least twice, but only 1.6% of students repeated the error in three parts of the question.

Nearly 70% of the students on at least one occasion, as measured by the NV Scale, signified a rigid outlook which implied that the symbols did not encompass the notion of variation. However, only 1.1% of the students consistently denied freedom for the symbols in seven or eight of the items registered by the scale.

Over one-tenth of students (11.9%) used a place value code at least once in the seven items included in the PV Scale. However, very few persisted with this idea, only 0.8% of the subjects making this form of error on 6 or 7 occasions.

The conjoining error, as measured by the CON Scale, was common. More than a quarter of the candidates (27.8%) confused conjoining with addition on at least one occasion. This misunderstanding was maintained across at least four items by just 2.3% and very few (0.6%) repeated the error six or more times.

Nearly 10% ignored symbols at least once in the five questions forming the IG Scale, but very few made this mistake repeatedly, as can be deduced from the zero values for the median and the "B%" in Table 6·7, and from the low mean per item.

A fairly large proportion (16.6%) of students used an alphabetic code at least once on the four questions included in the AL1 Scale. Thirteen percent made the error in Question 2 to decide on an ordering for '*a*' and '*d*' but, as the zero entry in the "B%" column shows, the error was not often repeated. Very few used an alphabetic approach to Question 14, as is shown by the data for the AL2 Scale, even amongst those in Year 7 who had experienced only about three weeks of algebra by the time they attempted Test 3.

Correlations within Level 1. The inter-correlations between the 12 basic errors scales are recorded in Table 6·8. The SC2 Scale scores correlated significantly and positively with seven other error scale scores, indicating that those who sought closure by substituting arbitrarily-chosen numbers for letters were likely to make other errors as well, namely, ignore the letter (IG Scale, $r = .483$, the strongest correlation in Table 6·8), show little understanding of the 17 items included in the PRE Prestructural Scale ($r = .408$), restrict the variable nature of the symbols (NV Scale, $r = .255$), regard coefficients as place-value holders (PV Scale, $r = .212$), declare that they could not proceed without being given values for letters (SC1 Scale, $r = .128.$), and give meanings for letters in terms of objects or people rather than numbers of objects or people (JFL Scale, $r = .120$, and OBJ Scale, $r = .114$).

Table 6-8
Correlations Between Scores for Basic Errors (Level 1)

r	OBJ	JFL	SC2	NV	AL2	SC1	IG	AL1	PRE	PV	CON	CF
OBJ	-	.118 **	.112 *	.212 ***	.107 *	N.S.	N.S.	N.S.	.217 ***	.143 ***	.124 **	N.S.
JFL	426	-	.120 **	N.S.	N.S.	N.S.	N.S.	N.S.	.101 *	N.S.	N.S.	.083 *
SC2	419	397	-	.255 ***	N.A.	.128 **	.483 ***	N.A.	.408 ***	.212 ***	N.S.	N.S.
NV	401	335	335	-	N.S.	N.A.	N.S.	.137 **	N.A.	.273 ***	.286 ***	N.A.
AL2	460	431	N.A.	349	-	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
SC1	413	392	400	N.A.	422	-	.172 ***	N.S.	N.A.	N.S.	N.S.	N.A.
IG	453	428	438	349	466	421	-	N.A.	.305 ***	N.A.	N.A.	.145 ***
AL1	371	362	N.A.	351	380	359	N.A.	-	N.A.	N.S.	N.A.	N.S.
PRE	321	N.A.	317	N.A.	329	N.A.	327	N.A.	-	.192 ***	.392 ***	N.A.
PV	460	431	429	411	470	424	N.A.	379	330	-	N.A.	N.S.
CON	453	420	438	410	466	421	N.A.	N.A.	327	N.A.	-	.096 *
CF	418	395	402	N.A.	427	N.A.	426	364	N.A.	427	426	-

Note. Years 7 to 12, using Test 3 for Year 7. Scales formed from dichotomous data. N.A. not applicable (common items). Figures below diagonal give no. of cases. *** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$. N.S. not significant.

The positive and significant correlations between the PRE scores and those from six other scales indicated that those who had little understanding of the 17 items included in the PRE Prestructural Scale were likely to replace symbols by arbitrarily-chosen numbers (SC2 Scale, $r = .408$), to simply ignore them (IG Scale, $r = .305$), or to express meanings of symbols in terms of objects rather than numbers of objects (OBJ Scale, $r = .217$, and JFL Scale, $r = .101$). They were also likely to misinterpret algebraic coefficients either by interpreting conjoining as addition instead of multiplication (CON Scale, $r = .392$) or by considering them as place-value indicators (PV Scale, $r = .192$).

Errors resulting from the use of an alphabetic code, as measured by the AL1 and AL2 Scales, did not correlate significantly with each other. However, there were two small but significant correlations with other error measures. Scores on Scale AL1 showed some tendency to follow the pattern of scores on the NV Non-Variable Scale ($r = .137$), and the trend in scores on the AL2 Scale was somewhat similar to that for the OBJ Objects View Scale ($r = .107$).

Scores on the two scales which measured tendency to think in terms of letters as objects or people rather than as numbers, the OBJ and JFL Scales, correlated positively and significantly with scores on the SC2 Scale, showing that those who tended to think of algebraic symbols as objects or people were the more likely to avoid working with symbols by replacing them with arbitrary numbers. Scores on the OBJ Scale tended to follow the same pattern as scores on the NV Scale ($r = .210$). Scores on the OBJ and JFL Scales correlated positively and significantly ($r = .114$) with each other, as those who thought in terms of objects on Questions 7 and 8 (a) were also similarly inclined in Question 4.

Those who tended to ignore letters, as measured by the IG Scale, also tended to be unable to work with letters when their values were unknown, in items included in the SC1 Scale, as is shown by the positive and significant correlation ($r = .172$) between scores on these two scales.

Interview Extracts

Two interview extracts are included here to instil a flavour of the reality of the cognitive challenges experienced by students, to illustrate some of their levels of thinking, and to relate their comments to their scores on some of the scales just discussed. Examples of extreme views were chosen, the first student using the highest level of understanding for symbols, and the second the lowest.

Interview Extract 1. (Student 'R' at Level 5, Year 7, School C, after Test 4, 26 Nov., 1990, then aged 12 years 7 months; the experimenter, 'E', was the interviewer)

This student had maximum scores on four of the five variables scales for Level 5 (Table 6·1) understanding of symbols, giving her 22 out of a possible 24. The extract shows that she worked with the variables notion in Question 15 (i) without needing to resort to trial numbers.

E All right. Number 15, the next page. You pick one of those parts and explain how you went about that one.
R The first one, $a + b + c = a + x + c$. I said it was true only when $b = x$.
E You were pretty fast on that. How did you know that?
R Um [Pause]. If they all equalled together ... since there was an 'a' in both, a 'c' in both, but one had a 'b' and one had an 'x', I worked out that if they were both the same.
E That's pretty easy isn't it? That's a good way to do it, I'd say.

Interview Extract 2. (Student 'R' at mainly Level 1, Year 7, School A, after Test 3, 27 April, 1990, aged 13 years 1 month)

This student was working mainly at Level 1, showing an inability to work with algebraic symbols as representing numbers. On the five scales for variable level (Level 5), the only scores registered were for partial answers to subparts of Question 15, and her score on the Level 4 scale, GNV, was a low 3 out of a possible 17. The strong tendency to seek closure was recorded as "5" on the SC2 Scale and "1" on the SC1 Scale.

E This one here the football match. How would you work out the total number of scores?
R You can't.
E Why can't you?
R Because you don't know what *'p'* is and you don't know what *'r'* is.
E But if you did know how would you work it out?
R Then you could add them up.
E Right. Can you add *'p'* and *'r'* ?..How would you write down "add *p* and *r*"?
R You just put '$p + r$', and you put the equals, and [Pause] like you've got, if 'p' = 5 and 'r' = 10 you can put '$p + r = 15$'.
E Now, see [Pause]. You're at a difficult stage. You've understood that the letters stand for any numbers.
R Yes.
E But you are not yet ready to say that '$p + r$' is the answer?
R No, because I ...

Section 3: Views of Algebraic Symbols and Success on Algebraic Tasks

Proposition 1

Several studies and papers have expressed the belief that progress in algebra is linked with a sound understanding of the meaning of the symbols used. The following assembly of quotations makes this point clear:

> The probability that a pupil will deal successfully with algebraic items therefore appears to depend ... upon the interpretation of the letter the pupil has available to him.
> (Harper, 1979, p. 272)

> We believe that it is essential that students be able to view variables as standing for *number*.
> (Rosnick & Clement, 1980, p.23)

> 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.
> (Bell, Costello & Küchemann, 1985, p. 144)

> Indeed, one could say that until a student does appreciate the use of letters as variables, or at least as 'generalised number', then algebra can have little real meaning.
> (Booth, 1986, p. 3)

> While the difficulties that students experience with algebra reveal themselves in the use of symbols and the rules that govern their use, it is a lack of acceptance of the symbols as legitimate mathematical entities in the first place that is the fundamental problem.
> (Booker, 1987, p. 279)

The above statements were expressions of opinions based largely on general overviews of research data. They were not supported by documenting statistical analyses which were directed specifically at the expressed relationships.

The opinions could be condensed into the proposition that

Students with better levels of understanding of the meaning of algebraic symbols are more likely to have higher degrees of success with algebraic tasks.

The major focus of the first investigation in this section is the statistical examination of this proposition, which will be referred to as *Proposition 1.* The obverse of this proposition is implied and is studied by association within this first investigation. It reads as follows:

Students with poorer levels of understanding of the meaning of algebraic symbols are more likely to have less success with algebraic tasks.

The Methodology: Correlation.

The bivariate Pearson correlation coefficient was chosen as an appropriate statistic for examining statements such as Proposition 1. The scales which have been established were designed to measure specific aspects of cognition or of ability to succeed with particular algebraic tasks. Although some of them shared common portions of the variance in the responses and correlated with each other, there was a sense in which they were measuring specific aspects of the data. Hence, a bivariate approach was deemed more appropriate than a multivariate one, at least for the investigation of Proposition 1 which concerned the relationship between the two aspects of levels of cognition regarding symbols and degrees of success with algebraic tasks.

Furthermore, different categories of responses to many of the test items were used in building up a variety of scales for different purposes. In other words, some test items provided data for more than one scale. Associating such scales within a common multivariate analysis would have given rise to spurious statistics.

Correlation coefficients provided the information sought, namely, measures of the consistency, or lack of it, with which students scored on pairs of scales:

> The Pearson Product Moment Coefficient of Correlation, or simply the coefficient of correlation, is a measure of the strength of the linear relationship between two variables.
> (Dietrich II & Kearns, 1986, p. 485)

The square of the correlation coefficient is, moreover, a measure of "the proportion of variance that the two variables share in common ... [or] ... the proportion of variance explained" (Kenny, 1987, p. 113). When dealing with variances, as Kenny (1987, p. 114) points out, we are really dealing with the squares of scores. Taking this into account, correlation coefficients can provide information about the extent to which there is an overlap in the variances for scores on cognitive measures and scores on task success, as is relevant in the case of Proposition 1.

<u>Correlations Amongst Test Scores and Variables-, Numbers-, and Objects- Views</u>

As mentioned earlier, correlations were used to investigate relationships between levels of understanding for the meanings of algebraic symbols and success with algebraic tasks. Perhaps the most telling and crucial correlation table is Table 6·9. Here is strong statistical evidence in support of the stated proposition and of the advantages of acquiring the understanding that letter symbols in early algebra represent numbers rather than objects. The table provides evidence about the implications of five different groups of relationships. These are now taken in turn.

<u>1. Variables notion and overall test score</u>. The five scales listed down the left side of Table 6·9 are all measures of the highest level of understanding of algebraic symbols, namely, the level which enabled students to work with the concept of a numerical variable. The correlations down the right side are correlations between the scores on the variables scales and the corrected test totals. To eliminate spurious correlations resulting from common items in the pair of variables being examined, the "corrected test totals" were obtained by subtracting from the test totals the scores for items which were components of the scale being examined.

All the listed correlations with corrected test totals are positive and highly significant (with $p \leq .001$), indicating that the more the students understood the symbols as representing numerical variables, the higher were their scores on test items not included in the particular scale being used in the correlation. To at least some degree, here is a case of the better thinkers scoring well more generally. However, the significance and the complexity of these interrelationships become clearer as this thesis progresses. Each of the five variable scales was built from different test items, without any overlap, and together they covered 24 items. The overall test scores were composed of 57 responses to the variety of tasks which challenged the students in the total test. In Table 6·9, we see statistical evidence for the likelihood that the better the numerical variable concept was understood, the better the resultant test score would be and, conversely, the weaker any student's grasp of the variable notion, the more likely would that student score poorly on the tasks involved in the test. Strong support is, therefore, given to Proposition 1 by these statistics. Taking the square of the largest of the five correlations, we find that 57.2% of the variance is common to scores on the VBL Scale and the corresponding corrected test totals.

Table 6·9

Correlations Amongst Test Totals & Variables-, Numbers-, & Objects- Views

VARIABLES View	Scale Description	NUMBERS View	OBJECTS View	Corrected TEST TOTAL
		NBR Scale Qq.6a v,7, 14i	OBJ Scale Qq.7, 8a	
VBL Scale	Variables Qq.6c,10,12, 15iii	.281 (385) ***	- .239 (386) ***	.756 (407) ***
EQN Scale	Equations Task Q.13	.309 (409) ***	- .231 (418) ***	.590 (441) ***
PL Scale	Parallel Lines Task Q.11	.159 (444) ***	- .112 (457) **	.472 (491) ***
BXBA Scale	$b = x$; $b = a$ Q.15 i, ii	.340 (434) ***	- .245 (441) ***	.619 (468) ***
CZ Scale	c = zero Q.15 iv	.249 (441) ***	- .183 (448) ***	.228 (470) ***
Corrected TEST TOTAL		.353 (461) ***	- .244 (476) ***	-

Note. Years 7 to 12, using Test 3 for Year 7 students. Scales from dichotomous data. Figures in brackets give the number of cases.

Corrected TEST TOTAL = TEST TOTAL *minus* marks gained from items in scale.

*** $p \leq .001$, ** $.001 < p \leq .010$.

2. Numbers view and overall test score. The NBR Scale was composed of four responses taken from three different test questions, each of which gave students the opportunity to select a meaning for algebraic symbols in terms of a number of objects or people rather than simply the objects or people. The correlation (.353) at the bottom of the column headed "NBR" in Table 6·9 shows that the greater a student's

inclination to select a number view of algebraic symbols, the greater is the chance that the student would score strongly on the overall test. Support for this statement is found as the correlation is positive and highly significant (with $p \leq .001$) between scores on the NBR Scale and scores on the adjusted overall test scores. Again, Proposition 1 is clearly supported by these figures.

3. Objects view and overall test score. In contrast, as the correlation (- .244) at the bottom of the column headed "OBJ" is negative and highly significant (with $p \leq .001$), the objects view was seen as a likely hindrance to scoring well on the variety of algebraic tasks spread throughout the test. The scores on the OBJ Scale were accumulated from responses to three items from two different test questions and they measured students' tendency to think of letters in algebra as standing for objects or people, as distinct from numbers of objects or people.

4. Numbers view and variable view. The first five correlations listed under the heading "NBR" in Table 6·9 were all correlations between scores on the NBR Scale and scores on scales designed to measure the degree of understanding of symbols as numerical variables. These, although not large numerically, were all positive and highly significant (with $p \leq .001$), providing evidence for the belief that the more clearly students see algebraic symbols as representing numbers rather than objects or people, the more likely it is that they will develop the desirable understanding of symbols as numerical variables.

5. Objects view and variable view. Under the heading "OBJ" in Table 6·9 are correlations between scores on the OBJ Scale and scores on the five variables scales. All of these were negative and, even though their numerical values were smaller than .25, all were highly significant (with $p \leq .001$), apart from one at a significance level between .010 and .001. These outcomes showed that the more the students thought of algebraic symbols as standing for objects or people rather than for numbers of these, the less likely were they to regard the symbols as representing numbers that can vary.

Caution. Although the correlations in Table 6·9 were generally found to be highly significant statistically (at the 1% level), this did not imply that all students fitted neatly into categories which supported the conclusions spelt out. Some correlation coefficients were fairly low in size and the corollary was to expect many exceptions to the general conclusions. Appendix 6A uses cross-tabulations of scores on the BXBA Scale against scores on the NBR and OBJ Scales to make the reality clear, tempering any tendency to exaggerate the extent of the above general conclusions, but without discrediting them in any way. Close attention is given in Chapter 9 to the details of relationships between variables in furthering the search for hierarchies of learning and sequences of development.

Additional data on objects view and numbers view. Appendix 6B gives a summary of correlations similar to that given in Table 6·9 but using measures of the numbers view and objects view which were based on the NFL and JFL Scales for responses to Question 4, about numbers of flowers. The outcomes supported implications similar to those derived from Table 6·9, even though some significance levels for correlations were lower than in the corresponding cases in Table 6·9.

Correlations Between Views of Symbols and Corrected Test Totals

An important aspect of the process of applying statistics to investigate Proposition 1 was to examine correlations between the test totals, corrected to avoid overlap of test items, and scores on scales designed to measure different levels of understanding for algebraic symbols. The point in these examinations was that, if scores on a particular scale correlated positively and significantly with corrected test totals, this indicated that the scale was a measure of a "Progress Indicator", while scales giving negative and significant correlations were "Hindrance Indicators". These terms, progress and hindrance indicators, were used in Chapter 5 when describing the establishment of various scales, but their use was justified at that time only by considerations about whether or not the scales tallied correct or incorrect responses or a mixture of these. The correlation study in this section analyses whether or not these terms have statistical support to justify their use as descriptors of those scales. This study bears directly on the investigation of Proposition 1. For a scale which merited the title "Progress Indicator", thinking of symbols along the lines measured by that scale was likely to be accompanied by successful performance of the various algebraic tasks included in the test, whereas regarding symbols according to the view measured by bona fide "Hindrance Indicator" scales was likely to be associated with poor achievement on the tasks.

Table 6·9 included correlations between several scales and their corresponding corrected test totals. These have already been discussed. All scales which measured views of symbols within the category of "Variables View" had positive and significant correlations with corrected test totals, showing that the more clearly students understood algebraic symbols as representing numerical variables, the more likely were they to be successful on the various algebraic tasks in the test. Table 6·9 and the associated discussion also gave support for describing the NBR Numbers View Scale as a Progress Indicator, since scores on this scale correlated positively and highly significantly ($p \leq .001$) with corrected test totals. Additionally, there is support for describing the OBJ Objects View Scale as a Hindrance Indicator because the correlations between scores on the OBJ Scale and corrected test totals were negative and highly significant ($p \leq .001$).

Table 6·10 summarizes the outcomes from correlational analyses for all other scales used for measuring different ways of viewing algebraic symbols and, for the sake of completion and ease of comparison, the NBR and OBJ Scales and all Variables Scales are included as well. These outcomes bear directly on the central issue of the research project and verify the opinions quoted at the start of this section about expectations regarding students' views of algebraic symbols.

Study of the information summarized in Table 6·10, in which entries are ordered by the size of the correlation coefficients, gave rise to the following six observations:

1. Twenty-five of the 26 scales gave correlations with corrected test totals which were statistically significant. This fact in itself verifies that student achievement on the algebraic tasks covered by the test items was related to their level of understanding of the meaning of algebraic symbols.

2. The positive and significant correlations for corrected test totals were with scales which measured how well students had developed sound views about algebraic symbols, extending across Levels 2 to 5. While Level 5 scales measured the extent to which students held the richest and most valuable concept of symbols, namely, that they represented numerical variables, the measures at Levels 4, 3 and 2 indicated how far students were along the road to attaining that concept. The fact that scores on all these scales correlated positively and significantly with degree of success in algebra strongly supported Proposition 1.

3. The negative and significant correlations with corrected test totals all involved scores on scales which measured the frequency with which students made certain fundamental errors associated with Level 1 views of algebraic symbols. These were views at the prestructural level (Collis & Campbell, 1987, p. 5; Collis, 1988, p. 71) and the fictitious level (Harper, 1979). They included errors such as regarding algebraic symbols as standing for objects rather than for numbers of objects, or as foreign to the possibility of variability, or as something to avoid by changing them into numbers. Thus, support is given for the obverse of Proposition 1, namely: *Students with poorer levels of understanding of the meaning of algebraic symbols are more likely to have less success with algebraic tasks.*

Table 6-10

Correlations Between Scale Scores and Corrected Test Totals

Scale	Description	Level of Understanding (Table 6-1)	Correlation with Corrected Test Total	No. of cases	Significance	Conclusion: Type of Indicator
VBL	Variable	5	.756	407	***	Progress
GNV	Gen.No. *or* Variable	4 or 5	.727	322	***	Progress
BXBA	$b = x$ or a	5	.619	468	***	Progress
EQN	Equation	5	.590	441	***	Progress
FZN	Not positive integer	4	.587	411	***	Progress
PL	Par'l Lines	5	.472	491	***	Progress
GN	Gen.No *not* Variable	4	.355	411	***	Progress
NBR	No.View	4	.353	461	***	Progress
INT	Integers only	4	.282	429	***	Progress
CZ	c = zero	5	.228	470	***	Progress
1REP	1 replacement	2	.148	429	***	Progress
12REP	1 or 2 replcmts	2 or 3	.129	429	**	Progress
NFL	No.flowers	4	.085	448	*	Progress
2REP	2 replacements	3	.080	441	*	Progress
PRE	Prestructural	1	- .685	335	***	Hindrance
CON	Conjoining	1	- .428	485	***	Hindrance
NV	non-variable	1	- .376	335	***	Hindrance
SC2	Seek closure	1	- .368	440	***	Hindrance
IG	Ignore symbol	1	- .305	485	***	Hindrance
OBJ	Objects view	1	- .244	476	***	Hindrance
AL1	Alphabetic	1	- .225	392	***	Hindrance
PV	Place value	1	- .203	489	***	Hindrance
CF	Coefficients	1	- .159	441	***	Hindrance
SC1	Seek closure	1	- .146	436	***	Hindrance
JFL	Obj.flowers	1	- .109	448	**	Hindrance
AL2	Alphabetic	1	N.S.	492	N.S.	Neutral

Note. Sorted by size of correlation coefficients. Scales from dichotomous data.
Corrected TEST TOTAL = TEST TOTAL *minus* marks gained from items in scale.
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$. N.S. not significant.

4. Some of the scales appeared to be weaker instruments than others for measuring student understanding or misunderstanding. The only one which failed to give a significant correlation, namely, the AL2 Scale, was a measure of whether or not students applied an alphabetic code to two subparts of Question 4. However, the correlations presented in Table 6·10 were based on responses given by students who had been studying algebra for at least three weeks, by which time most of them had relinquished the use of an alphabetic code. In Chapter 8, responses are examined in relation to time and age, and Figure 8C·9, in Appendix 8C, shows that very few students scored more than zero on the AL2 Scale, even in Year 7.

Of the three scales for measuring the tendency to replace algebraic variables with numerical examples, the 1REP Scale had the most significant correlation. It was, therefore, decided to use mainly this scale in analyses reported later, while keeping in mind the possibility that the other two scales could be especially relevant in the early stages of learning algebra.

The NFL, JFL and AL2 Scales were all derived from responses to Question 4, about numbers of flowers, and all three scales gave rise to correlations which were less statistically significant than any other scales. The wording of the question may have been a difficulty. Only 7.7% of the 517 students succeeded in responding correctly to all three subparts of Question 4, and some of these did not score well on other important aspects tested, as was indicated by the correlation between scores on the FL Scale, which tallied correct responses to Question 4, and the corrected test total (a low .130, even though significant at the 5% level). These outcomes reinforced a degree of wariness, as expressed on page 186, with respect to the use of Question 4 for any major conclusions.

5. The table gives a first testing of the hierarchical order reported in Table 6·1 from previous research. The entries were sorted according to the size of the correlation coefficients. Such an ordering gave an approximate indication of the relative strengths of the relationships between the scale measures and the corrected test totals. It is only approximate because the significance of the size of a correlation coefficient is dependent upon the number of cases used and the numbers varied from scale to scale because of variations in the number of missing values. Furthermore, the corrected test totals were determined by subtracting scores on the items involved in the particular scale being correlated and did not supply a constant distribution of scores for all of the correlation coefficients. In Chapter 9, another set of correlational analyses (Tables 9·2 and 9.3) is presented which reduced the effect of these limitations by using a selection of test items (parts of Questions 10 and 12) as a means of measuring the degree to which students held views of symbols at the different levels.

Keeping these limitations in mind, it can be argued that the Level 5 scale scores were the best indicators of success on the algebraic tasks assessed by the test. Level 5 scales generally produced the largest correlation coefficients. Thus, students with the better understanding of algebraic symbols as numerical variables were the most likely to have a higher degree of success with the algebraic tasks tested. The CZ Scale had a rather low correlation by comparison with the other Level 5 scales, probably because very few students (4.3%) succeeded with Question 15 (iv), the basis of the CZ Scale.

Level 4 measures were apparently the next best predictors of achievement on the algebra tasks. Those who had a more strongly developed view of symbols as capable of taking non-integral values (FZN Scale), and/or as representing generalized numbers (GN Scale), or as standing for numbers of objects or numbers of people (NBR Scale) were more likely to score well on test items. Wariness was expressed above with respect to the low correlation for the Level 4 NFL Scale, built on responses to Question 4, about numbers of flowers.

Understanding algebraic symbols at development levels 3 or 2 related in a positive way with success on test items but the relationship was noticeably weaker than for understanding symbols at higher levels of development.

Higher scores at Level 1 were associated with poorer performances on the algebraic tasks tested.

Overall this rather limited assessment method, by means of ordering correlation coefficients, resulted in support for the hierarchical order given in Table 6·1. Further examinations of the proposed hierarchical order are taken up in Chapters 8 and 9.

6. The Chapter 5 categorizations of scales as Progress Indicators or Hindrance Indicators were all justified by the signs of the correlations with corrected test totals, except for the inconclusive situation for Scale AL2. The only scales over which uncertainty was felt prior to the calculation of the correlations were those measuring the Levels 3 and 2 approach of replacing algebraic symbols by numerical examples. From one point of view, this procedure hindered the student in attaining the more general algebraic solutions to the particular problems. On the other hand, students who gave correct examples had thereby proved that they understood the nature of the problem and were more advanced than those who responded at some fictitious or prestructural level. The fact that the correlations in Table 6·10 were positive and significant for these scales, even though they were a little low, indicated that the scales should be considered as Progress Indicators.

Correlations Between Views of Symbols and Specific Algebraic Tasks

The previous investigation examined associations between the extent to which students had developed certain views about algebraic symbols and their overall performance on the algebraic tasks encompassed by the test questions. The issue is probed further in this subsection by studying correlations between views about symbols and degrees of success with particular tasks.

Frequency Statistics for Achievement Scales. Frequency data are presented in Table 6·11 for seven scales which measured levels of achievement on selected algebraic tasks. These achievement scales are as follows:

SUBS, the Substitute and Solve Scale - writing numerical answers by substitution into algebraic expressions and solving the equation '$3a = 36$';

C2, the Two 'c' Values Scale - accepting '3' and '7.4' as values for 'c' in the equation '$c + d = 10$';

AD, the a, d Scale - realizing that two unrelated symbols are also unordered;

AR, the Arithmetic Scale - interpreting and solving '$3 * 4 = 6 * y$' when * represented plus (+) and when it represented multiply (x);

SYM, the Symbols Scale - writing algebraic symbols correctly in answers;

PS, the Professors-and-Students Scale - interpreting the equation '$S = 6P$'; and

EQL, the Equality Scale - showing an understanding of the meaning of "equals".

Table 6·11

Frequency Statistics for Achievement Scales

Scale	Scale max.	No.of valid cases	A%	B%	C%	Median Score	Mean per item	S.D. per item
SUBS	7	489	93.3	88.3	80.2	7	0.838	0.290
C2	2	482	93.8	93.8	78.4	2	0.861	0.286
AD	2	407	73.5	73.5	69.0	2	0.713	0.441
AR	4	471	87.7	81.7	58.6	4	0.730	0.363
SYM	6	440	84.3	65.9	45.2	4	0.598	0.372
PS	2	507	31.2	31.2	23.1	0	0.271	0.422
EQL	9	385	90.4	47.5	9.9	4	0.451	0.290

Note. Years 7 to 12, using Test 3 for Year 7 students. Scales from dichotomous data. Scale maximum = no.of items in scale. A% = percentage of valid cases who scored at least 1 on the scale. B% = percentage of valid cases who responded at Level 5 to at least 50% of items. C% = percentage of valid cases who responded at Level 5 to at least 80% of items. Entries ordered by values of C%.

More than three-quarters of the students achieved at least 80% mastery of the tasks tested by items in the SUBS Substitute and Solve Scale and the C2 Two 'c' Values Scale. More than half the candidates attained a similar level of performance with the AD and AR Scales. Almost half the students (45.2%) recorded correct answers for 5 or 6 of the items in the SYM Symbols Scale, but the remaining two tasks were found to be much more difficult. Less than one-quarter (23.1%) were able to respond correctly to both parts of Question 7, as the scores on the PS Professors-and-Students Scale recorded. Less than one-tenth (9.9%) managed to attain at least 80% mastery on the EQL Equality Scale, indicating that the cognitive demands for at least some of the items in this scale were high. Kieran (1981a) also found that the concept of equality was difficult and elusive for some students.

Correlations. Table 6·12 summarizes data on correlations between degrees of success on these tasks and strength of views of symbols at Levels 5, 4, 3, and 2. The Level 3 measure, the 2REP Scale, was discounted in favour of the 1REP Scale by the analysis reported on page 202 and had no significant correlations with the achievement scales given in the table. The NFL Scale, which has also been given a low rating for usefulness, was omitted from the table. Scores on the NFL Scale correlated significantly and positively with only one of the achievement scales, namely, the PS Professors-and-Students Scale.

It was found that, in general, scores for understanding symbols at the variables level (Level 5) correlated positively and highly significantly ($p < .001$) with scores measuring success on the tasks listed. This finding strongly supported Proposition 1, corroborating similar outcomes with respect to success on the test as a whole. Those with the more-developed variables view of symbols tended to be those who were more successful at the tasks selected.

The correlations for scores on the SYM Symbols Scale showed that those who were more likely to write answers correctly in symbols were also more likely to have higher levels of understanding for the meaning of the algebraic symbols they were using. The two highest correlation coefficients were with the Level 5 VBL Variable Scale ($r = .607$) and the Level 5 or 4 GNV Generalized Number or Variable Scale ($r = .627$). Amongst the questions contributing to the SYM Scale were the three parts of Question 9, taken from Küchemann (1980), which could be regarded as tests of the skill of manipulating symbols (e.g., "Multiply $n + 4$ by 5"). Table 6·11 gives evidence that the more strongly a student viewed algebraic symbols as numerical variables, the more likely was it that the student had success in manipulating symbols. Although only about 40% of the variance was common to the SYM Scale and each of the Scales VBL and GNV (the square of the r values was nearly .40), it seems that at least some of the student sample could manipulate variables and understood what they represented, unlike the majority of 11th grade students who participated in the Fourth U.S.A. National Assessment of Mathematics who "could manipulate the variables, but they did not understand what they represented" (Carpenter & Lindquist, 1989, p. 165). The interaction between readiness to write answers in symbols and levels of understanding of the use of symbols is examined further when Propositions 8 and 9 are discussed in the Chapter 8. In Küchemann's terms (Table 6·1), those students with a high score on the SYM Symbols Scale were operating at the level of at least a "Specific Unknown" view of symbols. The correlations indicate that many of them were operating at the higher Generalized Number and even the Variable level.

Table 6-12

Correlations Between Symbols Views at Levels 5, 4, 3, & 2 and Achievement

SCALE	SCALE	AR Scale	SYM Scale	SUBS Scale	EQL Scale	PS Scale	AD Scale	C2 Scale
(Level)	description	Arith. Ops.	Write symbol	Subs-titute & solve	Equals	Prof-Student	a, d any value	$c =$ 3 or 7.4
VBL (Level 5)	Variables Qq.6c,10,12,15iii	.453 (374) ***	.607 (381) ***	.400 (401) ***	N.A.	.148 (403) ***	.367 (347) ***	.370 (393) ***
EQN (Level 5)	Equations Task Q.13	.305 (406) ***	.516 (402) ***	.285 (427) ***	N.A.	.206 (437) ***	.231 (372) ***	.209 (422) ***
PL (Level 5)	Parallel Lines Task Q.11	.248 (451) ***	.370 (428) ***	.274 (467) ***	N.A.	.078 (483) *	.270 (396) ***	.271 (463) ***
BXBA (Level 5)	$b=x$; $b=a$ Q.15 i, ii	.390 (428) ***	.432 (419) ***	.320 (451) ***	N.A.	.183 (464) ***	.329 (378) ***	.273 (444) ***
CZ (Level 5)	$c =$ zero Q.15 iv	N.S.	.083 (425) *	N.S.	N.A.	.167 (472) ***	.091 (383) *	N.S.
GNV (Level 4 or 5)	Generalized No. *or* Variable	.414 (302) ***	.627 (309) ***	.397 (320) ***	N.A.	.178 (324) ***	N.A.	.392 (325) ***
GN (Level 4)	Generalized No. *not* Variable	.203 (381) ***	.218 (382) ***	.200 (402) ***	.314 (373) **	N.S.	.126 (349) **	.326 (411) ***
FZN (Level 4)	Fractions, Zero, Negatives	.365 (376) ***	.470 (382) ***	.257 (404) ***	.581 (353) ***	.146 (407) **	.269 (345) ***	N.A.
NBR (Level 4)	Numbers View	.131 (423) **	.324 (422) ***	.152 (444) ***	.348 (366) ***	N.A.	.192 (377) ***	.128 (461) **
INT (Level 4)	Integers only	.201 (396) ***	.174 (394) ***	.161 (418) ***	.230 (385) ***	N.S.	.106 (364) *	.153 (411) ***
1REP (Level 2)	One Replace-ment value	.134 (396) **	.118 (394) **	.111 (418) *	N.A.	N.S.	.096 (364) *	N.S.

Note. Years 7 to 12, using Test 3 for Year 7. Scales formed from dichotomous data. N.A. not applicable (common items). Figures in brackets give no. of cases. *** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$. N.S. not significant.

Table 6·12 shows that those students whose view of symbols was well-developed even at Level 4 were likely to be those who also were successful with the given tasks. The high positive correlations associated with the Level 4 scales established this conclusion. Those, for example, who had a better grasp of the idea that letters stand for generalized numbers which included non-integers (assessed by the FZN Scale) were more likely to succeed with problems involving the concept of equality (measured by the EQL Scale).

Those whose view of symbols was more firmly at Level 3 or 2 and were most inclined to give single replacement examples to general problems were, nevertheless, likely to do better than some others on most of the tasks selected for Table 6·12. The correlation coefficients were weaker and less significant than the corresponding coefficients for scales measuring higher levels for understanding symbols. Hence it is probable that the "some others" were the students at Level 1. These latter are the focus of the following stage of this correlational study of achievement and views of symbols.

Table 6·13 records the correlations between views of symbols at Level 1 and achievement on the selected algebraic tasks. Of the 68 pairings of scale scores which avoided shared common items, 44 (or 64.7%) were statistically significant and all of these were negative. Here is strong support for the obverse of Proposition 1:

Students with poorer levels of understanding of the meaning of algebraic symbols are more likely to have less success with the selected algebraic tasks.

Comments are confined to the three Level 1 Scales which produced at least one negative correlation with a numerical value larger than .300.

1. Those who showed weaker understanding of the problems assessed by the PRE Prestructural Scale were less likely to succeed in using symbols in answers (SYM Scale, giving the most negative correlation in the table, viz., - .632). This is evidence that those with little understanding of the meaning and use of algebraic symbols are unlikely to succeed with questions requiring the manipulation of symbols and then writing answers in symbols (SYM Scale), again showing the importance of developing an understanding of symbols together with the development of skills generally classified as symbol manipulation. Those who were more firmly at the Prestructural level were also less likely to have success with substituting and solving an equation (SUBS Scale, r = - .402), answering problems involving arithmetical operations (AR Scale, r = - .370), choosing '3' and '7.4' as possible values for '*c*' in Question 6 (a) (C2 Scale, r = - .346), and interpreting the symbols in the professors-and-students problem (PS Scale, r = - .151).

Table 6·13
Correlations Between Symbols Views at Level 1 and Achievement

SCALE (Level)	SCALE Description	AR Scale Arith. Ops.	SYM Scale Write symbol	SUBS Scale Subs-titute & solve	EQL Scale Equals	PS Scale Prof-Student	AD Scale a, d any value	C2 Scale $c =$ 3 or 7.4
OBJ (Level 1)	Objects View Qq.7, 8a	N.S.	- .195 (419) ***	- 094 (460) *	- .277 (370) ***	N.A.	- 130 (382) **	N.S.
JFL (Level 1)	Objects View Q.4	N.S.	- 054 (397) *	- .142 (431) **	- .118 (354) *	N.S.	- .087 (373) *	N.S.
SC2 (Level 1)	Seek Closure Qq.5,9,14	- .364 (404) ***	N.A.	- .215 (429) ***	- .403 (365) ***	- .100 (437) *	- .263 (363) ***	- .191 (425) ***
NV (Level 1)	Non-Variable View	N.S.	- .237 (389) ***	- .266 (411) ***	N.A.	- .143 (418) **	- .110 (358) *	- .114 (404) *
AL2 (Level 1)	Alphabetic Code Q.4	N.S.	N.A.	N.S.	N.S.	N.S.	N.S.	N.S.
SC1 (Level 1)	Seek Closure ∴ can't do Qq.12,13	N.S.	- .137 (400) **	N.S.	N.A.	N.S.	N.S.	N.S.
IG (Level 1)	Ignore Symbol	- .186 (446) ***	N.A.	N.A.	- .261 (376) ***	- .081 (477) *	- .168 (389) ***	- .225 (456) ***
AL1 (Level 1)	Alphabetic Code Qq.2i,9	- .142 (366) **	N.A.	N.S.	- .201 (326) ***	N.S.	N.A.	- .149 (374) **
PRE (Level 1)	Prestructural level on 19 items	- .370 (311) ***	- .632 (317) ***	- .402 (330) ***	N.A.	- .151 (334) **	N.A.	- .346 (322) ***
PV (Level 1)	Place value Qq.3, 8b	N.S.	- .282 (429) ***	N.A.	N.A.	- .081 (481) *	- .141 (392) **	N.S.
CON (Level 1)	Conjoining Qq.3, 9	- .256 (446) ***	N.A.	N.A.	- .364 (376) ***	N.S.	- .183 (389) ***	- .229 (456) ***
CF (Level 1)	Coefficients Q.13	- .093 (406) *	- .116 (402) **	N.S.	N.A.	N.S.	- .106 (372) *	- .093 (422) *

Note. Years 7 to 12, using Test 3 for Year 7. Scales formed from dichotomous data. N.A. not applicable (common items). Figures in brackets give no. of cases. *** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$. N.S. not significant.

2. Those students who were more inclined to replace symbols by arbitrary numbers (as measured by the SC2 Scale) were more likely to be those who had little success with questions which involved correctly understanding equality (EQL Scale, $r = -.403$), arithmetic operations (AR Scale, $r = -.364$), allowing unrelated symbols to remain unordered in size (AD Scale, $r = -.263$), substitution and solving an equation (SUBS Scale, $r = -.215$), admitting certain values for symbols in an equation (C2 Scale, $r = -.191$), and the professors-and-students problem (PS Scale, $r = -.100$).

3. There was a significant tendency for those inclined to confuse conjoining with an addition convention (as measured by the CON Scale) to score poorly on items which tested the concept of equality (EQL Scale, $r = -.364$), the use of arithmetic operations in an algebraic setting (AR Scale, $r = -.256$), the acceptance of certain values for symbols in an equation (C2 Scale, $r = -.229$), and the realization that two unrelated symbols cannot be ordered in size (AD Scale, $r = -.183$).

Review and Forecast

The first two sections of the chapter have considered various measures of levels of understanding for algebraic symbols. A start has been made on sorting the levels into hierarchical order of cognitive difficulty. Such hierarchical ordering will be further explored in Chapters 8 and 9 when changes in viewpoints are examined. The last section has presented a correlational study of the wide range of views expressed by students about the meanings of algebraic symbols. The correlational study has given empirical support for the expectations quoted at the start of the section that the views students hold about the meaning of symbols are related to their levels of success with algebraic tasks. Information about cause and effect has not been presented, but the correlations discussed so far have been united in support of those quoted expectations and of the proposition given in the section, namely, Proposition 1, which read: *Students with better levels of understanding of the meaning of algebraic symbols are more likely to have higher degrees of success with algebraic tasks.*

Not all data from the research project sang in harmony. In particular, responses to two subparts of Question 6 (a) jarred discordantly with the general trend in the findings. These responses have been deliberately omitted from the discussion so far and they are the centre of attention in the last section of the next chapter. Chapter 7 takes up the question of whether or not students viewed algebraic symbols as representing objects or numbers of objects. The answer to such a question is shown to be elusive when responses to Questions 7 and parts of 6 (a) are discussed in detail.

CHAPTER 7

SECOND STUDY: NUMBERS VIEW AND OBJECTS VIEW

Overview

All previous presentations concerning the research data have been from what could be termed a static view. The data have been considered in a global form, using responses from all students given at one particular time, and without subdivision.

In this chapter the view is more dynamic. Account is taken of changes in responses by Year 7 students as they answered the same test items three times during their first three weeks or so of algebra and again after about six months. Differences are noted between responses by classes which differed by experience with algebra or by ability level.

This chapter focuses on two aspects of the meanings for symbols which were discussed in Chapter 6. The Numbers View and the Objects View of algebraic symbols are investigated more closely in terms of student responses to Questions 7 and 6 (a) of the test, a copy of which is in Appendix 3I. The background to measuring the Objects View and Numbers View of symbols is presented first.

Section 1 explains the statistical procedure of using Odds Ratios to compare frequencies within different categories of responses to pairs of variables and to assess the level of significance of any association between these. Odds Ratios are used in this and later chapters in the thesis.

Section 2 presents insights obtained from the further investigation of a test item which has been used by various researchers over the past decade, namely, Question 7, the professors-and-students problem. The study pursues the major objective of investigating difficulties students have in understanding symbols. The use of a multiple-choice version of this problem provided raw data on very specific aspects of the ways students viewed the symbols '*P*' and '*S*' in the context given. The frequency counts of these data were categorized into response types and analysed to find out if student pre-algebra views persist or if changes with experience were evident. Reasons for choosing certain descriptors for the symbols were probed by interviews. Possible influences of classroom experiences were investigated and a link between the reversal error and the way one views symbols was identified.

Section 3 shows how responses to Question 6 (a) gave rise to a paradox. Some students appeared to believe that algebraic symbols could stand for objects such as cabbages or pears and, in other test items, could work with symbols at the sophisticated level of understanding them as numerical variables. Evidence is presented of the prevalence and persistence of the Objects View. Interviews were undertaken in 1991 with the specific aim of trying to elucidate the paradox and the outcomes are presented here.

Background to Measuring Objects View and Numbers View of Symbols

The discussions presented in Chapter 6 have indicated that the concept of a numerical variable and the recognition that algebraic symbols can represent numerical variables are of paramount importance in the development of expertise with the algebra of generalized arithmetic. Evidence was given that an Objects View of symbols was likely to hinder success with algebraic tasks and the development of the Variable View of symbols, while a Numbers View was more likely to assist students to succeed both with algebraic tasks and the development of an understanding of symbols as variables. This chapter looks at the Numbers View and the Objects View more closely.

In the trialling stages, it was found that beginning algebra students were probably at a formative stage and had not settled on a clear view of algebraic symbols. This awareness ensured that the final test would not concentrate unduly on trying to categorize students according to their view of symbols, since some tended to use one view for one question and a different view for another (Chapter 3, p. 62). The final test included items which attempted to identify whether or not students tended towards an Objects View or a Numbers View of symbols (Table 3·3, p. 66). As this chapter will show, this was not a simple task and, maybe in some future research, better items might be devised for assessing these contrasting viewpoints.

Item 8 (d) was written as a combination of two items found to be inconclusive in trialling (Chapter 4, p. 115). It was hoped that this item would be able to classify students as having a Numbers View or an Objects View. While it seemed to achieve this objective within its own context, it was noted that students did not consistently stay in the one category when they answered the other items designed to register their viewpoint. The item did not qualify for admission to any of the Numbers View or Objects View scales established according to the standards described in Chapter 5. Scores for favouring the Objects View in the item did, however, correlate significantly with some of these scale scores, namely, the NFL Numbers of Flowers Scale (r = .113, p = .009, N = 440), the OBJ Objects View Scale (r = - .078, p = .047, N = 463), and the NJCP No Cabbage or Pear Scale (r = .076, p = .050, N = 467). Although the numerical values of these correlations were small, they did show that there was at least some minimal tendency for students to be consistent across items. Appendix 7A presents some cross-tabulations of responses to show that inconsistencies were, nevertheless, noticeable.

Other items for registering a Numbers View qualified for inclusion in scales.

The NBR Numbers View Scale incorporated the following four items from three different test questions: 6 (a) (v), 7 (i), 7 (ii), and 14 (a). Scores on this scale were found useful in several analyses discussed in Chapter 6 (e.g., pp. 197, 201, 206). The NFL Numbers of Flowers Scale was composed of two parts of one question, namely, 4 (b) and 4 (c), and gave support, although not strong, for the outcomes derived from the NBR Scale (Appendix 6B).

Three scale measures for the Objects View were established, as explained in Chapter 5. The OBJ Objects View Scale was comprised of three items from two questions, namely, 7 (i), 7 (ii) and 8 (a), and this scale contributed to several analyses in Chapter 6 (e.g., pp. 197, 201, 208). The JFL Objects Notion for Flowers Scale combined responses to Items 4 (b) and 4 (c) and was found to be of limited value, as was its counterpart, the NFL Scale (p. 202). The JCP Cabbage & Pear Scale, with constituent items 6 (a) (vi) and 6 (a) (vii), was described as an Unclassified scale in Chapter 5 (p. 170). Section 3 below reports on the investigations used to interpret student responses to this scale.

Section 2 gives special attention to two items which were incorporated in both the NBR and OBJ Scales by choosing different categories of answers. The items were the two parts of Question 7 which was a multiple-choice version of the well-documented Professors-and-Students problem.

Section 1: Odds Ratios

A statistical procedure involving odds ratios was used in several analyses with the purpose of comparing frequencies within given categories. The odds for success in some event is the value of the ratio of the frequency of success to the frequency of failure. An odds ratio is, as the name implies, a ratio of the odds for success on one variable to the odds for success on another variable. To illustrate the statistical technique, an example is presented which is relevant to the subject matter of the chapter.

Example. The odds ratio method identified significant differences between the Textbook and Manipulatives Groups on Item 8 (d) in Test 2. This item placed students in the category of thinking of algebraic symbols either in terms of numbers (correct) or objects (incorrect). Of those in the Textbook Group (Classes 2, 4, and 7), 38 were correct and 26 were incorrect, giving:

Odds for being correct, if in the Textbook Group

$$= \frac{\text{No.correct}}{\text{No.incorrect}} \qquad = \frac{38}{26}$$

Of those in the matched Manipulatives Group (Classes 1, 5, and 8, taken from the same schools as the Textbook Group and matching them on ability rankings), 48 were correct and 10 were incorrect, giving the following:

Odds for being correct , if in the matched Manipulatives Group

$$= \frac{\text{No.correct}}{\text{No.incorrect}} \qquad = \frac{48}{10}$$

The odds ratio for comparing the success rate of the Manipulatives Group to that of the Textbook Group is the ratio of the above two odds, namely,

Odds Ratio (Manipulatives : Textbook)

$$= \frac{\text{odds for success if in Manipulatives Group}}{\text{odds for success if in Textbook Group}}$$

$$= \frac{48\ /\ 10}{38\ /\ 26} = 3.28$$

As the ratio in this case is greater than 1, it indicates that membership of the Manipulatives Group gave a student more probability of preferring to regard symbols as representing numbers rather than objects than if the student were a member of the Textbook Group. Similarly, if the odds ratio had been less than 1, the probability of success would have favoured those in the Textbook Group.

The odds ratio quoted above could be given a different interpretation by writing it as

$$\frac{48/38}{10/26}$$ which has the same value, 3.28, as before.

In this form, it could be considered as a ratio expressed as follows:

Odds Ratio (successful : non-successful)

$$= \frac{\text{odds that successful students were in Manipulatives Group}}{\text{odds that non-successful students were in Manipulatives Group}}$$

Again the outcome supports the view that members of the Manipulatives Group were more likely to succeed on Question 8 (d) than members of the Textbook Group.

Odds ratios can always be re-written in ways analogous to this example. Such flexibility can add to their usefulness.

Discussion of the process. In the literature, the odds ratio is also referred to as "relative risk" (Goodman, 1964, p. 88; McPherson, G., 1990, p. 255), "cross-ratio" (A. W. F. Edwards, 1963, p. 109) and "cross-product ratio" (Fleiss, 1973, p. 45).

A. W. F. Edwards (1963) presented arguments to support the view that the measure of association in a 2 x 2 table, in which the marginal totals are not fixed, should logically be some function of the cross-ratio. He pointed out "the association is complete" (p. 111) if only one of the cell frequencies is zero.

Fleiss (1973) claimed that "The measure of association based on [two odds] that is currently in greatest use is simply their ratio" (p. 45). He also pointed out that, for fourfold tables of the type involved here, "the simplest and most frequently applied statistical test of the significance of the association indicated by the data is the classical chi square test" (p. 14). Hence, when odds ratios are presented in this study, the corresponding chi square values and their significance estimates (*p*) are also reported, unless there was reason for using the Fisher Exact Test instead, as explained below.

There has been some debate, "with firmly held views on either side" (Howell, 1982, p. 102), about whether or not to use a correction for continuity, such as the method suggested by Yates (1934, p. 222), when applying a chi square significance test to 2 x 2 tables. McNemar (1969, p. 262) claimed that a correction should be used if any expected (or theoretical) frequencies were from 5 to 10. Grizzle (1967) argued against using a correction for continuity when applying the chi square test to 2 x 2 tables if the sample size is small because it results in "a test that is so conservative as to be almost useless" (p.29). He based his decision on a series of 500 simulations for each of two cases, Case I in which either the row or the column totals were known in advance (as is true for the 2 x 2 tables for comparing performances by the Manipulatives and Textbook Groups), and Case II in which only the sample size is determined beforehand (as in analyses to study changes in scores).

Mantel and Greenhouse (1968), in debating Grizzle's paper, took as fundamental that "essentially we wish to make a Fisher exact test" (p. 28) and then discussed whether or not the corrected or the uncorrected chi square test gets closer to the Fisher exact test outcomes. The Fisher exact test, then, would seem to be the most appropriate to use if the calculation is available. The SPSSX User's Guide (SPSS Inc., 1986, p. 346) states that "Fisher's exact test is computed ... when there are fewer than 20 cases in a 2 x 2 table." In none of the reported 2 x 2 table analyses was the total number of cases less than 20, and so the Fisher test was not available on the printouts for any of these analyses.

Dixon and Massey (1957) rather dogmatically stated that "the minimum theoretical frequency for the 2 x 2 table should not be less than 5" (p. 226) if the chi-square test were to be useful. This view was supported by such writers as Lewis and Burke (1949, p. 463), McNemar (1969, p. 262), and A. L. Edwards (1973, p. 139).

The decision was made that all reports on tests of significance for 2 x 2 contingency tables would follow the three recommendations made by Cochran (1952, p. 334) that

1. the chi-square test, corrected for continuity, be used if the number of cases was greater than 40 and the smallest expected frequency was less than 500;

2. Fisher's exact test be used if the number of cases was less than 20; and

3. if the number of cases was from 20 to 40, the uncorrected chi-square test be used, unless the smallest expected frequency was less than 5, in which case Fisher's exact test was used.

As Siegel (1956, p. 99) pointed out, the Fisher test can become computationally very tedious if the smallest cell value in the contingency table is even moderately large. This is a consequence of the need to calculate and operate with a total of nine factorials for each stage of the test, progressing until a zero frequency is reached. Siegel published a table (1956, pp. 256 - 270) to alleviate the burden, when the significance level based on the Fisher test was required. Siegel's table did not cover cases in which both a row and a column frequency were larger than 15, and so a hand calculator was used in such cases if the Cochran recommendations required the use of the Fisher test. When a zero was in the observed frequencies the Fischer exact probability was calculated with a hand calculator and reported, as the calculation involved only one step.

Example revisited. In the case used above as an example, the fourfold, or 2 x 2, table was as shown in Table 7·1 and the chi square value, using Yates' correction, was 6.9 with a significance level $0.001 < p \leq 0.010$, indicating that the odds ratio favouring the Manipulatives Group was statistically significant.

In reporting Odds Ratio analyses in the pages which follow, tables will be presented to summarize the relevant features in the format shown in Table 7·2. The example discussed above is used in this table.

Table 7·1

2 x 2 Cross-tabulation Table for Responses in Item 8 (d)

Item 8d Frequencies	Correct	Incorrect	Row Totals
Manipulatives	48	10	58
Textbook	38	26	64
Column Totals	86	36	122

Note. Year 7 Manipulatives and Textbook Groups matched by ability ratings.
$\chi^2 = 6.9$, $df = 1$, $.001 < p \leq .010$.

Table 7·2

Odds Ratio Analysis for Success on Question 8 (d)

Variable		Odds Ratio for success on Q.8 (d) $= \frac{\text{odds if Manip.}}{\text{odds if Textbook}}$		Comment	χ^2 Test of Significance ($df = 1$)	
Item	Description	Calculation	Value		Value	p
8 (d)	Numbers view, not objects view	$\frac{48/10}{38/26}$	3.28	Manip. more success	6.9	**

Note. Test 2 responses from Manipulatives (Manip.) and Textbook Groups
** $.001 < p \leq .010$.

Section 2: The Professors-and-Students Problem

Question 7 was a revised version of Rosnick's (1981) multiple-choice format for the professors-and-students problem. This problem has been in the research literature for over a decade (e.g., Rosnick & Clement 1980; MacGregor 1991). Data assembled in the present study were able to throw more light on the cognitive processes used by the students when dealing with this problem. Statistics for performance by all secondary student participants on Question 7 were presented in Table 4·13 and Figure 4·11 within Chapter 4. The analyses reported in this section build upon the work of earlier researchers and complement those reported in the following section on the paradox which was derived from responses to Question 6 (a).

Categories of Responses to Question 7

The question asked students to interpret algebraic symbols in a real-life context and involved the complexities associated with understanding the meaning of the question and analysing the mathematical implications, as were identified by MacGregor (1991). The responses were useful in determining the extent to which students regarded letters in algebra as representing numbers of people or simply people. Regardless of whether or not they indicated reversal errors, they were divided into four categories for both parts of the question as follows:

1. Category 'N': *Number of persons.* Options (iii) or (vi). For students in this category, the letter stood for a number of persons.

2. Category 'N+P(s)': *Number of persons + Person(s).* Two or more options including (iii) or (vi), e.g., (ii) and (iii), or (iv) and (v) and (vi). Students here chose options which not only described letters as standing for numbers of people, but also as standing for a person (e.g., "Professor") or for persons (e.g., "Professors"). This implied that they thought of algebraic symbols as representing both numbers and people.

3. Category 'P+Ps': *Person + persons.* Options (i) and (ii), or (iv) and (v). Students in this category chose two options, one singular and one plural, with both showing that they thought of letters as standing for people.

4. Category 'P(s)': *Person or persons.* One of options (i), (ii), (iv), (v). Students in this category also recorded that they accepted letters as representing people, saying, for instance, that '*S*' stood for "student" or "students".

Category 'N' will be referred to as the "Numbers View" of algebraic symbols, 'N+P(s)' as a "Conglomerate View", and the other categories as the "People View" (i.e., Objects View). Table 7·3 gives the percentages of students in each of the four categories, showing separately the Year 7 responses (on Test 3 if they did the test more than once), the responses for Years 8 to 12, and the overall totals.

Table 7·3

Percentage Frequencies of Response Categories for Question 7 Parts (a) & (b)

Q.7 category	Year 7	Years 8 to 12	Years 7 to 12
N	26.2 (23.6)	38.6 (36.3)	33.1 (30.7)
N+P(s)	8.6 (8.2)	17.0 (15.1)	13.3 (12.0)
P+Ps	3.2 (3.2)	3.2 (2.9)	3.2 (3.0)
P(s)	62.0 (65.0)	41.2 (45.7)	50.4 (54.2)
No.of cases	220	277	498

Note. Numbers in brackets are for part (b). Year 7 responses from Test 3.

Table 7·3 draws attention to the following three differences between response patterns from students in Year 7 classes and those from students in higher classes:

1. A greater proportion of students in higher classes chose Category 'N', the Numbers View of symbols;

2. A greater proportion of students in higher classes selected Category 'N+P(s)', the Conglomerate View; and

3. A smaller proportion of students in higher classes chose categories 'P(s)' or 'P + Ps', the People View, although the proportion was still high at over 44%.

A method of analysis involving Odds Ratios was explained in Section 1 above. The method was appropriate for investigating whether or not the differences just noted were statistically significant. Table 7·4 reports the odds ratios in terms of whether or not students were in Year 7 when choosing different categories of options from Question 7 (a). Test 3 scores were used for the Year 7 data.

According to Table 7·4, the Years 8 to 12 students were significantly more inclined to choose options in the Number View category, 'N', than the Year 7 students. Perhaps that view was too sophisticated for the majority of the beginning students. The students with more experience of algebra were significantly more inclined to choose options in the Conglomerate View category, 'N + P(s)', than the Year 7 students, and they were significantly less inclined to choose the Objects View category, 'P(s)', which was chosen by the majority (over 60%) of the less-experienced students. Many of the latter may not have thought in terms of possibilities other than letters-for-objects (or persons) when presented with algebraic symbols in a real-life context, whereas the students in the higher grades had become more flexible in their outlook. Very few chose two options in the 'P + Ps' category and there was no significant difference between the Year groups on this aspect.

Table 7·4

Odds Ratios for Choosing Categories in Q.7 (a) in Terms of Year Levels

Categories	ODDS RATIO for choosing category $= \frac{\text{odds if in Yrs.8 to 12}}{\text{odds if in Yr.7}}$		Comment	χ^2 Test of Significance	
Q.7 (a)	Calculation	Value		Value	p
N	$\frac{107\ /\ 171}{54\ /\ 146}$	1.69	Yrs.8 to 12 more Number View	6.87	**
N + P(s)	$\frac{47\ /\ 231}{18\ /\ 182}$	2.06	Yrs.8 to 12 more mixed view	6.19	*
P + Ps	$\frac{9\ /\ 269}{6\ /\ 194}$	1.08	No significant difference	.02	N.S.
P(s)	$\frac{115\ /\ 163}{122\ /\ 78}$	0.45	Yr.7 more People View	17.94	***

*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

For more detailed information about these differences in viewpoint, percentages for each response in Question 7 (a) are graphed in Figure 7·1 for each Year grouping of classes. Figure 7·2 presents similar information for Question 7 (b).

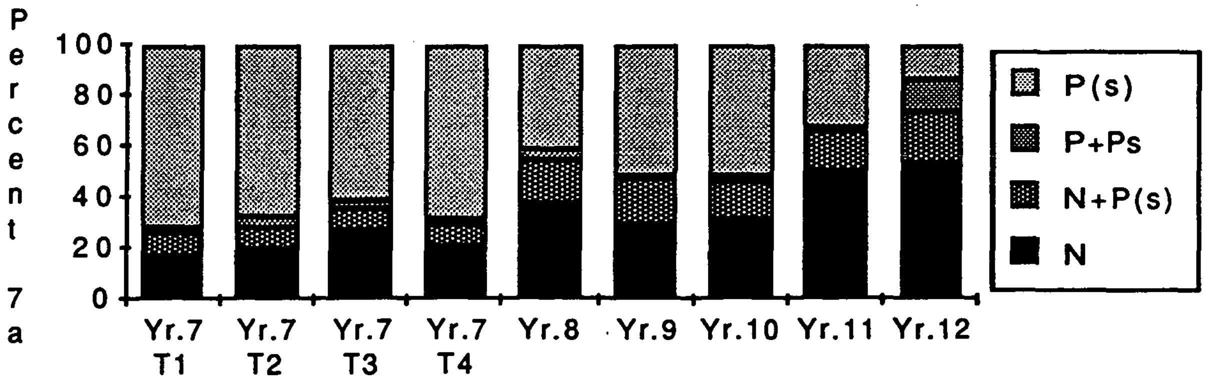


Figure 7·1. Percentage frequencies for response categories for Question 7 (a) (as described in the text)

The patterns of responses were similar for the two parts of the question but were not identical as 13.2% of students changed their choice of the type of option from Part (a) to Part (b). The Year 10 group recorded the largest discrepancies between responses to the two parts, showing less inclination for the number view (Category 'N') in Part (b) than in Part (a). Closer investigation revealed that it was the lower ability students of Year 10 that accounted for most of this inconsistency, as shown in Figures 7·3 and 7·4 below.

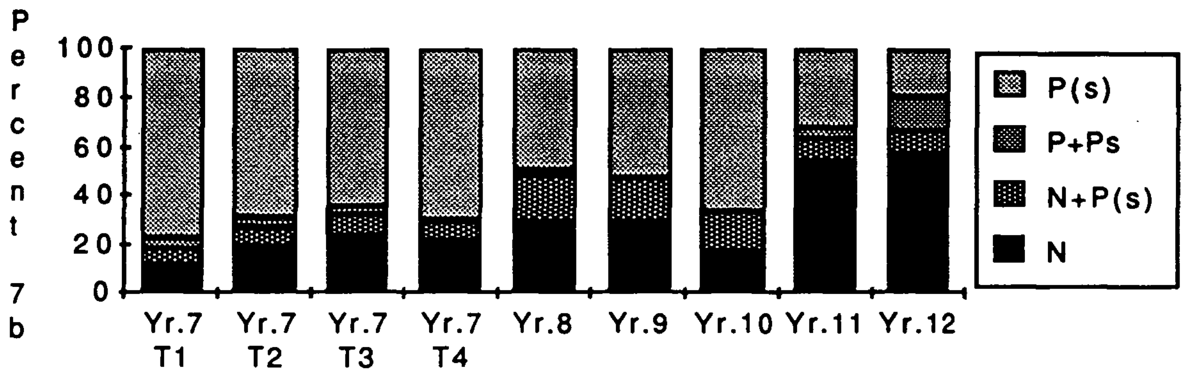


Figure 7·2. Percentage frequencies for response categories for Question 7 (b) (as described in the text)

Generally speaking, there was a trend towards the numbers view (shown by black shading in the figures) with more experience in algebra. This was so for the beginning students in Year 7 according to responses on the first three tests, but on Test 4 they showed a regression towards a people view.

Numbers View. Figures 7·3 and 7·4 show the percentages choosing the number of persons options in Parts (a) and (b) respectively. The percentages are shown separately for Advanced classes and for other classes in each Year grouping.

Figure 7·3. Percent choosing number of people in Q.7 (a)

Figure 7·4. Percent choosing number of people in Q.7 (b)

The most distinct differences were in the case of the most senior classes, namely, Years 11 and 12, where the advanced groups clearly favoured the numbers view of the symbols whereas the other groups did not. This could be an indication that the acceptance of the numbers view of symbols, which was correct in the case of Question 7, may characterize the better mathematicians, especially after considerable experience with algebra. Perhaps younger and more inexperienced algebra students tend not to perceive the importance of the distinction between letters as representing numbers and letters as representing objects. In the case of the Year 7 students, the Advanced groups outscored the other groups by making the distinction in most testings, particularly in Test 3.

<u>People View (or Objects View)</u>. As recorded in Table 7·4, the Year 7 students were significantly more inclined to choose a People View of symbols than were those in higher classes. Combining responses in categories 'P(s)' and 'P + Ps' gave the figures for the People View. This combination produced a result comparable to that tabulated for the 'P(s)' category alone, with a significant Odds ratio of 0.45. From Table 7·3, the proportion of beginning students who regarded the symbols as representing people rather than numbers of people was 65.2% (= 62.0 + 3.2) in part (a) and 68.2% (= 65.0 + 3.2) in part (b). These outcomes show that the majority of these students had misunderstood the meaning for the symbols in the professors-and-students problem after about three weeks of algebra. The influence of the teaching sessions on these views during that three weeks is considered later in this section. This form of error was not confined to beginning students. Nearly half of those in Years 8 to 12, all with at least one year's experience with algebra, held a People View of the same symbols: 44.4% for part (a) and 48.6% for part (b).

<u>Undifferentiated conglomerate view</u>. Across all the Year groups, there were 66 students (13.3%) who chose the 'N + P(s)' category in Part (a), and 60 (12.0%) who did the same in Part (b). These students demonstrated that they held a view of algebraic symbols that incorporated both a people view and a numbers view. They seemed to have an ambivalent approach to symbols which failed to differentiate between interpreting them correctly as standing for numbers and interpreting them as standing for, in this case, persons. For them, alphabetic symbols were seen as "undifferentiated conglomerates". Rosnick used this term to describe those who changed their view of the meaning of symbols, using them "in an imprecise, inconsistent, paradoxical and over-associative manner" (1982, p. 25). He gave an example (p. 8) of a student who considered the symbol '*B*' to stand for "books", "number of books", "a price", "one book", and "total amount of money spent on books", all in the one problem.

The total number of students exhibiting an undifferentiated conglomerate view was calculated by adding the number who kept the 'N + P(s)' option from Part (a) to Part (b) and the number who changed response categories from one part to the next. There were 50 in the first group and another 65 who changed their point of view from Part (a) to Part (b), making a total of 115 students. Thus 23.3 percent, nearly one quarter, of the students recorded their confusion about this fundamental issue in basic algebra, namely, the meaning of the symbols being used. The most common change of categories from one part to the next was from Category 'N' to Category 'P(s)', a change which accounted for 40.0% of all changes of viewpoint. Changes in the reverse direction were the next most popular, accounting for another 18.5% of the changes. Students making such combinations of choices in Question 7 indicated that they did not appreciate the fact that the symbols used in the form of algebra they were studying were numerical variables and could not, therefore, represent people as such. The intriguing fact was that the undifferentiated conglomerate view was exhibited mainly by the older and even the more able students, as the following statistics recorded:

1. More of the Years 8 to 12 students than the Year 7 students expressed a mixture of views in both parts of the question, as was reported in Table 7·3. These differences were statistically significant, as detailed for part (a) in Table 7·4;

2. For Part (a), groups that gave the greatest percentages of responses at the undifferentiated conglomerate level were: Yr.11 average stream (23.5 %), Yr.9 Advanced (21.8 %), Year 12 lowest stream (20.0 %) and Yr.12 top stream (20.0%);

3. For Part (b), the greatest percentages were from: Yr.8 Advanced (19.2 %), Yr.9 Advanced (19.2%), Yr.8 low stream (18.2%), Yr.10 Advanced (18.2%) and Yr.11 average stream (17.6%).

These outcomes parallel somewhat the findings from the "cabbage and pear paradox" which arose from Question 6 (a). As will be mentioned in the discussion of this paradox in the next section of this chapter, it is possible that some of the participating students were consciously or unconsciously moving beyond the numerical variable concept for algebraic symbols and becoming aware of the possibility,that such symbols might, in some circumstances, represent something other than numbers, such as lines, or slices of solids. On the other hand, the choosing of non-numerical options for '*S*' and '*P*' in Question 7 could have been simply an indication that the candidate was not thinking clearly about these symbols in the context of the given question.

Sample extracts from interviews illustrate the ambivalence exhibited by some students regarding the meaning of algebraic symbols.

Interview Extract 1. (Nov. 1990, after 4th Test - Year 7 student, 'M', chose options (i) and (iii) respectively)

E In No.7 you've said the letter '*P*' stands for some people - "professors".
M Mm.
E And in the next bit you've said '*S*' stands for "number of professors". You've got people, then you've got numbers.
M It says '*S*' equals six professors, so I put "number of professors", because '*S*' is 6. That's why.

Interview Extract 2. (June 1990, after 2nd Test - Year 7 student, 'P', chose options (i) and (iii) in Part (a))

E You got "Professors" and "number of professors".
P Yes, you were saying that students equals 6 professors, and students equals 6, like the number of professors.
E What's the number of Professors?
P 6.
E There's 6 Professors ?
P I just wrote it as the number of Professors, because that's what P stands for, the number of professors.
E You also said that '*P*' stands for "Professors".
P Yes.
E To me that's two different things. In one case you say it stands for people, and in the other you say it stands for a number of people. Can you see a difference?
P Yes.
E Are you saying that you can't make your mind up?
P Yes, I can't really. I suppose I go for the number of Professors. That's a better answer.

Plural Option Favoured

Those who chose people rather than numbers of people as the meaning for the symbols favoured the plural options rather than the singular options. As Table 4·13 recorded for all 517 students, 33.8% chose "students" and only 6.0% chose "student" in part (b). Several students explained in interviews that they regarded the plural version as acceptable since it allowed for a variable number of people, whereas the singular form was fixed at simply one person. They did not see the need to choose the options which specifically said "number of ...". The importance to students of whether or not an option was singular or plural was a particular finding of this study, not referred to in earlier studies listed in the references. Similarly, in the analysis of Question 6 (a) it was found that some students rejected "an object like ... " on the grounds that '*c*' should be plural.

The following two interview extracts exemplify student awareness of the need for a plural option. The first student keeps to plural, choosing options in terms of people rather than numbers of people although he shows considerable mathematical ability and presented his rationale in terms of numbers. The second student accepts on equal footing the options "students" and "number of students" for '*P*', keeping to plural again, although making the reversal error.

Interview Extract 3. (Year 12 student, 'S', School C, 24 July, 1991)

E Have a look at Number 7.
S Well, there are fewer professors than there are students. So you'd have to multiply the number of professors by another number larger than one to get the number of students. So '*P*' has to equal "professors" and 'S' would have to equal "students".
E Um. There's some other options that were available. You picked part (i) there I see.
S Yeah! Actually I think what it could have easily been ... well, no it couldn't ... because in 7 it said a certain university has 6 times as many students (plural!) as there are professors (plural!) and so it can only be "professors" and "students".
E Alright. Keep to plural.

Interview Extract 4. (Year 10 Advanced student, 'B', School B, 24 July, 1991)

E What about the next one?
B Six times as many students as there are professors. '$S = 6P$'. Yes. '*S*' is the sign for the professor 'cos there's 6 students to every one professor and '*P*' must be the students. And in this equation what does the letter '*P*' stand for? '*P*' stands for "students" or "the number of students". It can't be "student" because there's more students than one. Can't be "none of the above" or any of the others.
E So, "students" or "number of students"?
B Yes.
E What are you thinking there?
B 'Cos the question says there are 6 students. It can't be ... it has to be more than one student.
E So, options ... ?
B iv and vi.

Influence of First Letter

There were a few instances of students who indicated in interviews that they had decided on the meanings for '*S*' and '*P*' simply by matching the symbols with the first letters in the available choices. Interview Extract 5 illustrates this mode of reasoning.

Interview Extract 5. (Year 7 student, 'J', School D, 6 June, 1990, after Test 1)

E All right, what did you think J?
J I don't know. I just put '*P*' for "Professors" and '*S*' for "Students".
E Any special reason?
J ... the same letter at the start.

Persistence of Pre-algebra Views

Test 1 recorded the pre-algebra views of the Year 7 students since it was administered just before they started their classroom study of algebra. Of the 208 who completed the test three times, 14.42% repeated their Test 1 answer to Question 7 (a) in all three tests (but not in the fourth test) and the percentage who did the same for part (b) was 13.94%, as shown in Table 7·5. Of the 186 who did the test four times, 19.35% used the same answer four times in part (a) and 22.04% did the same in part (b). By far the most persistent pre-algebra views were option (i) in part (a) and option (iv) in part (b), which were respectively '*P*' as "professors" and '*S*' as "students". In both these cases, the students were thinking of the algebraic symbols as representing people rather than numbers of people, and their classroom experiences did not alter their perception. This combination of responses was maintained by 33 students for either the first three tests only (15 cases) or for all four tests (the other 18 cases). Only 6 students consistently gave the correct pair of responses, namely, option (iii) followed by option (vi), 3 of them for the first three tests only and the other 3 for all four tests. In responding to part (a), a total of 66 students did not change their mind for either the first three tests or for all four tests, and 70 students did likewise in responding to part (b). These outcomes underline the difficulty of communicating to beginning algebra students the true meaning of algebraic symbols. In the context of the question, both '*P*' and '*S*' represented numbers of people, yet students who started with some other view, as listed in Table 7·5, did not change to the correct view.

Table 7·5

<u>Persistence of Pre-algebra Views in Questions 7 (a) and 7 (b)</u>

		% who chose same option			
	Option	in 1st 3 tests only[#]		in all 4 tests[##]	
Code		7 (a)	7 (b)	7 (a)	7 (b)
1	professors	9.62	0.48	14.52	
2	professor	0.96	0.48		1.61
3	number of professors	2.88		1.61	
4	students	1.92	7.69	1.08	16.67
5	student		0.96	0.54	
6	number of students		1.44		1.61
7	none of the above	0.48	0.48		
8	2 or 3 of i, ii, iii	2.88		1.61	
9	2 or 3 of iv, v, vi	0.48	2.40		2.15
TOTAL	Percentages	14.42	13.94	19.35	22.04
	Numbers of students	30	29	36	41

[#]$N = 208$, [##]$n = 186$.

Numbers View and Teaching Approach

For responses in Test 3 on both parts of Question 7, students in the Textbook group (Classes 2, 4 and 6) were significantly less likely to choose a numbers view than those in the Manipulatives group (Classes 1, 5, and 6), the groups being matched on the ability ratings given for the classes by the school mathematics departments. Table 7·6 shows the relevant odds ratios for success in choosing a "number of people" option, depending on whether or not the student was introduced to algebra by a concrete approach (in a Manipulatives group) or a more traditional one (in a Textbook group). A similar significant difference was reported in Table 7·2 on the basis of odds ratios applied to Question 8 (d) responses, manipulatives classes again showing a numbers view more strongly. A possible explanation for this difference is that teachers using manipulative models may have been more aware of the need to stress that the algebraic symbols being used actually stood for numbers rather than objects. When using the objects-and-containers model, for instance, they emphasized that it was the variable number of objects in a container that represented a symbol such as 'y' rather than the container itself representing 'y'. It could have been a case of what you teach is what you get from students.

Table 7·6

Odds Ratios by Teaching Approach for Choosing a Numbers View in Q. 7

ITEMS	ODDS RATIO for being correct $= \frac{\text{odds if Manipulatives group}}{\text{odds if in Textbook group}}$		Comment	χ 2 Test of Significance	
Q.7	Calculation	Value		Value	p
(a)	$\frac{19 / 39}{11 / 54}$	2.39	Manipulatives better	4.17	*
(b)	$\frac{15 / 42}{8 / 57}$	2.54	Manipulatives better	3.90	*

Note. Test 3 responses. Year 7 groups matched on ability ratings.
* $.010 < p \leq .050$.

Observations on classroom activities and learning algebra. As was made clear at the start of this thesis (pp. 15 - 17), the challenges involved in mounting a research experiment to compare teaching approaches were beyond the scope this project. The inclusion of classes taught by differing approaches was found to be advantageous in teasing out the stages of the learning processes involved when being introduced to the basic concepts of algebra. While investigating rates of cognitive development, some differences were observed between responses given by students who had different experiences in class. The comparisons made in Tables 7·2 and 7·6 have been included as there was the possibility that classroom environmental factors were related to the outcomes. A selection of other observations which may have been related to classroom experiences has been assembled in Appendix 7B. Statistically significant differences were noted between the rates of progress of Manipulatives and Textbook groups during their first few weeks of algebra. However, very few differences were observed over the long term that might have been accounted for by different classroom activities or approaches.

Reversal Error

To complete the account of insights derived from responses to Question 7, the following two aspects of the reversal error are now addressed:

1. Does a relationship exist between the level of understanding that students have of symbols and their tendency to make the reversal error?

2. Does the setting of a multiplicative relationship in an abstract or real-life context affect the likelihood of the reversal error being made?

Firstly, the data did show that students were significantly more inclined to make the reversal error in part (a) of Question 7 if their responses were in the 'N + P(s)' or 'P + Ps' or 'P(s)' categories rather than in the 'N' category. Table 7·7 summarizes the odds ratio analysis which supports this claim.

Table 7·7
Odds Ratio for Reversal Error in Question 7 (a)

Odds Ratio for Making Reversal Error $= \frac{\text{odds if people view}}{\text{odds if number view}}$		Comment: Favours	χ^2 Test of Significance (df = 1)	
Calculation	Value		Value	p
$\frac{72/267}{21/144}$	1.89	Reversal if People view	4.79	*

Note. Years 7 to 12, using responses from Test 3 for Year 7 students.
* $.010 < p \leq .050$.

Of those who responded as if the symbols represented people or both people and numbers of people, 72 (or 21.2% of them) made the reversal error while 267 did not, as is shown by the figures in the numerator of the calculation in Table 7·7. Of those who viewed the symbols correctly as representing numbers of people, 21 (or 12.7%

of these) made the reversal error while 144 did not. The odds ratio was significant, supporting the contention that students were more likely to make the reversal error if they had a people view of the symbols rather than a number view. The chi-square value for corresponding analysis for part (b) was 3.50. This outcome was not statistically significant, indicating that the trend shown in Table 7·7 was not one that could be spoken of in completely generalized terms. Nevertheless, the outcome for part (a) underlined the importance of the level of understanding one has of the algebraic symbols being used in an equation such as 'S = 6P'. In the Chapter 2 summaries of the work of Rosnick and Clement (1980) and MacGregor (1989), it was noted that they had pointed out the influence on performance in such questions of the way students viewed algebraic symbols.

2. A comparison of reversal versus no reversal in Items 2 (ii) and 7 (a) using the odds ratio approach produced a non-significant outcome. Both items presented students with an equation in which a multiple of one variable was equal to another variable. In item 2 (ii), 61 students made the reversal error while 393 did not and, in Item 7 (i), the corresponding figures were 86 and 368. In percentage terms, 13.4% and 18.9% respectively made the reversal error in Items 2 (ii) and 7 (a). As explained in Chapter 3, Item 2 (ii) was a re-drafting of a question used by MacGregor (1989). When she used her version of the item with 235 students in Year 9, she found that 13.2% made the reversal error (1989, p. 122), a percentage which is similar to that recorded by the Years 7 to 12 students in the present study. MacGregor also reported that "changing an abstract content to a concrete one did not affect the tendency to reversal" (1991, p. 100) in the case of three other items. The fact that a significant difference was not inherent in the reversal data for Items 2 (ii) and 7 (a) corroborated MacGregor's finding. Although Item 2 (ii) was set in an abstract context (about two positive numbers) and Item 7 (a) was set in a real-life context (about professors and students), there was no significant influence from the contexts on the frequency of the reversal error. However, as the analysis of Proposition 10 in Chapter 8 indicates, the contexts may have contributed to the hierarchy of difficulty in the case of these questions.

Conclusions From an Investigation of Question 7

The results of the examination of the responses to the two parts of Question 7 have extended the previous research based on the professors-and-students problem. They have highlighted the cognitive challenge enshrined in distinguishing between symbols as standing for people (Objects View, or People View) and symbols as standing for numbers of people (Numbers View). Those who made the distinction were generally more experienced with algebra and/or were in Advanced level classes.

Older and more able students were the main ones to exhibit an undifferentiated conglomerate view of symbols. Options which were in terms of people in the plural were confused by some with options in terms of numbers of people, both being given a variables meaning. The most persistent pre-algebra views were that '*P*' stood for "professors" and '*S*' stood for "students". Those taught with the aid of concrete models were more inclined to take a Numbers View than those taught from a traditional textbook. Those with an Objects (or People) View were more likely to make the reversal error than those who had the Numbers View.

Section 3: The Cabbage and Pear Paradox

Unexpected outcomes were obtained from the last two options in Question 6 (a), which read as follows:

6. (a) If c + d = 10, tick ALL the meanings that c could have:

3 10 12 7.4 the number of apples in a box

an object like a cabbage an object like a pear

When tallying total test scores, a score of "1" was allocated for Q.6 (a) each time one of the first five choices was marked, and another "1" was allocated each time either of the last two choices was left unmarked. Students were considered to have omitted the question only if all seven options were unmarked.

Expectations. In the form of algebra being studied, that of generalized arithmetic, every letter-symbol must always stand for a number. To think of the symbols as representing objects (like a cabbage or a pear) rather than numbers or numbers of objects was expected to be a handicap to developing a true variables concept and to succeeding with the algebraic tasks covered by the test items (cf. Harper, 1979, pp. 151, 228; Booth, 1983, p. 153).

The paradox. The reality, as far as could be ascertained from the written responses, was exactly the opposite to what was expected. Practically every correlation between the variety of measures used in the investigation so far and responses to the last two parts of question 6 (a) had the opposite sign to the one expected, and many of them were statistically significant. The cabbage and pear options were sometimes chosen even by students from the top mathematics classes who merited high scores on the overall test. It seemed that choosing these objects options for '*c*' was not, at least for the better candidates, the hindrance to progress that it had been expected to be.

The paradox that these data appeared to enshrine is that students, who apparently regarded algebraic symbols as representing physical objects in Item 6 (a), were also able to work with the notion that algebraic symbols represented numerical variables when responding to other items. While viewing a symbol as standing for an object such as a cabbage or a pear, a student would seem to eliminate the notion that the symbol was a representation of a number which could vary. A pear is a piece of fruit and a cabbage is a vegetable, and neither of these physical things is explicitly a representation of a number, or so it would appear.

Data From Written Responses

Here was a case where the statistics based on data obtained from the written responses to the problem about the equation '$c + d = 10$' caused perplexity. As will be explained shortly, further data obtained by interviews were able to elucidate the problem and lead to a better understanding of how the students had been thinking when they selected these options in the written test situation. Before the interview information is to be presented, the reality of the paradox needs to be delineated.

Table 7·8 summarizes the frequencies of valid responses for the JCP Cabbage & Pear Scale. For this scale, a score of "1" was allotted each time an object option was selected.

Table 7·8
Frequencies of Responses on Scale JCP

Response	Scale Score	Frequency	Percentage#
reject cabbage and pear	0	365	75.7
accept cabbage only	1	9	1.9
accept pear only	1	7	1.5
accept both	2	101	21.0

Note. Responses from Years 7 to 12, using Test 3 responses for Year 7 students. $N = 482$, as there were 35 missing cases. # % of valid cases.

Just over one-fifth of the students (21.0%) chose both the cabbage and the pear option, and another 3.4% (= 1.9% + 1.5%) chose one or the other. A tendency noted by Booth (1983) was for students to interpret an algebraic letter-symbol as the initial letter of whatever the symbol represented. Those who chose only one of the objects options in Question 6 (a) did not reveal a preference for applying 'c' to "cabbage" rather than "pear" because of the initial letter, as they were fairly evenly divided between the two choices. Only two of the students who were interviewed showed the initial letter preference in Question 6 (a). As reported in Section 2 above, the initial letter also seemed to have little influence in the case of the professors-and-students problem.

Correlations. When scores on the JCP Scale were correlated with measures of levels of understanding for symbols, the statistically significant coefficients listed in Table 7·9 were obtained.

Table 7·9

Correlations Between Scores for JCP Scale and Understanding of Symbols

Scale	Level (Table 6·1)	Correlation with JCP	No.of cases	Significance	Conclusion
NBR	4	.191	461	***	Unexpected
BXBA	5	.163	444	***	Unexpected
GN	4	.145	325	**	Unexpected
VBL	5	.133	425	**	Unexpected
EQN	5	.110	422	**	Unexpected
PL	5	.098	463	*	Unexpected
FZN	4	.090	369	*	Unexpected
PV	1	- .126	462	**	Unexpected
SC2	1	- .122	425	**	Unexpected
JFL	1	- .104	427	*	Unexpected
NV	1	- .103	404	**	Unexpected
AL1	1	- .099	374	**	Unexpected
CON	1	- .096	456	*	Unexpected
CF	1	- .095	422	*	Unexpected
IG	1	- .092	456	*	Unexpected

Note. Responses from Years 7 to 12, using Test 3 responses for Year 7 students. Sorted by sizes of correlation coefficients.

*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Every one of these correlations bore a sign which was contrary to what would be expected if students who selected the cabbage and pear options had been thinking that the symbol '*c*' represented an object rather than a number which could vary. The correlation coefficients tabulated were rated significant on statistical grounds but the fact that their sizes were quite small was an indication that the problem may not have been very deep. The largest correlation was with the NBR Numbers View Scale and had a value of .191, indicating less than 4% common variance.

One further set of correlations completes the factual picture from the point of view of the written responses. Although numerically small, these are the significant correlations between achievement measures and scores on the JCP Scale, as listed in Table 7·10. The signs of the correlation coefficients, being this time all positive, were contrary to the expectation that the more students were inclined to regard symbols as standing for objects, the less likely were they to succeed with algebraic tasks. Scores on the JCP Scale always seemed to behave in ways that contradicted the pattern of outcomes (cf. Tables 6·9, 6·13, and 6B·1) from scores on the OBJ and JFL Scales which were other measures of how strongly students were inclined to take an objects view of symbols.

Table 7·10

Correlations Between JCP Scale Scores and Achievement Measures

Scale	Correlation with JCP	No.of cases	Significance	Conclusion
Corrected Test Total	.205	482	***	Unexpected
AR	.121	443	**	Unexpected
SYM	.157	425	***	Unexpected
SUBS	.179	462	***	Unexpected
EQL	.159	373	***	Unexpected
C2	.163	482	***	Unexpected

Note. Responses from Years 7 to 12, using Test 3 responses for Year 7.
*** $p \leq .001$, ** $.001 < p \leq .010$.

Frequencies. Table 7·11 shows frequencies of response patterns for Item 6 (a). There were 365 (or 75.7%) who chose their answers from only the five number options, the most common response being to choose three of these. However, 101 other students (or 21.0%) chose both of the objects options in addition to one or more number options. Only 16 students (or 3.3%) chose one object, while 2 of these did not choose any of the number options.

Question 6 (a) did not test the extent of students' ability to succeed with the more complex algebraic tasks, such as those which required the use of the variable concept. This was left to other test items. However, it is informative to look at the relationships between the patterns of options chosen in this question and the total test scores on all the other test items. Figure 7·5 displays the average scores on all test questions except the last two parts of Question 6(a) for students grouped according to their responses to those two parts of the question.

Table 7·11
Frequencies of Responses to Q.6 (a)

No.of NUMBERS (from first 5 options)	No.of OBJECTS (from last 2 options) 0	1	2	TOTAL
0	0	2	0	2
1	46	2	1	49
2	90	6	8	104
3	125	3	22	150
4	63	2	38	103
5	41	1	32	74
TOTAL	365	16	101	482

Note. Years 7 to 12, using Test 3 responses for Year 7 students. $N = 482$.

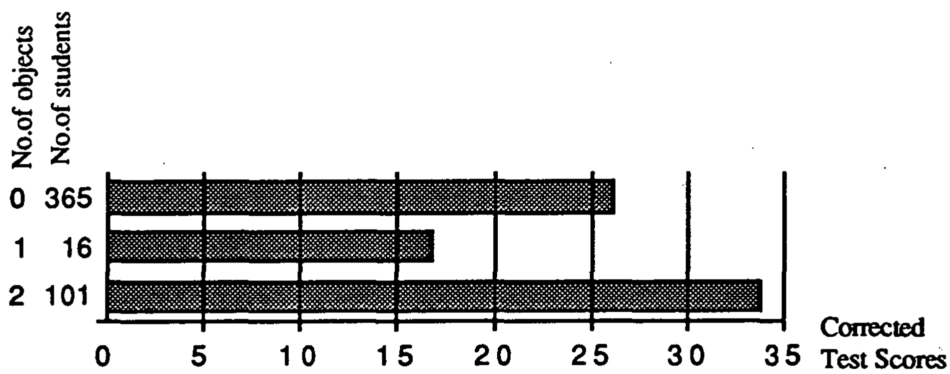


Figure 7·5: Test scores according to choices on 6a vi and vii ($N = 482$) (e.g., the bottom bar gives the corrected average score for the 101 students who chose both "an object like a cabbage" and "an object like a pear")

Figure 7·5 clearly shows that those students who chose both the cabbage and the pear options were the group with the highest average adjusted test scores. The 16 students who chose only one of the objects options averaged the least on the adjusted scores. Those who chose more of the objects options were likely to score better overall on the test than those who avoided the objects options, as was indicated by the fact that the correlation between the adjusted test score and score on the JCP Scale was highly significant ($p < .001$) and negative ($r = -.205$).

Figure 7·6 qualifies the view given by Figure 7·5 by showing that those who chose all of the first five options (all numbers) from Question 6 (a) but not the last two options (both objects) attained the best average on all the other questions in the test. They were followed closely by the group who not only chose the first five options but also chose the two last options which indicated that they accepted objects as meanings for '*c*'. Possibly these were students who had been thinking in terms of number when they selected the objects options, or they had the broader view of symbols as capable of representing not only numbers but also objects. Those who chose at least three of the five number options did better than the others, regardless of whether or not they also chose both cabbages and pears as meaningful options for '*c*', although also choosing only one of the objects options was associated with lower scores.

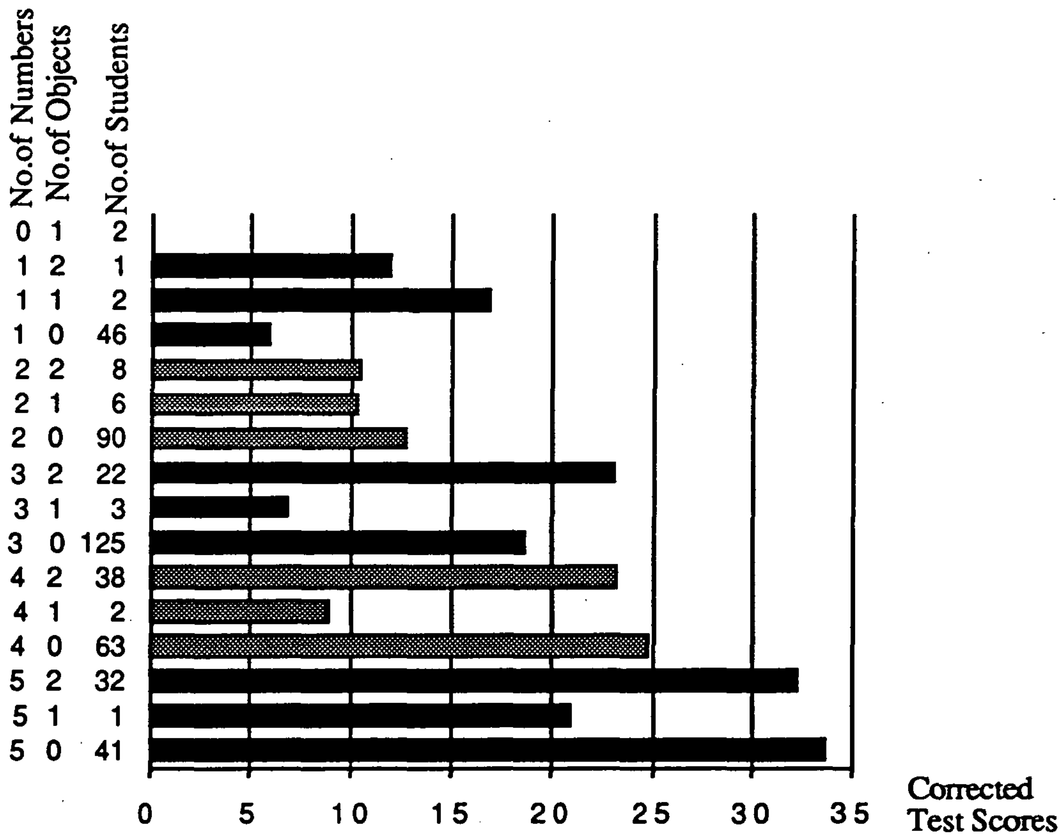


Figure 7·6. Test scores according to number choices and objects choices - numbers from 6 a i to v, - objects from 6a vi and vii ($N = 482$) (e.g. the bottom bar gives the corrected average score for the 41 students who chose all five numbers choices and neither of the objects choices)

While the choice of objects raises some questions, the better performances by those who incorporated more classes of number in their thinking illustrates the importance of generalizing algebraic statements across different aspects of number fields according to what Dienes (1963, p. 158) called the "mathematical variability principle". This principle is supported by the research data when presented in the format of Figure 7·7, which portrays the benefits associated with broadening one's range of acceptable number types.

The graphs in Figure 7·7 are for the average score on all test questions except the first five parts of Question 6 (a) for groupings of students determined by the number of options chosen from the five number options in the question. Scores out of 5 on the first five parts of Q.6 (a) were positively and significantly ($p < .001$) correlated with total scores on all the other test questions ($r = .586$), indicating that those who chose more of the five number options were likely to score better on the overall test that those who selected fewer options.

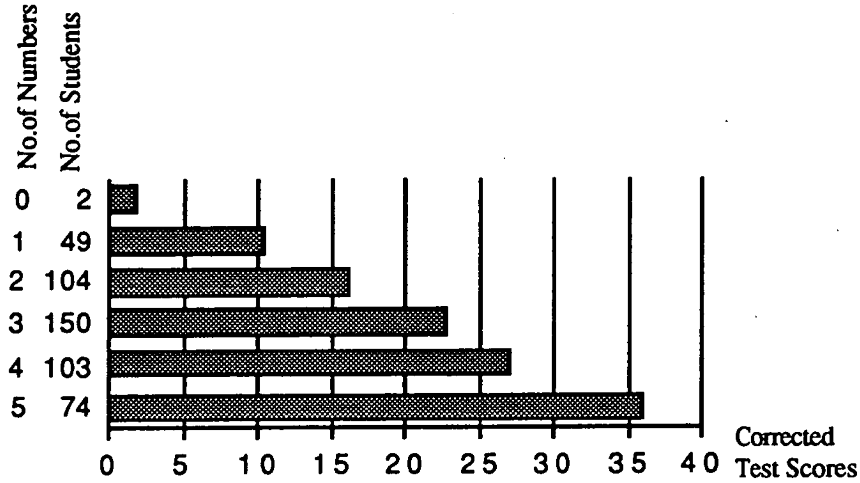


Figure 7·7. Test scores according to choices on 6a parts i to v ($N = 482$) (e.g., the bottom bar gives the corrected average score for the 74 students who chose the five number options given in 6a i to v)

Interview Data

Because the written responses were producing implications which were quite discordant with the expectations and the outcomes from the other test data, further enlightenment on the students' views about the meaning of symbols was sought from interviews. There was the possibility, for instance, that students had followed some interpretation other than the literal one of taking the letter '*c*' to stand for an object.

Interviews With Year 7 Students in 1990

The last two options of Question 6 (a) were discussed in 49 of the 173 interviews conducted in 1990. With the exception of one Year 8 student, interviewees were from Year 7. Because more than 20% of all the students in the study had incorrectly chosen one or both of these options, the interview data were important in the search for some explanation of the prevalence of these choices. Twenty of those interviewed had rejected the objects options in the written test and confirmed that rejection in the interviews. Table 7·12 summarizes the categories of thinking which were identified in interviews with the other 29 students who had accepted one or both of the cabbage and pear options.

Table 7·12

1990 Interviews: Frequency Distribution of Reasons for Choosing Cabbage & Pear

Category	Frequency	Percentage
NUMBER interpretation	8	27.6
clear OBJECT interpretation	6	20.7
vague OBJECT interpretation	5	17.2
no clear reason e.g., "guess"	10	34.5
TOTAL	29	100.0

Note. Percentages are of the 29 students interviewed.

The interviews provided the mechanism for finding out how students were viewing the symbol '*c*' when they gave a written record of acceptance that it could mean "an object like a cabbage" and/or "an object like a pear". Some of the information gathered was as follows.

1. Objects Rejected

Of the 20 interviewees who rejected the objects options, some gave specific reasons, such as the argument that an object does not have a numerical value, as in the first extract below. Others rejected the objects options intuitively, commenting that "It sounded really weird", or "I didn't understand that very much". Extract 2 shows a rejection at an intuitive level.

Interview Extract 1. (Year 7 students, 'K' and 'M', June 1990, after Test 3.)

E What do the letters mean to you at this stage, how do you think?
K I don't know.
E Perhaps question 6 makes it clearer. You've picked various numbers ... what '*c*' could be, but you didn't pick "an object like a cabbage" or anything. So what's that saying?
K Well, with an object you can't really give it a value.
E You're thinking values?
K Yes.
E What do you think, M, about what the letters mean?
M I thought that the letters were representing numbers, like '*c*' represented one number and '*d*' was another, I didn't think they could represent a table or anything.

Interview Extract 2. (Year 7 student, 'M', Dec., 1990, after Test 4.)

E Question 6 is further down the page here. '*c* + *d*' = 10'. What are some meanings that '*c*' could have?
M I think 3, 10. Because the only ones ... it has to be plus ... so it's either 3 + 7 ... or 10 + 0.
E Aha. And did you like any of these other choices like "an object like a pear"?
M No.
E Don't like that one?
M No, they didn't make any sense to me.
E What do you think the letters stand for in algebra, generally?
M Any number, you just don't know which one.
E Righto.

2. Object Given Number Implication

Of the 48 interviewees, eight were thinking in terms of numbers when they selected the objects options and were not making the mistake of equating an algebraic symbol with an object. Extracts from three of these interviews are given here and extracts from others are recorded in Appendix 7C so that these students could speak for themselves on this issue. The extracts show, however, that most of them were probably operating in the ikonic mode and were insecure in explaining their thinking, which was understandable for students who were just beginning algebra. In Extract 3 below, the student was thinking of "an object" as the number "one". The student in Extract 4 spoke in terms of weight as a way of introducing an "amount", an idea which Harper regarded as restrictive as it implies that each letter is a unique object with a unique content (1979, p. 224). Another student (in Extract 5), while speaking of the letters as standing for a number of objects, revealed a preference for using the first letters of the names of the objects, yet she avoided the confusion noted by Booth (1983, p. 268) between objects and numbers of objects.

Interview Extract 3. (Dec. 1990, after Test 4 - student 'M' is trying to clarify her thinking, which is unmistakably in terms of numbers, and ends up saying that if '*c*' = 1 cabbage, '*d*' = 9 cabbages.)

M Yes. Well '*c*' can be anything under 10 because '*c*' can equal 3 and '*d*' can equal 7, and it can be 10 and '*d*' would equal nothing, and it can't be 12, it can be 7.4 and then '*d*' can be 2.6, I think, something like that, and "the number of apples in a box" because it could be about 2 or 3 apples in a box, and then there's lots of them there, and then "an object like a cabbage", which equals one plus nine equals ten, then "an object like a pear" because that can be one pear plus nine equals ten.
E Now I'm interested in what you said then ... "an object like a cabbage, which is one"?
M Yes, well because a cabbage is just one.
E It stands for the number one do you think?
M Well, one cabbage is one cabbage.
E What would '*d*' equal in that case?

M Nine.
E Nine what?
M ..... [Pause]
E Just nine?
M Yes, it could be nine cabbages, nine whatever.

<u>Interview Extract 4</u>. (Dec.1990, after Test 4 - student 'N' provided the one detected occurrence of the idea that Harper wrote about in 1979, namely, thinking in terms of the weight of an object rather than just the object.)

E You've said it could also stand for an object like a cabbage and a pear. What did you mean by that?
N Because '*c*' could be an object like a pear that would weigh a certain amount, and plus '*d*' equals ten.

<u>Interview Extract 5</u>. (June.1990, after Test 1 - student 'Ka' is keen to have numbers of things, but is inclined to have the algebraic letter match the name of the things counted.)

Ka And it could be an object like a cabbage, and the '*c*' could stand for 2 cabbages and the '*d*' could stand for 3 doors, and an object like a pear.
E An object like, you're saying 2 objects or 3 objects, not just the object?
Ka Well it could still be, like, could be the '*c*' could stand for 1 and the '*d*' for 9, so it could be an object.
E Nine what?
Ka Nine cabbages, the '*c*' could be 1 cabbage and the '*d*' could stand for 9 doors or something.
E All right, I hear you saying that the letter stands for a number of things?
Ka Yes.

3. <u>Objects Accepted as Such</u>

Only six of the interviewees clearly claimed that '*c*' could stand for objects, as shown in the following two extracts:

<u>Interview Extract 6</u>. (June 1990, after Test 2 - student 'R' says '*c*' could stand for the object')

E You ticked 10 too?
R Yes, because it could be 10 plus 0, and the number of apples, '*c*' could stand for "the number of apples in a box", and for the objects. I ticked them because '*c*' could stand for the object.

<u>Interview Extract 7</u>. (Dec.1990, after Test 4 - student 'B' is quite clear in saying '*c*' could be "an object")

B Um ... '*c*' could have any number less than 10, or it could have 10, because '*d*' could equal 0, and it could be "the number of apples in a box", and "an object like a cabbage" and "an object like a pear", because '*c*' can have the value of anything.
E '*c*' could just be anything?

B Yes.
E What would you mean by '*c*' could be "an object like a pear"? What are you thinking of?
B Well, just '*c*' could be an object, just like that.

4. Objects Accepted for Other "Reasons"

Another 15 students indicated that they believed that '*c*' could stand for an object for a variety of other reasons, or for no clear reason at all. One said, rather vaguely, that it could be a pronumeral. A second said '*c*' could be a pear "because it's smaller". A third accepted the letters as objects even though he acknowledged that they then did not have meaning in the given equation. Another simply said letters could stand for anything. The fifth student confirmed that he was thinking in terms of the first letter, as recorded in Extract 8 below. Extract 9 is from an interview with a Year 8 student who had scored 55 on the test, a score which was outstanding, since only 16 of the 517 students from Years 7 to 12 scored 55 or better. She had succeeded with the harder test questions requiring her to apply the abstract concept of a variable, yet had selected the cabbage and pear options in Question 6 (a). Of the remaining nine students, there were six who ticked the objects options but were unsure of why they had done that, saying they just guessed an answer or that "I can't remember", as in Extract 10. Another two students were not sure the options were part of the question during Test 4 and one of these opted in the interview for accepting objects. Another had ticked objects in Test 3 but cancelled that view in favour of only numbers when interviewed. Such interviews highlighted the difficulty of identifying cognitive processes when an intuitive mode was probably being used. Extract 11 exemplifies a case of an intelligent student who could have been intuitively imagining that algebraic symbols can stand for numbers and for other possibilities as well, but could not logically make sense out of this suspicion in the context of the question.

Interview Extract 8. (June 1990, after Test 1 - Student 'R' was the only one, apart from the student quoted in Extract 4 above, who spoke of using the first letter of words. This was a difficulty discussed in Booth (1983).)

E R, an object like a cabbage, that '*c*' could be that?
R Yes, because '*c*' could stand for cabbage.
E Oh, just the first letter?
R Yes.

Interview Extract 9. (Dec.1990 - a Year 8 student, 'M', who scored 55 on the test yet she was confused about letters as representing numbers)

E Could it stand for "an object like a pear"?
M Yes.
E What would that mean then?
M It would be the pear plus '*d*' equals ten.

E What could '*d*' be in that case?
M It would have to be 10 or 9
E So if '*c*' were an object like a pear, you mean just a piece of fruit?
M Yes.
E plus '*d*' equals 10. You didn't tick that last June. You're changing your mind on that?
M Yes, probably.

Interview Extract 10. (June 6, 1990, after first test - student 'R' is unable to explain his decision-making process for the last few options in the question.)

E What about 10? What did you have in mind there?
R '*c*' could be 10, because '*d*' could be zero.
E Right. Now, you ticked "the number of apples in a box". What did you mean by that, R?
R I'm not sure. I forget.
E You ticked "an object like a cabbage", what did you have in mind for that?
R I don't know. I forget.

Interview Extract 11. (Dec 1990, after Test 4, a rather lengthy extract from a Year 7 student, 'M', who wavered in his acceptance of objects, yet accepted them while acknowledging that they did not have any meaning in the given equation.)

M Yes. "An object like a cabbage". That isn't a number either, and neither is "an object like a pear".
E Well you ticked them off last week.
M Yes. So '*c*' plus a pear, you can't add the pear, so if '*c*' was 10, it would still equal 10.
E Are you saying today [Pause] Are you changing your mind or do you still want it?
M No, because it's got there "an object like a cabbage"... if you added 10 plus an object like a cabbage, so the answer would still be 10.
E Can '*c*' stand for an object like a cabbage?
M Maybe...
E Or can '*c*' stand for an object like a pear?
M Oh, I probably went wrong there
E No. I'm just asking what you think. That's what you ticked off. Are you saying that ... the '*d*' is 10, so this cabbage doesn't matter. Is that what you're trying to say?
M Yes.
E I was just trying to interpret what you said a minute ago. Would you like to just explain it once more to get it clear?
M Well, so if '*c*' was a cabbage, and '*d*' was 10, if you added them together you'd get 10, and you'd also get a cabbage, but a cabbage isn't a number.
E So it wouldn't count in the equation?
M No.

Outcomes From 1990 Interviews

The 1990 interviews attested that 29 out of the 49 who discussed the last two parts of Question 6 (a) had chosen objects. Of these, 8 (or 27.6 percent of the 29) interpreted their response as some sort of number. Taking account of a possible numbers interpretation for options (vi) and (vii) offers some explanation for the success on the abstract test items by students who had selected these options. The question of their success otherwise is perplexing in the light of Harper's logical view (1979, p. 151) that thinking of the symbols as standing for objects severely restricts the possibility of variation, and consequently should restrict one's ability to work with the true concept of a variable. This perplexity is modified by the possibility that some of these beginning algebra students were starting to sense that symbols might be used to represent entities other than numbers. Such a possibility certainly becomes real in more advanced algebra.

Interviews With Years 10 and 12 Advanced Students in 1991

Background to 1991 Interviews. Analysis of the responses obtained in 1990 to Item 6 (a) from Years 9 to 12 students who had been graded into Advanced mathematics streams revealed that about 30% of them selected "an object like a cabbage" and "an object like a pear" as options for the meaning of '*c*' in '$c + d = 10$'. These Advanced students generally gained high scores on the overall test and many of the same students had succeeded with questions such as Items 10, 12, 13, and 15, which required the application of the concept of a numerical variable.

To shed more light on this paradoxical outcome, in July 1991 the researcher returned to the three Sydney schools involved in the research project. A short test headed "New Test 1991" was assembled, consisting of Items 6 (a), 7, and 15 from the test instrument, as shown in Appendix 3R. This test was administered to the Year 12 Three Unit Mathematics classes in Schools C and D, and to a class of Year 10 Advanced Mathematics in each of the Schools B, C, and D. Students in these five classes had contributed to the emergence of the paradox as they had responded to the complete test instrument in 1990 and they belonged to the Advanced stream. A total of 115 students were tested and 52 of these were interviewed.

Table 7D·1 in Appendix 7D records the numbers interviewed for each class and the frequencies of their responses to the last two subparts of Item 6 (a).

Students who Chose the Objects Options

Of the 115 students tested, 32 (or 27.8%) chose both the cabbage and the pear options, while none chose just one of these options. All but 4 of those who chose the objects options were interviewed. A resolution of the paradox for most of these students emerged from these interviews. As shown in Table 7·13, it was found that the number "one" was associated with the options by 11 of these students, while another 10 said that they were thinking in terms of one or more objects. Only 2 clearly thought of '*c*' as possibly representing an object, while 5 (or 17.9%) were unable to clarify their thinking about '*c*'.

Table 7·13

<u>1991 Interviews: Frequency Distribution of Reasons for Choosing Cabbage & Pear</u>

Category	Yr.12 3U School C	Yr.12 3U School D	Yr.10 Adv. School C	Yr.10 Adv. School D	Yr.10 Adv. School B	TOTALS
Number who chose objects	9	2	7	11	3	32
Interviewed	8	2	7	8	3	28
NUMBER interpretation c = 1	1 (12.5%)	2 (100%)	3 (42.9%)	5 (62.5%)	0	11 (39.3%)
NUMBER interpretation c ≥ 1	3 (37.5%)	0	3 (42.9%)	2 (25.0%)	2 (66.7%)	10 (35.7%)
OBJECT interpretation	1 (12.5%)	0	0	1 (12.5%)	0	2 (7.1%)
NOT CLEAR	3 (37.5%)	0	1 (14.3%)	0	1 (33.3%)	5 (17.9%)

<u>Note</u>. $n = 28$. Percentages are of those interviewed.

Three-quarters (21 students) of this group of interviewees had chosen the objects options with a numerical meaning in mind rather than the literal object meaning. Two sample extracts from interviews with this subgroup of students are now reported. In Extract 12 the student had taken "an object" to mean "one", and in Extract 13 the student took the meaning to be any number of objects.

<u>Interview Extract 12</u> (Year 12 student , 'S', School C - one object)

S An object - well, that's still one.
E If you said "an object like a pear" - back to the equation, what would the equation be saying?
S It might just say ... um ... Like you might have, say, one object, because it says "an object", so that's one. That might be like plus 9 ... um, I don't know ... um, I'm not really sure actually. That would be like something ... 9 ... maybe 9 other objects or something like that.

<u>Interview Extract 13</u> (Year 10 student, 'A', School B - number of objects)

A ... Then I think it could be "the number of apples in a box". It doesn't say any specific thing, and the same with the cabbage and the pear.
E Just explain a bit more about those last two.
A I thought, because it ... for algebra it could be names of ... sort of any object - apples or cabbage. Could be '*c*' number of apples or number of

cabbages or '*c*' number of pears ... numbers.

E Numbers?

A Yes. '*c*' amount.

A minority of the Advanced students (7 students, or 25.0% of those who had chosen the objects options and were interviewed) were confused about the fact that, for high school algebra, algebraic symbols always represented numbers. One Year 12 student and one Year 10 student said that the letter could simply stand for an object. The other five students in this minority group seemed to be confused about the meanings algebraic symbols could take. Extract 14 is an example of a student who regards the symbol as representing something other than a number and, in Extract 15, we find a student confused about the meaning of '*c*'.

<u>Interview Extract 14</u> (Year 12 student , 'M', School C - "an object" means something which is not a number)

M It could have been "an object like a cabbage" or "an object like a pear".

E Could you just tell me what you're thinking about when you say those last couple there?

M Um .. Instead of ... um, I know it's like ... it's not numbers ...um, it did say "all meanings" ... It could be like that.

E Not really numbers, you're thinking?

M No.

<u>Interview Extract 15</u> (Year 12 student , 'K', School C - confusion about the meaning of '*c*')

K I just figured '$c + d = 10$' ... '*c*' is a variable, so really it could be anything. You have the option of making it into whatever the question asks to make it more specific.

E Alright. Would you like to comment on some of the choices?

K Um .. Well, like usually of course I just assume it would be a number value like 3, 10, 12 and "the number of apples in a box" - that's ... I guess an object - that seems to be it, but because it's a variable I figure that it could still be something like that. Like if you put '6*c*' then it would be like 6 cabbages or something like that.

E If '6c' were 6 cabbages what would '*c*' be?

K An object like a cabbage.

E If '6*c*' equals '6 cabbages', '*c*' would be an object like a cabbage?

K Mm.

E Or a pear? It doesn't really matter?

K Mm.

E Um ... So you make ... I'm just trying to get inside your head. See, you said at first you assume it's a number value but it could be anything.

K Mm ... Because it's not specific in the question. Like it equals 10 but ... Oh! If it equals 10 then maybe it does have to be a number value. But I think it's ambiguous because it doesn't specifically say.

E Alright.

It is true that the paradox remained of how all but one of this minority of interviewees could nevertheless succeed (Table 7D-3 in Appendix 7D) with at least one part of Item 15, a question which required the sophisticated notion of a numerical variable. Perhaps they somehow compartmentalized their thinking and worked with the variable notion when it was appropriate. Alternatively, they had an intuitive sense that letters can stand for anything.

Students who did not Choose the Objects Options

Interviews were conducted with 24 of the 83 students who did not choose the objects options, and sample interview responses from this subgroup are given in Interview Extracts 16, 17, and 18. To the Year 12 student in Extract 16, "an object" has no significance in the equation '$c + d = 10$'. For the Year 10 student in Extract 17, '*c*' was not really a "thing". Extract 18 draws attention to a finding from the interviews that provided some extension of our understanding of the way students might think about algebraic symbols: Because "an object" was singular and not plural was sufficient reason for some to reject such an option. This finding underlined the importance these students placed on leaving the interpretation of '*c*' open to values of more than one so that it would be truly a variable.

Interview Extract 16 (Yr.12 student, 'N', School C - objects had no significance in the equation)

N Well, '*c*' ... I thought '*c*' could be any number, because '*c*' and '*d*' are both, like, numerals .. um ... like letters that can have numerals. I mean you can't ... "the number of apples in a box" ... that ... 'c' could stand for that.
E Yes?
N But "an object like a pear". I mean that would just have no significance in the equation.

Interview Extract 17 (Yr.10 student, 'D', School C - '*c*' is not really a thing)

D Well, because it's algebra you know that ... um ... because it's an equation that equals a number, that '*c*' must equal a certain number ... because it can be any number or number of objects.
E What about the last two choices?
D Um... Well, it can't really equal an object. It has to equal a number. It's not really a thing.
E Not really a thing?
D No!

Interview Extract 18 (Yr.10 student, 'J', School D - '*c*' always plural)

J What 'c' could be? Basically any number because ... um ... 'd' can be a negative number so if '*c*' is above 10 then '*d*' could reduce its overall answer.
E Good.
J Um.. I didn't put "an object like a cabbage" or "an object like a pear" because ... um ... '*c*' I felt could always be a plural ... um ... or more than one. Um ... so ... I didn't put ... an object means one to me and so I didn't write them. I just wrote "the number of apples in a box" because ... um ... that's a number of things ... um, more than one.
E I follow what you're saying.

The students in this group all indicated that they expected '*c*' in the equation '$c + d = 10$' to be a numerical variable. They rejected the objects options for one of two reasons: 18 (or 75.0 percent) of them indicated that they thought that the options did not allow '*c*' to represent a number and the remaining 6 students said that each option was expressed in the singular rather than the plural and did not allow for any variation. Table 7·14 summarizes these findings.

Table 7·14
1991 Interviews: Frequency Distribution of Reasons for not Choosing Objects

Category	Yr.12 3U School C	Yr.12 3U School D	Yr.10 Adv. School C	Yr.10 Adv. School D	Yr.10 Adv. School B	TOTALS
Number rejecting objects	16	9	24	14	20	83
Interviewed	5	7	7	3	2	24
OBJECTS are NOT numbers	4 (80.0%)	7 (100%)	5 (71.4%)	0	2 (100%)	18 (75.0%)
'c' needs to be PLURAL	1 (20.0%)	0	2 (28.6%)	3 (100%)	0	6 (25.0%)

Note. $n = 28$. Percentages are of those interviewed.

Data Regarding Variable Notion

Cross-tabulations of responses by the 115 students interviewed in 1991 to Items 6 (a) and 15 are presented in Tables 7D·2 and 7D·3 of Appendix 7D. These data show that a subgroup of these students had contributed to the paradox of being able to work with variables and yet accept that an algebraic symbol could represent "an object", according to their written test responses. All but 8 students (or 7.0 percent) succeeded with at least one part of Item 15, showing that they had at least some ability in working with the concept of a numerical variable. Almost half (49.5 percent) of the interviewees showed considerable skill with this concept by successfully answering three or even four parts of Item 15. The number of correct answers attained in Item 15 was not statistically related to whether or not the objects options were chosen in Item 6 (a), an outcome which could have been expected once the interviews had revealed that most of those who had selected "an object" as a meaning for a symbol had actually been thinking in terms of numbers at the same time.

Comparisons of 1990 and 1991 Responses From six students

Data were available on test responses in 1990 and 1991 from six Advanced students who, in their 1991 interviews, had either spoken of '*c*' as representing an object or were uncertain of their interpretation of '*c*'. Table 7·15 compares their responses in 1991 with those of 1990 for the items used in the short 1991 test.

Table 7·15
Comparison of Responses for Six Advanced Students

Student	1990			1991			Comments	
	Q.6a	Q.7	Q.15	Q.6a	Q.7	Q.15	on Q.6a	on Qq.7,15
A	3,10, 7.4	correct	2 parts correct	all	correct	all correct	moved to objects	improved
B	3,10,12 7.4	correct	all correct	all	correct	all correct	moved to objects	no change
C	3, 7.4	no. or people; reversal	2 parts correct	all	no. or people;	3 parts correct	moved to objects	improved
D	3 only	no.; reversal	2 parts correct	all	people	nil correct	moved to objects	regressed
E	all	people sing.or plural	3 parts correct	all	no. or people sing.or plural	2 parts correct	objects both tests	regressed
F	all	no.or people	3 parts correct	all	no. or people	3 parts correct	objects both tests	no change

Of the students listed in Table 7·15, only Student 'A' said in the 1991 interviews that she thought the letters could stand for objects as such ("I know it's like ... it's not numbers, but it did say all meanings, so it could be like that"). She improved on Questions 7 and 15 from 1990 to 1991 and had chosen the objects in only the 1991 test. The other five students listed appeared to be confused about why they chose the objects options and could have been moving towards a broad understanding of algebraic symbols as capable of representing numbers and also other things, such as objects. Three of them (Students 'B', 'C', and 'D') chose the objects in only the 1991 test. Student 'B' (whose comment was " '*c*' is a variable, so really it could be anything. You have the option of making it into whatever the question asks, to make it more specific") continued to have both Questions 7 and 15 correct from 1990 to 1991. Student 'C' (who explained " '*c*' can represent anything so I just said 'a number of apples in a box' and everything, because '*c*' can stand for anything") improved somewhat on Questions 7 and 15, and Student 'D' (who could not make up his mind in the interview, favouring numbers in Question 6, but people in Question 7) regressed. Students 'E' and 'F' had chosen objects in both tests. Student 'E' (who said " '*c*' represents any number or any object or whatever") registered a decline in performance on Question 15 as well as moving towards an all-inclusive view of the symbols in Question 7, while Student 'F' (who said "Well, 'a pear' - it could equal that. You could put '*x*' and it could mean anything. Like it's a variable") gave the same answers to Questions 7 and 15 in both tests.

The changes recorded in Table 7·15 did not show a pattern to support an argument either for or against any advantage from the acceptance of the two options involving objects.

Relation to Previous Research

Harper (1979), Rosnick and Clement (1980), Küchemann (1980), Booth (1983) and MacGregor (1989) all discussed the implications of regarding the letters in algebra as objects. The test questions they used, either in written form or in interview mode, did not directly measure whether or not students regarded algebraic symbols as representing objects. The information they obtained about students viewing letters as objects or numbers was derived mainly from inferences based on written answers or from interviews. Item 6 (a) was a valuable means of identifying a problematic aspect of students' understanding of symbols, but it was not self-sufficient in the written testing mode. Interviews were invaluable in the process of trying to clarify the ways students were thinking, since students reported having a variety of views when selecting the options which included the idea of a letter standing for "an object". The previously-mentioned researchers did not make use of any questions of the same type as Item 6 (a). Better items might well be devised for more work in this area.

Harper (1979) and Collis (1975a) used several items to measure the extent students understood the concept of algebraic symbols representing numerical variables, and some of these measures were applied in the present research to achieve the same end. However, neither researcher obtained data about how students interpreted the suggestion that an algebraic symbol might possibly represent "an object". Hence, they did not investigate the paradox. Küchemann (1980) and Booth (1983) used only one or two items which they regarded as measures of the notion of a variable, and they did not discover the paradox. Rosnick and Clement (1980) and MacGregor (1989) did not measure the variable concept at all.

A new factor in the thinking of students was identified in the 1991 interviews. This was the importance that several had placed on the fact that the option "an object" was singular and was, for that reason, unacceptable as a meaning for an algebraic symbol in the context of the equation '$c + d = 10$'. As far as can be ascertained, the significance of singular or plural in such a context has not been identified in previous research papers. This aspect was also noted above when discussing responses to Question 7.

Conclusions From the Cabbage and Pear Paradox

1. The wisdom of using interviews to clarify communication between subjects and researcher was confirmed clearly in this investigation. Item 6 (a) was the outstanding example where the written responses were ambiguous in their role of communication. The limitations of the interviews themselves were also brought to notice in the cases of students who had made choices while working in the ikonic or intuitive mode (Biggs & Collis, 1991). These students did not communicate clearly how they had arrived at their decisions because an analysis of one's cognitive processes is not likely in the intuitive mode. In interview extracts given in this section, examples may be found of students who were very likely operating intuitively and so gave incomplete communication about their views of algebraic symbols. Extracts numbered 2, 3, 10, 11, 12, and 15 give some indication of the difficulty some students had in explaining their interpretation of symbols in Question 6 (a). Other techniques need to be explored to identify clearly how such students were really thinking.

2. In most of the 1991 interviews, communication was clear. It was confirmed that the majority of the Advanced students saw that the symbols they were using in algebra indeed represented numbers. Some were even thinking in terms of number when they allowed options such "an object like a pear" as a meaning for a symbol. During the 1990 interviews, many of the beginning Year 7 students were unable to speak clearly about their understanding of algebraic symbols, probably because they were responding intuitively. However, more than one-quarter of those who were interviewed after they had selected one or both of the objects options explained in various ways that they were really thinking in terms of numbers rather than objects as the meaning for the symbol 'c'.

3. Interviews brought to light the fact that some students rejected the cabbage and pear options, believing that 'c' should not be limited to being associated with one thing. They applied the requirement, perhaps intuitively, that symbols should be able to represent variable numbers and found the singular notion suggested by "an object" unacceptable. In response to this finding, Question 6 (a) might be modified for future use by including a bracket of options such as:

an object like a pear   objects like pears   the number of pears in a bag.

4. There were only a minority of able mathematicians who could work with the

sophisticated concept of a numerical variable while being unsure that the algebraic symbols in the questions stood for numbers. Possibly some of these students were advancing towards some broad generalization whereby alphabetic symbols could represent more than just numbers but this was not entirely clear, even from the interviews, as explained in the first conclusion.

5. Overall, the data gave some support for seeing wisdom in the advice given in the National Statement on Mathematics for Australian Schools that, for beginning secondary students at least, "It is essential that students understand that letters stand for numbers, not for objects, and that clear distinctions are made between the use of letters in algebra and other uses" (Australian Education Council, 1990, p. 194). The research data implied a modification of this advice because it seemed possible that some students had at least begun to consider that symbols could stand for "variables" in a wider sense than simply numbers, as is explained in the next conclusion.

6. An alternative explanation for the paradox could be that some students had chosen the cabbage and pear options because they were intuitively thinking beyond numerical variables to consider the possibility that algebraic symbols could stand for something other than numbers. In the courses of study in mathematics followed by students in the early years of secondary school - Years 7 and 8 - algebraic symbols always represent numbers. However, in the senior years - Years 11 and 12 - sometimes alphabetic symbols are used to stand for non-numerical entities. Some teachers of the middle stage - Years 9 and 10 - may, perhaps, use symbols to represent more than simply numbers. The following are examples of such usage, taken from State-wide Higher School Certificate (H.S.C.) examination papers for senior students:

<u>Test Items from</u> <u>Tasmanian 1988 H.S.C. Algebra and Geometry Level III</u>
Question 5: "The line l_1 is defined by the equation $2x - y + 2 = 0$ ..."
Question 9 (b): "A, B and C are the non-zero matrices ..."
Question 10 (a) (i): "If A and B are mappings of the plane ..."

<u>Test Items from</u> <u>N.S.W. 1989 H.S.C. 2/3 Unit Mathematics</u>
Question 2: "The line l passes through ..."
Question 6 (b) (i): "Sketch the parabola P, whose focus is ..."

<u>Test Items from</u> <u>N.S.W. 1989 H.S.C. 4 Unit Mathematics</u>
Question 3.(a) (i): "Sketch the ellipse E ..."
Question 5 (b): "Consider a slice S of the pyramid ...", and
"Suppose ... that n identical pyramids ... are arranged ... to form a solid C."

These are examples in which alphabetic symbols are used to represent a variety of "objects" rather than numbers: lines, matrices, mappings, conic sections, and even a slice of a solid pyramid, and another solid made of a collection of pyramids. The last two are cases of using a letter to stand for solid, three-dimensional objects. It is possible that some students could extend the meaning of symbols in algebra beyond the limitation of standing for numerical variables to meanings which allow them greater powers for symbolizing generalizations. In the given research test item, however, it was stated that '$c + d = 10$' and the symbols were, therefore, intended to be limited to meanings within an arithmetic equation and, if so, could represent only numerical variables. Such a limitation to the range of meanings seemed to be overlooked by some who could have been intuitively envisioning more general possibilities for algebraic symbols.

7. Further research on this issue seems advisable.

Review and Forecast

This chapter reported analyses based for the first time on responses from subgroups of participating students. The research data have been used in a more dynamic way than in previous chapters. Consideration has been given to differences between responses of groups such as those in different Years (or grades) at school, or in classes categorized differently according to their general mathematical ability. Changes in understanding have been examined for Year 7 students (and for six more advanced students) who were tested more than once.

The question of whether or not students viewed algebraic symbols as representing objects or numbers of objects was pursued in relation to Questions 7 and 6 (a). It was found difficult, even with the aid of interviews, to identify the thought processes of some students with regard to their views of symbols. Devising suitable test items is one challenge. Another is to create mechanisms for probing the thoughts of those operating intuitively.

Possible interactions of student views with aspects of performance in algebra were investigated. Conclusions were detailed at the end of each section.

Chapter 8 presents material on applications of the research data to investigations of propositions about hierarchies of cognitive difficulties derived from understandings based on psychological principles.

CHAPTER 8

THIRD STUDY: A STUDY OF HIERARCHIES: CHANGES IN LEVELS OF UNDERSTANDING OF SYMBOLS

Overview

The uniting reference point throughout the chapter is the level of understanding exhibited by students for the meaning and use of symbols in the algebra of generalized arithmetic. The challenges involved and the changes in levels of understanding are analysed in terms of theories of learning as well as by statistical procedures.

The chapter is divided into three sections.

The first section describes a statistical method devised by the writer for identifying hierarchies of cognitive difficulty and, consequently, of possible sequences of learning.

The second section discusses the hierarchical levels of understanding of the meanings of algebraic symbols as detailed in Chapter 6. These levels are examined within this section in the following four stages:

Part 1. Generalizations about levels of understanding of algebraic symbols are drawn from an overview of the trends revealed by comparison of responses from different Year levels and the progressive test responses of the Year 7 students, keeping in mind the different mathematics ability levels of the various classes. The trends are summarized graphically for each of the five levels described in Chapter 6.

Part 2. The question of whether or not psychological analyses point to the existence of hierarchical levels of difficulty is examined. Psychological theories, outlined in Chapter 2, are applied to assess the cognitive processes involved in responding to selected test items with different levels of understanding of the meaning of symbols.

Part 3. This part examines the question of whether or not statistical analyses point to the existence of hierarchical levels of difficulty by using the available measures of the degree of understanding of algebraic symbols as numerical variables.

Part 4. Comparisons are made between particular outcomes of this study and those of Harper (1979) and of Küchemann (1980).

The third section analyses nine propositions about hierarchies of learning in early algebra. These propositions appeared logical in terms of psychological reasoning. The degree of empirical support for each of these propositions is discussed in terms of statistical analyses of the relevant research data.

Section 1: Statistical Method for Analysing Hierarchies of Learning

To find evidence for the existence of hierarchies of cognitive difficulty and possible sequences of learning for the concepts and skills tested, batteries of cross-tabulations of student responses were examined. For the subgroup composed of only the Year 7 students who were tested more than once, cross-tabulations of their responses were studied for each testing. Similar cross-tabulations were considered for the cohort with the largest number of students (517), taking the Year 7 students' responses to Test 3. These cross-tabulations presented the matrices which showed the frequencies of scores for two variables at a time. Three examples are offered in the discussion which follows, and bar charts, equivalent to Venn Diagrams, are displayed in Figure 8·1 to clarify the principles underlying the method developed by the investigator for analysing hierarchies of learning.

When scrutinizing any two variables A and B, interest focused on the frequencies for attaining the highest score possible for each. The term ***high*** is used throughout to designate the attainment of highest possible scores: correct responses for variables in the form of individual items, or correct responses for each item in the appropriate scale for those variables in the form of scale scores. Thus, "high" means 100% correct on an item or group of items. Two percentages were calculated:

First, "%AB", which was the percentage of those who scored high on B who also scored high on A. This gave a measure of A as a *prelude* (explained below) to B.

Second, "%BA", which was the percentage of those who were fully correct (scored "high") on A who also were completely correct (scored "high") on B. This gave a measure of B as a prelude to A.

Table 8·1 shows the relevant statistics for scores on Question 5 (variable 'A') and Question 12 (variable 'B').

Table 8·1

<u>Evidence for Success on Variable A as Prelude to Success on Variable B</u>

Variables		Frequencies			Percentages		Ratio	Correlations	
A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%AB}{\%BA}$	*r*	*p*
Q.5	Q.12	111	316	123	90.2	35.1	2.57	.402	***

<u>Note</u>. $\%AB = \frac{\text{No.high on A \& B}}{\text{No.high on B}} \times \frac{100}{1}$. $\%BA = \frac{\text{No.high on A \& B}}{\text{No.high on A}} \times \frac{100}{1}$.
$N = 454$ (i.e., 517, less 63 missing values). [Loevinger H = 0.7321] *** $p \leq .001$.

There were 316 students who scored top marks (1) on variable A, 123 who scored top marks (4) on variable B, and 111 students who scored top marks on both variables. Thus,

$$\%AB = \frac{111}{123} \times \frac{100}{1} = 90.2\%, \text{ and } \%BA = \frac{111}{316} \times \frac{100}{1} = 35.1\%.$$

Here, %AB is 2.57 times %BA, giving strong evidence for the view that success on variable A is a prelude to success on variable B. Over 90% of those who succeeded on B had high scores on A, leaving less than 10% of those who scored high on B to score low on A. In contrast, less than 36% of those who succeeded on A also scored high on B, the other 64% being spread across lower scores on B.

The writer set the following criteria for accepting the view that success on variable A was a prelude to success on variable B:

Criterion 1 %AB was greater than 70%,

Criterion 2 %AB was at least twice %BA, and

Criterion 3 the correlation coefficient between variables A and B was positive and statistically significant ($p \leq .050$)

Criteria for accepting the view that success on variable B was a prelude to success on variable A were similar. (See Table 8·3.)

Criterion 1 allowed a measured margin for exceptions to the rule (such as those due to sampling error) by the percentages imposed and did not demand 100% of cases with full marks on one variable being required for having everything correct on the other variable. Hence, the word "prerequisite" was not appropriate as it implied 100% conformity to a rule, and the word "prelude" was chosen to convey the notion of some leniency within the rule. The use of 70% as a boundary condition was thought to be reasonable and was not the result of "judicious manipulating of the cutting points" (Nie et al, 1975, p. 533), a process available when applying the Guttmann scalogram approach (Küchemann, 1984, p. 118). In fact, previous researchers had used percentages of this order to distinguish success from failure: Collis (1975b), used 60% (p. 117) or 70% (p. 133) and the CSMS study used about two-thirds (Hart, 1981b, p. 215). The Guttmann approach to hierarchy analysis was not applicable as it was relevant to ordering "three or more items" (Nie et al., 1975, p. 529) and the intention here was to examine the order of difficulty of cognitive levels of tasks, or groups of tasks, taken in pairs.

Criterion 2 was imposed to identify which variable was a prelude to the other. If both %AB and %BA were greater than 70%, a decision could not be made regarding the existence of a hierarchy. A ratio of two-to-one between the percentages was considered sufficiently large to be discriminating. Criterion 2 simply requires that one facility level be at least twice the other (excluding missing values), as the ratio used equals the ratio of frequencies for "High" on one variable to "High" on the other.

Criteria 1 and 2 considered only the numbers of students who were 100% correct on either variable, simplifying the comparison of success rates for different items or groups of items (in cases of scale score variables). By focusing on achievers, comparisons of cognitive difficulty could be validly made even if facility levels were low, as in Table 8·3. This directness eliminates possible problems (Küchemann, 1980, Chapter 9) inherent in using an association measure such as the Loevinger H coefficient which is defined in terms of those who "pass the harder item and fail the easier" (Küchemann, 1980, p. 94). The H value approaches unity for greater differences in facility between pairs of items and so tends to emphasize the most trivial relationships - "the instances when children who could answer an extremely difficult item also coped with an easy one" (Küchemann, 1980, p. 90). For interest, the H values are shown in Tables 8·1 to 8·3. In the Quinlan approach, the overlap of successes and the facility ratios are examined in the first two criteria and the full range of scores on each variable is incorporated in the separate third criterion which requires the Pearson Correlation Coefficient to be statistically significant. This requirement ensures that the identified hierarchy matches the overall pattern of relationship between the two variables. Otherwise, criteria 1 and 2 might hold as the outcome of spurious statistics dependent upon a set of percentages about frequencies for high achievers that did not reflect the general learning trends. The latter situation was possible, particularly when there were very few high achievers for either variable.

When a variable is a scale score, only those who scored 100% on all the items in the scale were considered in criteria 1 and 2. All scales were composed of at least two items, the items in any scale forming a cognitive unity (pp. 151 - 152 and summaries pp. 171 - 174). Identifying success on one scale or item as a prelude to success on another indicates, for the particular student group, the existence of a hierarchy of difficulty for the cognitive processes measured.

All three of the criteria were satisfied in support of the contention that success on Question 5 was a prelude to success on Question 12, as shown in Table 8·1. This outcome indicates that a hierarchical order of difficulty exists, with Question 5 being rated as less challenging cognitively than Question 12: writing the correct answer in symbols (viz., '$p + r$') for Question 5 was cognitively less difficult than working with the variable concept to compare the possible values of two algebraic expressions (viz., '$2n$' and '$n + 2$') in Question 12.

As an example of a failure to meet the criteria, success on Question 5 was tested as a prelude to success on Question 10 (which tested ability to use the variable concept to compare the values '$t + t$' and '$t + 4$'). It was found that, as Table 8·2 shows, the first and third criteria are matched but, because of the failure to meet the second criterion, the data were not regarded as supplying sufficient evidence to support the claim that success on Question 5 was a prelude to success on Question 10.

Table 8·2

Evidence for Deciding if Variable A is Prelude to Variable B

Variables		Frequencies			Percentages		Ratio	Correlations	
A	B	High A & B	High A	High B	%AB	%BA	%AB / %BA	r	p
Q.5	Q.10	149	316	173	86.1	47.2	1.83∗	.366	***

Note. $N = 451$ (i.e., 517, less 66 missing values). ∗ ratio too low (less than 2). *** $p \leq .001$. [Loevinger H = 0.5365]

A third example is now given for a case in which the statistics are displayed with Variable B as the prelude variable. In Table 8·3, the data from Year 7 students in Test 4 show that success on Question 10 was a prelude to success on Question 12. The ratio of %BA to %AB is greater than 2. All three criteria are met, leading to the conclusion that success on Variable B was a prelude to success on Variable A.

Table 8·3

Evidence for Success on Variable B as Prelude to Success on Variable A

Variables		Frequencies			Percentages		Ratio	Correlations	
A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%BA}{\%AB}$	r	p
Q.12	Q.10	20	22	58	34.5	90.9	2.64	.672	***

Note. Test 4 for Year 7 students. $n = 175$ (i.e., 186, less 11 missing values).
*** $p \leq .001$. [Loevinger H = 0.8640]

For conciseness, the symbol '>' is used to mean "was a prelude to", so that the statement 'A > B' means high success on variable A was a prelude to high success on variable B. The term "prelude" denotes an order of difficulty and does not impose a sequence of learning whereby success with one variable is the pathway to success with another.

To clarify the method of analysis, the three examples used in the explanation are summarized by the percentage bar charts in Figure 8·1.

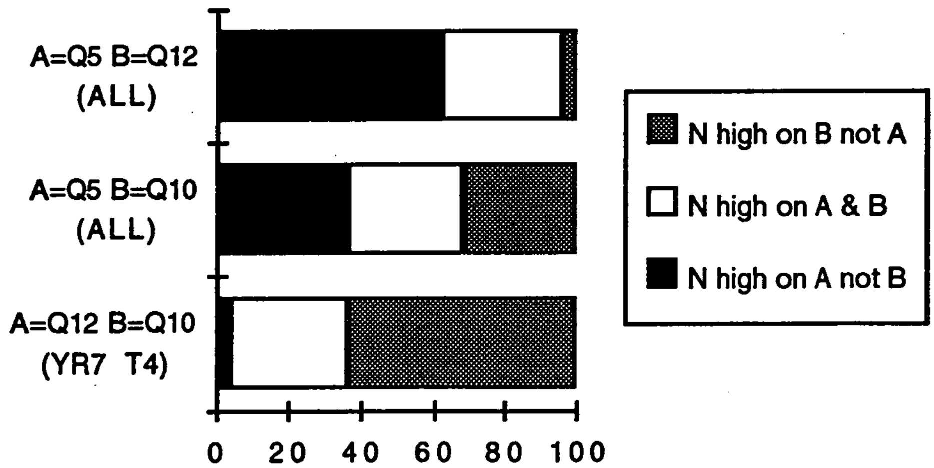


Figure 8·1. Percentage frequency chart for testing hierarchies
Showing Q.5 as a Prelude to Q.12, or Q.5 > Q.12;
Q.5 & Q.10 Indeterminate;
Q.10 as a Prelude to Q.12, or Q.10 > Q.12.

In terms of the diagram,

$$\%AB = \frac{\text{No.high on A \& B}}{\text{No.high on B not A + No.high on A \& B}}, \text{ and}$$

$$\%BA = \frac{\text{No.high on A \& B}}{\text{No.high on A not B + No.high on A \& B}}.$$

The conclusions are listed below the figure. The bars may be interpreted as Venn Diagrams, with the white patch representing the intersection of two sets, one set enclosing all those high on A, and the other those high on B. As this chapter unfolds, this method of analysis is used in the search for hierarchies of cognitive difficulty.

For a discussion of confidence intervals with regard to the criteria given, see Appendix 8D, page 486.

Section 2: Hierarchical Levels of Understanding of Symbols

Part 1: Generalizations From an Overview of Trends

The available data. The school mathematics departments had independently categorized the participating classes according to their perceived mathematics ability levels, as listed in Tables 3·6 and 3·7 of Chapter 3. The levels are referred to in this chapter as "Low" for groups of slow learners, "Mixed" for groups with average or mixed mathematics ability, and "Advanced" (Adv.) for the highest mathematics ability groups. The Year 12 class (Class 61) that studied the Mathematics in Society course was classified as "Low". One class (Class 3) of Low level Year 7 students was tested four times, and two other Low level Year 7 classes (Classes 14 and 15) were tested only once. In the graphs which follow, the latter are identified by "LOWx2". Advanced level classes were tested for all the Year levels, Mixed level classes for all but Years 8 and 12, and Low level classes for all but Years 10 and 11.

Despite the incompleteness of cover, the data were sufficient for identifying general trends and for providing grounds for discussion of relationships between increased experience with algebra and changes in levels of understanding of a variety of aspects of algebra. Longitudinal test data were obtained from Year 7 students over six or seven months on the dates shown in Table 3·10 (p. 83). Data from the single testing of students across Years 7 to 12, as listed in Table 3·7 (p. 78), provided a snapshot record of test results for these different groups. The two sets of data were combined to produce a series of graphs of the type shown in Figure 8·2. The longitudinal data blended well with the overall patterns of outcomes from the other students and identified the Year 7 students as truly beginners in the quest for mastery of the basic concepts of the algebra of generalized numbers. The graphs located the Year 7 beginners in the continuum of learning. In the CSMS study, outcomes from a single testing of students in different Year groups were used in a similar way (e.g., Küchemann, 1984, p. 123; Hart, 1981a, p. 186) to draw comparisons between levels of performance by students at different stages along the learning process.

Total Test Scores

Figure 8·2 graphs the averages of the total test scores attained by student subgroups. Overall, an improvement in scores was found to accompany more experience with algebra. The subgroup which produced the nearest to a continuous growth pattern was that composed of Advanced classes. The Low ability classes showed a regular improvement except for the slightly poorer performance of the Year 8 Low group. The pattern of growth for the Mixed ability classes was present but for depressed average scores by the Year 7 (in their third test) and Year 10 representatives of this subgroup. In School C, the latter Year 10 students were in a class ranked third of those following the Intermediate mathematics syllabus and so were close to being classified as General or Low level. In School D, they were tested while not in their regular mathematics groupings and included two General level students.

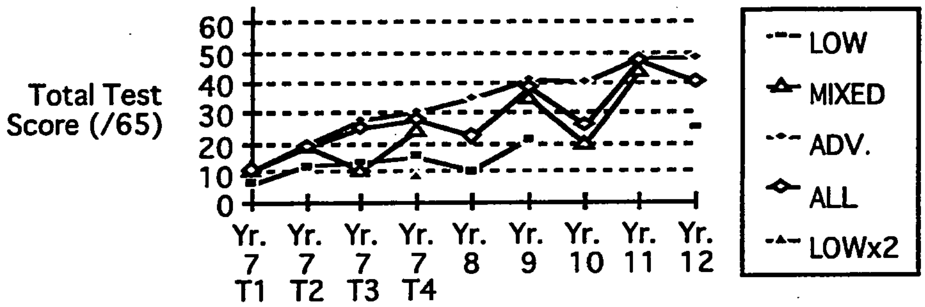


Figure 8·2. Averages for total test scores

The trend for scores to improve with experience supported the view that the test items were measuring aspects of algebraic thinking which were far from superficial and were probing cognitive challenges associated with early algebra. The average for the top ability class in Year 12 was 48.45, or 74.54% of the maximum possible test score of 65, indicating that the test items were not trivial. It was noted that only one student scored over 60 on the test. The items were able to differentiate between the more able thinkers and the less able, as indicated by the divergent degrees of success revealed by the graphs for the Advanced and Low ability level classes.

Some Year 7 students may have benefited from practice on the test. However, they were not drilled in the test items by teachers (an exception is discussed on pages 305 to 307) nor were they coached by the researcher during interviews between tests. Growth rates varied from student to student and from class to class. Improvement on many items was dependent on growth in understanding. There were reasonable indications that particular classroom experiences may have been important factors in some resultant test response changes, as in the instances discussed on pages 212 - 213, 225 - 226, 305 - 307, and 470 - 475.

Patterns of change in levels of understanding for the meanings of algebraic symbols are now discussed, taking in turn the five levels designated in Table 6·1.

Variable (Level 5)

VBL Scale. The graphs in Figure 8·3 summarize the average scores on the VBL Scale for the various groups of classes in the research sample. A comparison of Figure 8·3 with Figure 8·2 shows that the patterns of change in understanding algebraic symbols as numerical variables, as measured by the VBL Scale, were closely similar to those for overall test scores. As was pointed out in Table 6·9, the correlation between scores on the VBL Scale and corrected test totals yielded a coefficient of .756, indicating a significant relationship between these two measures of performance. The shared variance was 57.2%. The correspondences between the patterns of dynamic development summarized by Figures 8·2 and 8·3 verify the prediction implicit in this static correlation coefficient that test total and understanding of the variable concept should grow hand in hand.

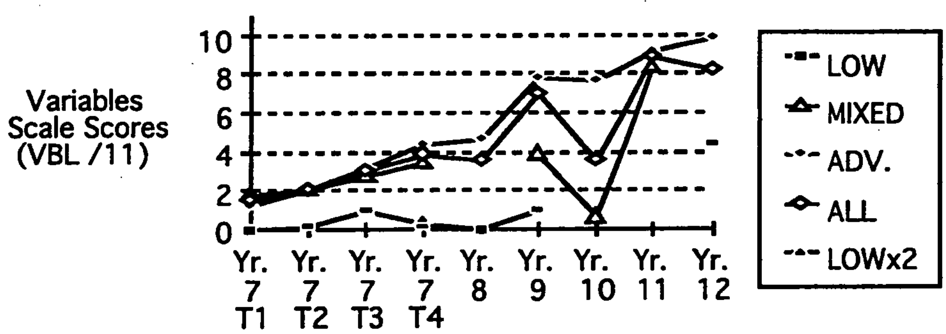


Figure 8·3. Average scores on the VBL Variables Scale, based on Questions 6 (c), 10, 12, and 15 (iii).

The Low ability classes found the variable notion to be very difficult to grasp. The best average achieved by the beginning Year 7 slow learners was 1.00 out of a possible 11 in the third test. After about six months with scarcely any work on algebra, this average dropped off to 0.20, and two other low ability Year 7 classes averaged 0.50 at the same period. Furthermore, none of the Year 8 Low ability class registered a positive score on any of the items in the VBL Scale.

The challenging nature of the variable notion, as measured by the items used in the VBL Scale, was further evinced by the rise in performance level from Year 8 to Year 9 (averages 4.63 and 7.86 respectively) for the Advanced ability classes. One interpretation of this feature is that growth in understanding of this difficult concept was developmental, in the Piagetian and neo-Piagetian sense. However, to counter an acceptance of this interpretation as implying that development is related simply to age and experience, it should be noted that the Year 12 Low ability class reached an average of only 4.38 on the VBL Scale, an average equivalent to that achieved by the Advanced Year 7 classes (4.37) after about seven months of algebra, and slightly lower than the Advanced Year 8 average (4.63), attained about a year after starting their study of algebra.

Other Level 5 scale measures. The four other scale measures of the extent to which students understood algebraic symbols as representing numerical variables are now considered. Graphical summaries of the averages attained by the different subgroups on these scales have been assembled in Appendix 8A.

The BXBA Scale assessed the readiness to allow conditions '$b = x$' and '$b = a$' in Questions 15 (i) and (ii) respectively. The overall trends in the patterns of change for average scores on this Scale (Figure 8A·1 in Appendix 8A) were closely allied to those for the VBL Scale, reflecting the existence of a high correlation (.574) between scores on these two scales, as recorded in Table 6·3. In all four tests, the Low level Year 7 class was slightly more successful on this scale than on the VBL Scale.

The EQN Scale was designed for assessing the ability to discuss relative sizes of two variables related by the equation '$2x + y = 9$'. Again, there were many similarities between the patterns of change for the average scores on this scale (Figure 8A·2 in Appendix 8A) on those on the VBL and BXBA Scales. The respective correlations (see p. 183) between these scales were positive and highly significant.

The CZ Scale assessed the readiness to allow the condition 'c = zero' in Question 15 (iv) and graphs recording the average scale scores are displayed in Figure 8A·3 in Appendix 8A. The difficulty of the question was explicitly reported by the low success rates on this scale. The best averages came from the Advanced senior students: Year 11 with an average of 22.7% and Year 12 an average of 37.5%. All other groups scored less than 12% except for the one occasion, on the second test, when the small group in the Low Year 7 class attained 16.7%. Interviews confirmed that some students worked through successive steps to considering the possibility that '$2c$' could equal '$4c$' but then did not think of the case '$c = 0$' and provided an incorrect answer. The format of the question did not allow for working to be shown and credit was not allocated for partial solutions to the challenging problem.

The PL Scale was used for assessing ability in using an algebraic argument about the comparative lengths of a pair of parallel lines. Several features of the graphs for this scale (Figure 8A·4 in Appendix 8A) were different from those for other scales measuring the degree of attainment of the variable notion:

1. Year 10 Advanced scored better than Year 11 Advanced;
2. Year 7 Low ability outscored Year 7 Mixed ability on Tests 2, 3, and 4; and
3. Year 12 Advanced scored less than Year 11 Advanced (as they did also on the EQN Scale, but not the other Level 5 scales).

It seemed that students either focused their attention on the geometry of the given sketch or on the significance of the algebraic labels used for the lengths of the lines in the sketch. For example, the university students who participated in the trialling of Question 11 (as reported in Chapter 3) tended to fall into one or other of these categories. The question was, nevertheless, included in the test as a means of probing students' ways of viewing symbols.

Generalized Number (Level 4)

The GN Generalized Number Scale was formed to keep account of those who worked at the generalized number level (Level 4) and not the variable level (Level 5). It was difficult to find sufficient statistical support for including all available measures of the Generalized Number concept in such a scale, so another scale was established to include credit for working at either the generalized number level or the variable level on selected test items. This was called the GNV Generalized Number or Variable Scale and it spanned Levels 4 and 5 for the understanding of the meanings for algebraic symbols. Four other scales at Level 4 were the INT Integers Scale, the FZN Fractions-Zero-Negatives Scale, the NBR Numbers View Scale, and the NFL Numbers of Flowers Scales. Details of the categories of responses allocated to all five scales are given in Appendix 5F.

As the graphs in Figure 8·4 record, the GN Scale scores were low for all groups, indicating either that the concept of generalized number was too difficult for most students or that operating with the generalized number view of symbols was simply a transitory state for those who advanced to the variable view.

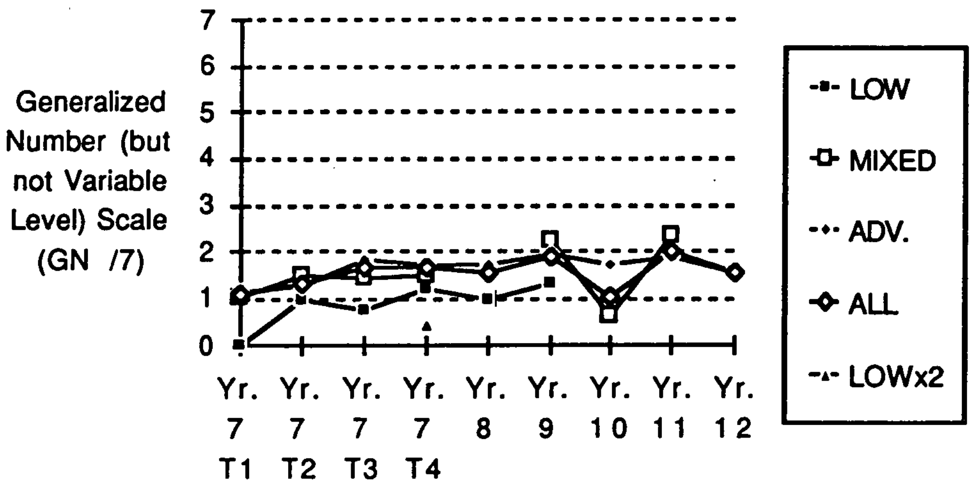


Figure 8·4. Average scores on GN Generalized Number Scale, based on Questions 6, 10, 12, and 13

When credit was given for both the Generalized Number level and the Variable level in the GNV Scale, the patterns of averages bore close similarities to those for the VBL Variables Scale score averages, as can be judged by comparing Figures 8·3 and 8·5. The implication seems to be that once students were able to think of algebraic symbols in terms of the generalized number concept, they were on the brink of moving to the next level of understanding, that of the numerical variable. One notable exception was the Low ability classes. Students in these classes rarely attained the variable concept but sometimes reached the generalized number concept.

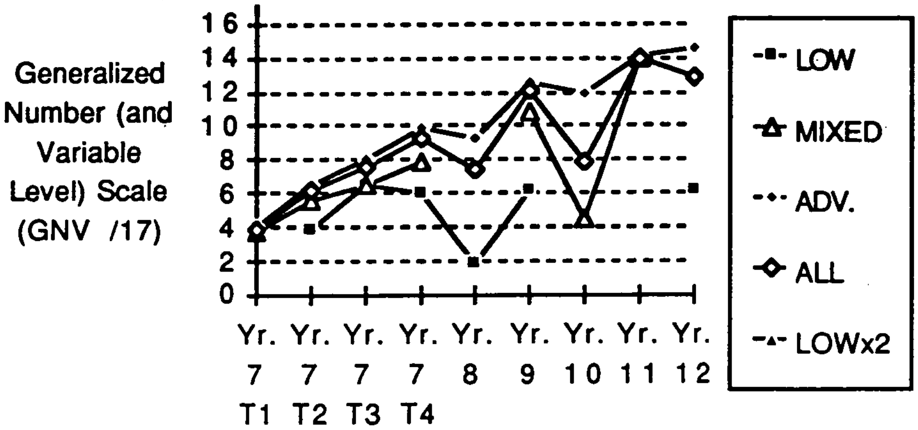


Figure 8·5. Average scores on GNV Scale, based on Questions 2 (i), 6, 10, 11, 12, and 13

Average scores were also low on the INT Integers Scale, as shown in Figure 8·6. Thinking in terms of integers only, another sign of Level 4 understanding of symbols, was not common amongst the population participating in the study.

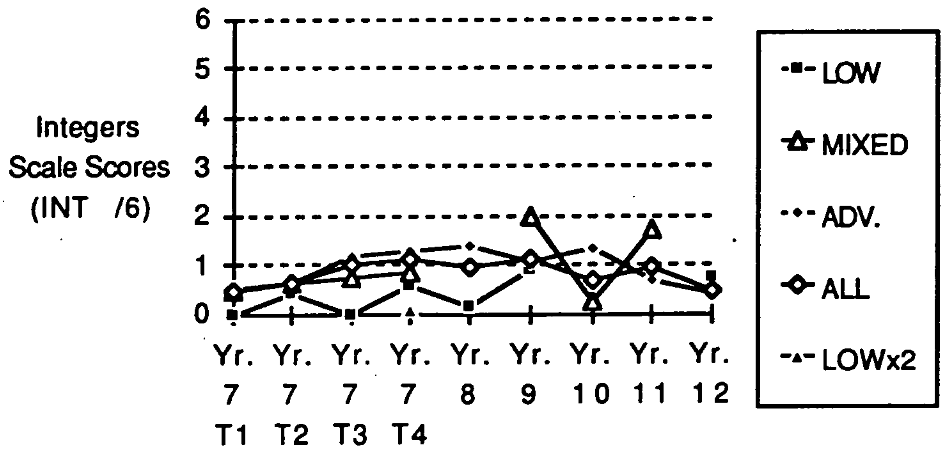


Figure 8·6. Average scores on the INT Scale, based on Questions 10, 12 and 13

Figure 8B·1 in Appendix 8B displays average scores on the FZN Fraction-Zero-Negatives Scale. Except for the more able Year 11 and Year 12 groups who were the only students to reach an average of more than 50% of the maximum for the scale, the scores were low for all groups.

Similarly, Figure 8B·2 shows that scores on the NBR Numbers View Scale were low except for the Advanced Year 11 and Year 12 groups, again the only students to reach an average of more than 50% of the maximum for the scale. The Low Year 7 class outscored all other Year 7 classes in Tests 2 and 3, showing that they were particularly conscious of the numbers view of symbols at the times that those tests were administered.

Figure 8B·3 indicates that the only group to reach the 50% mark for the NFL Numbers of Flowers Scale was the small Mixed ability Year 9 group although, on Test 3 for Year 7, all but the Low ability group attained almost the same score.

Replacement Values and Specific Unknown (Levels 3 & 2)

Figure 8·7 records that average scores on the 12REP Replacement Value(s) Scale were very low for all groups of students. The option of operating at the level of "specific unknown" (Küchemann, 1980) or "discovered content " (Harper, 1979) by using replacement examples instead of giving general solutions to problems was not popular. Graphs for the 1REP and 2REP, the One- and Two- Replacement Scales are not presented since the low scores on these are implied in the graphs of Figure 8·7.

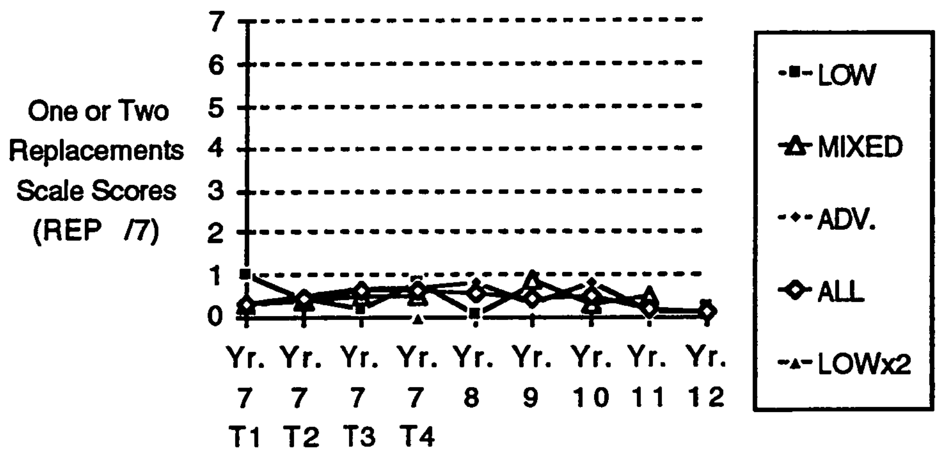


Figure 8·7. Average scores on the 12REP Scale, based on Questions 10, 12 and 13

Prestructural Views (Level 1)

Beginning Year 7 students were inclined to misunderstand the significance of test questions and to respond in ways that were classified as "prestructural" (Collis & Watson, 1989, p. 181). As Figure 8·8 shows, this inclination diminished quickly with continued exposure to algebra in the classroom, except for some of the students in Low ability classes, even in Year 8 and Year 12. The Year 10 Mixed ability class also registered a high average on the PRE Prestructural Scale, reaffirming, as do other graphs in this section, the impression expressed on page 256 that a proportion of these students could well have been classified as Low ability level.

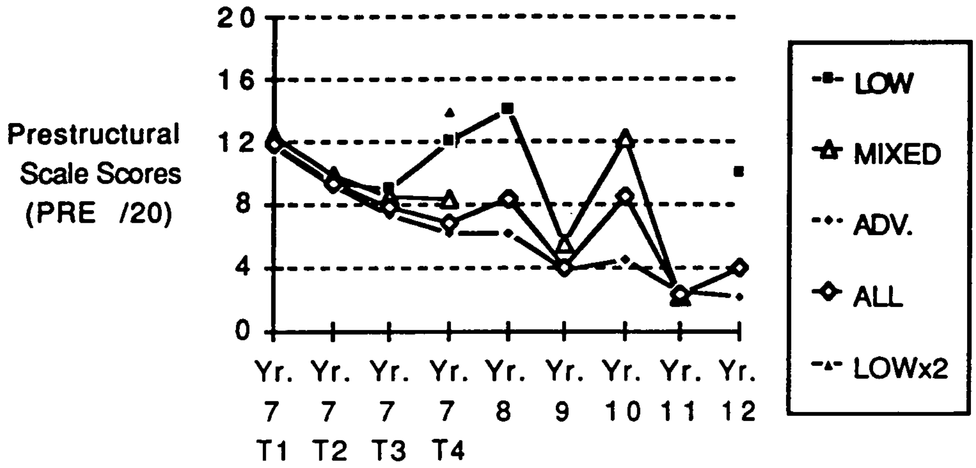


Figure 8·8. Average scores on the PRE Scale, based on Questions 2, 6, 10, 11,12, and 13

The graphs in Figure 8·9 indicate that, in most cases, the tendency to seek closure by the avoidance of writing symbols to express general answers dissipated rapidly with experience.

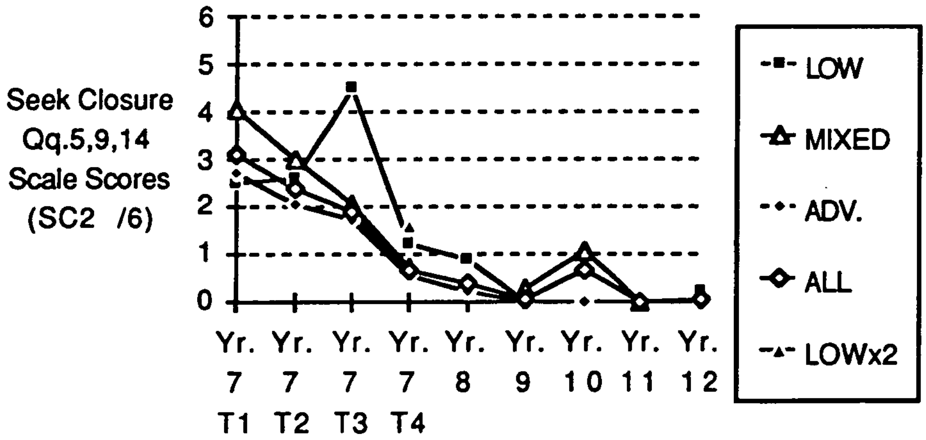


Figure 8·9. Average scores on the SC2 Scale, based on Questions 5, 9, and 14 ii, iii

The most persistent misunderstanding was that of viewing algebraic symbols as representing objects or people rather than numbers of objects or people. The graphs for the OBJ Objects View Scale given in Figure 8·10 report a general decrease in average scores across the year groups. The dissipation rate for the errors measured here is slower than for those measured by other Level 1 scales, including the nine scales reported in Appendix 8C.

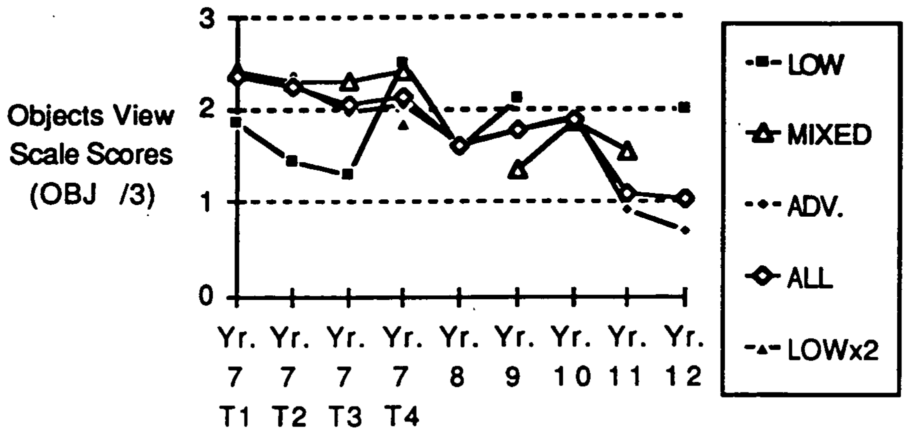


Figure 8·10. Average scores on the OBJ Scale, based on Questions 7 and 8 a

Part 2: Psychological Analyses of Cognitive Challenges in Question 10

In this part of Section 2, the theories of cognition outlined in Chapter 2 are applied to analyses of the task difficulties involved in the test items used in one of the questions for measuring the degree of success at Level 5 understanding of algebraic symbols. The test data have provided the opportunity to look closely at degrees of difficulty rather than simply dismiss groups of items in some general classification, such as "they all measure the variable concept". The existence of hierarchies of cognitive difficulty within such a classification will be exposed by the statistical analyses presented in the third part of this section. A deeper probing of the challenges posed by the items used for measuring the degree of development of the variable concept from a psychological point of view is undertaken first. Question 10, an item of comparatively little difficulty within the VBL Variable Scale, has been chosen as a suitable example for illustrating the process.

Cognitive Challenges Posed by Question 10

Question 10 on the test instrument provided data which shed light on students' levels of thinking about algebraic symbols. It was in four parts.

10. This question is about t + t and t + 4.
(a) Which is larger, t + t or t + 4 ? WHY?
(b) When is t + t larger?
(c) When is t + 4 larger?
(d) When are they equal?

(Harper, 1979)

The question may appear simple to those who have become familiar with the

The question may appear simple to those who have become familiar with the principles of basic algebra. However, on reflection, the reality is that successful responses are the fruit of complex cognitive processes. The letter '*t*' needs to be considered three times in each part of the question. In each case it represents a variable, the value of which may be any real number ranging from large positive numbers through zero to large-sized negative numbers and these possibilities include non-integral cases. The value is quite fluid but all three instances of '*t*' must reflect the same value at any particular moment. The expression '$t + t$' means to add the value of '*t*' to itself, giving twice the value of '*t*', while the expression '$t + 4$' means to add 4 to whatever value '*t*' happens to have. Each part of the question requires the respondent to take all of this into account, and then to proceed to carry out a comparison step in which relationships between the possible values of the two expressions are to be considered. Not only is there the demand for making comparisons but also the elements to be compared have unrestricted values. Moreover, the addition process in each expression cannot be closed to give a unique, known result unless, of course, '*t*' is given some particular numerical value. In part (d), the latter procedure could be successful in determining that '$t = 4$' designates the condition for the two expressions to be equal, each of them then having the value 8. In the other parts of the question, it is necessary to accept lack of closure to proceed beyond a strategy of using replacement values for '*t*'. Students who were astute enough to realize that '*t*' could be cancelled from each expression were left with the simplified problem of comparing the values of '*t*' and '4'. However, considerable cognitive processing is required before reaching this simplified form unless, of course, drill and practice had led to cancelling becoming a regular routine. Of the participating students, at least the Year 7 beginners were not taught this cancelling routine.

Analysis of Question 10 in Biggs and Collis Terms

Different types of responses to Question 10 may be categorized in terms of the Biggs and Collis (1991) modes, hierarchical levels, and stages. Consider the following examples taken from written test answers or from interviews.

Example 1. (Year 7 boy, first test):

In (a), which is larger? " 't+t' because it is double."

This is more a unistructural than a multistructural answer, as mainly one relevant piece of information is used. The expression '$t + 4$' was probably given just a cursory consideration. It could be in the Ikonic mode if the student was regarding the letters simply as objects, of which there are two, but if they were considered as symbols for numbers, the mode would be Concrete-symbolic.

Example 2. (Year 7 girl, second test):

In (a), which is larger? " '$t+t$' because 't' is larger than '4'."

This response would be in the Concrete-symbolic mode at the Multistructural level, as some recognition is given of the two expressions. However, an incorrect assumption is brought into the argument, namely, that 't' is larger than '4'. The same girl was happy to leave a contradictory, and correct, statement in the same test for (c) 't + 4' is larger "when '4' is larger than 't' ". The contradictory outcomes were consistent with a multistructural level of operation in which judgements are made in isolation. In the following extract from an interview soon after the second test, she showed that she regards the letter more at Harper's Discovered Content category than at the level of Variable:

> When is '$t + t$' larger? Well, would it be when 't' is smaller than '4'? Because when 't' is ... or no! ... When 't' is larger than '4', because ... umm ... if just, say, t stands for 5, well I think, ... well 5 plus 5 equals 10 and 5 plus 4 equals 9. So it would be.

This use of substitutions could be an indicator of Multistructural level activity and an unwillingness to accept lack of closure, an acceptance needed in a general argument for the parts of Question 10. Further evidence from written answers to three other test questions confirmed that she had not yet grasped the notion of variable. In questions 5, 9, and 14, she simply made up numbers as values for the letters involved and so avoided having to give answers in the form of expressions incorporating letters. She continued to show this rejection of lack of closure in her third test, even though she then correctly answered each part of question 10.

Example 3. (Year 7 boy, second test):

Now consider this set of written answers to Question 10:

> (a) Neither because t can equal 4
> (b) when t is 5 or larger
> (c) when t is 3 or smaller
> (d) when $t = 4$

Parts (a) and (d) are acceptable answers, while (b) and (c) were classified as "Algebra nearly correct", because they are correct for integers but disallow fractional values for 't'. Here we have a Relational response in the Concrete-symbolic mode. The fact that the student is thinking only in terms of integers suggests that he is not yet using the notion of 't' as a true variable. In an interview following the second test, he clearly repeated his use of integers:

> 't" has to be the same number in each equation ... er ... each statement so ... umm ... neither of them are [*sic*] larger because 't' could stand for '4' right? Then where it's got (b) when is '$t + t$' larger? ... '$t + t$' is larger only when it is 5 or above because 5 plus 5 equals 10, and 5 plus 4

> equals 9. When is 't + 4' larger? That's only when 't' is '3' or less, and when are they equal is when 't' is '4'. Then you have 4 plus 4.

Confirmation that he had not yet attained the more abstract level of thinking in which 't' would be regarded as a variable was given in his responses to questions such as Question 9: In the interview, he said:

> "Add 4 onto $n + 5$"? Well you can't tell because you don't know until ... what 'n' is. It's like trying to say add 4 to 7 + 5 or add 4 onto 6 + 5 or something like that and you'd always get a different answer.

By Test 3, he had all parts of Question 10 correct, writing:

> (a) t+t is larger when t equals over 4 . t+4 is larger when t equals under 4
> (b) when t is over 4
> (c) when t is under 4
> (d) when t equals 4.

Nevertheless, in the same test he wrote "What's n" [*sic*] for each part of Question 9. Here we have student who provides a Relational answer in Question 10 with acceptance of lack of closure but who shows an inability to handle lack of closure in another question.

These examples illustrate the truth of Harper's (1979) comment that the Literal Numbers Tasks alone could not clearly identify those using the species notion for algebraic symbols. They also show how such a task can reveal much about how a subject thinks about letters used as algebraic symbols.

Analysis of Question 10 in Halford Terms

The basis for ordering the levels of complexity of a task in the Halford (1987) Structure-Mapping approach was the number of elements that had to be considered in the mapping required for that task: "Complexity is assessed by the number of independent dimensions in the structure" (Halford, personal communication, 4 July, 1990). Such a theoretical stand is strongly reminiscent of the emphasis given to "short term storage space" by Case (1987a, p. 605) and Pascual-Leone's "attentional power", which "was defined as the maximum number of independent schemes that can be brought to full activation simultaneously" (Case, 1985, p. 33). Pascual-Leone regarded as central to his theory "the idea of a developmentally-growing mental capacity that conditions the processing complexity a subject can handle, and determines the transition from one stage to the next" (Pascual-Leone, 1987, p. 28).

As the number of elements involved in cognitive processes increased from one through to four, the relevant mappings, in Halford's terms, moved progressively from Element Mappings through Relational Mappings to System Mappings to Multiple-System Mappings. As pointed out in Table 2·1 in Chapter 2, the interpretation of an algebraic expression containing one arithmetic operation was regarded as a System Mapping, since three elements were being considered. The three elements involved, for instance, in interpreting 't + 4' in Question 10, were 't' and '4', which were related by the addition operation, and the value or meaning of 't + 4' itself. Question 10 could have become very difficult for students who needed a System Mapping process to struggle over simply interpreting each of the *two* expressions in the question ('t + t' and 't + 4'), because they then had a heavy cognitive load before even attempting to compare possible values of these unclosed algebraic entities.

The following recommendation of Halford and Boulton-Lewis (1989, p. 31) appears to apply here: To avoid an excessive cumulative load, the load imposed by one structure mapping needs to be reduced to zero before the next structure mapping is undertaken. It seems that there could be a new cycle of cognition built upon mastery of the stage of interpreting the meaning of an algebraic expression. If interpreting an algebraic expression containing one operation is mastered as a basic skill, then Question 10 is concerned with two elements. In Question 10 (d), conditions for the equality of these two elements becomes the task, and parts (b) and (c) of the question extend the task by asking when one element is greater than the other. In each of these cases, as two elements are being considered, the Halford level would be one of Relational Mappings.

The list of "independent dimensions" for Question 10 part (d) is:

1. recognizing that 't' is a numerical variable, capable of representing any member of the relevant number field, which was designated as that of the rational numbers when assessing the outcome to be at the variable level;

2. interpreting the expression 't + t' and realizing that its value depends upon the value that the letter 't' happens to have at any particular time;

3. interpreting the expression 't + 4' and realizing that its value depends upon the value that 't' happens to have at any particular time;

4. recognizing that at any particular time the value of 't' is the same for each expression; and

5. finding the value of 't' which gives both expressions the same value. For some, this would entail a process of trial an error and, for others, a solution of the equation '$t + t = t + 4$', either by some formal algebraic method or some more intuitive approach.

Students who had mastered the first four of these steps would, it could be expected, have had less cognitive load than those who had to work specifically at establishing each step along the way.

Part (b) of Question 10, by asking when 't + t' was larger than 't + 4', added to the list of "independent dimensions" in one of two ways, as follows.

First possibility:

6. realizing that as the value of 't' increases so does the value of '$t + t$ ';

7. realizing that as the value of 't' increases so does the value of '$t + 4$';

8. realizing that the two given expressions were covariants: Both increased as 't' increased but the first expression increased faster than (at twice the rate of) the second; and

9. concluding that if 't' is greater than '4' then '$t + t$' is greater than '$t + 4$'.

A second possibility was:

6. comparing '$t + t$' and '$t + 4$' to realize that, with 't' in common to both expressions, the problem dwindled down to simply relating 't' to '4'; and

7. concluding that if 't' is greater than '4' then '$t + t$' is greater than '$t + 4$'.

This latter method of solving the problem is the more meritorious mathematically as it shows a facility to work with abstract algebraic concepts. It avoids the requirement of seeking closure for each expression as a pathway towards resolving the question of which is greater. If the strategy of subtracting 't' (or cancelling a 't') from each side had been taught and had become a standard algorithm, then the difficulty of the problem would be reduced. In the case of the beginning Year 7 students in the study, such a procedure had not been taught. If they were to apply this efficient method, they would need to identify this way of recasting the problem to convert it into a case of merely comparing the values of 't 'and '4'. To identify the method, these students were faced with a demanding task, one which involved what Halford called a system mapping as summarized in Figure 8·11, although the presence of four elements in the mapping could be an argument for regarding it as a multiple-system mapping in accordance with Table 2·1 of Chapter 2.

$$\begin{array}{ccccccc} X & + & Y & = & X & + & C \\ \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\ t & + & t & = & t & + & 4 \end{array}$$

(X & Y are variables, C is a constant)

Figure 8·11. System Mapping for simplifying the problem in Q.10
(from Halford, personal correspondence, 19 Nov.,1991)

Several steps were needed in using this system mapping:

1. recognizing that the '4' was a constant term and this required a mapping of the element '4' to the more abstract element, the constant, or 'C', in the relationship given in the first line of Figure 8·11;

2. recognizing that each 't' term represented a variable and mapping them to abstract "variable" elements, 'X' and 'Y' in the first line of the figure;

3. realizing that the two variable terms marked 'X' corresponded on each side of the equals sign and could be cancelled out of the relationship by a subtraction

process.

After cancellation, the problem simplified to one requiring only a relational mapping, as shown in Figure 8·12. The student now only had to compare the magnitudes of the variable and the constant.

$$\begin{array}{ccc} Y & = & C \\ \updownarrow & & \updownarrow \\ t & = & 4 \end{array}$$

Figure 8·12. Relational Mapping for simplified Question 10

After the cancellation step, the equality case can lead to the inequality cases and can be completed by means of a relational mapping.

What these analyses make clear is that, without a knowledge of a set routine, the question imposes a fairly heavy cognitive load due to the number of independent dimensions to be considered. Such outcomes support the view that more advanced strategies, once understood, "require less processing capacity than the strategies they displace" (Küchemann, 1980, p. 146). Furthermore, judging by the number of dimensions, each of the inequality problems demands more cognitive load than does the equality problem in part (d).

Similar judgements apply to the corresponding parts of Question 12. The argument can be taken also to the equations task in Question 13 in which the equality case was set as part (c) and thus theory suggests that part (c) should have been easier than parts (b) and (d) in Question 13.

Analysis of Question 10 in Fischer Terms

When applying Fischer's (1980) Skill Theory to Question 10, the first step is to locate the understanding of simple algebraic expressions within the range of the cognitive levels that he identified, and which were summarized in Tables 2·2 and 2·3.

A Representational Mapping (Representations 2 Level, or Level 5 on the earlier numbering system) would be the minimal level of cognitive process for accepting an algebraic symbol such as '*t*' to stand for a numerical variable so that mappings could be made between the symbol and any member of the whole number field which is appropriate in the given context. This can be expressed as $[{}^5N - {}^5A]$, in Fischer symbols, where '*N*' stands for a number element (e.g., 4.2), '*A*' stands for an algebraic symbol (e.g., '*t*'), and '5' refers to the fact that this is the fifth of the 10 Fischer levels. In the case of Question 10, that number field would consist of all the rational numbers.

To understand an algebraic expression, the student has to be able to map arithmetic operations on elements of the number field to similar operations on elements of the algebraic symbolic field. If the value(s) of the algebraic symbol(s) in the expression is(are) known, then this step is not far removed from arithmetic. In Question 3, for example, students were told that 'y' had a value of '3' and were asked to evaluate some linear expressions, each of which contained 'y' as the one variable. If students understood the conventions used in the algebraic expressions, this question took them from algebra back to arithmetic. The skill required was that of translating the conventional algebraic expression into the corresponding arithmetical procedures with the appropriate order of operations if more than one arithmetic operation was involved. Closure was required in each part of the question. For instance, in the second part, to evaluate '$2y + 5$', students simply had to multiply '2' by '3' to reach a numerical value of '6' and then add '5' to give an answer of '11'. The process, nevertheless, involved the need for intercoordinating the skills of accepting that a letter could stand for a number and for applying the conventions used in writing the given expression so that the appropriate arithmetical calculation could be completed. This level of processing would be required for evaluating '$t + t$' or '$t + 4$' for arbitrary numerical values of 't'. The cognitive level for Question 10 is then Level 6, or Representations 3, at what Fischer and Lamborn (1989, p. 40) call the "concrete operations" level.

The following equation, in Fischer form (Fischer, 1980, p. 498), summarizes such a process:

$$[^5N_1 - {}^5A_1] \,.\, [^5N_2 - {}^5A_2] \;=\; [^6N_{1,2} \leftrightarrow {}^6A_{1,2}].$$

When algebraic expressions are treated more generally and closure to give a numerical outcome is made inappropriate, then the level of difficulty is higher. In Question 10, for example, to interpret the expression '$t + 4$' the process of adding '4' to another number has to be mapped to the process of adding '4' to an algebraic symbol which represents a number which could be any member of the appropriate number field. This involves an understanding of the general definition of addition and thus would be at the Fischer level Abstractions 1 (or Representations 4), the former Level 7 (Fischer, Hand & Russell, 1984, p. 48). This is a system of representational systems and may be designated as [^{7}E]. Fischer and Lamborn (1989, p. 40) referred to this level as "formal operations".

Question 10 extends the student still further by asking for a comparison of the values of two algebraic expressions. The skill level described for understanding one algebraic expression now has to be used twice and the two outcomes have to be compared so that conditions for equality and for inequality of the two expressions can be identified. The comparison step is a mapping of one outcome onto the other and so the cognitive task is at the next level, that of Abstract Mapping which can be described as Abstractions 2 or Level 8, and designated in Fischer symbols as $[{}^8E_1 - {}^8E_2]$.

There is a need to compound two skills in the case of identifying precisely the condition for the inequality of two expressions. One skill is that of defining when the expressions are equal. This, logically, must be decided before conditions for inequality can be clarified. The second skill is that of identifying a range of possibilities for the expressions to have differing values. For example, to answer Question 10 part (d), the student must realize that it is only when 't' equals '4' that '$t + t$' can equal '$t + 4$'. This could be designated as $[{}^8E_{1,t=4} - {}^8E_{2,t=4}]$. Then, for part (b) the student needs to consider values of 't' other than '4' and to note which class of numbers other than '4' cause the value of '$t + t$' to be greater than the value of '$t + 4$', rather than less. At this stage, numbers such as 5, 6, and higher might be considered but, to give the best answer to part (b), the student has to come back to the equality condition and recognize that, as long as 't' is even slightly larger than '4', then '$t + t$' is larger than '$t + 4$'. Hence the cognitive process for identifying precisely when '$t + t$' is larger could be described by the compounding summarized in an adaptation of the equation form given by Fischer (1980, p. 499) as:

$$[{}^8E_{1,t=4} - {}^8E_{2,t=4}] + [{}^8E_{1,t\neq 4} - {}^8E_{2,t\neq 4}] = [{}^8E_{1,t>4} - {}^8E_{2,t>4}].$$

Similarly, the cognitive process for identifying precisely when '$t + 4$' is larger could be described by the compounding summarized in equation form as:

$$[{}^8E_{1,t=4} - {}^8E_{2,t=4}] + [{}^8E_{1,t\neq 4} - {}^8E_{2,t\neq 4}] = [{}^8E_{1,t<4} - {}^8E_{2,t<4}].$$

Applications of Psychological Analyses

The above psychological analyses of the cognitive challenges posed by Question 10 exposed several hierarchies of difficulty. These insights prompted interest in further investigating such hierarchies, not only from a theoretical point of view but also by taking into account the empirical data collected in the field. Section 3 of this chapter investigates nine propositions regarding such hierarchies and, in so doing, indicates ways in which theoretical and empirical analyses concur. Chapter 9 reports on a further investigation involving Question 10 in which not only hierarchies but also sequences of learning are explored.

Further general considerations about cognitive hierarchies are presented in parts three and four of this section.

Part 3: Statistical Analyses Identify Hierarchies of Learning Within Level 5

Cross-tabulation Analyses

Application of the cross-tabulation approach, detailed in Section 1 of this chapter, succeeded in identifying several hierarchies of cognitive difficulty in terms of the measures used for the Level 5 understanding of symbols as numerical variables. Table 8·4 summarizes the statistics for analyses involving all 517 students. Evidence is presented for success on the first-mentioned of the following pairs of measures to be a prelude for success on the second:

Q.10 (part of VBL Scale) > Q.15 (iv) (CZ Scale)
Q.12 (part of VBL Scale) > Q.15 (iv) (CZ Scale)
Q.15ii (part of BXBA Scale) > Q.15 (iv) (CZ Scale)
Qq.15 (i) and (ii) (BXBA Scale) > Q.15 (iv) (CZ Scale)
Q.15 (i) (part of BXBA Scale) > Q.13 (EQN Scale)
Q.15 (iii) (part of VBL Scale) > Q.13 (EQN Scale)
Q.10 (part of VBL Scale) > Q.13 (EQN Scale)
Q.12 (part of VBL Scale) > Q.13 (EQN Scale), and
Q.15 (i) (part of BXBA Scale) > Q.15ii (another part of BXBA Scale)

Table 8·4
Evidence for Hierarchies Within Level 5

Variables		Frequencies			Percentages		Ratio	Correlations	
A	B	High A & B	High A	High B	%AB	%BA	%AB / %BA	*r*	*p*
Q.10	CZ	20	171	22	90.9	11.7	7.77	.149	***
Q.12	CZ	18	122	22	81.8	14.8	5.55	.180	***
Q.15ii	CZ	17	103	22	77.3	16.5	4.68	.236	***
BXBA	CZ	17	95	22	77.3	17.9	4.32	.283	***
Q.15i	EQN	47	202	59	79.7	23.3	3.42	.349	***
Q.15iii	EQN	53	182	60	88.3	29.1	3.03	.358	***
Q.10	EQN	50	170	60	83.3	29.4	2.83	.464	***
Q.12	EQN	47	123	60	78.3	38.2	2.05	.562	***
Q.15i	Q.15ii	95	207	103	92.2	45.9	2.01	.513	***

Note. $N = 517$, using Test 3 for Year 7. Ordered by size of ratio %AB to %BA.
*** $p \leq .001$.

To follow the investigation further, a similar procedure was carried out for Year 7 beginners' responses on each of the four times they completed the test, and also for responses from all the Advanced classes across Years 7 to 12. Relevant statistical details appear in Tables 8D·1 and 8D·2 of Appendix 8D. As expected, the frequencies for complete success on some of the items under scrutiny were low for the beginning Year 7 students (in 15 of the 21 reported sequences there were less than 10 students correct on one or other of the variables). This meant that these outcomes were all the more valuable and instructive as they pointed to the order in which some of these difficult concepts were first mastered. In Year 7 students' responses to Tests 2 and 3, the frequency for complete success on both aspects of Question 15 part (i) was the highest, indicating that mastery of the task tested by this question was achieved before mastery of the other tasks listed. The next most-frequently mastered tasks were those assessed by Question 10.

A comparison of the frequencies for the Advanced students (Table 8D·2) and those for all students (Table 8·4) confirms the predictable, namely, that it was mainly the Advanced students who succeeded in scoring full marks on those items measuring aspects of the highest level of understanding for algebraic symbols.

By synthesizing all these outcomes into summary form, the following hierarchies of learning in terms of mastery of the concepts assessed by the given test questions were identified:

Q.15i > Q.15ii > Q.15iv
Q.15i > Q.6c > Q.15iv
Q.15iii > Q.6c > Q.15iv
Q.10 > Q.6c > Q.15iv
Q.15iii > Q.12 > Q.15iv
Q.10 > Q.12 > Q.13
Q.11 > Q.13
Q.15i > Q.13, and
Q.15iii > Q.13.

These outcomes supported explicitly two hierarchies in terms of Level 5 Scales, namely:

BXBA ($b = x$, $b = a$) Scale > CZ (c = zero) Scale, and
PL Parallel Lines Scale > EQN Equation Scale.

Because the VBL Scale was based on Questions 6 (c), 10, 12, and 15 (iii), they also implied the hierarchies:

VBL Variable Scale > CZ (c = zero) Scale, and
VBL Variable Scale > EQN Equation Scale.

Two Additional Statistical Analyses

Two further statistical methods were used to throw more light on the differences in difficulty level between the tasks required by the items in the five scales at Level 5 for understanding the meaning of algebraic symbols.

The first was to compare success in the five scale measures in terms of the percentage of valid cases who attained at least 50% of the maximum scale score. The comparisons are shown in Figure 8·13 for the Year 7 beginners using their responses on each of the four times they completed the test. Here we have data on the same group of students over a period of time and can probe the evidence for sequences of learning. For any substantiated sequences, we can examine whether or not they complement the earlier findings about hierarchies of cognitive difficulty.

Figure 8·13 graphs the percentage of valid responses that recorded scale scores equal to or greater than half the maximum possible score on each of the Level 5 scales.

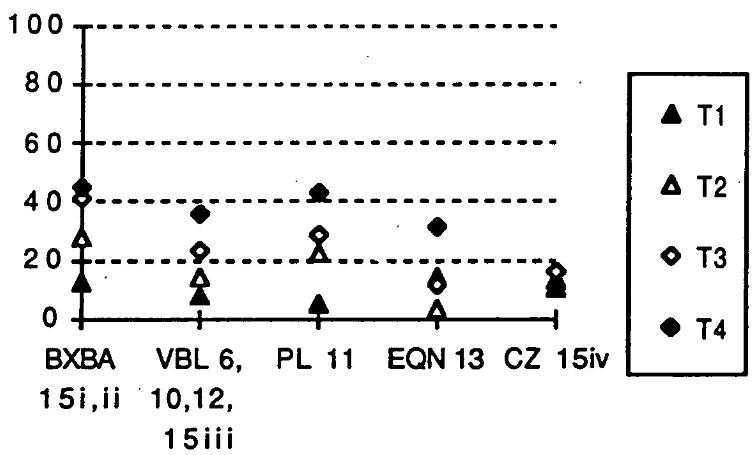


Figure 8·13. Year 7 percentage frequencies for 50% mastery at Level 5 (scores greater than or equal to half the maximum scale score)

The results indicate that progress was made first with the BXBA Scale scores, as noted by the growth from Test 1 to Test 2, and again from Test 2 to Test 3. Little change was recorded from Test 3 to Test 4. Next came better understanding of the ideas tested in the PL Scale. A rapid development from Test 1 to Test 2 was found but then only a few more students joined the 50% or over group in Test 3. Another fairly large improvement came between Tests 3 and 4. Growth reflected in the VBL Scale scores was fairly consistent but did not reach the levels of the BXBA and PL Scale results by Test 4. Development on the EQN Scale was uncertain over the period of intervention teaching from Test 1 to Test 3 and it was only by Test 4 that there were signs of secure advance. With the CZ Scale, scarcely any change was recorded across all four tests. Analysing the scale data in this fashion points to the sequence of development as occurring in the hierarchy:

BXBA (easiest), PL, VBL, EQN to CZ (hardest).

The second additional method of analysing the difficulty levels of the Level 5 scales was to plot the average score per item for each scale. Figure 8·14 shows the outcomes for each of the four tests completed by Year 7 students and for all other students.

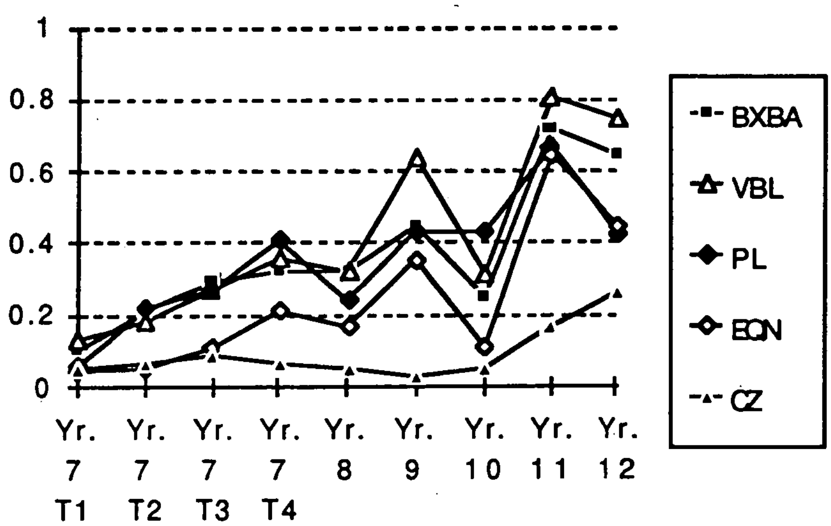


Figure 8·14. Average score per item on Level 5 scales for all students

The graphs in Figure 8·14 indicate that Scale CZ was the most demanding for all student groups, with the EQN Scale being second in difficulty. Year 7 students improved gradually and at approximately the same rate on each of the other three scales for Level 5. Achievement levels for the older classes were somewhat varied but there were indications that the VBL Scale posed the least challenge, followed by the BXBA Scale. It was difficult to rank the PL Scale from these graphs. Thus, this analysis gave the following hierarchical order:

VBL (easiest), BXBA, EQN to CZ (hardest).

A similar analysis was carried out for each of the two extremes of the ability range, namely, the Advanced level classes and the Low ability classes. The outcomes are displayed in Figures 8·15 and 8·16.

For the Advanced students, as Figure 8·15 portrays, success levels with the PL Scale were varied, whereas the other four scales showed a general pattern of improvement which indicated that the hierarchy of difficulty was the same as that derived from the graphs for all students and ranged from:

VBL (easiest), BXBA, EQN to CZ (hardest).

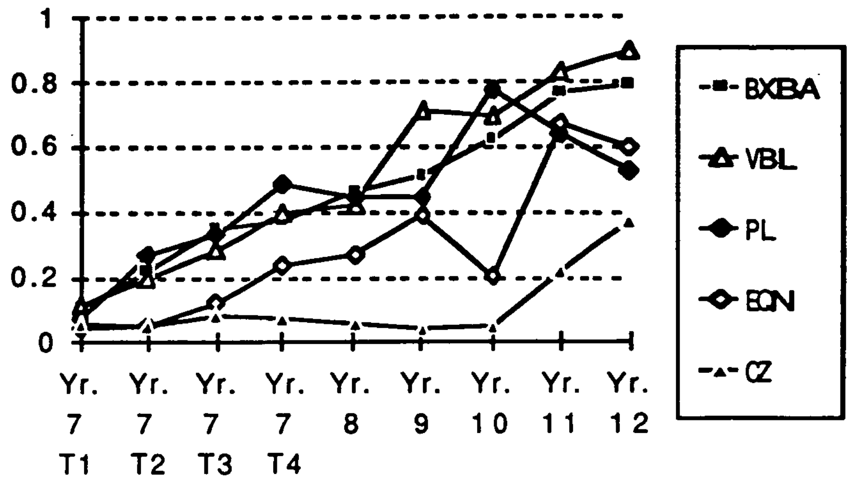


Figure 8·15. Average score per item on Level 5 scales for Advanced students

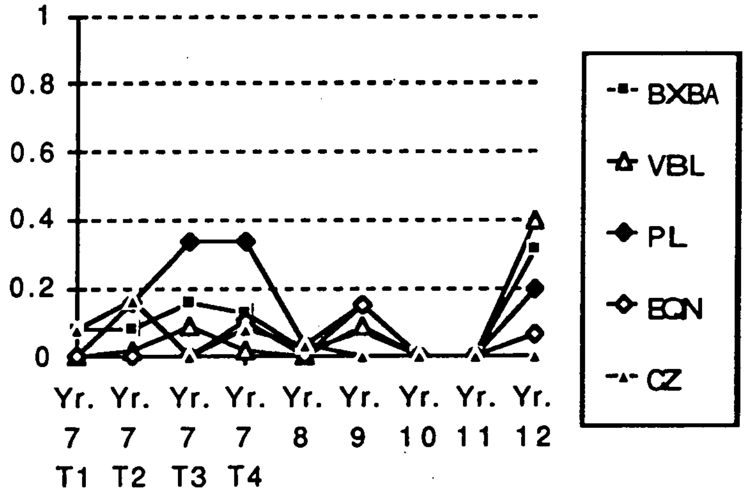


Figure 8·16. Average score per item on Level 5 scales for Low Ability students

The main message from the graphs in Figure 8·16 for the Low ability classes is that, but for occasional exceptions, they found it difficult to score on any of the Level 5 scales. The small group of eight Low ability students in Year 7 did remarkably well on the Parallel Lines Scale in Tests 3 and 4. The order of difficulty for the Low ability Year 12 students in this group was similar to that for the Advanced students although, of course, the level of success was much less. The averages for this Year 12 class reflected the hierarchical order derived from Figures 8·14 and 8·15, and included the ranking of the PL Scale in the centre, giving the following order of difficulty, an order which was compatible with the conclusions drawn:

VBL (easiest), BXBA, PL, EQN to CZ (hardest).

The outcomes from the Figure 8·13 method of analysis showed that the beginning students in Year 7 tended to follow a sequence of learning which placed achievement on the VBL Scale at the middle of the sequence rather than at the start. The same middle ranking showed up clearly in Figure 8·15 for the averages attained by the Advanced Year 7 students in Test 2 and Test 3. Only with further algebra experience did the concepts measured by the VBL Scale rank as the easiest to handle of all those measured by Level 5 scales.

Theoretical viewpoints. Hierarchies for scale measures of Level 5 understandings of algebraic symbols have been discussed above from the point of view of statistics. Further discussion occurs later in this thesis for several of these hierarchical findings. In Section 3 of this chapter, for instance, when a series of propositions is examined, the investigation of Proposition 6 leads to a discussion of the cognitive challenge of Question 13 as compared with that of Questions 10 and 12. The discussion on Proposition 7 considers the memory load for separate parts of Question 15 in terms of the number of steps required for solving each part.

To understand the middle-ranking of the difficulty of the Parallel Lines Task, as measured by the PL Scale, one needs to consider the number of steps required, the degree of cognitive challenge imposed, and that respondents were not familiar with algebra problems in such a geometric context. Students needed to overcome the strong geometric distraction provided by the sketched lines if they were to realize that the symbols '*a*' and '*b*' designated arbitrary values for the number of centimetres of length of the respective lines. Once they came to this realization, the challenge dwindled to simply comparing the values of two arbitrary numerical variables. Hence, it seems that the context made the processing more difficult than comparing, say, the two expressions given in Question 10. Had the latter expressions been used as descriptors for the lengths of two sketched lines, it could be expected that students would then have had more difficulty comparing them. Students found it easier to succeed on the PL Scale than the EQN Equation Scale. The cognitive challenges of Question 13, which supplied the components of the EQN Scale, are discussed in relation to Proposition 6 in Section 3 of this chapter.

Part 4: Comparisons With Other Studies

Divergence from Harper

Harper (1979) provided some information about hierarchies of his subjects' learning with regard to two measures which were included in the 1990 Level 5 group of scale measures. He pointed out that responses in his interviews about the Literal Numbers tasks (e.g., comparing values of '$t + t$' and '$t + 4$') could not, on their own, identify those students who operated at the variable or species level of understanding of algebraic symbols. He found that 100% of those who avoided fictitious ordering of the lengths of the lines in the Parallel Lines tasks gave algebraic answers to the Literal Numbers tasks (Harper, 1979, p. 189). This was a claim that success on the Literal Numbers tasks was a prerequisite for success on the Parallel Lines tasks. Harper applied his finding by using the Parallel Lines tasks for filtering out the non-species users. The 1990 scales incorporated these two types of tasks, thus providing a means for comparison. Question 10 (a contributing item for the VBL Scale) was the first subtask of Harper's Literal Number tasks, and Question 11 (which provided the PL Scale) was the third subtask of his Parallel Lines tasks. Question 12 was available as a further means of comparison, as it was modelled on Harper's Literal Numbers tasks, taking the expressions '$2n$' and '$n + 2$'.

Harper's definition for those who were at the species level was derived after complex explorations of phi-coefficients and included specific achievement levels on the four types of tasks in his study. In the present study, students were not categorized by levels of understanding of symbols. Rather, a variety of scale measures was retained for each student and, from the ordinal scores on each of these scales, a profile was constructed of the types of response levels used. This preserved flexibility in the use of data and allowed for the fact, noted by Harper himself, that "the nature of the tasks cause pupils to switch between different interpretations of the letter" (Harper, 1979, p. 377).

It was found that, as a whole group, the 1990 students did not fit into the clear-cut categories described by Harper with respect to responses to the Parallel Lines tasks and the Literal Numbers tasks. Only 57.3% of those who scored full marks on Question 10 (for comparing values of '$t + t$' and '$t + 4$') had scored full marks for Question 11 (for comparing algebraically the lengths of two parallel lines). To test whether or not the reason for obtaining less than 100% here was due to Harper having selected for his study only those who passed a preliminary task, responses to Questions 10 and 11 were compared for the Advanced students only. Of those who had Question 10 fully correct, 62.4% of the Advanced students had Question 11 fully correct. Smaller percentages were obtained for the beginning students in Year 7, whichever of their four tests was examined. The students did not fit the pattern observed by Harper that success on a Literal Numbers task was a prerequisite for, or even a prelude to, success on a Parallel Lines task. A similar outcome was found using responses to Question 12, which was adapted to the style of a Harper Literal Numbers task. The application of the cross-tabulation method for analysing hierarchies of learning, as explained earlier in this chapter, did not support the contention that high achievement on Question 10 or Question 12 was a prelude to high achievement on Question 11. There was only one significant outcome which included the PL Scale, namely, data from Year 7 students on Test 4 supported the view that success on the Parallel Lines task (PL Scale) was a prelude to success on the Harper Equations task (EQN Scale) given in Question 13. (See Table 8D·1 in Appendix 8D.)

Divergence from Küchemann

The 1990 students differed in their response patterns from the subjects in Küchemann's (1980) study with respect to the following three questions:

(a) "Which is larger, $2n$ or $n + 2$? Explain." (Adapted for the four-part Question 12 and incorporated into the VBL Scale);

(b) "$m + n + q = m + p + q$ is true (a) Always, (b) Never, (c) Sometimes, when ..." (Used with different letters as Question 15 (i), and part of the BXBA Scale); and

(c) "If $c + d = 10$, and c is less than d, what can you say about c?" (Re-worded for Question 6 (c) and given lead-up subparts; used as another part of the VBL Scale).

Harper reported that he had been informed, through personal communication with Küchemann, that "all pupils who answered item (a) correctly also answered items (b) and (c) correctly" (Harper, 1979, p. 283). Using the corresponding items from the 1990 research test instrument, it was found to be untrue that success on Question 15 (i) or question 6 (c) was a prerequisite for, or even a prelude to, success on Question 12 for students in the Quinlan study. The only exception was an interesting case in the testing of Year 7 students before they started their classroom study of algebra: One student had Question 12 completely correct and that student also was among the 11 who had Question 15(i) correct (Table 8D·1 of Appendix 8D). With the Advanced classes, for instance, 77.9% (and not the 100% forecast by the Küchemann data) who had Question 12 correct also had Question 15 (i) correct, and only 54.8% of those who had Question 12 correct also had Question 6 (c) correct, giving four or more values for 'c'. A possible reason for the discrepancies is the contrast between the success rates for Questions (a), (b) and (c) in the Küchemann study, of 6%, 25%, and 30% respectively for Third Year secondary students (aged about 14 years), and the success rates on the corresponding questions in the 1990 data which were, respectively, 27.3%, 40.4%, and 65.4% for all 517 students in the study, and 40.4%, 54.2%, and 37.8% for those in Year 9, who were in their third year of secondary school and also were aged about 14 years. The re-editing of the questions as well as the use of a new batch of subjects (in a different country) could have contributed to the variations in achievement patterns.

Comment. The divergence noted between results obtained from the 1990 students and from Harper's and Küchemann's subjects are reminders that some research results are not generalizable across different student samples. That the particular outcomes mentioned were not replicated in the present study indicates that the outcomes should be treated with caution in the total field of research data. The application of a criterion with a 100% requirement imposed a limitation on the generalizability of the earlier studies and was avoided in the analyses discussed in this chapter.

Proportion who Gain Variable Concept

Harper (1979, p. 275) claimed that "it is probable that the majority of pupils complete secondary school mathematical studies devoid of the symbolic conception of number." This claim was put to the test in the 1990 study. Levels of success on the VBL Variables Scale were taken as a first measure of the degree to which students had developed the concept of a numerical variable (or the symbolic conception of number) in algebra. Three success levels are reported in Table 8·5, namely, 100% success (with all 11 items in the scale correct), greater than 80% (with 9 or more items correct), and greater than 50% (with 6 or more items correct). Percentage frequencies have been assembled in the table for each Year Group from Year 7 to Year 12, subdivided according the rating given by the school to the various classes for mathematical ability.

Table 8·5

Percentage Frequencies for Three Levels of Success on the VBL Variables Scale

CLASS ABILITY Rating	VBL Scale SUCCESS (%)	CLASS YEAR GROUP / TEST								
		7 T1	7 T2	7 T3	7 T4	8	9	10	11	12
ADV	100	0	0	1.74	1.87	0	9.86	11.1	29.8	52.6
	> 80	0	2.13	6.09	15.9	16.7	56.3	33.3	74.5	73.7
	> 50	6.25	13.8	24.3	37.4	33.3	81.7	88.9	93.6	100
MIXED	100	0	0	2.00	5.77	-	0	0	11.8	-
	> 80	4.00	6.00	16.3	13.5	-	12.5	0	47.1	-
	> 50	12.0	16.0	24.5	34.6	-	25.0	0	94.1	-
LOW	100	0	0	0	0	0	0	-	-	0
	> 80	0	0	0	0	0	0	-	-	12.5
	> 50	0	0	0	0	0	0	-	-	50.0

Students in the Low ability classes, apart from those in Year 12, did not succeed on the VBL Scale at any of the chosen levels and would, it seems, be likely to leave school without the variable concept.

Except for the Year 10 group, at least some students in Mixed ability classes showed signs of progress with an understanding of algebraic symbols at the variable level. At the 100% success level in responding to the items in the VBL Scale, there were nearly 6% of the Year 7 Mixed ability students and almost twice this percentage of the Year 11 Mixed ability students. Nearly half of the Year 11 Mixed ability students attained the over-80% success level and all, except about 6%, reached the 50% success level. The vast majority of these Year 11 students, then, would leave school with at least a reasonable understanding of the variable concept. Just over one-third of the Year 7 Mixed ability classes had reached the over-50% success level after about six months of algebra, as shown by their Test 4 result.

As regards the Advanced level classes, Table 8·5 details the rise noted earlier in the development of the variable concept between Years 8 and 9. The figures for Year 8 Advanced are much lower than those for the older Advanced classes. The Year 7 Advanced group performed at a slightly better rate overall than the Year 7 Mixed ability group, as the graphs of scale score averages in Figure 8·3 testify. However, Table 8·5 records that, in Year 7, the Mixed ability classes did slightly better than the Advanced classes at the 100% success level.

When additional criteria were imposed in the form of the success rate on the other variables scales at Level 5, it was found that the percentages decreased compared with those for the VBL Scale on its own but they still identified a reasonable proportion of Advanced students who understood the variable concept.

Harper used his set of Parallel Lines tasks as a filter for identifying those who developed the concept, because he found that he could not identify them from the Literal Numbers tasks alone. The VBL Scale was composed largely of questions similar to the Literal Numbers tasks. Thus cross-checking against the Parallel Lines task (Question 11) was the most direct way to compare data from the two studies on the point at issue. The outcome was as shown in Table 8·6.

Table 8·6

Percentage Frequencies for Success on VBL Variable and PL Parallel Lines Scales for Advanced Classes from Years 9 to 12 (n = 146)

Scale Success (%)	Scale VBL only	Scale VBL & Scale PL
100	21.9	9.59
≥ 80	63.0	31.5
≥ 50	88.4	58.2

Nearly one-tenth (9.59%) of the subgroup of Advanced students achieved 100% success on both scales and could be confidently classified as having a well-developed sense of algebraic symbols as numerical variables. Nearly one-third succeeded on both scales at the 80% mark, showing a strong grasp of the variable notion. Over half of the group scored at least half of the scale maximum for both measures because they had at least a reasonable grasp of the concept.

The percentages were closely similar to those in Table 8·6 when the EQN Equations Scale was superimposed on the VBL Variables Scale responses, as indicated in Table 8E·1 of Appendix 8E. When the combination of the VBL Scale and the BXBA Scale was used, as summarized in Table 8E·2, the percentages were higher than the corresponding percentages in Table 8E·1. For the Advanced students from Years 9 to 12, the SYM Symbols Scale, which tallied the number of times students correctly used symbols in answers to Questions 5, 9, and 14, was called into this investigation. Leaving an answer as an unclosed operation on symbols for which the values were not known gave a different type of measure of the students' acceptance of algebraic symbols as mathematical entities representing numerical variables. The percentage frequencies (tabulated in Appendix 8E, Table 8E·3) were equal to or close to those obtained when only the VBL Scale was used as a measure. It was, therefore, unnecessary to adjust the above comments made on the basis of responses to the VBL Scale alone.

Harper's comment that the majority of school leavers probably have not learnt the symbolic conception of number could still apply to the population represented in the present study. However, it has been shown that a considerable number of the students were learning the concept before they left school. Judging from the information in Table 8·6 and Appendix 8E, about 60% of the Advanced students from Years 9 to 12 in the 1990 study achieved at least half the maximum score on two or more scale measures of the variable notion. As regards the 39.5% of students who were in classes other than Advanced level, Table 8·5 points out that differing proportions of subgroups gained a reasonable grasp of the concept, especially those with more years of secondary schooling.

The evidence presented points to the likelihood that the majority of the Advanced students, as well as some of the others, would leave school with at least a reasonable understanding of the concept that symbols in the algebra of generalized arithmetic represent numerical variables. Relevant to this context is the knowledge that almost all Australian school students complete at least Year 10, and about 25 to 30% of Years 9 and 10 students in N.S.W. take Advanced level mathematics.

Section 3: Nine Propositions About Hierarchies of Learning

Consideration of the cognitive difficulties associated with various problems included in the test led to the formulation of nine propositions, the first of which is called Proposition 2, as Proposition 1 was introduced in Chapter 6. The test data were applied to examine whether or not the propositions could be supported empirically. If they could, then the research project would have provided support for the psychological reasoning which, from a theoretical point of view, had led to the propositions. The following are the propositions to be examined:

Proposition 2: *Success in interpreting algebraic expressions is a prelude to success in comparing the values of two expressions or comparing the values of two variables within the one equation.*

Proposition 3: *Success in solving a simple linear equation is a prelude to success in identifying the conditions for equality of two algebraic expressions or for the equality of two variables within the one equation.*

Proposition 4: *Success in identifying the conditions for equality of two algebraic expressions (or two variables in the one equation) is a prelude to success in identifying the conditions for one expression (or variable) to be greater than the other.*

Proposition 5: *Success in one question is a prelude to success on a similar question which entails more arithmetic operations than the first.*

Proposition 6: *Success in comparing values of two simple linear expressions is a prelude to success in comparing values of two variables within a linear equation.*

Proposition 7. *Success on a question is a prelude to success on another question of a similar type but requiring more cognitive steps.*

Proposition 8. *Success in interpreting algebraic expressions is a prelude to success in using symbols correctly in problems requiring algebraic answers.*

Proposition 9. *Success in using symbols correctly in problems requiring algebraic answers is a prelude to success in understanding the meaning of algebraic symbols at the level of generalized number or at the level of variables .*

Proposition 10. *Success in ordering two variables confined to an abstract multiplicative relationship is a prelude to success in interpreting the meanings of symbols describing a multiplicative relationship in a real-life context.*

These propositions will now be examined in turn.

Proposition 2. *Success in interpreting algebraic expressions is a prelude to success in comparing the values of two expressions or comparing the values of two variables within the one equation.* Question 3 tested students' ability to interpret a set of six linear expressions, all functions of 'y', by asking them to substitute the value '3' in place of 'y', an exercise found to be one of the easiest tasks in the test. From a cognition point of view, the tasks led students away from symbols back to arithmetic once they had made the substitution for 'y'. They were more familiar with arithmetic and the main challenge was to be able to interpret the conventions for writing the given algebraic expressions. According to Proposition 2, students should first have the ability to succeed on Question 3 if they were then to be able to succeed on questions such as 6 (c), 10, 12, 13, and 15, each of which demanded the ability not only to understand the conventions for writing algebraic expressions, but also to compare the relative sizes of either two expressions or two variables. Statistical support for this logical conclusion was abundant. Table 8·7 illustrates the strength of the support from responses by all students, as can be gauged by the high values of %AB and the ratio of %AB to %BA.

Table 8·7
Evidence for Proposition 2 from responses by all students

Variables		Frequencies			Percentages		Ratio	Correlations	
A	B	High A & B	High A	High B	%AB	%BA	%AB / %BA	r	p
Q.3ii	Q.13b	70	393	71	98.6	17.8	5.54	.128	**
SUB	EQN	54	317	60	90.0	17.0	5.28	.263	***
Q.3ii	Q.13d	75	393	75	100	19.1	5.24	.154	***
SUB	VBL6c	61	319	81	75.3	19.2	3.94	.093	*
SUB	VBL	84	319	92	91.3	26.3	3.47	.385	***
SUB	BXBA	85	322	95	89.5	26.4	3.39	.304	***
3ii	Q.12c	153	421	153	100	36.3	2.75	.258	***
SUB	Q.12	112	331	123	91.1	33.8	2.69	.368	***
3ii	Q.13c	153	393	153	100	38.9	2.57	.248	**
3ii	Q.12b	167	421	168	99.4	39.7	2.51	.263	***

Note. $N = 517$. Test 3 responses for Year 7. Ordered by size of ratio %AB to %BA. *** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Two measures for students' level of understanding of algebraic expressions are used in Table 8·7, namely, "3ii", or Question 3 part (ii), which was a record of whether or not the students could correctly substitute into the expression '$2y + 5$', and "SUB" which was the total score on all six parts of Question 3. Table 8·7 presents evidence that success in interpreting an expression such as '$2y + 5$' was a prelude to success with Questions 12 (b), 12 (c), 13 (b), 13 (c), and 13 (d), all of which required students to interpret algebraic expressions as a first step towards solving a more difficult problem. Evidence is also tabulated that success on all parts of Question 3 was a prelude to success on the BXBA '$b = x, b = a$' Scale (based on part (iv) of Question 15), the VBL Variables Scale (and its subsidiaries, Question 12 and Question 6 (c) when answered at the variables level), and the EQN Equations Scale, which tallied scores on all four parts of Question 13.

Table 8·8 displays further statistical support for Proposition 2 from data provided by the Year 7 students on Tests 2, 3, and 4. Amongst the 23 outcomes which met the criteria established earlier, several merited mention three times. In all three tests, these beginning students showed that, for them, success on Question 3 (ii) was a prelude to success on Questions 10 (b) and 10 (c), two questions about conditions for the inequality of two algebraic expressions. Furthermore, in all three tests, their success on Question 3 (ii) was a prelude for their success with Questions 12 (d), 15 (i) and 15 (iii), all questions in which conditions for the equality of two algebraic expressions were to be identified. Question 3 (ii) success, in the three tests, was a strict prerequisite (as %AB was 100%) for success in deciding when equality held for the two variables in the equation used for Question 13 (c). Only the last one of these outcomes appeared in Table 8·7 as all the others did not meet the stated criteria when the analysis included the more experienced students who generally performed well on the other questions previously mentioned. In the very early stages of learning, success on 3 (ii) was a prerequisite for, not merely a prelude to, success on 15 (ii), as shown by %AB at 100% in the outcomes reported for Tests 2 and 3 responses from the Year 7 beginners.

It is instructive to compare the frequencies and the values of the ratio of %AB to %BA for the Variable B entries which occur for more than one test. These statistics indicate the likelihood of sequential learning paths. As the ratio of %AB to %BA is actually the ratio of the frequencies for "High on A" to "High on B", it provides a readily-available comparison of the degree of success with the two variables under consideration: the larger the ratio, the greater the difference between the degrees of success. Changes in the value of this ratio from one test to the next record the comparative rates of development of the concepts and/or skills being measured by the variables A and B. In the case of Item 13 (c), the ratio of %AB to %BA underwent the following changes from Test 2 to Test 4: 9.10 to 5.46 to 2.68. These figures show that, as time progressed, the students' level of expertise with Item 13 (c) was gradually catching up with that for Item 3 (ii). They had developed the ability to interpret an algebraic expression (as measured by Item 3 (ii)) before the ability to decide when two variables in one equation were equal (as measured by Item 13 (c)).

Table 8·8

Evidence for Proposition 2 from Year 7 responses

Source	Variables		Frequencies[a]			Percentages		Ratio	Correlations[b]	
(Test)	A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%AB}{\%BA}$	*r*	*p*
T2	3ii	10b	33	101	42	78.6	3207	2.41	.261	***
T3	3ii	10b	58	166	61	95.1	34.9	2.72	.189	**
T4	3ii	10b	68	160	71	95.8	42.5	2.25	.159	*
T2	3ii	10c	30	101	35	85.7	29.7	2.89	.302	***
T3	3ii	10c	51	166	52	98.1	30.7	3.19	.219	***
T4	3ii	10c	62	160	65	95.4	38.8	2.46	.138	*
T2	3ii	12d	26	104	28	92.9	25.0	3.71	.316	***
T3	3ii	12d	49	166	50	98.0	29.5	3.32	.208	**
T4	3ii	12d	62	161	62	100	38.5	2.60	.243	***
T2	3ii	13c	10	91	10	100	11.0	9.10	.227	**
T3	3ii	13c	28	153	28	100	18.3	5.46	.166	*
T4	3ii	13c	56	150	56	100	37.3	2.68	.227	**
T2	3ii	15i	35	102	45	77.8	34.3	2.27	.275	***
T3	3ii	15i	55	175	59	93.2	31.4	2.97	.124	*
T4	3ii	15i	68	162	71	95.8	42.0	2.28	.181	**
T2	3ii	15ii	7	98	7	100	7.14	14.0	.188	**
T3	3ii	15ii	26	170	26	100	15.3	6.54	.155	*
T2	3ii	15iii	18	101	18	100	17.8	5.61	.304	***
T3	3ii	15iii	33	174	34	97.1	19.0	5.12	.139	*
T4	3ii	15iii	46	159	47	97.9	28.9	3.38	.177	**
T2	3ii	Q.10	20	101	22	90.9	19.8	4.59	.330	***
T3	3ii	Q.10	38	166	38	100	22.9	4.37	.251	***
T4	3ii	Q.10	57	160	58	98.3	35.6	2.76	.218	**

Note. Multiple entries for Variable B ordered by size of ratio %AB to %BA.
[a] Frequencies tally only numbers of students giving *correct* responses (See p. 251).
[b] Correlations take account of *all* valid responses, correct and incorrect (See p. 253).
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Table 8·8 similarly records that development of expertise with Item 15 (iii) lagged behind that for Item 3 (ii), with the ratio for success changing from 5.61 to 5.12 to 3.38 over the interval from Test 2 to Test 4. With Item 10, a similar progression was recorded for the "Q.10" entries, which reported the overall score on the four parts of the item, whereas the progressions for parts (b) and (c) of the item were not so clearly ordered, perhaps indicating that, for at least some students, the degrees of difficulty were somewhat similar for Item 3 (ii) and parts of Item 10.

In general, the ratio at Test 4 was smaller than for earlier tests, the only exception being the case of Item 15 (i), when the Test 2 ratio was 2.27 and the Test 4 entry was 2.28. The conclusion is that, although the table detailed only cases in which the rate of success on Item 3 (ii) was at least twice as great as for other items, the level of success on the latter items was gaining on that for the former.

The outcomes imply that the beginning students were helped along the road to progress in algebra by learning how to substitute in simple algebraic expressions. Once the success rate on interpreting a simple algebraic expression at this task level was improved, the students were more likely to manage problems which involved such expressions. This illustrates the effectiveness of reducing the cognitive load involved in simply understanding the expressions so as to free up mental space for dealing with problems about algebraic expressions. As Halford and Boulton-Lewis (1989) would argue, understanding an algebraic expression implies mapping to the corresponding arithmetic structure. Once such mappings are learned "they no longer impose a processing load" (p. 30). Recoding such correspondences in an abstract form "reduces the processing load once the abstraction is achieved, but ... processing loads can be high during acquisition because of the correspondences that must be recognized" (p. 30).

Substitution in an algebraic expression is at Level 6 on Fischer's scale (Fischer, 1980, p.494) and success here opens the way to Level 7 and then Level 8, the optimum level for solving problems involving unclosed algebraic expressions, as in the inequality questions within Questions 3 (c), 10, 12, and 13, and the equality problems in Questions 15 (i) and (ii). In Biggs and Collis' terms (1991, p.65), mastery of the meaning for algebraic expressions forms the first unistructural stage of a new learning cycle through multistructural to relational levels. The relational level was certainly involved in Questions 6 (c), 10, 12, 13, and 15. A comparable example of two cycles of learning was identified by Campbell, Watson, and Collis (in press): Mastery of the formula for the volume of a rectangular prism was found to be the first stage of a new cycle of learning when developing the ability to solve problems about the volumes of objects with more complex shapes than that of a rectangular prism. As this latter document suggests, the number of learning cycles identified in any sequence of learning depends on the magnification being used, an idea comparable to the fractal notion so commonly used in the currently popular Chaos Theory: "Fractal means self-similar. Self-symmetry is symmetry across scale. It implies recursion, pattern inside pattern" (Gleick, 1987).

The empirical evidence summarized in Tables 8·7 and 8·8 strongly supported the view expressed in Proposition 2. The two graphs in Figure 8·17 further support this proposition, showing that progress with substitution skills (as measured by the SUB Scale) was more rapid than progress with Question 10, which required students to compare the values of two algebraic expressions. To achieve a suitable level of comparability, the graphs are given for the Advanced classes. Corresponding graphs for all students are given in Figure 8F·1 of Appendix 8F. For the Advanced classes, the graphs show that it was not until about Year 9 that students had similar levels of success on the two measures. As a group, the younger beginners in algebra needed to consolidate their ability to interpret algebraic expressions singly before they could manage problems about comparing two expressions.

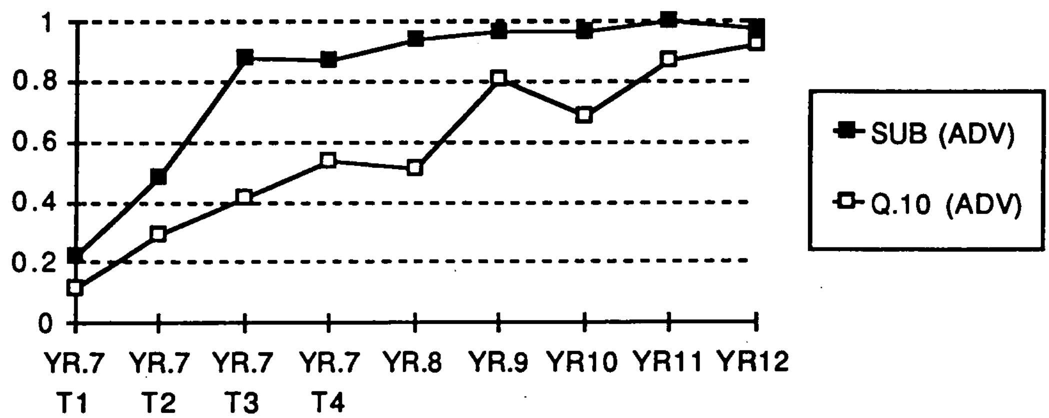


Figure 8·17. Average scores per item by Advanced Classes on SUB Scale and Q.10

Proposition 3. *Success in solving a simple linear equation is a prelude to success in identifying the conditions for equality of two algebraic expressions or for the equality of two variables within the one equation.* To test Proposition 3, responses to Question 8 (b) were used as a yardstick for assessing students' ability to solve a simple algebraic equation. The equation given in the question was '$3a = 36$'. The proposition that success at this level was a prelude to being successful with the equality problems in Questions 12 (d), 13 (c), and the four parts of Question 15 was supported by the analyses registered in Table 8·9. For the beginning Year 7 students in Tests 2, 3, and 4, success on Question 8 (b) was a prelude for success on Questions 15 (iii) and 12 (d). The latter two questions were parallel forms of the same problem, namely, to find out when expressions of the form '$2n$' and '$n + 2$' were equal. As the first six entries under the heading "High B" show, the somewhat bland equation format of Question 15 (iii) proved to be harder in each of the tests than the more wordy form of Question 12 (d) in which an equation was implied but not stated explicitly. Within the first two weeks or so of algebra, success on Question 8 (b) was also a prerequisite to success on Questions 15 (ii) and 13 (c), as shown by the 100%

value for %AB in the reported outcomes from Test 2. It was a prelude to success on the same two questions and on Question 15 (i) at the time of Test 3. The analyses for all students showed success on Question 8 (b) to be a prelude to complete success on the equality problems covered by Question 13 (c) and the CZ and BXBA Scales.

Table 8·9
Evidence for Proposition 3

Source	Variables		Frequencies			Percentages		Ratio	Correlations	
(Test)	A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%AB}{\%BA}$	r	p
T2	8b	15iii	17	75	18	94.4	22.7	4.17	.36	***
T3	8b	15iii	29	127	34	85.3	22.8	3.74	.197	**
T4	8b	15iii	43	145	47	91.5	29.7	3.09	.155	*
T2	8b	12d	22	80	28	78.6	27.5	2.86	.300	***
T3	8b	12d	42	126	50	84.0	33.3	2.52	.227	***
T4	8b	12d	56	144	62	90.3	38.9	2.32	.163	*
T2	8b	15ii	7	72	7	100	9.72	10.3	.247	***
T3	8b	15ii	22	125	26	84.6	17.6	4.81	.164	*
T2	8b	13c	10	74	10	100	13.5	7.40	.280	***
T3	8b	13c	26	119	28	92.9	21.9	4.25	.226	***
T3	8b	15i	44	127	59	74.6	34.7	2.15	.138	*
ALL	Q.8b	CZ	22	357	22	100	6.16	16.2	.102	**
ALL	Q.8b	BXBA	90	352	95	94.7	25.6	3.71	.281	***
ALL	Q.8b	Q.13c	145	347	153	94.8	41.8	2.27	.265	***

Note. Multiple entries for Variable B are ordered by size of ratio %AB to %BA.
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Table 8·9 summarizes analyses which supported Proposition 3. As explained in the discussion on Table 8·8, the values of the ratio of %AB to %BA for those items which qualified for multiple entries to such a table provide a record of the comparative rates of success on pairs of items from one test to the next. In every case reported in Table 8·9, there was a gradual narrowing of the differences in the rates of success with Item 8 (b) and with the other Variable B items that are reported for more than one test. For instance, the ratio of success rates for Item 8 (b) compared with Item 15 (iii) changed from 4.17 to 3.74 to 3.09 from Test 2 to Test 4. These ratios showed that, although Item 8 (b) was consistently handled the more successfully, the students were gradually becoming successful with both Item 15 (iii) and Item 8 (b). The sequence of learning was mastery of the equation given in Item 8 (b) and then mastery of the challenge in Item 15 (iii) to identify conditions for the equality of two expressions.

The equation '$3a = 36$' was classified as "simple" because the unknown appeared in the left side of the equation and the right side gave simply the result of a single arithmetic operation on the unknown. The correct written answer did not reveal the method used. The equation could be solved by, as Kieran (1981b, p. 162) worded it, "plugging-in [trial and error] ... number facts ... [or] undoing" , or "backtracking" as Lowe and Lovitt (1984, Lesson A7#7) called it, or by using the more sophisticated notion that division is the inverse of multiplication, as explained by Collis and Biggs (1979, pp. 97 - 101). Success was possible if the equals sign was regarded as an operator, leading to the statement of a result. At least one algebraic variable appeared on the right side as well as on the left side of the equations explicitly given in the four parts of Question 15 and those implied in Questions 10 (d) and 12 (d). As Kieran (1981a, p. 321) pointed out, this factor "would add to the cognitive strain", particularly as it was likely that students had not been prepared for equations that had something other than a result on the right side. For success, students now needed to regard the equals sign as a symbol for equality and not simply as an operator symbol. The outcomes reported in Table 8·9 support such an analysis of the hierarchy of difficulty. They also show that solving the simple equation posed less cognitive challenge than did Question 13 (c) which required the identification of the condition for equality of the two variables in the equation '$2x + y = 9$.'

Figure 8·18 shows that the Advanced classes, especially in Year 7, improved in their ability to solve a simple equation (viz., '$3a = 36$' in Question 8 (b)) more rapidly than they improved in the cognitive skills needed to solve equality problems involving two expressions or two variables within one equation. The graphs record variations in outcomes recorded by the Advanced classes for Question 8 (b) and for the EQL* Adjusted Equality Scale. A student's EQL* Scale score was calculated as that student's score for the EQL Equality Scale minus the student's score for Question 8 (b). Corresponding graphs for all classes were found to be similar and are given in Figure 8F·2 of Appendix 8F.

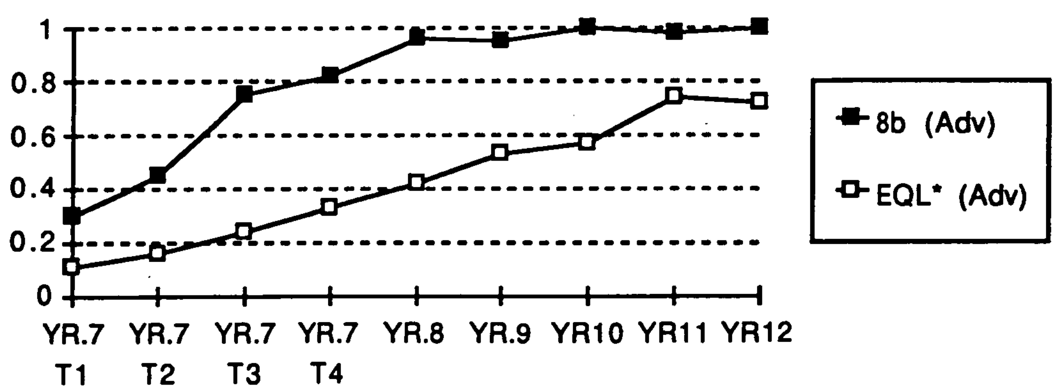


Figure 8·18. Average Scores/ Item by Advanced Classes on Q.8 (b) and EQL* Scale

Proposition 4. *Success in identifying the conditions for equality of two algebraic expressions (or two variables in the one equation) is a prelude to success in identifying the conditions for one expression (or variable) to be greater than the other.* This proposition was tested by examining cross-tabulations using responses to parts of the same question. These allowed a comparison of success rates for identifying equality conditions with success rates for identifying inequality conditions within the same problem context. Suitable questions were Questions 10, 12, and 13. Table 8·10 summarizes the outcomes which succeeded in meeting the statistical criteria given early in the chapter.

Table 8·10
Evidence for Proposition 4

Source	Variables		Frequencies			Percentages		Ratio	Correlations	
(ALL, or Test)	A	B	High A & B	High A	High B	%AB	%BA	%AB / %BA	r	p
T4	13c	13b	10	56	12	83.3	17.9	4.67	.292	***
T3	12d	12c	19	50	21	90.5	38.0	2.38	.516	***
ALL	13c	13b	64	153	71	90.1	41.8	2.16	.510	***
ALL	13c	13d	71	153	75	94.7	46.4	2.04	.570	***
T4	12d	12c	27	62	31	87.1	43.6	2.00	.505	***
T2	12d	12c	13	28	14	92.9	46.4	2.00	.622	***

Note. ALL = the whole student sample, using Test 3 responses for Year 7. Ordered by size of ratio %AB to %BA.
*** $p \leq .001$.

The results for Question 12 supported the view that success with part (d), which tested the ability to identify when two algebraic expressions were equal, was a prelude to success with part (c), which asked for the identification of when the same expressions were ordered so that '$n + 2$' was larger than '$2n$'. Part (b) did not qualify for inclusion in Table 8·10 as the success rate for ordering them the other way around was more than half the success rate for part (d).

The results for Question 13 showed that success in identifying when 'x' and 'y' were equal in the equation '$2x + y = 9$' was a prelude to the identification of when the two variables were ordered one way or the other.

In the case of Question 10, the success rate for part (d) was consistently higher than for parts (b) and (c), showing that the equality aspect of the problem was easier than the inequality aspect. Since the rate for part (d) was not at least twice that for the other parts, Question 10 did not meet the imposed criteria for inclusion in the table. However, the analysis of Question 10 according to Fischer's Skill Theory, as set out on pages 270 to 272 above, applies in similar fashion to Questions 12 and 13. To succeed with the inequality problem in each question, the skill of defining the conditions for the equality case needs to be compounded with the skill of identifying the range of possibilities for inequality cases. This detailed analysis helps to clarify the difference in cognitive difficulty between solving the equality and inequality cases. The compounding process in the case of deciding conditions for 'x' to be greater than 'y' in Question 13, for instance, could be summarized in Fischer's symbolism as:

$$[^8E_{x=3;\, y=3}] + [^8E_{x \neq 3;\, y \neq 3}] = [^8E_{x>3;\, y<3}].$$

Figure 8·19 displays graphs which record that, for Advanced classes, the degree of success on Question 12 part (d), the equality case, was consistently greater than success rate on part (c), the inequality case. Corresponding graphs for all classes appear in Figure 8F·3 in Appendix 8F.

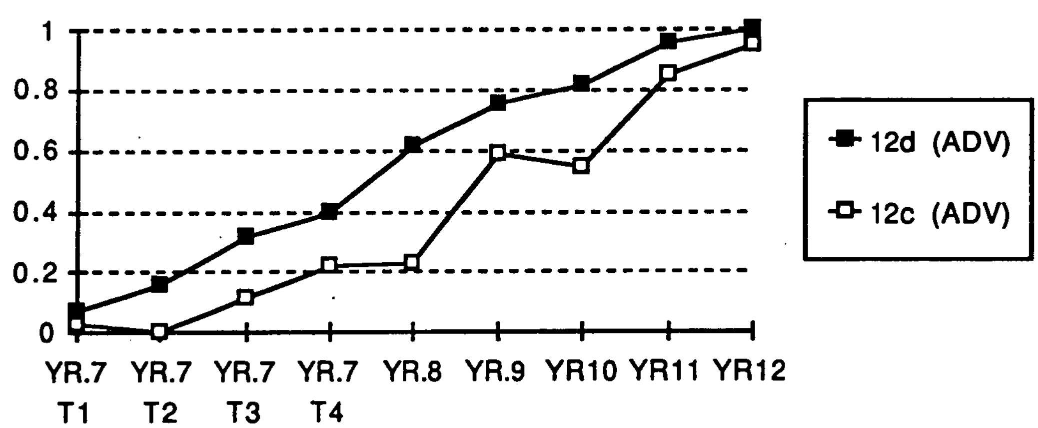


Figure 8·19. Average scores by Advanced Classes on Questions 12 (d) and 12 (c)

The outcomes gave empirical support for the proposition that the identification of the boundary case of equality is a step which is logically necessary before the identification of some inequality condition in the types of questions examined. Once the equality condition is identified, further steps are required for deciding the inequality cases. Thus demand for short term storage space (Case, 1985, p. 290) is greater for the inequality problems than for the equality problems. The cognitive load for the early step needs to be reduced if success is to be attained at the inequality stage (cf. Halford & Boulton-Lewis, 1989, p. 31). The detailed examination of the cognitive processing required for Question 10, given in Part 2 of Section 2 above, is relevant here.

Proposition 5. *Success in one question is a prelude to success on a similar question which entails more arithmetic operations than the first.* The evidence that success on Question 10 was a prelude to success on Question 12 was strong as it came repeatedly from the Year 7 students. Their responses on Tests 2, 3, and 4 affirmed that these beginning students were more likely to attain success on Question 12 if they had already attained success on Question 10. Such a finding deserved investigation. Had '$2n$' in Question 12 been written as '$n + n$', then the question would have paralleled Question 10 by replacing a comparison of '$t + t$' and '$t + 4$' with a comparison of '$n + n$' and '$n + 2$'. However, none of the written responses and none of those interviewed expressed '$2n$' in this way. The difference in difficulty is, therefore, credited to the use of '$2n$'. Table 8·11 lists the statistical details.

Table 8·11
Evidence for Proposition 5 from Year 7 responses

Source	Variables		Frequencies			Percentages		Ratio	Correlations	
(Test)	A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%AB}{\%BA}$	r	p
T2	Q.10	Q.12	5	22	6	83.3	22.7	3.67	.544	***
T2	10b	12c	10	42	14	71.4	23.8	3.00	.329	***
T3	Q.10	Q.12	12	37	14	85.7	32.4	2.65	.602	***
T4	Q.10	Q.12	20	58	22	90.9	34.5	2.64	.672	***
T2	10c	12c	10	35	14	71.4	28.6	2.50	.380	***
T4	10b	12c	29	71	31	93.6	40.9	2.29	.501	***
T2	10d	12d	26	59	28	92.9	44.1	2.11	.547	***
T4	10c	12c	30	65	31	96.8	46.2	2.10	.573	***

Note. Ordered by size of ratio %AB to %BA.
*** $p \leq .001$.

Collis (1975b) carried out an experiment which controlled for level of abstractness of the elements in given mathematical tasks and the level of structure represented by the operations on the elements. He found evidence for "operations being the prime cause of difficulty of items" (p. 91) although he drew attention to the interaction between the two dimensions (cf. Collis, 1978, p. 231). In Questions 10 and 12, the nature of the elements used was the same, but Question 10 kept to addition and Question 12 introduced multiplication with addition. As reported in Tables 4·20 and 4·25 of Chapter 4, just over 40% chose one or other expression as the larger in Question 12 (a) but less than half that number did the same in Question 10 (a). A common argument in Question 12 (a) was that multiplication is larger than addition, pointing to the influence of the mixture of operations as the clue to the difference in difficulty levels for these two questions.

Figure 8·20, which displays a comparison of the Advanced classes' success rates on Questions 10 and 12 gave further empirical support for Proposition 5. Graphs for all classes appear in Figure 8F·4 of Appendix 8F. Both sets of graphs recorded that, prior to Year 10, the success rate was higher for Question 10 than for Question 12. Only after Year 9 did Proposition 5 lose its relevance.

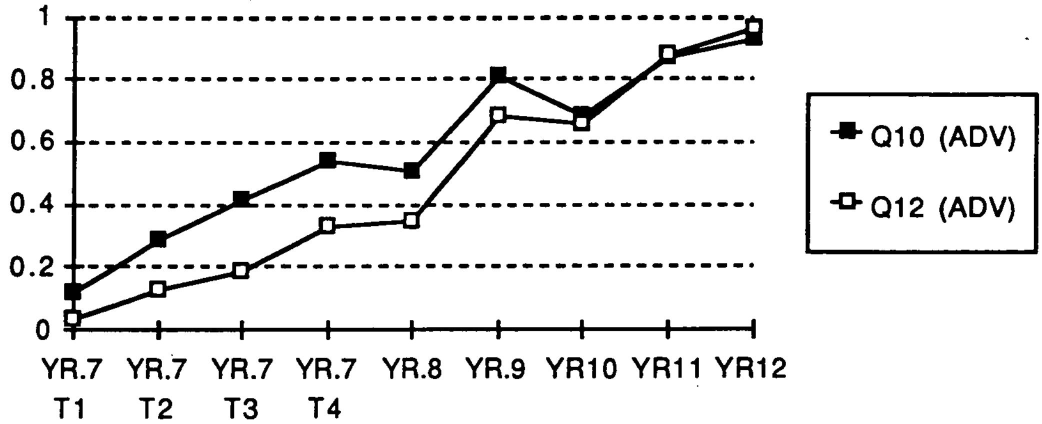


Figure 8·20. Average scores per item by Advanced Classes on Questions 10 and 12.

Proposition 6. *Success in comparing values of two simple linear expressions is a prelude to success in comparing values of two variables within a linear equation.* Data from the Advanced classes (Table 8·12 below) indicated that success on Question 10 was a prelude to success on Question 13 and Question 6 (c) when, for each question, answers were expected to include the possibility of values other than positive integers. All three questions included a decision about inequalities.

In Question 10, the inequality had to be considered between two given algebraic expressions whereas in Question 6 (c), it was between two variables for which the sum was restricted to '10' and in Question 13 it was between the two variables in the equation '$2x + y = 9$'. Although students had been alerted to the implications within the Question 6 (c) problem by the preceding parts of that question, they still found it more difficult than Question 10. Students were able to reason more easily with relationships concerning two covariant expressions than those concerned with two covarying variables within one equation.

The cognitive challenges involved in responding to Question 10 have been discussed at length in Part 2 of Section 2 above. Biggs and Collis (1991) would classify correct algebraic responses to both Questions 10 and 13 at Relational level in the Concrete-Symbolic, or even Formal, mode of functioning. Halford's analysis of Question 10, as explained on page 269 above, shows how the processing can be simplified to a comparison of 't' and '4'. There is no comparable way to simplify the processing of either Question 6 (c) or Question 13. This could well account for their level of difficulty being greater than that for Question 10, if it is assumed that at least some students made use of the simplified approach to Question 10.

Table 8·12
Evidence for Proposition 6

Source (ADV, or Test)	Variables A	Variables B	Frequencies High A & B	Frequencies High A	Frequencies High B	Percentages %AB	Percentages %BA	Ratio %AB/%BA	Correlations r	Correlations p
T2	10b	13b	1	40	1	100	2.50	40.0	.137	.042
T3	10d	13c	20	80	28	71.4	25.0	2.86	.219	***
T4	Q.10	EQN	9	55	9	100	16.4	6.11	.404	***
T4	10b	13b	11	66	12	91.7	16.7	5.50	.293	***
T4	10c	13b	11	60	12	91.7	18.3	5.00	.320	***
T4	10b	13d	13	66	15	86.7	19.7	4.40	.299	***
T4	10c	13d	14	60	15	93.3	23.3	4.00	.372	***
ADV	10b	6c	44	175	58	75.9	25.1	3.02	.165	**
ADV	10c	6c	46	169	58	79.3	27.2	2.91	.219	***
ADV	10b	13c	52	177	63	82.5	29.4	2.81	.237	***
ADV	10c	13b	56	172	63	88.9	32.6	2.73	.321	***
ADV	10b	13d	54	177	66	81.8	30.5	2.68	.236	***
ADV	Q.10	EQN	45	144	54	83.3	31.3	2.67	.363	***
ADV	10c	13d	57	172	66	86.4	33.1	2.61	.303	***
ADV	Q.10	6c	41	143	58	70.7	28.7	2.47	.255	***
ADV	12b	13b	56	141	63	88.9	39.7	2.24	.430	***
ADV	12b	13d	59	141	66	89.4	41.8	2.14	.448	***
ADV	12c	13b	54	130	63	85.7	41.5	2.06	.438	***

Note. Ordered by ratio %AB to %BA. ADV. = Advanced classes, Years 7 to 12.
*** $p \leq .001$, ** $.001 < p \leq .010$.

The application of Fischer's Skill Theory (1980) to Question 10 leads to the following equation, as given on page 271 above, to sum up the compounding of two skills so that the condition for the expression '$t + t$' to be greater than '$t + 4$' can be identified:

$$[{}^8E_{1,t=4} - {}^8E_{2,t=4}] + [{}^8E_{1,t\neq 4} - {}^8E_{2,t\neq 4}] = [{}^8E_{1,t>4} - {}^8E_{2,t>4}].$$

Similarly, as shown on page 292, the compounding needed in Question 13 to identify when 'x' is greater than 'y' is summarized by the equation:

$$[{}^8E_{x=3;\,y=3}] + [{}^8E_{x\neq 3;\,y\neq 3}] = [{}^8E_{x>3;\,y<3}].$$

Comparing these two analyses does not immediately explain why the second process is more difficult than the first. The basic difference, as expressed in Proposition 6, is simply that, in the second case, two variables within one equation are being compared whereas, in the first case, the comparison is between two separate expressions. The data indicate that students found the second comparison more demanding than the first. To identify the equality case in Question 13, for instance, students needed to recognize that '*x*' and '*y*' were numerical variables and that it was possible for them to be equal. They could arrive at the conclusion that each equalled '3' either by solving the equation '$3x = 9$ (= $3y$)' or by trial and error. The steps required for identifying the equality case in Question 10, by either of the methods described on pages 268 and 269, were found to be less difficult. As regards the inequality problem, in Question 10 the expressions both increased or decreased with '*t*' but in Question 13 '*x*' decreased as '*y*' increased. Data were not available to test the association of the difference in difficulty with changes which were in the same or reverse directions. Testing this aspect of cognitive difficulty could well be addressed in some future research project.

Table 8·12 summarizes cross-tabulations which supported Proposition 6. The frequency columns in the table show that success on the measures tabulated was relatively low for the Year beginning students but was higher for the group composed of all the Advanced classes from Years 7 to 12. Nevertheless, in all cases listed, the success rate on Variable A, which was a measure of ability to compare two expressions, was twice the success rate on Variable B, which was taken from the Equations task in Question 13 or from similar type of task in Question 6 (c). All three previously-set criteria were met in each case. A similar table for all students from Years 7 to 12 is presented in Appendix 8F, Table 8F·1.

Figure 8·21 indicates that the success rate on Question 10 was consistently greater than the success rate on Question 13 for the Advanced classes, thus giving more empirical support for Proposition 6. Corresponding graphs for all classes (Appendix 8F, Figure 8F·5) also gave support.

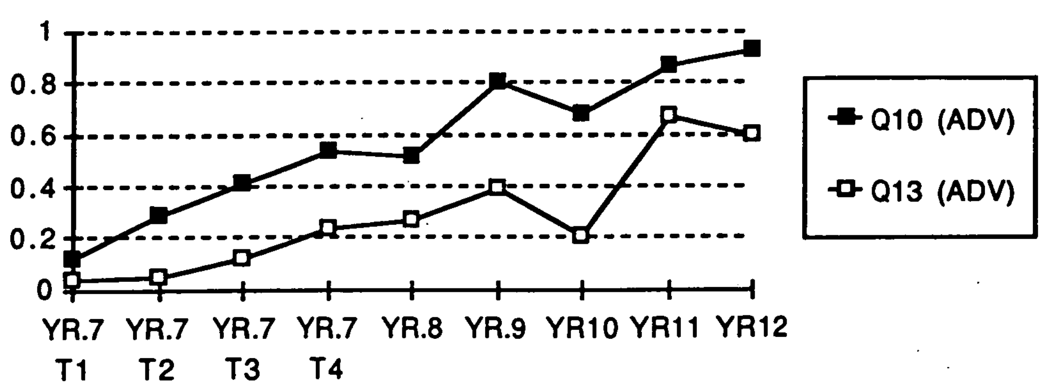


Figure 8·21. Average Scores per Item by Advanced Classes on Questions 10 and 13

Proposition 7. *Success on a question is a prelude to success on another question of a similar type but requiring more cognitive steps.* This proposition is explored in relation to Question 15.

The four parts of Question 15. Question 15 asked students to consider whether the following equations were true always, never, or sometimes and, if the latter, they had to state when they were true (given that the letters stood for whole numbers or zero):

(i) $a+b+c = a+x+c$, giving rise to the "BX" or "$b=x$" measure;

(ii) $2a+3b+7 = 5a+7$, giving rise to the "BA" or "$b=a$" measure;

(iii) $2a = a+2$;

(iv) $a+2b+2c = a+2b+4c$, giving rise to the "CZ" or "c = zero" measure.

The data which supported Proposition 7 in terms of the four parts of Question 15 have been presented amongst other outcomes in Section 2 Part 3 of this chapter and in Appendix 8D.

Evidence that scores on Scale BXBA were a prelude to scores on Scale CZ was recorded in Table 8·4 for all students and in Table 8D·2 of Appendix 8D for the Advanced students. Scale BXBA was formed from responses to the first two parts of Question 15 and Scale CZ from responses to the last part of the same question.

To concentrate attention on testing Proposition 7, the outcomes for relevant hierarchies are assembled here in isolation.

1. All students ($N = 517$, taking Year 7 Test 3 responses):

   Qq.15 (i) and (ii) > Q.15 (iv)

   Q.15 (i) > Q.15ii.

2. Year 7 Test 3:

   Qq.15 (i) > Q.15 (iv)

   Q.15 (i) > Q.15ii.

3. Year 7 Test 4:

   Q.15 (iii) > Q.15 (iv)

   Q.15ii > Q.15 (iv)

   Q.15 (i) > Q.15ii

   Q.15 (iii) > Qq.15 (i) and (ii).

4. All Advanced level classes:

   Q.15 (iii) > Q.15 (iv)

   Q.15ii > Q.15 (iv).

These results indicate that the overall hierarchical order of difficulty for the four parts of Question 15 was:

15 (iii) > 15 (i) > 15 (ii) > 15 (iv).

This order corresponds with the number of steps required (when using an efficient method) for each of parts (i) to (iii), as can be judged by the following analyses.

In each part of Question 15, students need to be aware that an algebraic symbol stands for a numerical variable and can take a range of values, but that if a symbol occurs more than once in any part of the question, it has to keep the same value at any instant. They also need to be able to interpret the meaning of the given pair of algebraic expressions before they can consider whether or not they could be equal. In part (iii), there is only one variable, 'a', and the equation could be solved in one step by subtracting 'a' from both sides, giving the single numerical value, namely '2', which makes it possible for the given expressions ('$2a$' and '$a + 2$') to be equal. Alternatively, it could be solved by trial-and-error or by intuition. Each of the other parts contains more than one variable, leading to the need for more cognitive steps.

In part (i), the steps are:

1. Subtract 'a' from both sides;
2. Subtract 'c' from both sides; and
3. Realize that the remaining equation, namely, '$b = x$', asks whether two numerical variables are able to have the same value, and recognize that it is possible for any value in the range allowed.

In part (ii), the steps are:

1. Subtract '7' from both sides;
2. Subtract '$2a$' from both sides;
3. Divide both sides of the remaining equation, namely '$3b = 3a$', by '3'; and
4. Realize that the remaining equation, namely, '$b = a$', asks whether two numerical variables can have the same value, and recognize that it is possible for any value in the range allowed.

Thus, there is one more step in part (ii) than in part (i), and step 2 is more challenging than the corresponding step for part (i).

In part (iv), a solution could have been obtained with the same number of steps as in part (iv), but an additional difficulty arises after the first couple of steps:

1. Subtract 'a' from both sides;
2. Subtract '$2b$' from both sides;
3. Realize that the remaining equation, namely, '$2c = 4c$', is not true for many values in the range allowed; and
4. Realize that it is true only if 'c' equals zero, a conclusion which eluded all but 22 students. To reach such a conclusion, intuition, trial-and-error, or routes such as the following were used:

4a. Divide both sides by '2', giving '$c = 2c$'; and

4b. Subtract 'c' from both sides, giving '$0 = c$'; and then conclude that 'c' equalled zero.

Alternative explanation. Rather than seek an explanation for the hierarchy of difficulty in the number of steps required, the cognitive processing necessary for the final steps is an alternative source of explanatory power. As Collis has pointed out: "The order of difficulty for me in 5 (i) > 5 (ii) > 5 (iv) lies in the processing in the last step" (personal communication, Dec., 1991). By the "last step", he meant, respectively, '$b = x$', '$3b = 3a$', and '$2c = 4c$'. In part (i), as explained above, the acceptance of '$b = x$' requires an acknowledgement that the symbols stand for any numbers from the given number field (of natural numbers) and that it is possible for both symbols to take the same value. A similar processing procedure is required for part (ii) once the equation '$b = a$' is reached, but the extra step of dividing through by '3' adds to the processing load and contributes to the added difficulty. It seems that the increased difficulty of Question 15 (iv) cannot be accounted for simply by listing the number of steps required. An alternative explanation comes from the consideration that one needs to change one's frame of reference to deal with "the last step", namely '$2c = 4c$'. The steps in part (iv) can be isolated as

'$a = a$' for all possible values of 'a',

'$2b = 2b$' for all possible values of 'b', but

'$2c \neq 4c$' for all possible values of 'c'.

At this point, a change of reference is required. One needs to consider single values of 'c' in the search for some circumstance(s) in which the two expressions would be equal. This leads to the solution '$c = 0$'. The added difficulty could lie in the need for such a change in thought pattern as the solution unfolds.

Conclusion. These analyses of the first three parts of Question 15 give unqualified support for Proposition 7. Varying degrees of the processing load were identified once the number of variables in the given equations was reduced. Increases in the cognitive load (Case, 1987a, p. 590; Halford, Wilson, Guo, Wiles & Stewart, in press) in terms of the number of steps required can account for the hierarchy of difficulty for the first three parts of the question. For placing part (iv) as the most difficult, a different factor, the need for a change in the frame of reference, was seen as a contributing factor.

Proposition 8. *Success in interpreting algebraic expressions is a prelude to success in using symbols correctly in problems requiring algebraic answers.* It would be expected that students needed to understand the conventions used in algebraic expressions if they were to make use of the conventions in formulating answers to problems. This proposition was supported by the data as assembled in Table 8·13. Success on the six substitution exercises in Question 3 (as measured by the SUB Substitution Scale) was found to be a prelude to writing correct algebraic expressions as answers in the five cases tallied by the SYM Symbols Scale. As explained when dealing with Proposition 2, the ability to substitute correctly into an algebraic expression and evaluate it successfully indicates that the conventions used in writing the expressions have been understood. Such an understanding has been shown by the data to be a prelude to applying such knowledge to problems requiring general answers in the form of algebraic expressions. This outcome gave statistical support to Proposition 8.

Table 8·13
<u>Evidence for Proposition 8</u>

Source	Variables		Frequencies			Percentages		Ratio	Correlations	
(ALL, or Test)	A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%AB}{\%BA}$	r	p
T2	SUB	SYM	4	51	4	100	7.84	12.8	.323	***
T3	SUB	SYM	7	113	10	70.0	6.20	11.3	.163	*
T4	SUB	SYM	19	116	23	82.6	16.4	5.04	.367	***
ALL	SUB	SYM	121	312	138	87.7	38.8	2.26	.406	***

*** $p \leq .001$, * $.010 < p \leq .050$.

The cognitive difficulties associated with accepting algebraic expressions as "answers" have been acknowledged in terms of "the process-product dilemma" (Kieran, 1989a, p. 41) or in terms of "acceptance of lack of closure" (Collis, 1975a, pp. 5 - 6; 1978, p. 223). As was indicated in Table 4·18 of Chapter 4, about 10% of students avoided using '*n*' in answers to Question 9 by substituting some number (cf. Firth's study, 1975, p. 40). Table 4·30 recorded that the writing of algebraic expressions as answers in Question 14 parts (ii) and (iii) was avoided by 12.0% and 11.2% respectively, and a further 2.7% and 2.1% respectively simply wrote that they could not proceed without knowing values for '*f*' and '*g*'. The outcomes summarized in Table 8·13 present empirical evidence for the importance of developing an understanding of the way algebraic symbols are used in algebraic expression if one is to make use of such expressions as the product or end-point of these types of problems.

Halford pinpointed a way of explaining why the use of algebraic conventions in problems is more challenging than simply learning to understand the conventions. When beginners are learning the conventions of using algebraic symbols in general expressions they need to hold sufficient information to check that such representations are consistent. Processing for consistency is no longer needed once the use of the conventions is understood and stored in long-term memory. The gateway is then open to proceed with problem solving that involves applications of algebraic conventions.

> The requirement to check for consistency of representation is a kind of 'gate' through which a person must pass when first acquiring a new concept or when first learning to represent a particular situation. It will not operate once the concept is acquired or the situation is understood.
> (Halford, 1982, p. 86)

Proposition 9. *Success in using symbols correctly in problems requiring algebraic answers is a prelude to success in understanding the meaning of algebraic symbols at the level of generalized number or at the level of variables.* It would appear logical to expect that students could show a reasonably well-developed understanding of the meaning of algebraic symbols as a prelude to making use of symbols in answers requiring a generalization to be stated in algebraic form. However, several researchers (e.g., Bell, Costello & Küchemann, 1985; Carpenter & Lindquist, 1989; Kieran, 1989b) have suspected that at least some of the students in the studies they examined showed the ability to use symbols, for instance, in manipulating expressions to give equivalent expressions, without understanding that the symbols stood for numerical variables.

Cross-tabulations were examined between scores on the SYM Symbols Scale and other scales which measured certain aspects of students' levels of understanding of symbols. The SYM Scale accumulated scores attained by correctly using symbols to answer general algebraic questions. The outcomes are reported in two tables. The first, Table 8·14, displays evidence that success in understanding some basic concepts about algebraic symbols was a prelude to successfully using symbols in answers. The second, Table 8·15, presents empirical evidence to support the suspicions of others that students succeed in using symbols even though they have not attained a well-developed notion of what the symbols represent.

The AR Arithmetic Scale, based on Question 1, measured success in interpreting and solving the relationship '3 * 4 = 6 * *y*' for the two cases in which '*' represented "add" and "multiply". Being successful in handling these arithmetic processes in an algebraic setting was found to be a prelude to success in using symbols to write answers, as measured by the SYM Scale. Question 1 sought specific values for the '*y*' in each of two cases and did not require students to think of '*y*' as a variable capable of representing many values at once. The level of understanding of symbols approximated that of Level 2 in the hierarchy proposed in Table 6·1 of Chapter 6. Attaining this level of understanding of symbols was seen to be a prelude to success on the SYM Scale.

Table 8·14
Evidence Related to Proposition 9

Source	Variables		Frequencies			Percentages		Ratio	Correlations	
(ALL, or Yr.7 Test)	A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%AB}{\%BA}$	r	p
T2	AR	SYM	4	58	4	100	6.90	14.5	.196	**
T3	AR	SYM	9	82	9	100	11.0	9.11	.201	**
T2	AD	SYM	4	64	4	100	6.25	16	.265	***
T3	AD	SYM	9	104	9	100	8.65	11.6	.215	**
T4	AD	SYM	19	111	21	90.5	17.1	5.29	.362	***
ALL	AD	SYM	109	264	122	89.3	41.3	2.16	.347	***
T3	C2	SYM	10	133	10	100	7.52	13.3	.189	**
T4	C2	SYM	22	142	23	95.7	15.5	6.17	.263	***
ALL	C2	SYM	128	347	135	94.8	36.9	2.57	.346	***
ALL	Q.2ii	SYM	128	341	135	94.8	37.5	2.53	.311	***

Note. ALL = Data from Years 7 to 12 students, using Test 3 for Year 7 students.
*** $p \leq .001$, ** $.001 < p \leq .010$.

The AD Scale measured success in recognizing that, given symbols '*a*' and '*d*' as representing "any two numbers" (part (i) of Question 2), there was freedom for the symbols to take on any value and that it was not possible, therefore, to decide the order of size of the two variables. This is an important quality of any pair of true variables: They are unordered until some restriction or relationship is placed upon them. This concept was found to be easier to attain than successfully using symbols in the five answers incorporated in the SYM Scale.

To succeed in the C2 Two '*c*' Values Scale, students were required to accept that '*c*' could take the values '3' and '7.4' if '$c + d = 10$'. Whether or not they considered these as separate possibilities or, more correctly, as merely two examples among many possibilities was not measured by the scale. Success indicated that students had moved towards regarding algebraic symbols as representing generalized numbers by at least allowing the symbol to have more than one value and by including a fractional possibility. This level of understanding of symbols was found to be a prelude to successfully using symbols in answers.

Question 2 (ii) asked students to decide which variable was the larger given two positive numbers related by the equation '$6y = d$'. This question tested students' understanding of the convention that conjoining meant multiplication, and it required them to consider the letters as standing for a range of possible values. This level of thinking, one that bordered on understanding symbols as numerical variables, was found to be a prelude to the cognitive task of correctly using symbols in answers as required by the items in the SYM Symbols Scale. Possibly the main contributing feature for this analysis was the need to understand the conjoining convention. A trial-and-error approach could have been used to solve the problem of which variable was the larger, without the need to operate at the more general variable level.

Hence, the material contained in Table 8·14 shows that the minimal understandings of symbols as measured by the item in Question 2 (ii) and three scales, namely, the AR, AD, and C2 Scales, and were supported as preludes to success with the use of symbols as measured by the SYM Scale. The support for Proposition 8 above shows that understanding basic conventions for writing algebraic expressions is also a prelude to success in the SYM Scale. The cognitive challenge of understanding the symbols at higher levels was, however, found to be greater than that of being able to accept symbols in answers, as Proposition 9 suggests.

Success on the SYM Scale proved to be a prelude to success on the VBL, EQN, EQL, CZ, GNV, FZN, and NBR Scales, as recorded in Table 8·15 for data based on responses of all students. The results listed in this table shed further light on the process of coming to an understanding of the meaning and use of algebraic symbols. In essence, the outcomes meant that the ability to use symbols correctly in writing algebraic expressions as answers to general problems can precede the development of a clear understanding of the meaning of algebraic symbols as representing numerical variables or at least as generalized numbers.

Table 8·15

Evidence for Proposition 9

Source	Variables		Frequencies			Percentages		Ratio	Correlations	
(ALL)	A	B	High A & B	High A	High B	%AB	%BA	%AB / %BA	r	p
ALL	SYM	FZN	10	128	10	100	7.81	12.8	.471	***
ALL	SYM	CZ	18	134	22	81.8	13.4	6.1	.080	*
ALL	SYM	NBR	21	134	27	77.8	15.7	5.0	.324	***
ALL	SYM	EQL	29	127	31	93.6	22.8	4.1	.665	***
ALL	SYM	GNV	25	116	32	78.1	21.6	3.63	.627	***
ALL	SYM	VBL	32	132	37	86.5	24.3	3.57	.606	***
ALL	SYM	EQN	42	132	60	70.0	31.8	2.20	.514	***

Note. ALL = Years 7 to 12 students, with Test 3 results for Year 7 students. Ordered by size of ratio %AB to %BA.

*** $p \le .001$, * $.010 < p \le .050$.

Logically, it could have been expected that students would need to be clear in their minds that the symbols represented numbers before they could consider using them in contexts where such a meaning was implied. This was not found to be the case for the student sample. The NBR Number Scale provided a measure of how well students were able to distinguish between symbols as standing for numbers of objects or people rather than just the objects or people themselves. The evidence is that developing this power of discrimination is not a necessary acquisition before using symbols correctly in the five items assessed by the SYM Symbols Scale. Similarly, students' ability to write symbolic answers is a prelude to their acceptance of non-integral values for algebraic symbols, as measured by the FZN Fractions-Zero-Negatives Scale. Two of the test questions incorporated in the SYM Scale, namely, Questions 5 and 14, dealt with only integral values of symbols, and the remaining question, Question 9, involved a symbol, '*n*', which could take any rational number value. However, success on the SYM Scale is a prelude to success on the FZN Scale. Both the FZN and NBR Scales were rated as measuring understanding of symbols at Level 4 according to Table 6·1 in Chapter 6. The statistics in Table 8·15 report that success on the SYM Symbols Scale is a prelude to success in understanding symbols at Level 4.

The other scales listed in Table 8·15 involved Level 5 understanding of symbols. The GNV Generalized Number and/or Variables Scale spread across Levels 4 and 5, and the remainder were Level 5 measures. Success on the SYM Symbols Scale is seen to be a prelude to success in understanding symbols at Level 5.

This outcome indicates that students found it more difficult to acquire the concept of algebraic symbols as representing numerical variables than to develop skill in correctly using symbols in answers. These analyses do not enable decisions to be made about cause and effect. The fact that success on one task or skill is a prelude to success on another indicates that the cognitive challenge posed for the students by the first was less than the challenge posed by the second. In the case of Proposition 9, the data presented in Table 8·15 simply give support to the view that students found it easier to write symbols in answers than to grasp the concept of a numerical variable in algebraic form. Table 8·14 has supported the view that at least some minimal understanding of the meaning of symbols is a prelude to success in writing answers in algebraic form, but Table 8·15 shows that student success in using symbols is a prelude to developing Level 4 and 5 understanding of symbols, that is, understanding symbols as representing generalized numbers and variables.

Beginning students. The cross-tabulation figures to correspond with Table 8·15 for Year 7 students as they completed the test four times did not produce outcomes which satisfied the criteria for preludes. However, as displayed in Table 8G·1 of

Appendix 8G, correlations between beginners' scores on the SYM Scale and the last four scales listed in Table 8·15 were all positive and statistically significant in all four tests, indicating that students' understanding of algebraic symbols as numerical variables develops as they progress with using symbols correctly.

Acceptance of Lack of Closure. In Table 8G·2 of Appendix 8G it is shown that students found it cognitively less demanding to write answers in symbols, either correctly or incorrectly, than to attain an understanding of symbols at Level 5 (variables) or Level 4 (generalized numbers). Proposition 9 specifically deals with the correct use of symbols. The ALC Acceptance of Lack of Closure Scale recorded the tendency to use symbols in answers, where appropriate, and included recognition for attempting to use symbols even when the expression written was not accurate. The figures in Table 8G·2 show that success on the ALC Scale is a prelude to success in attaining Levels 4 or 5 understanding of algebraic symbols, thus giving supplementary support to the line of thought expressed in Proposition 9.

Lessons from Classes 8, 9, and 10. Classes 8, 9, and 10, matched-ability Advanced Year 7 classes from School D, provided further elucidation of the relative importance of the ability to write answers correctly in algebraic form and the acquisition of the concept of a numerical variable. Between Test 1 and Test 2, students of Class 9 were told the correct answer to Question 5 by their teacher, with hints about answers to the other items in the SYM Scale. The teacher decided to help the students with test items on aspects of algebra which they had not directly been taught, despite the fact that such testing was integral to the research plan. Meanwhile, Class 8 students were taken through an exercise designed specifically to encourage them to accept algebraic expressions as answers. This was the first trialling of this particular activity, a copy of which is included in Appendix 3N Part II, as Question 6. During the same period of time, students in Class 10 were not given particular help towards the acceptance of algebraic expressions in answers since they had spent all their class time on forming generalizations from geometric patterns and had reached the stage of including letter symbols within sentences to describe the generalizations.

There were 25 students in each of Classes 8 and 9 and 26 in Class 10, making comparisons of the frequencies of correct answers simple. Table 8·16 records the frequencies of correct answers for selected questions on Tests 1 and 2. The first three required single answers in symbols, respectively '$p + r$', '$n + 9$', and '$3f + 4g$'. Question 10 was the four-part Harper Literal Number Task testing the concept of a variable, and the table shows how many students had all parts correct.

Table 8·16
Frequencies of Correct Answers by Classes 8, 9, & 10

Question	Test	Class 8 ($N = 25$)	Class 9 ($N = 25$)	Class 10 ($N = 26$)
5	T1	3	3	1
	T2	9	23	4
9 i	T1	3	2	1
	T2	14	8	3
14 ii	T1	3	3	1
	T2	11	11	5
10	T1	0	1	0
	T2	5	5	1

In Table 8·16, the only case in which there was a statistically significant difference between Classes 8 and 9 was for Question 5 on Test 2, as recorded by the chi-square test in Table 8G·3 of Appendix 8G. Of the 24 students in Class 9 who had been told the correct answer, 23 dutifully wrote it down in Test 2 (the other student was absent in the previous mathematics lesson and did not succeed on the question in Test 2). In Class 8, of the 25 who had been taught more indirectly about using symbols in answers, 9 wrote '$p + r$' for Question 5. The technique of telling the class the answer to Question 5 dramatically improved the success rate of class members on that question but did not produce a transfer of similar success to Questions 9 (i) or 14 (ii). For the latter question, the success rate was the same (11 correct) for both classes and, for Question 9 (i), Class 8 (with 14 correct) performed better than Class 9 (with 8 correct). The difference was not statistically significant, the chi-square value for the appropriate two-by-two table being less than 3. The fact that 23 students from Class 9 could correctly use symbols in Question 5 was no guarantee that they were, thereby, showing that they understood the meaning of the symbols. In Test 2, only 5 of the Class 9 students scored full marks on Question 10, a question which provided a measure of their level of development of the concept of a variable. The same proportion of Class 8 students also had all of Question 10 correct. In Test 3, as shown in Table 8G·4, Class 9 maintained their significant advantage over Class 8 for success on Question 5.

There were two statistically significant differences involving Class 10 (chi-square tests are given in Appendix 8G). Class 10 was significantly less successful than Class 9 on Question 5 in Test 2 and had significantly poorer results than Class 8 on Question 9 (i) in Test 2. The difference on Question 5 is understandable as a result of one class being told the answer whereas the other was not. The other outcome may be because Class 8 had covered class activities that were more relevant to Question 9 (i) than had Class 10.

Conclusions. The outcomes from the three classes supply some support for claims made in Chapter 2 that the learning environment can influence learning. All three of the theories of cognition chosen to underpin this study gave credence to such an influence. Tables 8·16 and 8G·3 to 8G·6 supply empirical evidence for this viewpoint. The outcomes from Classes 8 and 9 reinforce acceptance of Proposition 9 that it is possible for students to use symbols correctly in answers without understanding that they represent numerical variables.

Proposition 10. *Success in ordering two variables confined to an abstract multiplicative relationship is a prelude to success in interpreting the meanings of symbols describing a multiplicative relationship in a real-life context.* Proposition 10 was examined by investigating the possibility of a hierarchy of difficulty for Questions 2 (ii) and 7. Success on Question 2 (ii) was identified as a prelude to success on Question 7. The statistical details are displayed in Table 8·17.

Table 8·17
Evidence for Success on Q.2(ii) as Prelude to Success on Q.7

Variables		Frequencies			Percentages		Ratio	Correlations	
A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%AB}{\%BA}$	r	p
Q.2ii	Q.7	99	379	111	89.2	26.1	3.41	.089	*

Note. Years 7 to 12, using Test 3 responses for Year 7 students. * $.010 < p \leq .050$.

Both questions were based on equations of similar form. Question 2 (ii) tested students' ability to select the larger of two positive numbers related in an abstract context by the equation '$6y = d$'. Question 7 asked for the meanings of the symbols in the equation '$S = 6P$' in the context of the professors-and-students problem, a task which could be clarified by identifying which symbol stood for the larger number. Question 7 involved more steps than did Question 2 (ii) and was set in a real-life context in terms of professors and students. The reasons why it proved to be more difficult than Question 2 (ii) could include the additional number of steps and the influence of the distracting nature of the context in Question 7. The latter possibility aligns with MacGregor's (1991) findings in relation to success with algebraic problems. She identified the importance of the influence of psycholinguistic factors such as language processing and comprehension, and the strong urge to represent meaning in a problem while disregarding the syntactic form of the statements given in the question.

Review and Forecast

A selection of algebraic tasks was scrutinized according to the theories of cognition described in Chapter 2 and levels of difficulty were identified in this way. The research data were explored in terms of the hierarchy of difficulty for levels of understanding of algebraic symbols described in Chapter 6. The data gave general support for that hierarchy. Some differences were noted between the findings from the 1990 data and those from earlier sources.

An original statistical method of identifying hierarchies of cognitive difficulty was applied to nine propositions derived from psychological analyses of algebraic tasks given in the test program. For all nine of these investigations, the research data gave empirical support to the theoretical ordering of tasks in terms of cognitive difficulty.

Table 8H·1 in Appendix 8H lists the sources of items used in analyses of all ten propositions treated in this thesis. Aspects were incorporated from previous research investigations of students' understanding of symbols, and interrelationships were considered between measures that were previously restricted to separate studies. In this way, the second objective of the study, as stated on page 17 above, was addressed. The blending of sources added rigour to the analyses and strengthened the credibility of the outcomes.

Chapter 9 presents a fourth study, one which examines the 1990 data more closely. Hierarchies of cognitive difficulty in understanding the meaning of algebraic symbols are pursued further by narrowing the focus to a small selection of key questions and by seeking evidence for sequential learning paths. Finally, by focusing on small subgroups of students, a search is conducted for differences in learning patterns between those beginning students who progressed with algebra and those who did not.

CHAPTER 9

FOURTH STUDY: A STUDY OF DIFFERENCES: DIFFERENTIAL DEVELOPMENT RATES FOR UNDERSTANDING SYMBOLS

Overview

Chapter 6 presented the First Study which explored the range of meanings for algebraic symbols which the students had revealed and the relationship between levels of meaning and degrees of success in algebraic tasks. Analyses and discussions were based on responses given by all students on just one occasion and the major statistical interest was correlation coefficients. In the two investigations reported in Chapter 7, subgroups of students were considered such as those in different Year groups, and consideration was given to variations in responses by Year 7 students from test to test. The Third Study, presented in Chapter 8, looked more closely at the available data. The focus of the analyses was the development across levels of understanding of symbols which were associated with the amount of algebraic experience students had accumulated, as well as the ability ratings of mathematics classes.

In this chapter, the order of magnification is increased still further. Subgroups within the Year 7 cohort are identified in terms of the rates of growth of their development of an understanding of algebraic symbols. The Fourth Study is presented as a search for possible factors which contribute to the differential rates of development of this understanding. An important aspect of the investigations is assessing the relevance of the propositions analysed in the previous chapter to an explanation of the differential growth rates.

Two investigations are reported.

Investigation 1. Responses to Questions 10 and 12 are applied to explicate sequences of learning along pathways towards an understanding of the meaning of algebraic symbols.

Investigation 2. Differences are identified between Year 7 students who progressed rapidly in algebra during the data collection period and those who made very little progress in the same period.

Investigation 1: Levels of Understanding Using Responses to Questions 10 and 12

To elucidate further the cognitive processes by which students develop their levels of understanding of algebraic symbols and to pursue the question of whether or not the hierarchy of levels of understanding of algebraic symbols (as given in Table 6·1 of Chapter 6) delineated a sequential learning path for the students, responses to parts (b) and (c) of Questions 10 and 12 were used as the basis of scales measuring the degree to which students recorded differing and different levels of understanding. These test items were chosen as important indicators of algebraic thinking as they were components of the cluster of variables identified within the factor (or principle component) which explained more of the variance in overall test scores than did any other, as was pointed out on page 156 in Chapter 5. The responses to the four items selected could, moreover, be categorized in terms of all five levels of understanding reported in Table 6·1.

Tables 9A·1 and 9A·2 of Appendix 9A summarize the way the responses to the four items were allocated to the formation of eight scales. The scales so formed were given statistical support of the type discussed in Chapter 5. This support is detailed in Table 9A·3 of Appendix 9A. As the scales were subsets of scales previously discussed, they are given corresponding names and an asterisk is used as a reminder that only four items were used in their formation. A correlational study based on these scales is presented next, followed by other analyses which pursue the issue of sequential learning paths.

Correlational Study of Hierarchies for Understanding the Meaning of Symbols

Table 9·1 records the significant correlations between test totals corrected for scores on all parts of Questions 10 and 12 and six of the scale scores derived from responses to Items 10 (b) and (c) and 12 (b) and (c). Neither of the two scales measuring the frequency of using one or more replacement values correlated significantly with corrected test totals and, for this reason, neither appears in the table. In contrast to Table 6·10 of Chapter 6, Table 9·1 displays correlations of scale scores with the *same* corrected test total for each correlation, since the same test items were used for each scale. The number of students is *constant* throughout the table as the number of missing cases was determined by the same two questions throughout. Hence, the correlation coefficients are more validly comparable with each other than were those reported in Table 6·10. The entries are arranged in order of size of the correlation coefficients to facilitate the legitimate use of the coefficients "in a comparative sense" (McPherson, 1990, p. 496).

Table 9·1

Correlations Between Scale Scores and Corrected Test Totals

Scale for Qq. 10,12	Description	Level of Under-standing (Table 6·1)	Correlation with Corrected Test Total	No.of cases	Signi-ficance
GNV*	Gen.No.or Variable	5 or 4	.755	468	***
VBL*	Variable	5	.655	468	***
GN*	Gen.No.not Variable	4	.223	468	***
INT*	Integers only	4	.213	468	***
PRE*	Prestructural	1	- .523	468	***
NV*	Non-Variable	1	- .315	468	***

Note. Sorted by size of correlation coefficients. Years 7 to 12, using Test 3 responses for Year 7.
Corrected TEST TOTAL = TEST TOTAL *minus* marks from Questions 10 & 12.
*** $p \leq .001$.

The two scales at the top of the list in Table 9·1, namely, the GNV* Generalized Number or Variable Scale and the VBL* Variable Scale had correlations which were much larger than the others. Taking the square of the correlation coefficient as a measure of the proportion of variance that two variables share in common, the GNV* and VBL* Scale scores are seen to explain respectively 57.0% and 42.9% of the variance in the test totals after correction for scores on Questions 10 and 12. The outcomes indicate that those who scored better overall on the test were generally those who had developed a clearer understanding of algebraic symbols as standing for variables.

The correlation for the GN* Generalized Number Scale was much smaller than the first two and accounted for only 5.0% of the corrected test total variance, indicating that those who stayed at this view of the meaning for symbols had only a slightly enhanced chance of scoring well on the other test items. Nevertheless, as the correlation (.223) for the GN* Scale was third in order of size and statistically very significant, attaining this level of understanding appeared to be at least somewhat helpful for dealing with the problems presented in the test. Similar comments apply to the attainment of an understanding of algebraic symbols as numerical variables while restricting one's view to integers only. The correlations support the ranking of the levels of understanding of symbols. This infers that the Variables view is the most sophisticated (at Level 5), and that the Generalized Number view and the Integer view are less sophisticated (at Level 4). As the Levels 3 and 2 scales for using replacement value(s) did not correlate significantly with corrected test totals, comment about their ranking could not be based on this correlation analysis which took account of responses by all 517 students, using Test 3 scores for the Year 7 students who did the test more than once.

Two measures correlated negatively with the corrected test scores. Both had been ranked as Level 1 in Table 6·1, and the correlations in Table 9·1 support this verdict. Those who were more immersed in prestructural views, as measured by the PRE* Scale which explained 27.4% of the corrected test score variance, were more likely to score poorly on the test items other than those in Questions 10 and 12. Those who more strongly tended to deny the symbols any variability, as measured by the NV* Scale, were similarly more likely to find it difficult to succeed with the other test items.

Tables 9·2 and 9·3 record similar correlational analyses for the subgroup of Year 7 students for each of the four times they completed the test.

But for one exception in Test 1 results, scale measures for Levels 5 and 4 ranked the highest of the positive correlation coefficients, although there were some variations in the order within this bracket of levels. Next, the Level 3 or 2 scale (for one or two replacements) followed by the Level 2 scale (for one replacement) registered significantly in each of the tests. Level 1 Prestructural Scale recorded a strongly negative correlation in each test. Overall, these tables of results supported the concept of ranking the variety of views of symbols into hierarchical levels according to the order listed in Table 6·1.

Variations were registered in Tables 9·2 and 9·3 for the ordering of the scale measures in terms of the correlations listed. Apart from the results from Test 1, when most students had little idea of the significance of many of the test items, the relative positions of the Levels 5 and 4 scales provide some tentative indication of a sequence of development as recorded by the responses from one test to the next.

The high ranking for the INT* Scale correlations throughout Tables 9·2 and 9·3 indicates that the use of integers only could be considered as closer to the Level 4 view of symbols as generalized numbers than the Levels 3 and 2 views which sought particular values for the symbols rather than allow them to take a more general range of values. As reported in Appendix 9A, a score of "1" was allocated to the INT* Scale for three types of partly correct answers, namely, giving one sample solution, giving more than one sample solution, and giving a general answer which overlooked the possibility of non-integer values. Thus, scores on the INT* Scale were equivalent to the sum of the scores on the 1REP* and GN* Scales, or, in equation form,

$$\text{INT*} = \text{1REP*} + \text{GN*}.$$

Table 9·2

Correlations Between Scale Scores & Corrected Test Totals for Year 7 Tests 1 & 2

Test	Scale for Qq. 10,12	Description	Level of Under-standing (Table 6·1)	Correlation with Corrected Test Total	No.of cases	Signi-ficance
T1	GNV*	Gen.No.or Variable	5 or 4	.648	103	***
T1	VBL*	Variable	5	.568	103	***
T1	INT*	Integers only	4	.504	103	***
T1	12REP*	1 or 2 replacements	3 or 2	.431	103	***
T1	1REP*	1 replacement	2	.377	103	***
T1	GN*	Gen.No.not Variable	4	.368	103	***
T1	NV*	Non-Variable	1	.219	103	*
T1	PRE*	Prestructural	1	- .323	103	***
T2	GNV*	Gen.No.or Variable	5 or 4	.564	177	***
T2	INT*	Integers only	4	.452	177	***
T2	VBL*	Variable	5	.418	177	***
T2	GN*	Gen.No.not Variable	4	.400	177	**
T2	12REP*	1 or 2 replacements	3 or 2	.305	177	***
T2	1REP*	1 replacement	2	.274	177	***
T2	PRE*	Prestructural	1	- .338	177	***

Note. Sorted by size of correlation coefficients within each Test.
Corrected TEST TOTAL = TEST TOTAL *minus* marks from Questions 10 & 12.
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

It was not until Test 4 that the scale measuring the degree to which the candidates had developed an understanding of symbols as numerical variables (VBL* Scale) had moved to the second position (after the GNV* Scale which combined the views for Variable and Generalized Number) in the ordering by correlation coefficients. The proportion of variance shared with the corrected test scores was then 35.8% and 20.1% respectively for the GNV* ($r = .598$) and VBL* ($r = .448$) Scales. Thus, it was not until Test 4 that students who were more likely to score well on items other than Questions 10 and 12 as better performers at algebraic tasks were those who had registered a stronger development of the Variables View in their responses to

Questions 10 and 12. They did not exhibit this characteristic so strongly in the earlier tests, and it could be argued that Tables 9·2 and 9·3 indicate that the concept of algebraic symbols as numerical variables develops later than some views such as the Integers Only view (the INT* Scale ranked second for Tests 2 and 3) and the Generalized Number view (the GN* Scale ranked before the VBL* Scale in Test 3). Thinking in terms of only integers appeared to be a strong tendency in Tests 2 and 3, and it was not until Test 4 (completed about seven months after starting classroom algebra) that the variable concept, which included non-integer values for symbols, was given more recognition by the student group.

Table 9·3

Correlations Between Scale Scores & Corrected Test Totals for Year 7 Tests 3 & 4

Test	Scale for Qq. 10,12	Description	Level of Under-standing (Table 6·1)	Correlation with Corrected Test Total	No.of cases	Signi-ficance
T3	GNV*	Gen.No.or Variable	5 or 4	.549	187	***
T3	INT*	Integers only	4	.406	187	***
T3	GN*	Gen.No.not Variable	4	.370	187	***
T3	VBL*	Variable	5	.351	187	***
T3	12REP*	1 or 2 replacements	3 or 2	.247	187	***
T3	1REP*	1 replacement	2	.200	187	**
T3	NV*	Non-Variable	1	- .270	187	***
T3	PRE*	Prestructural	1	- .267	187	***
T4	GNV*	Gen.No.or Variable	5 or 4	.598	175	***
T4	VBL*	Variable	5	.448	175	***
T4	INT*	Integers only	4	.405	175	***
T4	GN*	Gen.No.not Variable	4	.380	175	***
T4	12REP*	1 or 2 replacements	3 or 2	.266	175	***
T4	1REP*	1 replacement	2	.196	175	**
T4	PRE*	Prestructural	1	- .440	175	***
T4	NV*	Non-Variable	1	- .246	175	***

Note. Sorted by size of correlation coefficients within each Test.
Corrected TEST TOTAL = TEST TOTAL *minus* marks from Questions 10 & 12.
*** $p \leq .001$, ** $.001 < p \leq .010$.

Conclusion from correlational study. Correlations supported the ordering of different levels of understanding for algebraic symbols into the hierarchy of cognitive difficulty which was derived from a synthesis of the work of earlier researchers (Collis, 1975a, Harper, 1979, Küchemann, 1980), as summarized in Table 6·1. Correlations using Year 7 responses to Test 4 supported the same hierarchical order, although the order of correlations varied for the Year 7 students during their first three weeks of algebra. The changes indicated that a common sequence of learning could be to restrict algebraic symbols to integral values before including other possibilities and attaining an understanding of symbols as true numerical variables. The identification of sequential progression through levels of understanding for symbols is pursued further by the following investigation.

Study of Sequential Progression Through Levels of Understanding

None of the three researchers mentioned above had investigated whether or not the hierarchical order of difficulty for levels of understanding of symbols was also a sequential order for learning, such that students might follow through the levels as their understanding developed. The 1990 data were suitable for carrying out such an investigation. Three approaches were employed. The first concentrated on sequences involving the replacement(s) strategy and the other two examined sequences within all five levels of difficulty.

Approach 1 involved studying the scale scores of those who had, in at least one of the four tests, recorded responses which reflected a Level 2 or 3 understanding of symbols. (See Table 9·4.)

Approach 2 centred on studying the average scale scores for each of the four tests. (See Figures 9·1, 9·2, and 9·3.)

Approach 3 compared test totals for students categorized according to those views of symbols for which they registered at least half the maximum total for the scales derived from responses to Questions 10 and 12. (See Figures 9·4 and 9·5.)

Sequencing - Approach 1.

There were 58 Year 7 students (27.9%) who used one or more replacements as their written responses to Question 10 and/or 12 on at least one of the four times they completed the test. The odds ratio analyses summarized in Table 9·4 show that it was significantly more likely that a student would score at least "1" or "2" (out of 4) on the VBL* Scale if that student had, at some time, used the replacement approach than if the student had not used that approach at any time.

Table 9·4
Odds Ratios for Developing Variable Notion in Terms of Use of Replacement Values

Measure		Odds Ratio for Variables View = odds if used replacements / odds if no replacements		Comment: Favours	χ^2 Test of Significance ($df = 1$)	
Name	Description	Calculation	Value		Value	p
VBL1*	VBL* ≥ 1	$\frac{39/19}{63/87}$	2.83	Use of Replacements	9.67	##
VBL2*	VBL* ≥ 2	$\frac{32/26}{55/95}$	2.13	Use of Replacements	5.15	#

Note. VBL* Variable Scale based on Questions 10 and 12.
$.001 < p \leq .010$, # $.010 < p \leq .050$.

The first calculation of Table 9·4 records that, of the 58 (= 39 + 19) Year 7 students who in at least one of the tests used a replacement strategy, 39 at some stage scored at least "1" on the VBL* Scale, while 63 of the 150 (= 63 + 87) who did not use replacements also scored at least "1" on the VBL* Scale. The odds ratio significantly favoured the notion that those who used replacements were more likely to score at least "1" on the VBL* Scale. The odds ratio in the second calculation supports a similar view regarding those who attained a score of at least "2" on the VBL* Scale.

From the record of responses across the four tests by the 39 students who scored at least "1" on the VBL* Scale and had, at some time, given written evidence of using a replacement strategy, there were 19 examples of improvement on VBL* *after* using some replacement, 14 examples of improvement on VBL* *simultaneously* (i.e., on the same test) with their first use of a replacement, and 18 examples of improvement on VBL* *before* they recorded any replacement strategy. (Some improved more than once.) Another 19 students did not make any progress on the VBL* measure although they had used replacements at some stage. Thus, this analysis approach was unable to detect a relationship between the sequencing of the use of a replacement strategy and the improvement in VBL* Scale score. The use of this strategy was shown, however, to be related in some way to the development of the variable notion.

Interview support for speedy learning sequence. As Table 4·22 in Chapter 4 recorded, only 3% of the Years 7 to 12 students wrote answers to Question 10 in the form of replacement values, while Harper (1979) reported more than 12% in the same category when he used an interview method of testing the same question. Similarly, Table 4·28 indicated that, of those who responded with numerical examples in Item 13, the proportion of those tested in the 1979 interviews was more than double that for those who were tested by the 1990 pencil and paper method. More so than the written test, the interview possibly gave students the opportunity to explain their thinking in terms of numerical examples, thus augmenting the proportion of Harper's subjects who were categorized within the "Numerical replacements" group. It was found that, during interviews after the 1990 written tests, the numerical replacements approach was used by several students to explain their thinking for Question 10. For example, in the two extracts which follow, both students spoke about numerical examples even though neither had given replacement values as their written answers. The second student stated specifically that he had used the technique of "scribbling some numbers" during the test but he had not recorded them on his test paper.

Interview extract 1. (Year 7 student, 'M', 19 June, 1990, after Test 2 in which she had registered a score of "0" on both the 1REP* and 12REP* Scales)

E Why did you think number 10 was a good one, M?
M Because it depended on what '*t*' was. If it was bigger than '4', '*t* + *t*' would be larger, or if it was smaller than '4' ....
E Can you explain that?
M Say it was 5. You write 5 + 4 is 9, 5 + 5 is 10, so that's larger. Then if it was '*t*' equals 3, that's 3 + 3 is 6, and 3 + 4 equals 7.

Interview extract 2. (Year 7 student, 'K', 6 June, 1990, after Test 1 in which he had registered a score of "0" on both the 1REP* and 12REP* Scales)

E OK. Over the page let's see what's interesting there. The next four questions are similar. Now how did you think about question 10, did you try different numbers for '*t*', or how did you go about this question?
K Oh... I think that they were the same, because if '*t*' were above 4 you'd get, like if it was 5 you'd get 10, and if '*t*' was 5 and you added 4 you'd get 9.
E Yes. Did you try some numbers like that in the test?
K Yes. I remember scribbling some numbers on paper about it. And if they're, like numbers, if '*t*' represented something like 3, 2 or 1, that '*t* + 4' would have to be larger. Like if it was 2 you'd have two 2's , that's 4.

Conclusion from Approach 1. Initially, the first approach simply identified the likelihood of some relationship between using the replacement strategy and scoring on the VBL* Scale. Sample interviews indicated that the use of replacement values might have been a popular strategy in the process of arriving at more general answers to some aspects of Questions 10 and 12. These considerations indicate that more students than those already identified by the analyses reported above probably made use of trial values and used the replacement strategy while completing the written tests.

If so, such students followed a learning sequence which took them from replacements to some higher level of understanding of the symbols involved. They moved through the stages quickly, avoided writing down replacement values as their solutions and, thereby, wrote more general conclusions.

Sequencing - Approach 2.

Figures 9·1, 9·2, and 9·3 display the average scale scores for the five levels for understanding of symbols for all Year 7 students, for the Advanced Year 7 students, and for the Low Ability Year 7 students respectively. All scales mentioned are those based on parts (b) and (c) of Questions 10 and 12 and were scored to have a maximum of "4". The three figures are printed on the same page for ease of comparison.

Figure 9·1 shows that, for all Year 7 students,

1. gradual and consistent, if slight, improvement was registered from test to test in the Level 5 understanding of algebraic symbols as numerical variables (VBL* Scale). The average VBL* Scale scores were 0.359, 0.633, 0.829, and 1.177 for Tests 1 to 4 respectively, and the corresponding averages for the GNV* Scale were 0.524, 0.848, 1.262, and 1.543;

2. considerably less progress was made with the Level 4 understandings which included the view of symbols as Generalized Numbers (GN* Scale) and in terms of integers only (INT* Scale). The progress was reasonably consistent from Test 1 to Test 3, with a slight drop to Test 4, the average scores being 0.165, 0.215, 0.433, and 0.366 respectively for the GN* Scale, and 0.272, 0.299, 0.578, and 0.560 for the INT* Scale;

3. there was scarcely any tendency to use the Levels 3 or 2 technique of replacement values. This was registered by 12REP* and 1REP* Scale scores; and

4. there was a strong tendency to retain incorrect ideas about symbols. This was revealed by the average scores on Level 1 measures (PRE* and NV* Scales). The average scores on the PRE* Scale were high, with a slight decline from Test 1 to Test 3 and a very slight increase in Test 4. The averages were 2.07, 1.91, 1.52 and 1.59, respectively. The Non-Variable NV* Scale scores fluctuated slightly, increasing for the first three tests and then decreasing (the averages were 0.67, 0.74, 1.01 and 0.79 respectively). These outcomes show the persistence of prestructural approaches and those that denied the symbols the freedom to vary.

The Figure 9·2 results, for Year 7 students from Advanced classes only, closely resemble those in Figure 9·1 for all the Year 7 students. Observations 1 to 4 above applied, in general, to the students from Advanced classes. The latter showed a slightly greater decline in scores for the prestructural view and slightly higher gains in the variables and generalized numbers views.

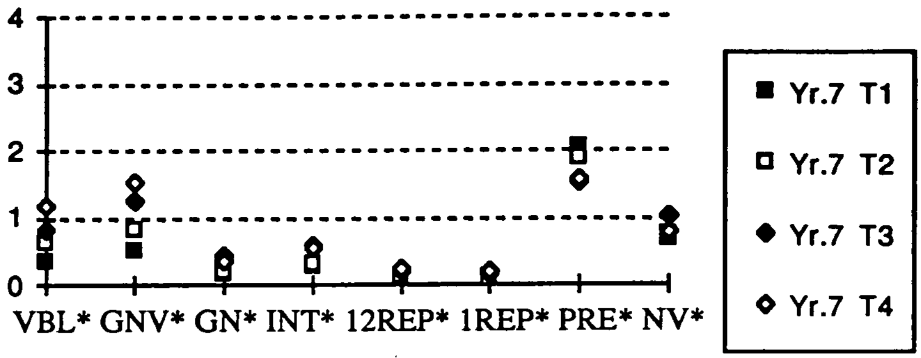


Figure 9·1. Average Qq.10 & 12 scale scores for all Year 7 students

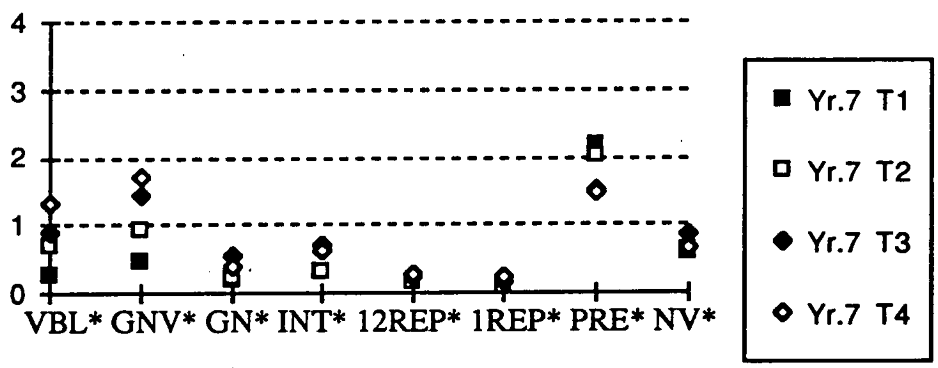


Figure 9·2. Average Qq.10 & 12 scale scores for Advanced Yr 7 students

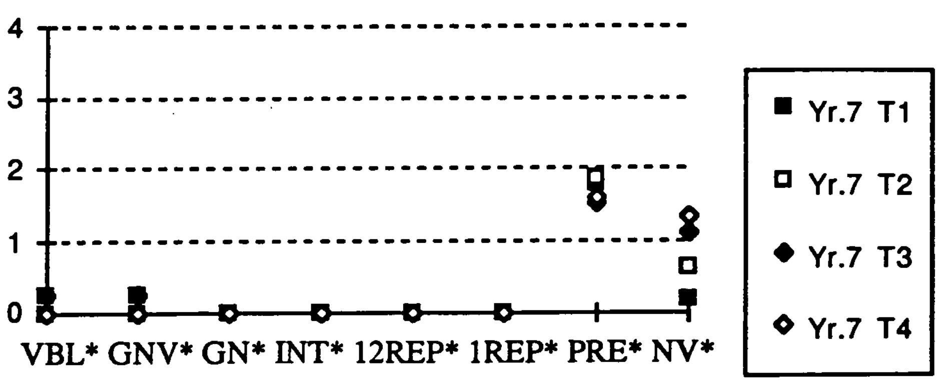


Figure 9·3. Average Qq.10 & 12 scale scores for Low Ability Yr 7 students

As displayed in Figure 9·3, the eight students in the Low Ability Year 7 class who participated in the study scarcely registered on scales for Level 5, did not score on scales for Levels 4, 3 or 2, and were firmly based in Level 1 views of symbols. Their average scores on the NV* Non-Variable Scale actually increased from test to test, showing a deterioration in understanding as time passed.

Conclusions from Approach 2. The second approach showed that, during the data collection period, the beginning Year 7 students registered their main improvements at the Level 5 understanding of symbols; they recorded a slight improvement followed by a slight decline at Level 4; they scarcely registered at all at Levels 3 and 2; and they tended to persist at Level 1, although some decline was noted. In other words, the main record of progress was at the top end of the hierarchical ladder. If those who improved made use of the intermediate rungs, they apparently did not linger on them long enough to leave a clear record of being at those levels.

Sequencing - Approach 3.

Students were categorized in terms of whether or not they attained scores of "2" or more on the scales built from responses to parts (b) and (c) of Questions 10 and 12. A score of "2" was half the possible maximum for each scale. For students in each category, average scores on all test items, other than all parts of Questions 10 and 12, were examined. These corrected average test scores are displayed in Figure 9·4 for the Years 8 to 12 subgroup and for the Year 7 students each time they were tested. Figure 9·5 records the percentage frequencies in each category for students who registered valid scale scores by responding to each of the four items common to all scales. The number of such valid cases is included in Tables 9·2 and 9·3 for each testing of the Year 7 subgroup. The number of valid cases for the Years 8 to 12 subgroup was 281. Figures 9·4 and 9·5 are presented on the same page for easy comparison.

Year 7 students. Figure 9·4 shows that, in general, overall corrected average test scores improved for the Year 7 students from one test to the next, regardless of the level of understanding for symbols registered by the scales based on parts (b) and (c) of Questions 10 and 12. A noteworthy exception is that the average scores of those who favoured the technique of offering one or more replacement values (those in the 12REP* and 1REP* categories) recorded a fairly consistent average test score of around 30%, even in the early stages. As the frequency graphs in Figure 9·5 indicate, apart from the 13 students (12.6%) who used the single replacement approach in Test 1, there were very few students who used replacements. Those who did so were amongst the highest scorers on all test items other than Questions 10 and 12. Using replacements correctly indicates an understanding of the problem in question, although the user is not obtaining a general algebraic solution.

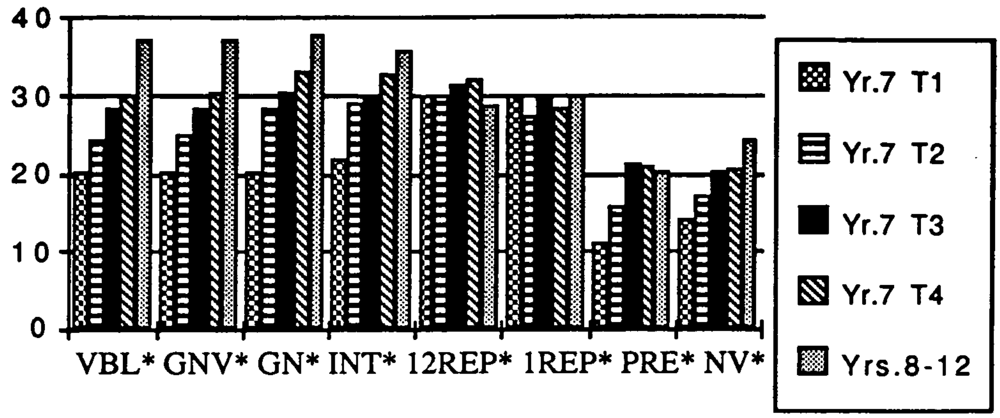


Figure 9·4. Corrected average test totals according to scale categories (Test totals corrected for Qq.10 & 12; Categories determined by scale scores of "2" or more)

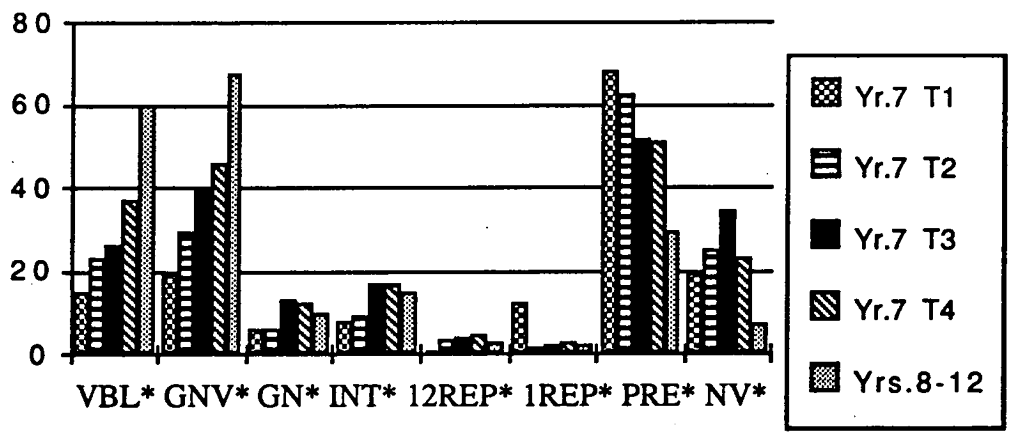


Figure 9·5. % frequencies of valid cases for scale categories (Categories determined by scale scores of "2" or more)

There were few Year 7 students who used Level 4 responses to qualify for the GN* and INT* categories. By Test 4, these few students averaged a slightly higher score than those who used the Levels 3 or 2 replacement approaches and about three marks better than the larger number in the VBL* Level 5 variables category. As the differences were not great, this form of analysis does not lay claim to definitive conclusions about the merits of one or other view of symbols across the Levels 2 to 5: These averages were all in the range from 28.2 to 33.1 for Test 4. However, the analysis supports the contention that completely misunderstanding Questions 10 and 12 (as measured by the PRE* Scale) or taking a view that the symbols in the questions were non-variables (as measured by the NV* Scale) was associated with poorer performance on the other test items. The average scores for those who made these errors was lower than for all the other students on every test. In Test 4, for instance, the averages were 21.0 and 20.6 respectively for those in the PRE* and NV* categories.

Years 8 to 12 students. The graphs in Figure 9·5 for 281 Years 8 to 12 students show that most (60.1%) were able to attain the VBL* Variables category for thinking at Level 5. When discursions to the Level 4 Generalized Number view of symbols were included to form the GNV* Generalized Number or Variable category, the proportion was larger, at 67.6%. Only 10.0%, however, were restricted to the GN* Generalized Number category at Level 4, while 14.6% qualified for the Level 4 INT* Integers category. Less than 4% were classified in the replacements categories for Levels 3 or 2. Clearly, the most popular interpretation of symbols was as numerical variables. Most of those who did not reach Levels 4 or 5 had given test responses which indicated that they had little understanding of the problems posed by Questions 10 and 12, resulting in their classification in the Level 1 PRE* Prestructural category (for 29.2%), with a few (7.5%) in the Level 1 NV* Non-Variable category.

The corrected average test totals in Figure 9·4 gave support for Proposition 1, which was discussed in Chapter 6. The Years 8 to 12 students with better understanding of the meanings of symbols (at Levels 5 or 4) performed better on algebraic tasks than those at lower levels. The averages for those in Levels 5 or 4 categories ranged from 35.7 to 37.6. The Levels 3 and 2 categories yielded lower averages (30.0 or less). Those in the Level 1 categories averaged less than 25, with the lowest result (20.1) being for those in the PRE* category. These outcomes gave clearer support for Proposition 1 than did the corresponding findings for the beginning Year 7 students.

The proportion of Years 8 to 12 students at the Level 5 variables standard was 60.1%. This statistic is nearly two-thirds more than the corresponding percentage (37.1%) for the Year 7 students at Test 4, giving a strong indication that the development of the variable view takes time. The students in the Year 7 group were on their way towards a Level 5 understanding of symbols as variables but, even after seven months of algebra, many still had a fair way to go.

Sequence clues from frequency graphs. The most rapid growth in percentage of valid cases, as displayed in Figure 9·5, was for those who registered "2" or more on the GNV* Scale, with most of the increase being due to those who improved on the VBL* Scale which registered the development of the variable notion. The number who favoured the Generalized Number view (as measured by the GN* Scale) was small compared with those who moved beyond this concept to that of a numerical variable. The growth in percentage frequency for the Variable notion had to be accounted for by a decrease in percentage frequency on some other measure(s), as all the scales in the discussion were based on responses to the same two questions. The PRE* Prestructural Scale lost favour from Test 1 to Test 2 to Test 3 and it was the NV* Non-Variable Scale that showed a decrease in popularity from Test 3 to Test 4. The other main contributor to the increase in the percentage who scored at least half the total score on the VBL* Scale was the 1REP* One Replacement Scale, which registered a rapid loss of numbers after Test 1. Hence, the percentage frequency graphs of Figure 9·5 indicate that likely sequential learning paths were to lead towards the variable notion after misunderstanding the problems by registering at first a prestructural outlook, or after accepting a non-variable view at some stage, or after using a replacement approach in the early stages. For a few of the Year 7 beginners, this process possibly stalled at the Level 4 stage: The number of Year 7 students in the Level 4 GN* and INT* categories approximately doubled after Test 2.

<u>Advanced classes</u>. In Appendix 9A, Figures 9A·1 to 9A·4 display similar types of data for students in Advanced classes and for those in Low Ability classes. The outcomes for the Advanced classes (Figures 9A·1 and 9A·2) show that those Year 7 students at Levels 2 and 3 tended to score the higher averages on test items other than those in Questions 10 and 12. On the other hand, the Years 8 to 12 students tended to score the better averages if they were at Levels 5 or 4. It seems that more than seven months is needed for a longitudinal study if sequences of learning are to be identified in terms of the levels of understanding of symbols. The Year 7 students in the study seem to have been moving slowly along the path towards the concept that algebraic symbols represented numerical variables. During the data collection period, the benefits of this progress were not explicit.

<u>Low ability classes</u>. The Low ability classes (Figures 9A·3 and 9A·4) showed a tendency to stay at the Level 1 Prestructural level, the level of misunderstanding the basic meaning of symbols.

<u>Conclusions from Approach 3</u>. Average test scores for beginners appeared to be independent of their level of understanding of symbols in Questions 10 and 12, except that those at Level 1 registered the lowest averages. Average test scores from Years 8 to 12 students were graded approximately in the order of the levels of understanding for symbols, thus supporting Proposition 1 which claimed that those with better understanding of algebraic symbols were more likely to score well on algebraic tasks. The beginning Year 7 students generally had difficulty attaining the variable notion, some stalling at Level 4 in their progress along the sequential learning path.

Investigation 2: The Key to Progress

An important question, especially from a teacher's point of view, is that of why some students show little, if any, progress despite attendance at the same lessons as other students who do progress. What is it that students need to learn before they can show signs of general progress? Do they need to reach certain steps in the hierarchy of concepts and/or skills suggested in Chapters 6 and 8 before they can make secure progress in other aspects of learning? One of the ways in which concern about the "slow learners" was addressed was to study their performance sequentially from one test to another in comparison with the "fast learners".

Table 9·5 presents descriptive statistics for the results of the four testing stages for the 208 Year 7 students who completed the test instrument three times and the 186 of these who completed it a fourth time some six months after the teaching intervention sessions.

Table 9·5
Test Results for Year 7 Students

Test	Mean	*S.D.*	Min.	Max.	*N*
1	10.76	6.43	0	35	208
2	18.49	9.30	3	48	208
3	25.02	10.10	2	54	208
4	27.37	11.34	3	52	186

The means show that there was a general improvement from test to test but the minimum scores show that some students did not score well even after the third test. In contrast, the maximum scores record that some others attained impressively high results, considering that the test was marked out of 65 and these students were only beginners at algebra.

Preliminary explorations. Octile rankings were established for students on each of the four testings. From one test to a later one, some students improved their ranking while others did not. Contrasts between the improvers and non-improvers were investigated by means of discriminant analyses and by odds ratios analyses. For the discriminant analyses, missing values were scored at zero to ensure that all students were included. For the odds ratios, categories for single items were success or failure whereas, for scores on each scale, two categories were determined by using the median scale score as a boundary. The outcomes of a large number of these types of analyses gave some indication of the key variables which might be crucial to the differential development rates. However, the standards method, discussed below, was chosen as a more reliable approach to the question of the key to progress. The major drawback with the octiles approach was that from test to test there were changes in the average scores attained by each octile grouping, so that students who actually had improved in their test scores could be classified as not improving their octile ranking. Furthermore, the membership of the various octile groupings changed from test to test, and setting missing values to zero left some doubts regarding the usefulness of the discriminant analyses.

Analyses Based on Standards.

To set criterion levels for following changes in performance from one test to the next, four equal ranges of test scores up to and including the maximum scored by these Year 7 students were chosen: 0 to 13, 14 to 27, 28 to 41, and 42 to 55. These ranges defined categories which were called Standards 1 to 4 and will be referred to as S1, S2, S3, and S4 respectively. Table 9·6 displays the frequency distribution of students in terms of these categories.

Table 9·6
Frequency Distribution of Year 7 Students in Standards Categories

Standard	Range of	Test			
	Scores	1	2	3	4
S1	0 - 13	158	71	24	23
S2	14 - 27	42	103	111	67
S3	28 - 41	8	30	59	76
S4	42 - 55	0	4	14	20
Totals		208	208	208	186

The number of students in S4 improved from zero to 20, and the number in S3 increased more dramatically. However, at the other end of the scale, there were over 20 students still in S1 after three weeks of algebra (Test 3) and a similar number were at the same standard after another six months (Test 4). An investigation was carried out to find out if any students stayed at this lowest level in all four tests. It was found that 15 actually did. There were two more students who were absent for the fourth test but who remained in S1 for the first three tests. These 17 students had shown little, if any, improvement during the period of the testing program, even though it included at least three weeks of introductory algebra teaching. Another 7 students stayed in S2 for all four tests. There were 18 students who, in contrast, had moved from either S1 (in 8 cases) or S2 (in 10 cases) to S4 during the same period. An answer was required as to whether or not some vital aspect of learning enabled the latter two groups of students to progress so markedly while the former two groups made little headway.

Developmental Contrasts Between Two Extreme Groups

An investigation of the contrasts in learning patterns of the two extreme groups shed light on the problem of why some improve and others do not. The groups were the 17 students who stayed at S1 (referred to as Group SS11) and the 8 who moved from S1 to S4 (Group SS14).

Choosing appropriate statistical analyses. To identify differences between the test performances of the two groups, it was impractical and inappropriate to try a discriminant analysis for three main reasons:

1. Many of the scale measures derived from the test results were highly correlated and some shared responses to common test items;

2. The numbers in each group were small (namely, 17 and 8, respectively); and

3. Data were missing from responses by some students, especially those in Group SS11, and this would have led to a very small number of cases being retained in the analysis, possibly resulting in zero cases, especially in Tests 1 and 2 where many questions were omitted.

Two alternative statistical analyses were chosen.

1. Odds Ratios based on 2 x 2 contingency tables were used, together with the appropriate test of significance, when there was an objective for comparing frequencies within certain categories. There were two such cases, both reported in Appendix 9B. The first records that Group SS11 omitted more test items in Tests 3 and 4 than did Group SS14. The second verifies that students from Advanced classes were more likely to be members of Group SS14 than of Group SS11.

2. For comparing average scores on any of the scale scores, *t*-tests were found to be most appropriate. They took into account the number of valid cases and whether or not the variances for the two groups differed significantly or not. If not, a pooled variance was used, otherwise separate variances were used. As will be seen when the *t*-test outcomes are reported, the fact that the differences between the groups were statistically significant for so many of the scales reduced the likelihood that the number of significant differences identified by repeated *t*-tests could have occurred merely by chance. Furthermore, these *t*-tests were independent of each other as they were used to compare the group means on the scales for each of Tests 1 to 4 separately.

Opening the Door to Progress

Analyses of differences between the means attained by the extreme groups, SS11 and SS14, provided an answer to the question of what it was that enabled some students to progress. Table 9·7 presents a summary of the relevant statistics for Tests 1 and 2.

Table 9·7
Summary of *t*-tests for Groups SS11 and SS14 on Tests 1 & 2 Responses

Test	Scale	Comment	Max.	Mean SS11	Mean SS14	*t* value	*df* ∗	*p*	Favours
1	Total	Test Total	65	5.12	9.25	3.86	23	***	SS14
2	GNV	Gen.No.and/ or Variable	17	1.33	5.83	5.49	5.70	**	SS14
2	Total	Test Total	65	8.88	20.50	4.51	8.08	**	SS14
2	SUBS	Substitute & Solve	7	0.38	5.00	4.41	7.22	**	SS14
2	EQL	Equals Scale	9	0.11	1.67	3.57	5.70	*	SS14
2	VBL	Variable	11	0.10	2.17	3.39	5.28	*	SS14
2	NBR	Numbers View	4	0.43	1.67	2.49	18	*	SS14
2	AD	"*a, d*" Scale	2	0.18	1.00	2.34	17	*	SS14
2	CON	Conjoining	7	2.31	0.43	- 3.38	21.00	**	SS14
2	PRE	Prestructural View	19	15.50	8.67	- 3.10	10	*	SS14

Note. Test 2 entries sorted in order of *t* values.
∗ *df*: decimal point if using separate variances; otherwise, pooled variance.
Max. = maximum possible score (= no. of items for scales).
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

In Test 1, Group SS14 scored significantly higher on test total than Group SS11, although all members of both groups were at Standard 1 (test scores less than 14) and the difference between the mean scores was only about 4 points. There were no other significant differences at this pre-algebra stage.

Table 9·7 targets aspects of learning that made the difference between these two groups after their first week and a half of algebra. The *t*-tests based on Test 2 responses identified that those who were destined to progress to Standard 4 level (Group SS14) significantly out-scored those who stayed at Standard 1 in the following ways:

1. Their average test score was more than double that attained by Group SS11 and was about 12 points better;

2. They had progressed significantly further in their understanding of algebraic symbols as generalized numbers and/or variables (GNV Scale);

3. They were more rapidly developing the skills for substituting a numerical value into algebraic expressions and for solving a simple equation (SUBS Scale);

4. They were more successful in identifying when two algebraic expressions were equal (EQL Scale);

5. They were developing the variable concept more rapidly (VBL Scale);

6. They had the greater tendency to regard the letter symbols in early algebra as standing for numbers rather than for objects (NBR Scale);

7. They were more able to allow the symbols '*a*' and '*d*' to stand for numbers without restriction in Question 2 (i) (AD Scale);

8. They were less inclined to conjoin symbols for addition (CON Scale); and

9. They were breaking away more rapidly from some of their prestructural views of algebra (PRE Scale).

All of findings 2, 4, 5, 6, and 7, to some degree, record that those in the group on the verge of greater improvement had started to develop the concepts that the alphabetic symbols of early algebra represented numbers and that those numbers could vary. Those students in Group SS11 lagged significantly behind the others and did not appreciate that the symbols that had been introduced to them were standing for numbers which could vary.

Findings 2, 4, and 5 indicate that those who were to make greater progress had started to understand that algebraic symbols represented numerical variables and they had successfully applied this concept to some of the problems presented in the test.

Finding 3 suggests that those on the way to higher scores were faster in developing the skills needed for solving a simple equation ('$3a = 36$'), and for substituting a given value for 'y' in expressions such as '$2y + 5$', showing that they had a clearer grasp of the meanings of algebraic expressions.

Findings 8 and 9 report that it is helpful to understand the convention that conjoining is used for multiplication in algebra, and to comprehend the implications of the test questions. Those who did not progress made very little intelligent headway at the time of Test 2 on most of the items in the Prestructural Scale, as shown by their high mean score (15.50 out of a possible 19) on the PRE Scale.

Data obtained in Test 3 allowed continuation of the investigation into the aspects of learning that discriminated between Groups SS11 and SS14. Differences between the means of the two groups were significant on 22 scales, the large number emphasizing the growing gap between the rates of development within the two groups. In every case, the differences favoured Group SS14. The outcomes are summarized in Table 9B·3 of Appendix 9B. At the time of responding to Test 4, there were significant differences between the means for 21 scales, as reported in Table 9B·4 of Appendix 9B.

Comparison of Groups SS11 and SS14 over four tests. Figures 9·6 and 9·7 summarize the scores per item for Groups SS11 and SS14 respectively on scales which registered significant differences, using *t*-tests, for Tests 2, 3, and 4. They are presented on the same page for ease of comparison.

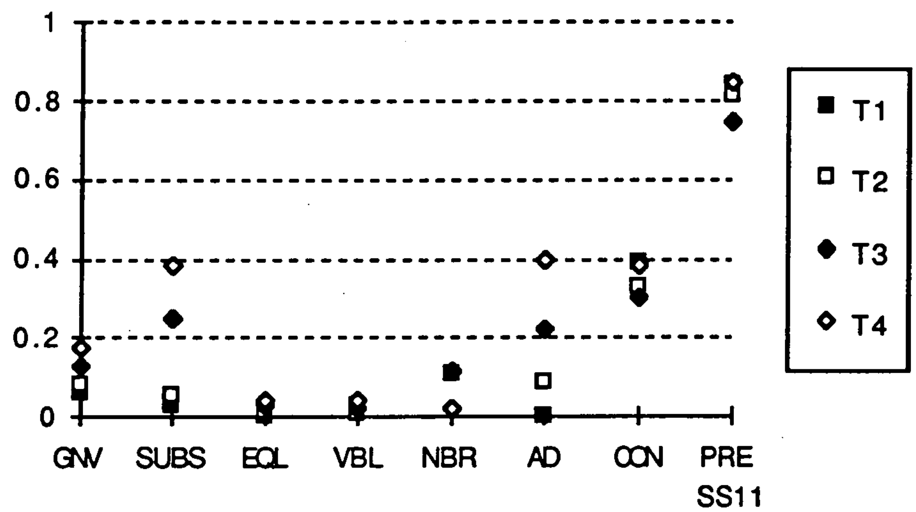


Figure 9·6. Average scores per item for SS11 Group

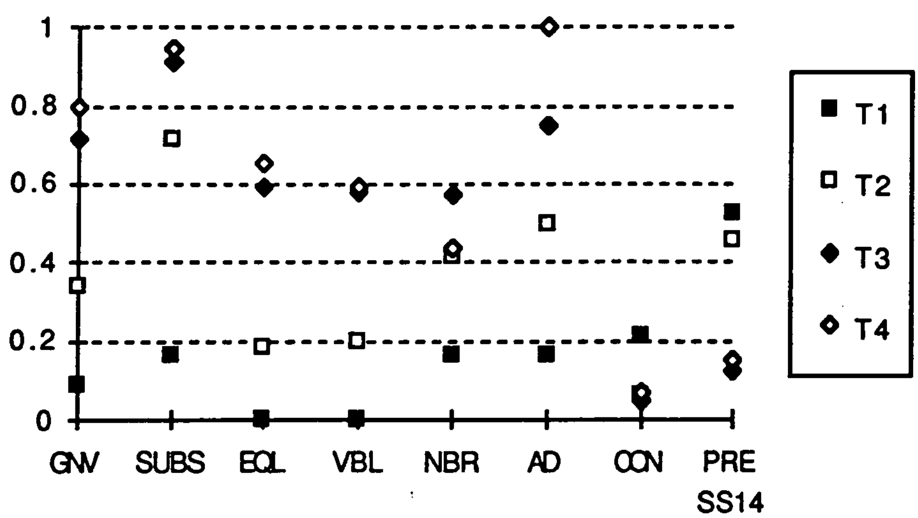


Figure 9·7. Average scores per item for SS14 Group

The graphs record that the views of students in Group SS11 changed least.

Their average scores in the four tests for each scale are close, indicating that there was little change from test to test, except for an improvement on the SUBS Substitute and Solve Scale and the AD '*a*, *d*' Scale. The group recorded minimal changes in their incorrect views, as measured by the CON Conjoin Scale and the PRE Prestructural Scale. On the other hand, Figure 9·7 records that students in Group SS14 showed a rapid improvement on the first six scales, especially the GNV Generalized Number and/or Variable Scale, the SUBS Substitute and Solve Scale and the AD '*a*, *d*' Scale. There was a growth in their ability to solve equality problems (EQL Scale), and in understanding algebraic symbols as variables (VBL Scale) and as representing numbers (NBR Scale). These students moved away from incorrect views, as recorded by the noticeable decrease in average scores per item on the CON Conjoin Scale and the PRE Prestructural Scale.

Discussion Regarding Groups SS11 and SS14.

The above analyses have shown that students in Groups SS11 and SS14 differed significantly in very few measured respects before they had started classroom work on algebra, as all knew very little, if anything, about algebra. Group SS11 was composed of students from 7 classes across Schools A, B, and D, and Group SS14 of students from 5 classes across Schools A, C, and D. Group SS14 scored significantly better on Test 1, before starting algebra, yet their average score was less than 10. Significantly, more SS14 students were members of Advanced classes. One could speculate on ways that could be associated with their rate of progress. The students had probably been placed in these classes because teachers had identified their aptitude for mathematics. Perhaps the flow of ideas was greater in the Advanced classes than in the other classes. The research data could not clarify these possibilities.

After the first few lessons on algebra over a period of about one and a half weeks, differences started to show. Besides acquiring the concept that algebraic symbols stood for numbers which could vary, students who reached Standard 4 acquired more quickly basic skills for substitution in simple expressions and solving a simple equation. SS11 students tended more to miss the point of the problems set, indicating that they were at the Prestructural stage and were not coping with the new ideas that were being placed before them. They were more persistent in regarding the conjoining process in algebra as a convention for addition, rather than multiplication.

The data used in the search for the key to progress were found to be applicable to two of the propositions which were supported by analyses of the frequencies of successful responses from all 517 students in the study. These are discussed in turn.

Proposition 2. In Chapter 8 empirical support was given to Proposition 2 which claimed that *success in interpreting algebraic expressions is a prelude to success in comparing the values of two expressions or comparing the values of two variables within the one expression.*

Of the scale measures listed in Table 9·7 and graphed in Figures 9·6 and 9·7, the Substitute and Solve SUBS Scale included the assessment of the level of students' success in substituting into several algebraic expressions. To make substitutions successfully, students had to know the conventions for writing the expressions so that they could interpret their meaning. An investigation was carried out to see how the two groups compared on a subset of the SUBS Scale, namely the SUB Substitution Scale which was available as a measure of skill in substitution. As expected, *t*-tests using scores on the SUB Scale showed that there was not a significant difference between the two groups in Test 1 but that the differences were significant in the other three tests, as was the case for the SUBS Scale. A summary of *t*-test analyses for the SUB Scale responses is presented in Table 9·8.

Table 9·8
Summary of *t*-tests for Groups SS11 and SS14 on SUB Substitution Scale Responses

Test	Scale	Max.	Mean SS11	Mean SS14	*t* value	*df* ✶	*p*	Favours
2	SUB	6	0.24	4.38	5.09	8.31	***	SS14
3	SUB	6	1.76	5.75	7.33	18.92	***	SS14
4	SUB	6	2.07	5.75	6.64	15.35	***	SS14

Note. Test 2 entries sorted in order of *t* values. ✶ *df* using separate variances.
Max. = maximum possible score (= no. of items for scales).
*** $p \leq .001$.

These analyses clearly showed that SS14 students progressed to almost 100% efficiency in substitution skills after only three weeks of algebra, as could be judged by their high average scores (5.75) on Tests 3 and 4. Those in the SS11 group achieved only about one-third of that success rate. If Proposition 2 applied to SS11 students, then their failure to interpret algebraic expressions sufficiently to enable them to substitute correctly should have been accompanied by failure with the test items that formed the VBL Variables Scale, a scale which gave a measure of their degree of success with items requiring them to compare the values of two expressions (in Items 10, 12, and 15 (iii)) or two variables within one expression (in Item 6 (c)).

It was found that, while Group SS11 tallied eight scores of more than "1" on the SUB Scale in Tests 1, 2, or 3, none in the group scored more than "1" on the VBL

Scale in the same tests, a result which did not challenge the proposition. In Test 4, six members of the group scored more than "1" (five scored more than "3") on the SUB Scale but none of these scored more than zero on the VBL Scale, again allowing the pattern outlined by the proposition. There was one anomaly: One student had three parts of Question 10 correct and tallied a score of "4" on the VBL Scale in Test 4, while having only one substitution example correct. However, the definitions of the criteria governing Proposition 2 had built in an allowance for departure from the 100% "prerequisites" situation, and the proposition had been deliberately stated in terms of "preludes". Taking the overall outcomes for Group SS11, the proposition had pointed to the likelihood that one of the reasons why members of the group had not progressed in the development of an understanding of algebraic symbols as numerical variables, as measured by the VBL Scale, was that they had not learnt to interpret algebraic expressions efficiently.

In Group SS11, the few (6 in number) who had begun to master the skill of substitution (as shown by scores of "4" or more on the Substitution Scale) were still unable to succeed with questions involving the concept of a variable.

1. One student scored "4", "6", and "4" on SUB Scale in Tests 2, 3, and 4 respectively but either scored zero or registered missing data on the variable measures embraced by the VBL Scale, except for having one part of Question 10 correct in the third test.

2. In Test 4, two other students scored "4" on the SUB Scale, but zero or "missing" on the VBL Scale - the latter registered zero scores for Questions 10 and 15 (iii), which were part of the VBL Scale.

3. Two more students scored "5" on the SUB Scale in Test 4 but scored zero for Scale VBL.

4. One student attained the score of "6" in Test 3 on the SUB Scale but returned "missing" for the VBL Scale, while recording zero for Questions 10 and 12, and "1" for Question 15 (iii), all components of the VBL Scale.

The conclusion from the SS11 results is that developing the skills needed for substitution is not sufficient for being successful with the concept of a numerical variable. Nevertheless, Group SS14 responses were helpful in support for the claim that development of the ability to interpret algebraic expressions sufficiently well to enable substitution to be successfully completed is a prelude to developing the ability to compare two expressions at a cognitive level which requires the notion of a numerical variable.

Over the four tests, it was found that any member of Group SS14 who scored "4" or more on the VBL Scale (which had a maximum possible score of "11") also scored "5" or "6" on the SUB Scale (with a maximum score of "6"). The growth was in conformity with Proposition 2.

1. In Test 1, none scored above zero on the VBL Scale and only two students scored above zero (respectively "3" and "4") on the SUB Scale.

2. In Test 2, one student scored more than "3" on the VBL Scale and "5" on the SUB Scale. Three others scored "6" on the SUB Scale and "1", "2", or "3" on the VBL Scale.

3. In Tests 3 and 4, all scored "5" or "6" on the SUB Scale and all scored from "4" to "10" on the VBL Scale in both tests, except for one student who scored "3" in Test 3 and another who scored "2" in Test 4.

Success in substitution in Question 3 was evidence that members of Group SS14 had developed an understanding of the conventions used when writing algebraic expressions. They were then able to progress with problems involving comparisons of two algebraic expressions. Thus the data from Group SS14 confirmed and exemplified the claim made by Proposition 2.

By graphing the average scores per scale item on the scales in question, Figure 9·8 illustrates the way the differential rates of development for Groups SS11 and SS14 followed the expectations expressed in Proposition 2.

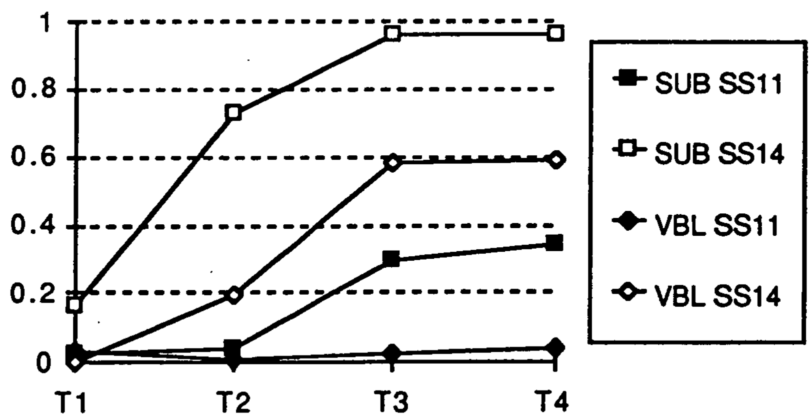


Figure 9·8 Average scores/item for SS11 & SS14 Groups on SUB & VBL Scales

The SS14 graphs seem to indicate that progress on substitution is a prelude to progress with problems involving the notion of a numerical variable, in the form required by the items in the VBL Scale. At the same time, the graphs for Group SS11 show that improvement on the VBL measure does not necessarily follow once there is some improvement on the SUB Scale. These outcomes support the reasoning that one needs to reduce the memory load required for simply understanding the meaning of an algebraic expression before one can carry out cognitive tasks involving relationships between algebraic expressions, as was required in items belonging to the VBL Scale. This point was made in Chapter 8 under the heading "Analysis of Question 10 in Halford terms".

Proposition 3. Proposition 3 was stated in Chapter 8 as: *Success in solving a simple linear equation is a prelude to success in identifying the conditions for equality of two algebraic expressions or for the equality of two variables within the one equation.*

The EQL Equality Scale was placed in Table 9·7 and Figures 9·6 and 9·7 as one of the measures which identified significant differences in the learning rates of the beginning algebra students in Groups SS11 and SS14 in their first one and a half weeks of algebra. This scale supplied a measure of success in dealing with the equality notion in the contexts mentioned in Proposition 3 and so provided a suitable assessment for testing the proposition with respect to the rates of cognitive development shown by the two groups. One item, Question 8 (b), measured ability to solve a simple algebraic equation and was chosen as the other measure needed for testing Proposition 3. As Question 8(b) was a subset of the EQL Equality Scale, the scale scores were amended to form the EQL* Adjusted Equality Scale by subtracting scores on Item 8 (b), so that there would be no overlap between the two measures being considered.

Figure 9·9 displays the differential rates of development on these two measures for each of the groups SS11 and SS14.

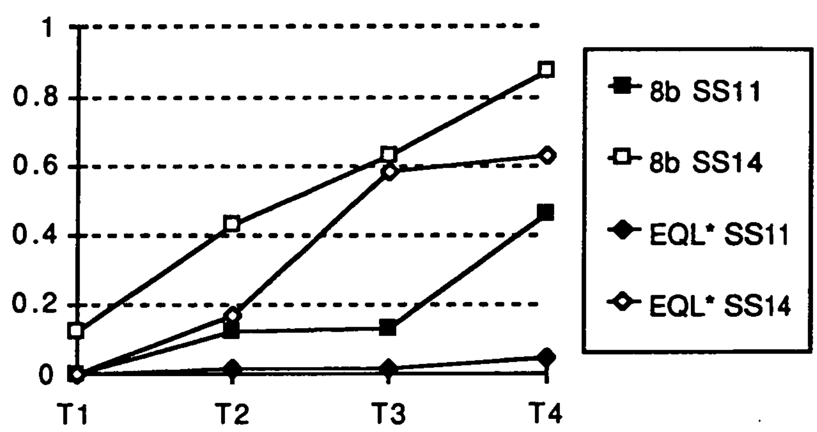


Figure 9·9 Average scores/item for SS11 & SS14 Groups on EQL* Scale & Q.8b

The graphs show how the average score per item changed over the period of the research testing program. The pair of graphs for Group SS14 record that these students, on average, improved first on their ability to solve the equation '$3a = 36$', as required by Question 8 (b), before improving markedly on the EQL* Scale. There were four exceptions. Three students out of this group of eight registered "3" or more on the EQL* Scale out of a possible score of "8" without having Question 8 (b) correct - one managed this at Test 3, another at Test 4, and the third on both Tests 2 and 3. However, considering that these were the only 4 exceptions out of the 24 instances

(17%) of valid records on EQL* Scale from the whole group, the word "prelude" still could be accurately and correctly applied to this group in support of Proposition 3. Success in solving a simple equation was a prelude to success in finding out conditions for the equality of two algebraic expressions, as measured by the EQL* Adjusted Equality Scale.

As regards Group SS11, Figure 9·9 indicates a gradual improvement in the ability to solve the equation in Question 8 (b), but that this did not necessarily result in increased success with the equality problems in the other items of the EQL Scale. It seems that the memory load needed for solving the easier cases of algebraic equality had to be reduced before more complex equality problems could be solved. Only one student in Group SS11 scored more than "0" on the EQL Scale, scoring "1" in both Test 2 and Test 3, and this was without having Question 8 (b) correct. There were 10 students who were correct on Question 8 (b) across all the tests. In 6 of these cases missing values were registered for the EQL* Scale while, in the 4 other cases, a score of "0" was registered for the EQL* Scale. Support was evident for Proposition 3. Those who improved their scores on the Equality Scale had, in most cases, previously improved their success with solving the simple equation in Item 8 (b). Solving an equation involving one arithmetical operation and one variable was apparently less demanding cognitively than finding the condition for the equality of two algebraic expressions. It was established, however, that success on Item 8 (b) was not a sufficient condition for success on the other Equality Scale items.

Persistence of incorrect ideas. Two other variables from Table 9·7 and Figures 9·6 and 9·7 deserve close attention in this exploration of the problem of why some students progress while others do not. They were scores on the CON Conjoining Scale and the PRE Prestructural Scale, both measures of incorrect thinking about early algebra. There were no test items in common between these two scales and, between them, they covered 26 items. The scale averages for the two groups, as graphed in Figure 9·10, indicated that members of Group SS11 were more persistent in their incorrect views than were their counterparts in Group SS14. This is an additional insight into the factors influencing the vastly different rates of development of these two groups.

Those in Group SS11 hardly changed, on average, in their acceptance of the incorrect view of conjoining. Only six of these 17 students kept their CON Scale score under "3" (out of 7) for Tests 3 and 4. They also showed very little change in their prestructural approaches to the 19 problems registered in the PRE Scale. Many had missing data classification for the scale and, except in two cases, all the registered scores were "11" or more. It seemed that the class activities were making little impact on their way of thinking.

In contrast, those who made rapid advances in mastering the basic concepts of early algebra, the members of Group SS14, distanced themselves from the misconceptions measured by these two scales. None of them, for instance, scored greater than "1" on the CON Scale after Test 2. Only two scored more than "4" on the PRE Scale after Test 2 and all but one of the others scored "2", "1", or "0" after Test 2.

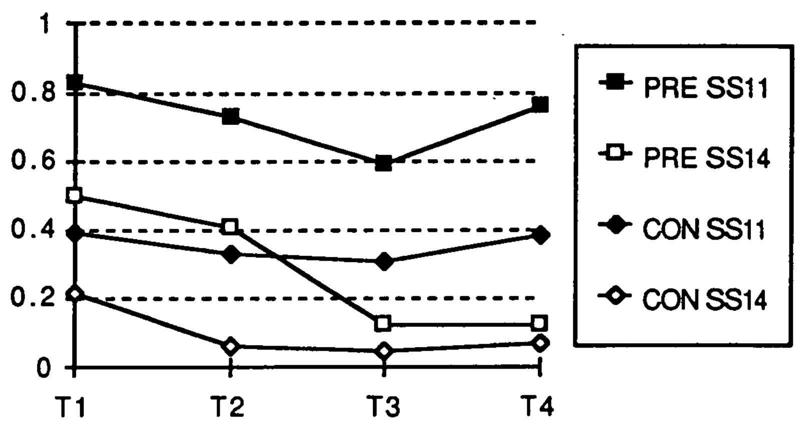


Figure 9·10 Average scores/item for SS11 & SS14 Groups on PRE & CON Scales

Numbers View versus Objects View. The remaining feature of Table 9·7 to merit comment is the entry which recorded that students in Group SS14 were significantly more inclined to view algebraic symbols as representing numbers than were the members of Group SS11. The contrast between the two groups in terms of their tendencies to regard the symbols as standing for numbers or for objects (or people) is brought out by the graphs in Figure 9·11. By Tests 3 and 4, the difference between the groups on the OBJ Objects View Scale was statistically significant, as reported in Appendix 9B, and the difference on the NBR Numbers View Scale continued to be statistically significant.

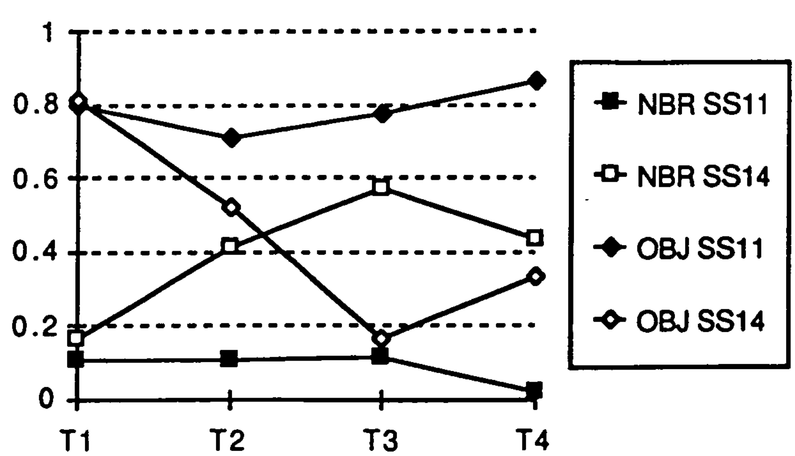


Figure 9·11. Average scores/item for SS11 & SS14 Groups on NBR & OBJ Scales

Figure 9·11 shows that SS11 students scarcely changed their point of view at all about what algebraic symbols basically represented. They mostly kept to the Objects View. For instance, for all valid test returns over four tests, 11 of this group of 17 students chose options in the both parts of the professors-and-students problem, Question 7, which indicated that they accepted that the letters '*P*' and '*S*' stood for people rather than numbers of people (e.g., '*S*' stands for "students"). Five more did the same for at least two of the tests and, in four of these cases, they chose people in both parts of the question in the last of the tests.

The graphs for Group SS14 display the dramatic drop in average item score for the OBJ Objects Scale and the corresponding rise in the preference for the Numbers View, as measured by the NBR Scale, over the three weeks of intervention teaching which introduced these students to algebra. In the six months following this period, there was some regression towards an Objects View. Only one member of Group SS14 kept to the people meaning (or Objects View) for the symbols in the professors-and-students problem for all four tests. Six of the eight in this group chose only number options in Test 3, but in the months between this stage and the final test, three of them became unsure of themselves on this point.

The contrast between the groups on whether or not they favoured an Objects View or a Numbers View of symbols emphasized that the development of an understanding of algebraic symbols as representing numbers rather than objects appeared to be beneficial for ensuring substantial progress in the algebraic tasks assessed by the research instrument.

Developmental Contrasts for Two More Groups.

There were 7 students who stayed at Standard 2 (S2) for all four tests, and another 10 who were at Standard 2 for first test but moved to Standard 4 (S4) on a later test. The first of these groups will be referred to as Group SS22, and the second as Group SS24. Six students in each group were from Advanced classes. There were no significant differences between the SS22 and SS24 groups on Test 1. However, as Tables 9B·5 and 9B·6 in Appendix 9B indicate, 7 statistically significant differences were recorded in Test 2, 12 in Test 3, and 10 in Test 4, all favouring Group SS14. Aspects of thinking which accounted for most of these differences were those related to an understanding of algebraic symbols as numerical variables and the readiness with which incorrect ideas were discarded. Hence, this supplementary analysis reinforced the outcomes derived from comparing the SS11 and SS14 groups. Table 9B·7 in Appendix 9B lists all scales which registered significant differences between Groups SS11 and SS14 and/or Groups SS22 and SS24.

Summary of Findings From Investigation 2

An important problem for many teachers of students beginning algebra is the question of why, in the same class, some students show little, if any, progress compared with other students. Throughout the data-collection phase, 17 students (Group SS11) remained at the lowest standard, S1, while 8 students (Group SS14) moved from S1 to S4, the highest standard. Investigation of the differential rates of progress for these two extreme groups, although small in size, shed some light on the problem.

In Test 1, Group SS14 scored significantly better on test total than Group SS11, although the difference between the mean scores was only about 4 points. There were no other significant differences at this pre-algebra stage. By Test 2, significant differences on test total and eight cognitive measures had developed. Group SS14 progressed more rapidly towards acquiring the concept that algebraic symbols stood for numbers which could vary than did their counterparts. Thus, empirical evidence indicated that it was the students' views of the meaning of the symbols which identified significant differences between those who were improving and those who were not. The SS14 students also acquired more quickly the basic skills required to substitute in simple expressions and to solve simple equations. Group SS11 misunderstood more persistently such conventions as the conjoining process in algebra, and tended more to miss the point of problems set. This indicated that they were not coping with the new ideas that were being placed before them. Significant differences between the groups were registered in each following test on more than 20 cognitive measures.

The failure of SS11 students in interpreting algebraic expressions (e.g., '$3y + 5$') to enable them to make substitutions correctly was accompanied by failure with test items requiring them to use the variable concept, as in comparing the values of two expressions (e.g., '$n + 2$' and '$2n$'). The overall outcomes for Group SS11 pointed to the likelihood that one of the reasons why they had not progressed was that they had not learnt to interpret algebraic expressions. A theoretical argument supporting this finding is that one needs to reduce the memory load required for simply understanding the meaning of a single algebraic expression before one can carry out cognitive tasks involving relationships between algebraic expressions.

Success in solving an equation with one operation on one variable (viz., '$3a = 36$') is a prelude to success in finding conditions for the equality of two algebraic expressions. It seemed, moreover, that the memory load needed to solve the easier case of equality had to be reduced before more complex equality problems could be solved. Group SS11 recorded a gradual improvement in solving the given equation but this did not necessarily result in increased success with the other equality problems.

The contrast between the groups also indicated that the development of an understanding that algebraic symbols represent numbers rather than objects appeared to be beneficial in ensuring progress in the successful completion of algebraic tasks.

A parallel investigation of two other groups, none of whom were at the lowest standard, S1, supported the finding that those who developed an understanding of the notion that algebraic symbols represented numerical variables were more likely to progress. Support was also given to the finding that those who were less inclined to change some of their early misunderstandings were more likely to make little, if any, progress in algebra.

Review and Forecast

This chapter reported two investigations based on a closer examination of the research data than the analyses discussed in earlier chapters. The focus was on differences between the ways various student groups developed understandings of the meaning and use of the algebraic symbols. Investigation 1 used responses to four test items and supported the previously-given hierarchies of understandings for algebraic symbols by means of correlations. Possible sequences of learning about algebraic symbols were considered in various other analyses, one of which also gave some limited support for Proposition 1. For Investigation 2, the Key to Progress, criterion levels were set so that students could be classified into levels or standards on their test scores. Those who stayed at the lowest standard were contrasted with those who progressed from that standard to the highest standard. Progress was characterized by early signs of learning to understand the conventions for writing first degree algebraic expressions, and the development of the concepts that algebraic symbols stood for numbers which could vary so that the symbols represented numerical variables. Propositions 2 and 3 from Chapter 8 were given further support. A similar investigation, carried out with more able students, produced similar outcomes.

Chapter 10 summarizes the main aspects of the thesis and includes suggestions for further research and for teaching.

CHAPTER 10

SUMMARY AND RECOMMENDATIONS

Summary

This thesis set out to contribute to the explication of several aspects of "Developing an Understanding of Algebraic Symbols." The project was designed to concentrate on learning more about the difficulties that algebra students experience when developing an understanding of the meaning and use of algebraic symbols, especially when beginning their study of the algebra of generalized number.

Amongst the stated reasons for undertaking this study were the importance of this branch of mathematics in terms of cultural heritage and of enabling the learner to enter the pursuit of higher mathematics, a role reflected in the common practice of including algebra in secondary school mathematics curricula. The research was a response to the well-documented fact that many students around the world have difficulty with learning algebra.

A series of investigations and analyses have been presented which were firmly-grounded in established theories of cognition and relevant research over the past two decades. While attempts were not made to revise or extend psychological theory, selected theories were applied to tease out the difficulty levels of the tasks, skills, and concepts examined, and to contribute to considerations of pathways to progress. The three theories of cognition chosen were those of Biggs and Collis (1982, 1991), Halford (1982, 1987), and Fischer (Fischer, 1980; Fischer & Lamborn, 1989). Previous researchers had identified major aspects of concern and had devised methods for measuring levels of cognitive progress in early algebra. A solid framework for the test instrument was built from the earlier work of Collis (1975a, 1975b), Harper (1979), Küchemann (1980), Booth (1983), Rosnick and Clement (1980), and MacGregor (1989, 1991).

Trialling of test items, although somewhat dependent upon the availability of respondents, was an essential step in producing a research test instrument which incorporated newly-devised items, revisions of items from earlier research, and earlier items that were not revised. The final test instrument proved to be useful as a tool for achieving the aims stated on page 17 of Chapter 1. The essence of its usefulness was its impregnation by the wisdom of the past two decades.

The research methodology provided avenues for examining cognitive processes involved in learning secondary school algebra. Obtaining written responses to the test instrument, teaching one group of students for three weeks, observing other lessons, and talking to a cross-section of students in interviews were means of gathering insights into the ways the students thought as they attended lessons on early algebra and attempted algebraic problems posed by the test items. The methodology of testing beginning Year 7 students four times and of testing students across the range of secondary classes to Year 12 enabled a picture to be built of hierarchical levels of understanding and possible sequences of learning.

Frequency tables obtained by processing the test responses from 517 students should make a useful contribution to the store of research information in the area. Comparisons at this level were possible with outcomes obtained by previous researchers.

Raw data were found suitable for establishing scales that tallied frequencies for categories of responses with similar cognitive content. Chapter 5 detailed the statistical support for grouping item responses into scales. Factor analysis and other statistical procedures used in establishing the scales helped attain the objective of identifying patterns of thought inherent in the array of student responses. Scale scores carried more information about student thinking than did single item responses. The scales produced ordinal measures of levels of understanding for algebraic concepts, some levels being correct, others partly correct, and still others incorrect. None of the previously-mentioned studies reported the use of scales.

The following is an example of scale formation which identified an underlying pattern of thinking. All the factor and principal component analyses reported in Tables 5·2 and 5·3 clustered together (as Cluster C) correct responses to solving a simple equation and substituting a given value in some algebraic expressions. At first glance, solving and substituting seemed to have little in common. Cognitive unity was, however, considered to reside within the cluster as, in each case, students had to interpret conventional algebra, use familiar arithmetic, and produce a numerical answer. The Substitute and Solve Scale was composed of responses to these items.

The formation of scale variables was one way of furthering the aim of considering interrelationships between measures previously restricted to separate studies or treated separately within one study. This was achieved in 14 scales by including within the same scale test items from more than one previous research project. For instance, the VBL Variables Scales was composed of items based on studies carried out by Collis (1975a), Harper (1979), and Küchemann (1980), and which were all measures of the degree to which students had developed the concept of an algebraic variable. Another 15 scales incorporated measures previously considered as isolated items within the one study. New items supplied the content for another 10 scales and new items were also incorporated in a further 14 scales alongside items based on previous studies. Tables at the end of Chapter 5 listed the sources of all test items used in scales.

Once the scale measures were established, they were applied to measure levels of understanding in algebra for the 517 students in the research sample. The frequency distributions of scale scores, as reported in Chapter 6 for the whole sample, provided a useful form for expanding the research data base about many aspects of learning the basics of the algebra of generalized arithmetic. These reports were organized according to difficulty levels based on previous work by Collis (1975a), Harper (1979) and Küchemann (1980).

Much of the research, in fact, dwelt on the question "What are the *levels of understanding* in algebra with respect to specific concepts/processes?" (Wagner and Kieran, 1989, p. 227). Some progress was made in identifying levels of understanding in psychological terms and in showing how the data provided empirical evidence for hierarchies of cognitive difficulty. The major challenge for secondary students appeared to lie in developing the concept of a numerical variable. Support was presented for ordering levels of understanding of algebraic symbols in the hierarchy of cognitive difficulty, the most challenging being the recognition that symbols could be regarded as numerical variables, in the sense of a species or class of number representing "simultaneously yet independently, many different numbers" (Wagner, 1983, p. 475). Of less difficulty was that of viewing algebraic symbols at the generalized number level where a general class of number was accepted, provided checking could readily be carried out with trial numbers. The category of generalized number was found to border more closely on the variable category than on the less challenging outlook which led to the replacement of symbols by specific sample values. The use of replacements was found to be beneficial in the learning continuum. It appeared to be a short-lived stage as very few of the participating students gave written responses in this form although, as interviews suggested, others used trial examples during test sessions to help them arrive at general algebraic answers. At the lowest cognitive level were a variety of categories of misunderstandings about symbols and conventions for their use, such as denying them any variation, removing them by substituting arbitrary numerical values, and regarding them as representations of objects or people rather than numbers of objects or people. Students were considered to be at a prestructural level when such misunderstandings precluded them from thinking algebraically.

The data from the Years 7 to 12 students were used in the correlational study reported in Chapter 6. Correlations, as listed in Table 6·10, between test scores and measures of the degree to which students were at different hierarchical levels of understanding for symbols were helpful in assessing the rank order of cognitive difficulty for various views of algebraic symbols. Some of the limitations noted (on page 202) regarding the use of these correlations for such an application were overcome by the Chapter 9 analyses based on responses to Questions 10 and 12 rather than on a variety of item groupings. The use of correlations (in Tables 9·1, 9·2, and 9·3) for ranking levels of understanding of symbols, as reflected in Year 7 students' responses to these two questions, further supported the classifications of different ways of viewing symbols into levels of difficulty. Comparisons of test scores, especially those for Years 8 to 12 students, in terms of categories of levels of understanding (in Figure 9·4) gave similar support.

As reported in Chapter 6, the first of ten propositions was investigated in terms of correlations between scale measures of the ways students viewed algebraic symbols and their test totals on items other than those in the particular scale being considered. These correlations included all test items. The main outcomes were that operating with symbols at prestructural levels generally seemed to be a hindrance to success on the algebraic tasks selected for testing, and, in contrast, as growth towards a view of algebraic symbols as numerical variables developed, success on algebraic tasks appeared to become more likely. The 1990 data gave support to Proposition 1, which stated that students with better levels of understanding of the meaning of algebraic symbols are more likely to have higher degrees of success with algebraic tasks. Prior to this, opinions had been expressed (as on pages 194 - 195) in support of such a proposition although empirical evidence had not been clearly documented.

It proved to be a challenge to devise test items for measuring the extent to which students regarded algebraic symbols as representing objects rather than numbers of objects or simply numbers. Chapter 7 discussed two investigations about the Numbers View and the Objects View of symbols. The first investigation extended reflections made by other researchers in relation to the well-known professors-and-students problem, with particular emphasis on students' levels of understanding of the symbols used in the context of the problem. The second helped resolve the paradox that some students who succeeded in using the sophisticated variable notion in some questions appeared to accept that algebraic symbols in other questions represented objects. Interview data were found helpful in interpreting students' responses to some of the items used, especially responses which only superficially implied that algebraic symbols could stand for objects. Some students seemed intuitively to imagine that, beyond generalized arithmetic, symbols may represent "virtually any object, person, place, or idea" (Wagner, 1981, p. 168).

In Chapter 8, information stored in the data was teased out further. Scale scores were reported for each of the four testings of the beginning algebra students and for the other students according to class groupings. Averages were also given for classes with different ability ratings. Analyses of algebraic tasks in terms of psychological understandings led to the formulation of the nine propositions discussed in the last section of the chapter. Each of these propositions forecast hierarchies of cognitive difficulty such that students would probably need the ability to succeed on one measure of cognitive processing before they could succeed on another. Three criteria were defined and applied to test the proposed hierarchies of cognitive difficulty. The criteria, devised by the researcher, proved efficient and provided an alternative to the approaches used by Harper (1979), Küchemann (1980, 1984), Hart (1981c), and others (Hart, 1981b) for a similar objective. The empirical data supported all propositions. While examining the process of moving from one level to a higher level, some constraints on the rate of development were identified, such as the need to become proficient in understanding simple algebraic expressions before problems involving the comparison of two expressions are likely to be mastered. Some sequential learning paths were tentatively supported by the data. In pursuing these investigations, interrelationships between measures based on items from different sources were again included. Appendix 8A lists the sources of items included in the reports on each proposition.

Rates of development were far from similar for all groups of students, as reported in Chapter 9. Further insights about cognitive hierarchies and sequences of learning were attained by conducting another investigation using categories of students determined by their responses to a small selection of test items. From the data, limited evidence was extracted for sequential learning paths in terms of the cognitive hierarchy for levels of understanding of symbols. It seemed that records for more than seven months would be needed for more appropriate data on sequences of learning as it was found that the Year 7 students in the study were still developing their notion of numerical variable by the end of the school year. Use of their responses made it difficult to distinguish clearly the benefits of moving towards the variable concept. The clearest support for such benefits was obtained from their Test 4 results in the correlational study based on responses to Questions 10 and 12 and was summarized in Table 9·3. It was not until Test 4 that some of the Year 7 students had advanced sufficiently to produce a high ranking for the correlation between scores on the VBL* Variable Scale and test scores on items other than Questions 10 and 12.

The final investigation centred on the contrast between the Year 7 students who recorded progress in algebra during the time of the data collection and those who did not progress noticeably. Comparisons were made of responses of the two extreme groups, namely, those who stayed at Standard 1, the lowest level of test scores for all four testings, and those who progressed from Standard 1 to Standard 4 (the highest level). This investigation found that characteristics revealed by those who made rapid progress in algebra, but not shown by the group who made very little progress, included noticeable gains in the acquisition of the concept of a numerical variable, and in developing the basic skills required for substitution in simple expressions and solving a simple equation. The students who made minimal progress tended more to miss the point of the new ideas that were being placed before them and were more reluctant to change views that contradicted some of the conventions of algebra. A parallel investigation did not include any of the students who were in the lowest performance group (Standard 1), but compared the group who stayed at Standard 2 with those who moved from Standard 2 to Standard 4. The features which discriminated between these groups were similar to those that distinguished the groups in the earlier investigation.

Limitations of the Study

Research such as this, which gathers its data in the field, is subject to many limitations. It is not intended to set out an exhaustive list of these here. However, it is appropriate to emphasize the following limitations which apply to the data and the analyses on which conclusions from this study are based.

1. The time period was not adequate for following the Year 7 students until, as a group, they had developed a sound understanding of the meaning and use of algebraic symbols. To make a deeper analysis of the progress of beginning algebra students, a longitudinal study extending for, say, two years or more would be necessary. The Year 7 students had not, on average, developed the concept of a numerical variable to any great extent in the seven months of the study. This produced difficulty in identifying sequences of learning on the basis of the Year 7 students' responses, although several sequences seemed to be appearing.

2. The collection of data from students across the secondary spectrum was beneficial to the study of the development of understandings in algebra. A more even spread of ability levels across the sample of students would have provided greater opportunities for comparing different rates of progress and for learning more about hierarchies of difficulty and sequences of learning. To obtain a more representative sample, the number of students involved could well be much greater than the 517 used in this study.

3. The numbers of students involved was small in some of the analyses, especially in Chapter 9. The numbers were small for some of the categories of understanding based on responses to parts of Questions 10 and 12, as graphed in Figure 9·5. Small groups were used also, as noted on page 326, when comparing the learning patterns of those who did not progress far in algebra with the patterns of those who moved to the top level of scores on the test. The results of such analyses needed to be regarded with caution and should not be broadly generalized due to the limitations of small samples for the statistical analyses used.

4. The classes studied in Year 7 were described in terms of the teaching approach as Manipulatives Classes or Textbook Classes. Having two modes of introducing algebra contributed to the study of the ways students developed their

understanding of algebraic symbols. To investigate thoroughly the variables involved, a much larger sample of classes would be needed, with an appropriate balance of subjects under each approach. Achieving a good balance of subjects would entail matching on such factors as ability, gender, time spent on algebra, and teacher characteristics. Steps would need to be taken to regulate teaching styles, content, and strategies used by all the teachers involved.

Recommendations

The complexities of both the learning process and the role of teaching in developing concepts were readily apparent as this project unfolded. There is much that is yet to be discovered about education and learning and much that needs to be clarified regarding understandings that have already been developed. The following recommendations for future research and teaching flow from the work done so far.

Meanings for symbols. Investigating the meanings students give to algebraic symbols was central to this study, as the title of the thesis emphasizes. Further research needs to be conducted on the degree to which students give varying meanings to algebraic symbols. For the algebra of generalized arithmetic, the paramount interpretation of symbols is that they stand for numerical variables.

It was found that some views held by students were elusive and difficult to measure. The creation of better evaluation techniques including, perhaps, better test items, appeared as a challenge for the future (cf. Collis, 1983; Collis, Romberg & Jurdak, 1986). Some questions used in this study had mixed levels of success, such as those designed to measure the extent to which students viewed symbols as representing objects. Amongst options given to students as possible meanings for an algebraic symbol, the following bracket might well be beneficial:

an object like a pear   objects like pears   the number of pears in a bag.

The possibility that students may be moving beyond the view of symbols as numerical variables needs to be kept in mind. Even though the algebra being studied is the algebra of generalized arithmetic, some students may be sufficiently astute to realize that symbols may be used to stand for a variety of options. In later secondary school, letters may be used to represent such entities as geometric shapes, processes, functions, and even solid objects.

Despite the limitations of the test items used, one of the misconceptions noted was that of regarding symbols as representing objects or people rather than numbers of these. This misconception seemed to be related to reduced progress in learning the algebra of generalized arithmetic. Perhaps students would benefit from teaching which more consciously emphasized that the symbols in this form of algebra, the form used in early secondary classes, always stand for numbers. It seems that carelessness with regard to the definition of symbols may be detrimental to progress in understanding algebra. If dealing with the professors-and-students problem, for example, it seems that it would be advisable to make clear during teaching that '*S*' stands for "number of students" and not for "students".

When computing assistance becomes more common in mathematics classrooms, one research focus could be the possible benefits of giving variables names which reflect their meaning. A comparison could be made, say, of the use of '*S*', or '*STUDENTS*', or '*NSTUDENTS*', and/or '*NSTUD*' as the name for the variable "number of students". The inclusion of the '*N*' as a reminder that the variable is, indeed, standing for a number of people and not the people themselves, might prove to be beneficial to beginning algebra students. Fisher (1988, p. 261), however, reported of the use of 'N_P' and 'N_S' in the professors-and-students problem that "the more explicit notation not only did not help problem solving performance but in fact may have hindered it."

Hierarchies of difficulty. Five levels of understanding for symbols were used throughout much of the research and the rationale for the ranking of these levels was shown to have at least some support from the data. A claim is not being made that the question of levels has been forever decided. Rather, a start has been made on this aspect of analysis and much more needs to be done. Questions such as the following could be researched further: Would data collected from another body of subjects endorse the finding that the development of the notion of generalized number seems to place students very close to attaining the concept of a numerical variable? Is use of replacement values (as sample solutions to a general problem) a beneficial technique and one that is short-lived in the process of developing clearer understandings of symbols as, at least, generalized numbers?

Empirical support for propositions about hierarchies of difficulty for algebraic tasks is a major outcome from this project. Analyses of task difficulties in terms of psychological understandings are supported by statistical analyses of the research data. Similar cohesion between theory and reality could be sought in terms of other propositions, such as:

1. When comparing two variables in one equation, success for equations in which one variable increases when the other decreases (e.g., '$c + d = 10$', and '$2x + y = 9$') is a prelude to success for equations in which both variables change in the same direction (e.g., '$2x - y = 9$'); and

2. Acquiring the ability to compare two expressions which both increase as the variable increases (e.g., '$t + t$' and '$t + 4$') is a prelude to acquiring the ability to

compare two expressions whose values change in opposite ways as the variable increase (e.g., '$t + 2$' and '$8 - t$').

Teaching approaches. Because teaching is presumed to make a difference (pp. 22 - 23), the methodology of incorporating two teaching approaches appeared useful for providing insights into the effects of learning environment on cognitive development. The pilot study summarized in Appendix 7B on comparing the two teaching approaches gave some indication that further research projects might profitably examine whether teaching approaches are related to rates of development of understandings in early algebra. Such a project could incorporate approaches involving the traditional textbook (cf., Margolinas, 1991), manipulatives (cf., Sowell, 1989; Quinlan et al., 1989; Quinlan, 1990, 1991; Quinlan & Collis, 1990), problem-solving (cf., Rachlin, 1987), or technology (cf., Sutherland, 1989; Rojano & Sutherland, 1991).

Biggs and Collis (1991), Collis and Biggs (1991), and Leinhardt (1988) direct attention to the possible advantages of learning which draws upon more than one mode of operation. Hiebert (1988) stresses that the first step in developing competence with written mathematical symbols is that of "connecting individual symbols with referents" (p. 335) and that "it is important that the referents are meaningful to the students" (p. 350). It would appear logical to undertake a research project to study the use of educationally suitable manipulatives in early algebra as a possible means for making use of the ikonic mode while leading to the concrete-symbolic, or as an available mode when difficulties arise because of uncertainties in remembering certain principles or when unfamiliar problem areas are first encountered.

Sowell (1989) examined 60 research studies on the effects of manipulative materials in mathematics instruction during the 1960s and 1970s. When comparing the concrete versus abstract instructional condition for effects on achievement, she reported that "when treatments lasted a school year or longer, the result was significant in favor of the concrete instructional condition. Treatments of shorter duration did not produce statistically significant results" (p. 502). Furthermore, when instructional conditions were randomly assigned, "attitude measures were significant in favor of the concrete instructional condition" (p. 502). She pointed out that, overall, the research studies produced mixed results, some favouring manipulatives, some not, and that answers were not readily available about when to use manipulatives and "which manipulatives are most appropriate in particular situations" (p. 504). Sowell's findings, together with experiences associated with the present project, suggest that further studies are needed in the area. Possible aspects to consider in such studies are as follows:

1. the characteristics of analogues in the form of concrete manipulatives if they are to assist students in forming their initial concepts about the algebra of generalized arithmetic (cf. Gentner, 1982, 1983; Gentner & Toupin, 1986; Boulton-Lewis, & Halford, 1991; Halford et al, in press). A reduction might follow in the use of analogues with "no empirical base ... to get over a temporary problem [leading] to further and more difficult problems at a later date" (Collis & Biggs, 1991, p. 199). Serious consideration might be given to the principle enunciated by Boulton-Lewis and Halford (1991, p. 37):

> The value of a concrete representation is that it mirrors the structure of the concept and the child should be able to use the structure of the representation to construct a mental model of the concept.

2. comparisons between the process of forming algebraic generalizations, such as '$2(y + 3) = 2y + 6$', from sets of arithmetic examples and from sets of examples presented by suitable models in the form of visual patterns. (See pages 31 - 32 and Appendix 3P.)

3. comparisons of the outcomes of learning activities that use concrete models in mappings only from algebra to models with those which use mappings from models to algebra. A suspicion generated in the study is that students drilled in the first mapping mode would tend to seek numerical answers to algebraic problems, while those familiar with the second mode would better appreciate the use of algebraic symbols for expressing generalizations. This needs further investigation.

4. long-term effects of manipulatives. Appendix 3P reported a pilot study on the long-term effects of introducing students to algebra by means of manipulatives. Interviews held after the delayed posttests, some six months after algebra was introduced, indicated that the trend was for a retention of the ability to use models profitably for self-correction, without a dependence on them. More highly-structured research on this issue is recommended.

5. the role that manipulatives play in developing discussion of mathematical concepts and processes as an integral part of the learning process and as a means of promoting metacognition. As Leinhardt (1988, p. 141) remarked,

> We need to explore more elegant ways of building consistent concrete representations that can serve as both an explanatory and exploratory system for children and to give them language tools for talking about such systems.

6. affective outcomes from the use of manipulatives. "It seems that a well organised inter-modal strategy influences children's attitude to, as well as their comprehension of, the content being taught" (Collis & Biggs, 1991, p. 202). Generally favourable reactions from teachers and students to the use of manipulatives for algebra were reported in Quinlan et al. (1989) following four years of action

research. Further investigation is needed involving a more representative group of teachers. Others could be included apart from enthusiasts given a special introduction to the approach. Affective outcomes of a variety of teaching approaches could be compared in a well-designed study.

Secondary mathematics teachers Recommendations to secondary mathematics teachers would be:

1. Deliberately spend time on strategies to develop in students an understanding of algebraic symbols as standing for numbers which can vary and which can be any real numbers, not just integers. This should be a focus of attention in the early stages of learning the algebra of generalized arithmetic and should be constantly reinforced over the years, especially by care in defining the meanings for symbols used in problems. Serious consideration should be given to Harper's (1979, pp. 272 - 273) suggestion that the early discussion of indeterminate equations (such as '$x + y = 10$') could avoid the conditioning of beginning algebra students to think of a "letter in arithmetic as an ordered entity with a unique determination" (Harper, 1979, p. ii);

2. Take time to ensure that students understand the meaning of algebraic expressions. They need to be able to interpret the usual conventions for writing algebra without inefficient use of cognitive space if they are to be able to cope with problems involving algebraic expressions. The careful use of suitable models can help here;

3. Develop the acceptance of lack of closure, so that students will write answers in symbols even though the values of the symbols are not known (and do not need to be known) and can carry out operations on unclosed algebraic expressions. The teachers need to be aware, however, that students can successfully write answers in algebraic form without thereby indicating that they have a sound understanding of the meaning of the symbols they are employing. Using symbols correctly could assist the gradual growth in understanding that they stand for numerical variables;

4. Use exercises which bring out the power of algebra to generalize from a pattern (cf. Stacey, 1989; Quinlan et al., 1989; Pegg & Redden, 1990a, 1990b) or across a number of examples of a structure. Examples could be in arithmetic form (cf. Halford & Boulton-Lewis, 1989), or could be presented by means of an appropriate model which could enable the general structure to be recognized within a visible pattern (cf. Appendix 3P); and

5. Be flexible and creative in teaching, keeping in mind a multi-modal approach (Collis, 1988; Biggs & Collis, 1991). Be careful of using "highly sophisticated techniques for helping students to avoid the necessity for using formal-operational level thinking" (Collis, 1975b, p. 53) and of "teaching them not to think" (Collis, 1988, p. 74) by imparting the view that mathematics consists of rules that appear to be

teacher-controlled and quite arbitrary. Experience with the Manipulative Classes in this study and a heightened awareness of the cognitive difficulties facing beginning algebra students point to the value of encouraging teachers to follow the advice given by Collis and Biggs (1991, p. 205):

> Teachers should take advantage of the concrete symbolic and earlier modal abilities already acquired when introducing material requiring formal mode functioning. To plunge students directly into the complete abstractions eventually required in a particular topic is likely to be found ineffective in practice.

Concluding Comment

Considerable progress has been made by scholars in understanding the psychology of learning and in gauging the influence of environmental factors such as teaching on cognitive development. This thesis has reported on the strategies used, the data obtained, and the implications of the findings in a project centred on the development of an understanding of algebraic symbols. The focus was valuable in terms of mathematics education research because the participating subjects included those who were being exposed to new and potentially very abstract mathematical concepts for the first time, thus providing a context for studying the process of concept development. The branch of mathematics under scrutiny was also seen to be valuable on account of the central role of the concept of a variable in the science of mathematics.

Projects conducted over the past two decades provided the soil and the seeds for the study to grow into a development and extension of some of the earlier ideas. The findings appear to have contributed to our store of knowledge about the variety of views students may hold for algebraic symbols, the ranking of these views according to levels of cognitive difficulty, relationships between views about symbols and degrees of success on algebraic tasks, hierarchies of difficulty for selected tasks in algebra, possible sequences of learning in algebra, and insights into factors relevant to differential rates of learning. The algebra of generalized number provides material for many more research endeavours.

The hope is that research into the teaching and learning of early algebra will ensure that, instead of becoming befuddled and insecure, beginning students will develop an understanding and appreciation of this important branch of mathematics and, through these, grow in self-confidence, both personally and as mathematicians.

REFERENCES

American Psychological Association (1990). Publication manual of the American Psychological Association (Third Edition). Hyattsville, MD: American Psychological Association.

Arndt, A. (1983). Al-Khwarizimi. Mathematics Teacher, Dec., 668-670.

Australian Education Council (1990). The national statement on mathematics for Australian schools. Carlton, Vic.: Curriculum Corporation for Australian Education Council.

Ball, W. (1960). A short account of the history of mathematics. New York: Dover.

Battro, A. M. (1973). Piaget: Dictionary of terms. New York: Pergamon Press.

Bell, A. W., Costello, J., & Küchemann, D. E. (1985). A review of research in mathematical education Part A: Research on learning and teaching. Windsor, Berks: NFER-Nelson.

Biddle, T. R., & Fischer, K. W. (in press). Cognitive development in educational contexts: Implications of skill theory.

Biggs, J. B., & Collis, K. F. (1982). Evaluating the quality of learning: The SOLO taxonomy (Structure of the observed learning outcomes). Sydney: Academic Press.

Biggs, J. B., & Collis, K. F. (1991). Multimodal learning and the quality of intelligent behavior. In H. A. H. Rowe (Ed.), Intelligence: Reconceptualization and measurement (pp. 57-75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Booker, G. (1987). Conceptual obstacles to the development of algebraic thinking. In J. C. Bergeron et al. (1987), Proceedings of the eleventh international conference of Psychology of Mathematics Education (Vol. 1, pp. 275-281). Montreal: Université de Montréal.

Booth, L. R. (1983). Misconceptions leading to errors in elementary algebra (Generalised arithmetic). Ph. D. Thesis, Chelsea College, University of London.

Booth, L. (1984a). Algebra: Children's strategies and errors, Windsor: NFER-Nelson.

Booth, L. (1984b). Misconceptions leading to error in elementary algebra. Journal of Structured Learning, 8, 125-138.

Booth, L. (1986). Difficulties in algebra. The Australian Mathematics Teacher, 42(3), 2-4.

Booth, L. (1988). Children's difficulties in beginning algebra. In A. F. Coxford, & A. P. Shulte (eds.). The Ideas of Algebra, K - 12: 1988 Yearbook (pp.20-32). Reston, VA: National Council of Teachers of Mathematics.

Boulton-Lewis, G. M., & Halford, G. (1991). Processing capacity and school learning. In G. Evans (Ed.), Learning and teaching cognitive skills (pp. 27-50). Melbourne: Australian Council for Educational Research.

Boyer, C. B. (1968). A history of mathematics. New York: John Wiley & Sons.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), Handbook of research on teaching, Third Edition (pp. 328 - 375). New York: Macmillan.

Brown, M., & Küchemann, D. (1976). Is it an 'add', Miss? (Part I). Mathematics in Schools, 5(5), 15-17.

Brown, M., & Küchemann, D. (1977). Is it an 'add', Miss? (Part II). Mathematics in Schools, 6(1), 9-10.

Campbell, K. J., Watson, J. M., & Collis, K. F. (in press). Volume measurement and intellectual development.

Carpenter, T. P., & Lindquist, M. M. (1989). Summary and conclusions. In M. M. Lindquist (Ed.). Results from the fourth mathematics assessment of the National Assessment of Educational Progress (pp. 160-169). Reston, VA: National Council of Teachers of Mathematics.

Case, R. (1985). Intellectual development: Birth to adulthood. Orlando, FL: Acadamic Press.

Case, R. (1987a). Structure and process of intellectual development. International Journal of Psychology, 22, 571-607.

Case, R. (1987b). Neo-Piagetian theory: Retrospect and prospect. International Journal of Psychology, 22, 773-791.

Cattell, R. B. (1978). The scientific use of factor analysis in behavioral and life sciences. New York: Plenum Press.

Clement, J. (1982). Algebra word problem solutions: Thought processes underlying a common misconception. Journal of Research in Mathematics Education, 13(1), 16-30.

Cochran, W. G. (1952). The χ^2 test of goodness of fit. The Annals of Mathematical Statistics, 23, 315-345.

Collis, K. (1972). A study of concrete and formal operations in school mathematics. Ph. D. Thesis, University of Newcastle.

Collis, K. (1975a). The development of formal reasoning. Newcastle, N. S. W. : University of Newcastle.

Collis, K. (1975b). A study of concrete and formal operations in school mathematics: A Piagetian viewpoint. Melbourne: Australian Council for Educational Research.

Collis, K. (1978). Operational thinking in elementary mathematics. In J. A. Keats, K. F. Collis, & G. S. Halford (eds.). Cognitive development (pp.221-248). Brisbane: John Wiley & Sons.

Collis, K. (1983). Development of a group test of mathematical understanding using superitem SOLO technique. Journal of Science and Mathematics Education in Southeast Asia, 6(1), 5-14.

Collis, K. (1986). Learning intellectual skills and school mathematics: a psychological viewpoint. In International Committee of PME, Psychology of Mathematics Education, Plenary lectures, PME 10 (pp. 1-23). London: University of London Institute of Education.

Collis, K. F. (1988). The 'Add up' or 'Take away' Syndrome. In J. Pegg (Ed.) Mathematical interfaces: Proceedings of the 12th biennial conference of AAMT (pp.66-75). Newcastle: Australian Association of Mathematics Teachers.

Collis, K. F., & Biggs, J. B. (1979). Classroom examples of cognitive development phenomena: The SOLO taxonomy. Hobart: Research Report, University of Tasmania.

Collis, K. F., & Biggs, J. B. (1991). Developmental determinants of qualitative aspects of school learning. In G. Evans (Ed.), Learning and teaching cognitive skills (pp. 185-207). Melbourne: Australian Council for Educational Research.

Collis, K. F., & Campbell, K. J. (1987). The SOLO taxonomy and the calculation of volume in primary school: A working paper. Hobart: University of Tasmania.

Collis, K. F., Romberg, T. A., & Jurdak, M. E. (1986). A technique for assessing mathematical problem-solving ability. Journal for Research in Mathematics Education, 17(3), 206-221.

Collis, K. F., & Watson, J. (1989). A SOLO mapping procedure. In G. Vergnaud, J. Rogalski, & M. Artique (Eds.), Proceedings of the thirteenth international conference of Psychology of Mathematics Education (Vol. I, pp. 180-187). Paris: Laboratoire PSYDEE.

Computing Department of Brisbane College of Advanced Education (1982). A concrete approach to algebra. Brisbane: Author.

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). Mathematical psychology: An elementary introduction. Englewood Cliffs, NJ: Prentice-Hall.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. Psychometrica, 16, 297-334.

Cureton, E. E., & D'Agostino, R. B. (1983). Factor analysis: An applied approach. Hillsdale, NJ: Lawrence Erlbaum.

Day, R., Webb, N. L., Nabate, M., & Romberg, T. A. (1987). Evaluation of quantitative literary series: Exploring Data and Exploring Probability (Program Report 87 - 5). Madison, WI: University of Wisconsin - Madison, Wisconsin

Center for Education Research. (ERIC Document Reproduction Service No. ED 289 707)

Demetriou, A. (Ed.) (1988). The Neo-Piagetian theories of cognitive development: Towards an integration. Amsterdam: North-Holland.

Dienes, Z. P. (1963). An experimental study of mathematics learning. London: Hutchinson.

Dietrich II, F. H., & Kearns, T. J. (1986). Basic statistics: An inferential approach. San Francisco, CA.: Dellen Publishing Company.

Dixon, W. J., & Massey, F. J. (1957). Introduction to statistical analysis. New York: McGraw-Hill.

Dubé, L. (1990). Modeling mathematical problem-solving behaviour. Journal of Science and Mathematics Education in S. E. Asia, XIII, 2, pp. 7-25.

Dunn-Rankin, P. (1983). Scaling methods. Hillsdale, NJ: Lawrence Erlbaum.

Dunteman, G. H. (1984). Introduction to multivariate analysis. Beverly Hills, CA: Sage.

Edwards, A. W. F. (1963). The measure of association in a 2 x 2 table. Journal of the Royal Statistical Society Series A, 126, 109-113.

Edwards, A. L. (1973). Statistical methods. New York: Holt, Rinehart & Winston.

Eves, H. (1969). An introduction to the history of mathematics. New York: Holt, Rinehart and Winston.

Firth, D. E. (1975). A study of rule dependence in elementary algebra. Master of Philosophy Thesis, University of Nottingham.

Fischer, K. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. The Psychological Review, 87, 477-531.

Fischer, K., Hand, H. H., & Russell, S. (1984). The development of abstractions in adolescence and adulthood. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), Beyond formal operations: Late adolescence and adult cognitive development (pp. 43-73). New York: Praeger.

Fischer, K., & Lamborn, S. D. (1989). Mechanisms of variation in developmental levels: Cognitive and emotional transitions during adolescence. In A. de Ribaupierre (Ed.), Transition mechanisms in child development: The longitudinal perspective (pp. 33-67). Cambridge University Press.

Fischer, K., & Pipp, S. L. (1984). Processes of cognitive development: Optimal level and skill acquisition. In R. J. Sternberg (Ed.), Mechanisms of cognitive development (pp. 46-80). New York: W. H. Freeman & Co.

Fischer, K., & Silvern, L. (1985). Stages and individual differences in cognitive development. Annual Review of Psychology, 36, 613-648.

Fisher, M. (1988). The students-and-professors problem revisited. Journal of Research in Mathematics Education, 19(3), 260-263.

Fleiss, J. L. (1973). Statistical methods for rates and proportions. New York: John Wiley & Sons.

Gentner, D. (1982). Are scientific analogies metaphors?. In D. S. Miall, Metaphor: Problems and perspectives (pp. 106-132). Sussex: The Harvester Press.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. Cognitive Science, 7, 155-170.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. Cognitive Science, 10, 277-300.

Gittleman, A. (1975). History of mathematics. Columbus, OH: Merrill.

Goodman, L. A. (1964). Simultaneous confidence limits for cross-product ratios in contingency tables. Journal of the Royal Statistical Society Series B, 26, 86-102.

Gleick, J. (1987). Chaos. London: Cardinal.

Grizzle J. E. (1967). Continuity correction in the χ^2 - test for 2 x 2 tables. The American Statistician, Oct., 28-32.

Halford, G. S. (1982). The development of thought. Hilldsale, NJ : Erlbaum.

Halford, G. S. (1987). A structure-mapping approach to cognitive development. International Journal of Psychology, 22, 609-642.

Halford, G. S., & Boulton-Lewis, G. M. (1989). Value and limitations of analogs in teaching mathematics. Paper for the Third Conference of the European Association for Research on Learning and Instruction. Madrid, Spain.

Halford, G. S., & Wilson, W. H. (1980). A Category Theory Approach to Cognitive Development. Cognitive Psychology, 12, 356-411.

Halford, G. S., Wilson, W. H., Guo, J., Wiles, J., & Stewart, J. E. M. (in press). Connectionist implications for processing capacity limitations in analogies.

Harper, E. (1979). The child's interpretation of a numerical variable. Ph. D. Thesis, University of Bath.

Harper, E. (1981). Psychological changes attending a transition from arithmetical to algebraic thought. In C. Comiti, & G. Vergnaud (Eds.), Proceedings of the fifth international conference of Psychology of Mathematics Education (pp. 171-176). Grenoble: Scientific and Medical University of Grenoble.

Harper, E. (1987). Ghosts of Diophantus. Educational Studies in Mathematics, 18, 75-90.

Harris, C. W. (1967). On factors and factor scores. Psychometrika, 32, 363-379.

Harris, R. J. (1975). A primer of multivariate statistics. New York: Academic Press.

Hart, K. M. (General editor).(1981a). Children's understanding of mathematics: 11 - 16. London: John Murray.

Hart, K. (1981b). Hierarchies in mathematics education. Educational Studies in Mathematics, 12, 205-218.

Hart, K. (1981c). The hierarchies. In K. M. Hart (General editor).(1981), Children's understanding of mathematics: 11 - 16 (pp. 187-207). London: John Murray.

Heath, T. L. (1910). Diophantus of Alexandria. Cambridge University Press.

Herscovics, N., & Kieran, C. (1980). Constructing meaning for the concept of equation. The Mathematics Teacher, 73(8), 572-580.

Hiebert, J. (1988). A theory of developing competence with written mathematical symbols. Educational Studies in Mathematics, 19, 333-355.

Hollingdale, S. (1989). Makers of mathematics. London: Penguin.

Howell, D. C. (1982). Statistical methods for psychology. Boston, MA: Duxbury Press.

Husén, T. (Ed.) (1967). International study of achievement in mathematics. New York: John Wiley & Sons.

Ifrah, G. (1988). From one to zero: A universal history of numbers. London: Penguin.

Italey, W. (Ed. in chief) (1969). Encyclodædia Britannica. Chicago, IL: Benton.

Kaput, J. J. (1987). PME Algebra papers: A representational framework. In J. C. Bergeron, N. Herscovics, & C. Kieran (Eds.), Proceedings of the eleventh international conference of Psychology of Mathematics Education (Vol. I, pp. 345-354). Montreal: Université de Montréal.

Kaput, J., & Sims-Knight, J. E. (1983). Errors in translations to algebraic equations: Roots and implications. Focus on learning problems in mathematics. 5, 63-78.

Kenny, D. A. (1987). Statistics for the social and behavioral sciences. Boston, MA: Little, Brown & Co.

Kieran, C. (1981a). Concepts associated with the equality symbol. Educational Studies in Mathematics, 12, 317-326.

Kieran, C. (1981b). Pre-algebraic notions among 12- and 13- year olds. In C. Comiti, & G. Vergnaud (Eds.), Proceedings of the fifth international conference of Psychology of Mathematics Education (pp. 158-164). Grenoble: Scientific and Medical University of Grenoble.

Kieran, C. (1989a). The early learning of algebra: A structural perspective. In S. Wagner, & C. Kieran (Eds.), Research issues in the learning and teaching of algebra (Vol. 4, pp. 33-56). Reston, VA: National Council of Teachers of Mathematics.

Kieran, C. (1989b). A perspective on algebraic thinking. In G. Vergnaud, J. Rogalski, & M. Artique (Eds.), Proceedings of the thirteenth international conference of Psychology of Mathematics Education (Vol. II, pp. 163-171). Paris: Laboratoire PSYDEE.

Kieran, C., & Wagner, S. (1987). Research Agenda Project Conference on the teaching and learning of algebra. In J. C. Bergeron, N. Herscovics, & C.

Kieran (Eds.), Proceedings of the eleventh international conference of Psychology of Mathematics Education (Vol. III, pp. 425-435). Montreal: Université de Montréal.

Kirk, R. E. (1982). Experimental design: Procedures for the behavioral sciences. Monterey, CA.: Brooks/Cole Publishing Company.

Kline, M. (1972). Mathematical thought from ancient to modern times. New York: Oxford University Press.

Küchemann, D. (1978). Children's understanding of numerical variables. Mathematics in Schools,7(4), 23-27.

Küchemann, D. E. (1980). The understanding of generalised arithmetic (algebra) by secondary school children. Ph. D. Thesis, Chelsea College, University of London.

Küchemann, D. (1981). Algebra. In K. M. Hart, (General editor). Children's understanding of mathematics: 11 - 16 (pp. 102-119). London: John Murray.

Küchemann, D. E. (1984). Stages in understanding algebra. Journal of Structured Learning, 8, 113-124.

Leinhardt, G. (1988). Getting to know: Tracing students' mathematical knowledge from intuition to competence. Educational Psychologist, 23(2), 119-144.

Lewis, D., & Burke, C. J. (1949). The use and misuse of the chi-square test. Psychological Bulletin, 46, 433-489.

Lin, F. (1988). Societal differences and their influences on children's mathematics understanding. Educational Studies in Mathematics, 19, 471-497.

Lins, R. C. (1990). A framework for understanding what algebraic thinking is. In G. Booker, P. Cobb, & T. N. de Mendicuti (Eds.), Proceedings fourteenth PME conference, (Vol.I., pp. 93-100). Mexico: Program Committee of 14th PME Conference.

Lockhead, J. (1980). Faculty interpretations of simple algebraic statements: The professor's side of the equation. Journal of Mathematical Behavior, 3(1), 29-37.

Lowe, I., & Lovitt, C. (1984). Reality in mathematics education (RIME) Lesson pack. Melbourne: Materials Production, Curriculum Branch, Education Department of Victoria.

Lynch, B., Parr, R., Picking,L., & Keating, H. (1980). Maths 8. Melbourne: Longman Cheshire.

MacGregor, M. (1989). Students' errors in constructing and interpreting algebraic equations. Ph.D. thesis, University of Melbourne.

MacGregor, M. (1991). Making sense of algebra: Cognitive processes influencing comprehension. Melbourne: Deakin University Press.

McKnight, C. C., Travers, K. J., Crosswhite, F. J., & Swafford, J. O. (1985). Eighth-Grade Mathematics in U.S. Schools: A Report from the Second

International Mathematics Study. Arithmetic Teacher, 32(8), 20-26.

McLeod, J. K., Ganderton, G. D., Creeley, T. A., & Tanti, L. (1988). Mathematics for Australian schools Year 7. Melbourne: MacMillan.

McNemar, Q. (1969). Psychological statistics. New York: John Wiley & Sons.

McPherson, G. (1990). Statistics in scientific investigation: Its basis, application, and interpretation. New York: Springer-Verlag.

McSeveny, A., Conway, R., & Wilkes, S. (1989). Signpost Mathematics Year 7. South Melbourne: Longman and Cheshire.

Mantel, N., & Greenhouse, S. W. (1968). What is continuity correction? The American Statistician, Dec., 27-30.

Margolinas, C. (1991). Interrelations between different levels of didactic analysis about elementary algebra. In F. Furunghetti (Ed.), Proceedings fifteenth Psychology of Mathematics Education conference (Vol. II, pp. 381-388). Genoa: Università di Genova.

Mason, O. M., & Broom, W. J. (1979). Introducing algebra with Math-Tiles. Perth: Achievement Mathematics Publications.

Mathematics and Computing Department of Brisbane College of Advanced Education (1982). A concrete approach to algebra. Brisbane: Author.

National Council of Teachers of Mathematics (NCTM) (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: Author.

N.S.W. Board of Secondary Education (1989). Memorandum No.150/89.

Newmann, F. M., & Thompson, J. A. (1987). Effects of cooperative learning on achievement in secondary schools: A summary of research. Madison, WI: University of Wisconsin - Madison, Wisconsin Center for Education Research. (ERIC Document Reproduction Service No. ED 288 853)

Nie, N. H., Bent, D. H., Jenkins, J. G., Steinbrenner, K., & Hull, C. H. (1975). SPSS: Statistical Package for Social Sciences. New York: McGraw-Hill.

Norris, E. T. (1973). Algebra of a sort - Mathematics for the majority. London: Chatto & Windus Educational.

Nunn, T. P. (1919). The teaching of algebra. London: Longmans.

Pascual-Leone, J. (1987). Organismic processes for neo-Piagetian theories: A dialectical causal account of cognitive development. In A. Demetriou (Ed.), 1988, The Neo-Piagetian theories of cognitive development: Towards an integratio.(pp. 25-64). Amsterdam: North-Holland.

Pegg, J. & Redden, E. (1990a). Student experiences in developing algebraic symbolism using number patterns. A paper presented at Mathematics Education Research Group of Australasia Conference. Bathurst, N.S.W.

Pegg, J. & Redden, E. (1990b). Procedures for, and experiences in, introducing algebra in New South Wales. Mathematics Teacher, May, 386-391.

Peterson, P. L. (1988). Teaching for higher-order thinking in mathematics: The challenge for the next decade. In D. A. Grouws, T. J. Cooney, & D. Jones (Eds.), Perspectives on research on effective mathematics teaching (pp. 2 - 26). Reston, VA: National Council of Teachers of Mathematics.

Prohhorov, A. M. (Ed. in chief) (1973). Great Soviet encyclopedia. New York: Macmillan

Quinlan, C. (1990). Addressing difficulties with early algebra. In M.Bezzina, & J.Butcher (Eds.), The changing face of professional education - Collected papers of The AARE Conference Sydney University (pp. 774-779). Sydney: Uni.of Western Sydney, Nepean.

Quinlan, C. (1991). Structure mapping with manipulatives for algebra. In F. Furunghetti (Ed.), Proceedings fifteenth Psychology of Mathematics Education conference (Vol. I, p. lxxv). Genoa: Università di Genova.

Quinlan, C., & Collis, K. (1990). Investigating a rationale for a concrete approach to algebra. In K.Milton, & H.McCann (Eds.) Mathematical turning points: Strategies for the 1990s.: Papers presented to the 13th Biennial Conference of AAMT (pp. 435-454). Hobart: Mathematical Association of Tasmania.

Quinlan, C., Low, B., Sawyer, E., & White, P. (1989). A concrete approach to algebra. Sydney: Mathematical Association of N. S. W.

Rachlin, S. (1987). Using research to design a problem-solving approach for teaching algebra. In Proceedings of fourth Southeast Asian Conference of Mathematics Education (ICME-SEAMS) (pp. 151-161). Singapore: Singapore Institute of Education.

Resnikoff, H. L., & Wells, R. O. (1984). Mathematics in civilization, New York: Dover.

Robitaille, D. F. (1989). Students' achievements: Population A. In D. F. Robitaille, & R. A. Garden (Eds.), The IEA study of mathematics II: Contexts and outcomes of school mathematics (pp.102-125). Oxford: Pergamon Press.

Robitaille, D. F., & Garden, R. A. (Eds.) (1987). Second international mathematics study: Appendices. Oxford: Pergamon Press.

Robitaille, D. F., & Garden, R. A. (Eds.) (1989). The IEA study of mathematics II: Contexts and outcomes of school mathematics. Oxford: Pergamon Press.

Romberg, T. A. (1977). Developing mathematical processes: The elementary mathematics program for Individually Guided Education. In H. J. Klausmeier, R. Rossmiller, & M. Saily (Eds.), Individually guided elementary education: Concepts and practices (pp. 77-109). New York: Academic Press.

Romberg, T. A., & Collis, K. F. (1987). Learning to add and subtract (Journal for Research in Mathematics Education Mongraph Number 2). Reston, VA: National Council of Teachers of Mathematics.

Rojano, T., & Sutherland, R. (1991). Symbolising and solving algbera word problems: The potential of a spreadsheet environment. In F. Furunghetti (Ed.), Proceedings fifteenth Psychology of Mathematics Education conference (Vol. III, pp. 207-213). Genoa: Università di Genova.

Rose, S. P., & Fischer, K. W. (in press). Constructing task sequences: A structured approach to skill theory.

Rosnick, P. (1981). Some misconceptions concerning the concept of variable. Mathematics Teacher, Sept., 418-420, 450.

Rosnick, P. (1982). Students' symbolization processes in algebra (Cognitive Development Project). Amerhest, MA: University of Massachusetts, Department of Physics and Astronomy. (ERIC Document Reproduction Service No. ED 300 230)

Rosnick, P., & Clement, J. (1980). Learning without understanding: The effect of tutoring strategies on algebra misconceptions. The Journal of Mathematical Behavior, 3(1), 3-27.

Sawyer, W. W. (1964). Vision in elementary mathematics. Harmondsworth, Middlesex: Penguin.

Schoenfeld, A. H., & Arcavi, A. (1988). On the meaning of variable. Mathematics Teacher, 81(6), 420-427.

Scott, J. F. (1969). A history of mathematics: From antiquity to the beginning of the 19th century. London: Taylor & Francis.

Shayer, M. (1987). Neo-Piagetian theories and educational practice. International Journal of Psychology, 22, 751-772.

Shepard, R. N. (1972). Introduction to Volume I. In R. N. Shepard, A. K. Rommey, & S. B. Nerlove (Eds.) Multidimensional scaling: Theory and applications in the behavioral sciences (pp. 1-23). New York: Seminar Press.

Siegel, S. (1956). Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.

Simpson, J. A., & Weiner, E. S. C. (Eds.) (1989). The Oxford English Dictionary (2nd ed.). Oxford: Clarendon Press.

Singleton, G. (1989). Progress Maths Year 8. Melbourne: Macmillan.

Sowell, E.J. (1989). The effects of manipulative materials in mathematics instruction, Journal for Research in Mathematics Education, 20(5), 498-505.

SPSS Inc. (1986). SPSSX user's guide (2nd ed.). Chicago, IL: Author.

Stacey, K. (1989). Finding and using patterns in linear generalising problems. Educational Studies in Mathematics, 20, 147-164.

Sternberg, R. J. (1987). A day at Development Downs: Sportscast for race #2 - Neo-piagetian theories of cognitive development. International Journal of Psychology, 22, 507-529.

Sutherland, R. (1989). Developing algebraic understanding: The potential of a computer based environment. In G. Vergnaud, J. Rogalski, & M. Artique (Eds.), Proceedings of the thirteenth international conference of Psychology of Mathematics Education (Vol.III, pp. 205-212). Paris: Laboratoire PSYDEE.

Swafford, J. O., & Brown, C. A. (1989). Variables and relations. In M. M. Lindquist (Ed.). Results from the fourth mathematics assessment of the National Assessment of Educational Progress (pp. 55-63). Reston, VA: National Council of Teachers of Mathematics, pp. 55-63.

Timm, N. H. (1975). Multivariate analysis with applications in education and psychology. Monterey, CA: Brooks/Cole.

Wagner, S. (1981). An analytical framework for mathematical variables. In C. Comiti, & G. Vergnaud (Eds.), Proceedings of the fifth international conference of Psychology of Mathematics Education (pp. 165-170). Grenoble: Scientific and Medical University of Grenoble.

Wagner, S. (1983). What are these things called variables? Mathematics Teacher, 76(6), 474-479.

Wagner, S., & Kieran, C. (1989). An agenda for research on the learning and teaching of algebra. In S. Wagner, & C. Kieran (Eds.), Research issues in the learning and teaching of algebra. (Vol.4, pp. 220-237). Reston, VA: National Council of Teachers of Mathematics.

Wigley, A., Rooke, D., Hart, M., & Bell, A. (1984). Algebra: Ideas and material for Years 2 - 5 in the secondary school - South Notts Project. Nottingham: Shell Centre for Mathematical Education, University of Nottingham.

Yates, F. (1934). Contingency tables involving small numbers and the χ^2 test. Journal of the Royal Statistical Society Suppl.(1), 217-235.

The 1989 Timetable for Preliminary Investigations

Fri.NOV.24	School X's PRETESTS: "BRAIN-BOX QUIZ No.1" (Appendix 3B) 1.25 - 2.15 p.m. Group I: Year 7M (Manipulatives) 2.15 - 3.00 p.m. Group II: Year 7T (Textbook) (Parallel classes, mixed ability).
Tues.NOV.28	9.30 - 10.15 a.m. TAUGHT School X's Year 7M "AREA model", as in Unit One (Quinlan et al., 1989) Worksheet Two
Mon.DEC.4	11.15 - 12.00 TAUGHT School X's Year 7M "OBJECTS-AND-CONTAINERS model", as in Unit One Worksheet Three page 3 and Unit Two Worksheet One - first few pages.
Tues.DEC.5	9.15 - 12.00 INTERVIEWED 10 School X's Yr.7M boys in pairs - mainly on selected questions from Pretest.
Thurs.DEC.7	School X's Year 8 TEST: "BRAIN-BOX QUIZ No.2" (Appendix 3C) 9.30 - 10.15 a.m. for BOTH Yr.8 classes (Group V).
Fri.DEC.8	PRETEST School Y: "BRAIN-BOX QUIZ No.2" 9.30 - 10.15 a.m. Group III: Yr.7M 10.15 - 11.00 a.m. Group IV: Yr.7T. 11.15 - 12.00 TAUGHT School X's Yr.7M - extended Area model to modelling fractions and subtraction, and extended Objects and Containers model to consider the case $2x + 4y = 6y$, from Unit 2 Worksheet 1, Q.4 POSTTEST School X: "BRAIN-BOX QUIZ No.1" 1.25 - 2.15 p.m. Group I: Yr.7M 2.15 - 3.00 p.m. Group II: Yr.7T.
Mon.DEC.11	9.30 - 10.15 a.m. TAUGHT School Y's Yr.7M - AREA model, including fractions and zero, and subtraction. 11.15 - 12.00 TAUGHT School Y's Yr.7M - OBJECTS-AND-CONTAINERS model, including fractions and zero, and case $x = y$.
Tues.DEC.12	POSTTEST School Y: "BRAIN-BOX QUIZ No.2" 9.30 - 10.15 a.m. Group IV: Yr.7T. 2.15 - 3.00 p.m. Group III: Yr.7M.

The 1990 Timetable for Preliminary Investigations

FEB.28 ... Started working twice a week for 30-40 minutes after school with a volunteer group of about 5 to 8 Yr.9 students from School X (part of Group V).
Trialling was carried out for a newly-written student worksheet:
Unit One Worksheet One Part B - Appendix 3M.
These sessions continued until the Easter break in early April.

MAR.2 20 students Yr.9 School X (part of Group V - included 8 from the volunteer group) trialled revised Professors-and-Students question, as in test "Algebra Project 1990" (Appendix 3E).

MAR.12, 14, 15 Three groups of UNIVERSITY students (Group VI) from Catholic College of Education Sydney (now Australian Catholic University - N.S.W.) completed revised Professors-and-Students question, as in test "Algebra Project 1990" (Appendix 3E).

28 students in Yr.2 B.Ed. (Primary), 18 in Yr.2 B.Ed.(Secondary - Mathematics), and 15 in Yr.3 B.Ed.(Secondary - Mathematics).

MAR.16 30 students Yr.9 School X (part of Group V) trialled Harper questions in written form and questions on Arithmetic processes, using question sheet headed "1990 Algebra Project" (Appendix 3F)

MAR.29 and APRIL 5 Two groups of UNIVERSITY students (Group VII) from University of Tasmania completed revised Professors-and-Students question and one Harper question (about two parallel lines):

17 students in Dip.Ed. (Primary), and 19 in Yr.2 B.Ed. (Primary).

APRIL 5 Re-tested 20 students from Yr.9 School X (part of Group V) on selected questions, headed "Yr.9 Test 1990" (Appendix 3G).

DELAYED POSTTEST for Yr.7 (1989) students from School Y (Groups III and IV), using "New Test 2 1990" (Appendix 3H), the students then being in Yr.8.

Trialled the FINAL RESEARCH INSTRUMENT as "Algebra Project New Test 1990" (Appendix 3I) with Yr.7 School X (Group VIII) - these Yr.7 students had not started classroom algebra.

BRAIN-BOX QUIZ No.1

NAME: . . . . M/F DATE: . . . .
CLASS: . . .DATE OF BIRTH: . . . . SCHOOL: . . . .

1. If x represents the number of sheep on a certain farm, could x equal:
   (CIRCLE either YES or NO in each case.)
   (i) 200 YES/NO (ii) 2 YES/NO (iii) 0 YES/NO (iv) 3.7 YES/NO
   (v) -11 YES/NO (vi) 3 728 YES/NO (vii) a sheep YES/NO ?

2. Numbers may be added in any order.
   For example, 3 + 2 = 2 + 3 (= 5), and 7 + 4 = 4 + 7 (= 11).
   Thus, we can write **a + b = b + a** , where **a** and **b** are numbers.
   (a) In **a + b = b + a**, could **a** equal
   (i) 3 YES/NO (ii) 16 YES/NO (iii) 578 YES/NO ?
   (b) In **a + b = b + a**, could **b** equal
   (i) 0 YES/NO (ii) 2.753 YES/NO (iii) - 6 YES/NO ?
   (c) In **a + b = b + a**, could **a** equal
   (i) an apple YES/NO (ii) the number of apples in a box YES/NO ?

3. In a football match, Essendon scored **p** goals and Geelong scored **q** goals.

   How many goals altogether were scored in the match? ....

4. A milkman can just fit n milk cartons in a milk crate. What number of cartons could he fit in 4 of the same sized crates? Tick ALL correct answers:

   4 + n **4 x n** **4n** **n + n + n + n** **nnnn**

   n+4 **n x 4** **3n + n** **2(n + n)** $\mathbf{n^4}$

5. (a) **a + 3a** can be written in a shorter way as **4a.**
   Write the following in a shorter way, if possible:

   (i) 2a + 5b + a ....

   (ii) 2a + 5b ....

   (iii) 3a - b + a ....

   (iv) 5a + 3b + 2a - 4b ....

   (b) Can a and b represent anything here, do they stand for anything? If so, what?

   ....

   (c) Could your answers be shortened any further? Explain.

   ....

6. If $y = 3$, what is the value of (i) $2y$? ....

(ii) $2y + 5$? ....

(iii) $2(y + 5)$? ....

(iv) $2y + y$? ....

(v) $3y - y$? ....

7. If d represents a number, could the following be true?

(i) $d + 6 = 10$ YES/NO. If YES, d = ...

(ii) $d + 6 = 6$ YES/NO. If YES, d = ...

(iii) $d + 6 = 6.8$ YES/NO. If YES, d = ...

(iv) $d + 6 = 2006$ YES/NO. If YES, d = ...

(v) $d + 6 = 4$ YES/NO. If YES, d = ...

8. For a school excursion, 3 buses take f students each and 4 cars take g students each.

(i) What does 3f tell us ?...

(ii) Give the total number of students taken by these buses and cars. ....

(iii) One car leaves early with g students. How many students remain? ......

9. (i) Add 4 onto $n + 5$. ....

(ii) Add 4 onto $3n$. ....

(iii) Multiply $n + 5$ by 4. ....

10. (i) In a school there are 15 times as many students as there are teachers. Using S = the number of students, and T = the number of teachers, which of these two equations is correct?
(a) $15S = T$ (b) $15T = S$ MY CHOICE: ....

(ii) "The number y is eight times the number z."
Write this information in mathematical symbols. (Write an equation.)

....

(iii) If y and d are any two numbers, which, if either, is the bigger? Give a reason for your answer.

....

(iv) If y and d are two positive numbers and $6y = d$, which is the bigger number, y or d? ....

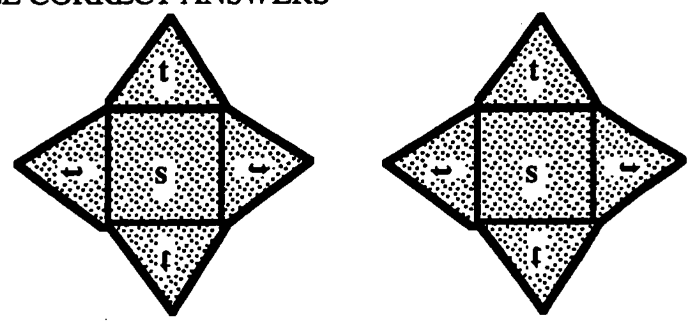
BRAIN BOX QUIZ No.1 page 3

(v) In a certain town the number of bicycles (x) and the number of motor bikes (w) are related by the equation **4x = w.**
What does this equation tell you about the numbers of bicycles and motor bikes in the town?

....

11. Suppose **s** represents a square and **t** represents a triangle as shown:

Then what would the following diagram represent?
TICK ALL CORRECT ANSWERS



t + t + t + t + s + t + t + t + t + s **4t + s + 4t + s**

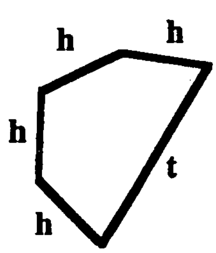
8t + 2s **2s + 8t** **ttttstttts** **4t2s**

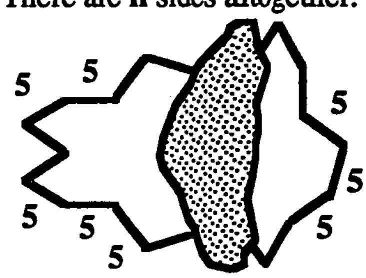
s + 4t + s + 4t **2(s + 4t)** **4(2t + s)**

12. Write down the perimeter (the distance around) for these two cases:

(a) The letters give the number of centimetres in the lengths of the sides.

(b) Part of this shape is hidden. All the sides are of length 5 cm. There are **n** sides altogether.





(a) Perimeter = ...

(b) Perimeter = ...

13. Which is bigger **2n** or **n + 2** ? Explain carefully.

....

14. (a) If **c + d = 10**, tick the values that **c** could have:

3 10 12 7.4 a cabbage

(b) If **c + d = 10**, what happens to **d** as **c** gets bigger?

....

(c) If **c + d = 10**, and **c** is always less than **d**, what values may **c** have?

....

15. (i) If the expression **4g + 8** represents a number of flowers, COULD **4g** represent the number of flowers in 4 equal bunches of flowers ? YES / NO.
If YES, the **g** represents ....
If YES, the **8** represents ....

(ii) If the expression **4f + 8** represents a number of fish, MUST **4f** represent 4 FISH? YES / NO.
IF YES, **f** represents ....and **8** represents ....
Could **4f** represent 12 fish ? YES / NO.
IF YES, **f** represents ....and **8** represents ....
Could **4f + 8** represent 24 fish? YES / NO.
IF YES, **f** represents ....and **8** represents ....

16. Decide whether the following statements are TRUE always, never or sometimes. Tick the correct answer.
If you tick "**true only when ..**", write **when it is true.**
All the letters stand for whole numbers or zero (0, 1, 2, 3, 4, ...)

(i) $a + b + c = a + x + c$
O true always
O never true
O true only when ....

(ii) $2a + 3b + 7 = 5a + 7$
O true always
O never true
O true only when ....

(iii) $2a = a + 2$
O true always
O never true
O true only when ....

(iv) $a + 2b + 2c = a + 2b + 4c$
O true always
O never true
O true only when ....

ACCEPTABLE ANSWERS (See Appendix 3D for further details about responses)
1i Y; **ii** Y; **iii** Y; **iv** N; **v** N; **vi** Y; **vii** N. **2a** Y; Y; Y; **b** Y; Y; Y; **c** N; Y. **3** p + q. **4** 4 x n; 4n; n+n+n+n; n x 4; 3n+n; 2(n+n). **5ai** 3a+5b; **ii** 2a+5b; **iii** 4a-b; **iv** 7a-b; **b** numbers; **c** No, as we don't know what numbers the letters stand for. **6i** 6; **ii** 11; **iii** 16; **iv** 9; **v** 6. **7** all Yes: **i** 4; **ii** 0; **iii** 0.8; **iv** 2000; **v** -2. **8i** The number on the buses; **ii** 3f + 4g; **iii** 3f + 3g. **9i** n + 9; **ii** 3n + 4; **iii** 4n + 20. **10i** b; **ii** y = 8z; **iii** Cannot tell as they may be any value; **iv** d; **v** There are 4 times as many motor bikes as bicycles. **11** t+t+t+t+s+t+t+t+t+s; 4t+s+4t+s; 8t+2s; 2s+8t; s+4t+s+4t; 2(s+4t). **12a** 4h+t; **b** 5n. **13** Depends - for n > 2, 2n is larger, but for n < 2, n + 2 is larger, and if n = 2, they are equal. **14a** 3; 10; 12; 7.4; **b** d gets smaller; **c** c < 5 (e.g., 4.9, 0. -1).**15i** Yes; no.of flowers in one bunch; 8 extra flowers; **ii** No, one fish, 8 more fish; Yes, 3 fish, 8 more fish; Yes, 4 fish, 8 fish. **16i** true only when b = x; **ii** true only when a = b; **iii** true only when a = 2; **iv** true only when c = 0.

BRAIN-BOX QUIZ No. 2

NAME: . . . . M/F DATE: . . . . . . CLASS: . . .

DATE OF BIRTH: . . . (month). . . .(year) SCHOOL: . . . .

1. If **d** represents the number of cows on a certain farm, which of the following could **d** equal Tick ALL correct answers.

   200 1 a cow a dog 0 -11 3.7

2. Numbers may be added in any order.
   For example, $3 + 2 = 2 + 3$ (= 5), and $7 + 4 = 4 + 7$ (= 11).
   Thus, we can write **a + b = b + a** , where **a** and **b** are numbers.

   (a) In **a + b = b + a**, which of the following could **a** equal?
   Tick ALL correct answers

   3 16 578 0 2.75 -6

   an apple the number of apples in a box a football

   (b) In **a + b = b + a**, which of the following could **b** equal?
   Tick ALL correct answers

   an apple the number of apples in a box a banana

3. In a cricket match, New Zealand scored **p** runs and Australia scored **r** runs

   How many runs altogether were scored in the match? ....

4. A milkman can just fit **n** milk cartons in a milk crate. What number of cartons could he fit in **4** of the same sized crates? Tick ALL correct answers:

   4 x n 4 + n 4n n + n + n + n nnnn

   n+4 n x 4 $\mathbf{n^4}$ **3n + n 2(n + n)**

5. **a + 3a** can be written in a shorter way as **4a**.
   Write the following in a shorter way, if possible:

   (i) 2a + 5b + a ....

   (ii) 2a + 5b ....

   (iii) 3a - b + a ....

   (iv) 5a + 3b + 2a - 4b ....

6. Which is bigger **n + 2** or **2n** ? Explain carefully.

   ....

   ....

BRAIN BOX QUIZ No.2 page 2

7. (a) If **3a** represented 3 apples, what would **a** represent? ....

(b) If **3a** = **36**, what would be the value of **a** ? ....

(c) Kay and Ray say that **3a** + **2b** could represent the total number of people seated in a restaurant, some at 3 large tables (the same number at each) and some at 2 smaller tables (the same number at each). **Tick ONE** of the following to show how strongly you agree or disagree with Kay and Ray:

I strongly agree.... I agree.... I disagree.... I strongly disagree....

(d) Jack and Jill say you must not add **3a** and **2b** because it would be like trying to add 3 apples to 2 bananas. **Tick ONE** of the following to show how strongly you agree or disagree with Jack and Jill:

I strongly agree.... I agree.... I disagree.... I strongly disagree....

(e) Joanna and Joshua say you must not add **3a** and **2b** because **a** and **b** stand for.numbers but you do not know what the numbers are. **Tick ONE** of the following to show how strongly you agree or disagree with Joanna and Joshua:

I strongly agree.... I agree.... I disagree.... I strongly disagree....

8. If y = 3, what is the value of

(i) 2y ? ....

(ii) 2y + 5 ? ....

(iii) 2(y + 5) ? ....

(iv) 2y + y ? ....

(v) 3y - y ? ....

9. If **d** represents a number, could the following be true?

(i) **d + 6 = 10** YES/NO. If YES, **d** = ...

(ii) **d + 6 = 6** YES/NO. If YES, **d** = ...

(iii) **d + 6 = 6.8** YES/NO. If YES, **d** = ...

(iv) **d + 6 = 2006** YES/NO. If YES, **d** = ...

(v) **d + 6 = 4** YES/NO. If YES, **d** = ...

10. For a school excursion, 3 buses take **f** students each and 4 cars take **g** students each.

(i) What does the value of **3f** tell us? ....

(ii) Give the total number of students taken by these buses and cars. ....

(iii) One car leaves early with **g** students. How many students remain? ......

BRAIN BOX QUIZ No.2 page 3

11. (i) Add 4 onto $n + 5$. ....
(ii) Add 4 onto $3n$. ....
(iii) Multiply $n + 5$ by 4. ....

12. (i) In a school there are 15 times as many students as there are teachers. Using S = the number of students, and T = the number of teachers, which of these two equations is correct?
(a) $15S = T$ (b) $15T = S$ MY CHOICE: ....
(ii) "The number y is eight times the number z."
Write this information in mathematical symbols. (Write an equation.)

....

(iii) If **a** and **d** are any two numbers, which, if either, is the bigger?
Give a reason for your answer.

....

(iv) If y and d are two positive numbers and $6y = d$,
which is the bigger number, y or d? ....
(v) In a certain town the number of bicycles (x) and the number of motor bikes (w) are related by the equation $\mathbf{4x = w}$.
What does this equation tell you about the numbers of bicycles and motor bikes in the town?

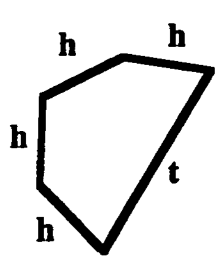
....

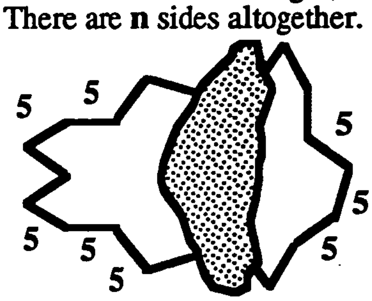
13. If the expression **4g + 8** represents a number of flowers, could **4g** represent the number of flowers in **4** same-sized bunches of flowers ? YES / NO.
If YES, the **g** represents ....
If YES, the **8** represents ....

14. Smarties are sold in two packet sizes, small holding **t** Smarties each, and large holding **s** Smarties each. TICK ALL CORRECT expressions which give the number of Smarties in **8** small packets and **2** large packets:

t + t + t + t + s + t + t + t + t + s		**4t2s**	**8s + 2t**
ttttstttts	**8t + 2s**	**2s + 8t**	**8t2s**
2(s + 4t)	**s + 4t + s + 4t**	**4(2t + s)**	

15. Write down the perimeter (the distance around) for these two cases:
(a) The letters give the number of centimetres in the lengths of the sides.
(b) Part of this shape is hidden. All the sides are of length, 5 cm. There are **n** sides altogether.





(a) Perimeter = ... (b) Perimeter = ...

16. (a) If **c + d = 10**, tick the values that **c** could have:

3 **10** **12** **7.4** **a cabbage** **a pear**

(b) If **c + d = 10**, what happens to **d** as **c** gets bigger?

....

(c) If **c + d = 10**, and **c** is always less than **d**, what values may **c** have?

....

17. Decide whether the following statements are TRUE always, never or sometimes. Tick the correct answer.
If you tick "**true only when ..**", write **when it is true.**
All the letters stand for whole numbers or zero (0, 1, 2, 3, 4, ...)

(i) **a + b + c = a + x + c**

O true always

O never true

O true only when ....

(ii) **2a + 3b + 7 = 5a + 7**

O true always

O never true

O true only when ....

(iii) **2a = a + 2**

O true always

O never true

O true only when ....

(iv) **a + 2b + 2c = a + 2b + 4c**

O true always

O never true

O true only when ....

ACCEPTABLE ANSWERS (See Appendix 3D for further details about responses)

1 200; 1; 0. **2a** 3; 16; 578; 0; 2.75; -6; the number of apples in a box. **3** p + r. **4** 4 x n; 4n; n+n+n+n; n x 4; 3n+n; 2(n+n). **5i** 3a+5b; **ii** 2a+5b; **iii** 4a-b; **iv** 7a-b. **6** Depends - for n > 2, 2n is larger, but for n < 2, n + 2 is larger, and if n = 2, they are equal. **7a** one apple; **b** 12; **c** I strongly agree; **d** I strongly disagree; **e** I strongly agree. **8i** 6; **ii** 11; **iii** 16; **iv** 9; **v** 6. **9** all Yes: **i** 4; **ii** 0; **iii** 0.8; **iv** 2000; **v** -2. **10i** The number on the buses; **ii** 3f + 4g; **iii** 3f + 3g. **11i** n + 9; **ii** 3n + 4; **iii** 4n + 20. **12i** b; **ii** y = 8z; **iii** Cannot tell as they may be any value; **iv** d; **v** There are 4 times as many motor bikes as bicycles. **13** Yes; no.of flowers in one bunch; 8 extra flowers. **14** t+t+t+t+s+t+t+t+t+s; 8t+2s; 2s+8t; 2(s+4t); s+4t+s+4t. **15a** 4h+t; **b** 5n. **16a** 3; 10; 12; 7.4; **b** d gets smaller; **c** c < 5 (e.g., 4.9, 0. -1).**17i** true only when b = x; **ii** true only when a = b; **iii** true only when a = 2; **iv** true only when c = 0.

Performance statistics from 1989 trialling of Brain-Box Quiz No.1 and No.2 with comments on the outcomes

Comments on Question 1. The item was not used in the final test instrument as the context restricted options to positive integers or zero and other items obtained similar data without such a restriction. (See Table 3·3 in the text of Chapter 3.) The change of format from Quiz 1 to Quiz 2 did not make any significant difference, as can be judged by the similarity in responses to the options which were common to both tests. The format from Quiz 2 was adopted in the corresponding question (Item 6 a) which was included in the final test.

The option "a sheep" or "a cow" was chosen by 64.4%. At first sight, this could be interpreted as indicating that students who selected this option regarded the letter as standing for an animal, which would be a case of considering that a letter could stand for an object and not just a number. However, interviews indicated that the response "Yes" was meant by at least some students to imply that 'x' could equal the number '1' as they thought of "a sheep" as "*one* sheep". Others may well have put "Yes" to indicate that they thought that 'x' could stand for an actual object, which was, in this case, a sheep. The "Yes" answer to the question, therefore, was ambiguous. In the final research instrument, an effort was made to avoid such ambiguity by changing the wording so that the corresponding question (Item 6 a) included options worded as "an object like a cabbage" rather than just "a cabbage". Figure 3D·1 summarizes the frequencies of correct answers to the parts of Question 1.

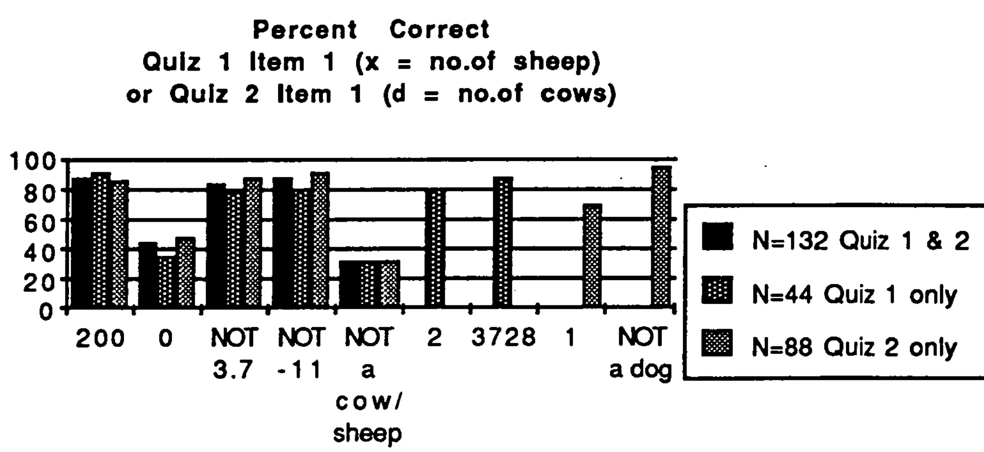


Figure 3D·1. Percentage correct on Item 1 Quiz 1/Item 1 Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Comments on Question 2. Item 2 duplicated information obtained from other items. (See Table 3·3 in the text of Chapter 3.) Setting the context in arithmetic by using the identity 'a + b = b + a' meant that any numbers were applicable as values for the symbols. This item was not used in the final test. The option "number of apples in a box", accepted by about 60% in the trialling, was retained in Item 6 in the final test, an item which also used an arithmetic context, viz., '$c + d = 10$'. Figure 3D·2 gives the percentage frequencies of responses to Question 2.

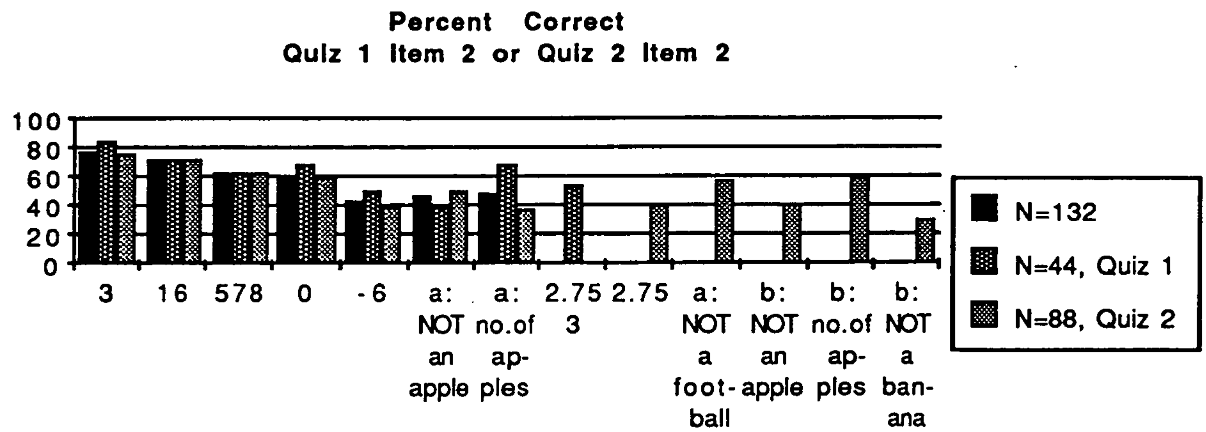


Figure 3D·2. Percentage correct on Item 2 Quiz 1/Item 2 Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Comments on Question 3. The percentages of students giving different responses are presented in Figure 3D·3. This item was retained in the final test as Item 5, after editing out any reference to particular sporting teams.

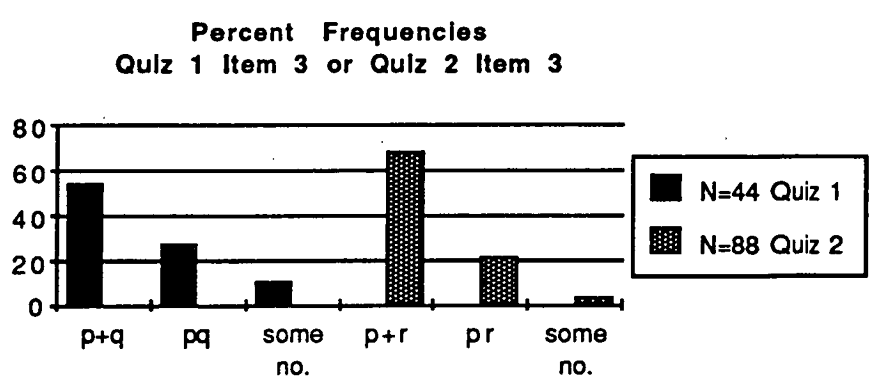


Figure 3D·3. Percentage responses on Item 3 Quiz 1/Item 3 Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Item 3 registered the number of students who gave numerical answers instead of accepting a response in symbolic form. Those who gave their answers in symbols showed a certain willingness to accept lack of closure by operating on symbols without knowing the numerical values that the symbols might have. A degree of seeking closure could have been shown by those who wrote '*pq*' as an "answer" for the sum of '*p*' and '*q*'.

Comments on Question 4. This item tested students' ability to recognize equivalent expressions from a given list. It was not included in the final test, mainly because it gave little indication of the ways students understood algebraic symbols, although it did measure their expertise with some of the conventions for writing algebraic expressions. Some of the options took students into expressions that went beyond the first degree, which would have been outside the experience of those just beginning algebra. The percentage success rates are graphed in Figure 3D-4.

Percent Correct
Quiz 1 Item 4 or Quiz 2 Item 4
(number of milk cartons in 4 crates)

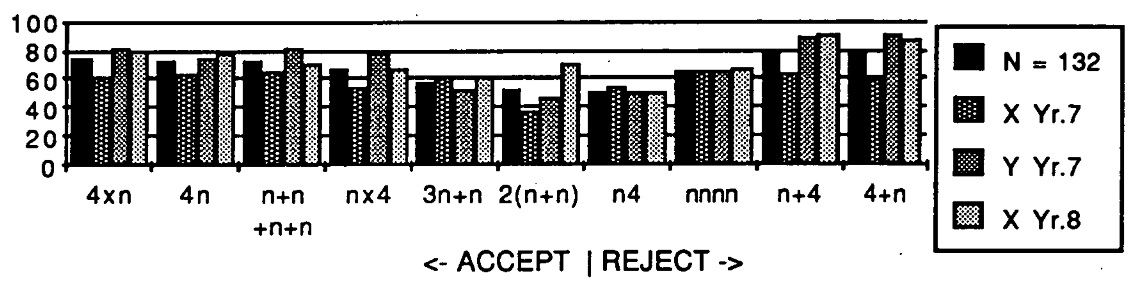


Figure 3D-4. Percentage correct on Item 4 Quiz 1/Item 4 Quiz 2
(*N* = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time.
Note that n4 was printed as n^4 in the test)

Comments on Question 5. This item measured students' ability to manipulate symbols in an abstract setting. Percentage frequencies of responses to part (a) are shown in Figure 3D-5.

Percent Frequencies
Quiz 1 Item 5 (a) or Quiz 2 Item 5
N = 132, for Quiz 1 & 2

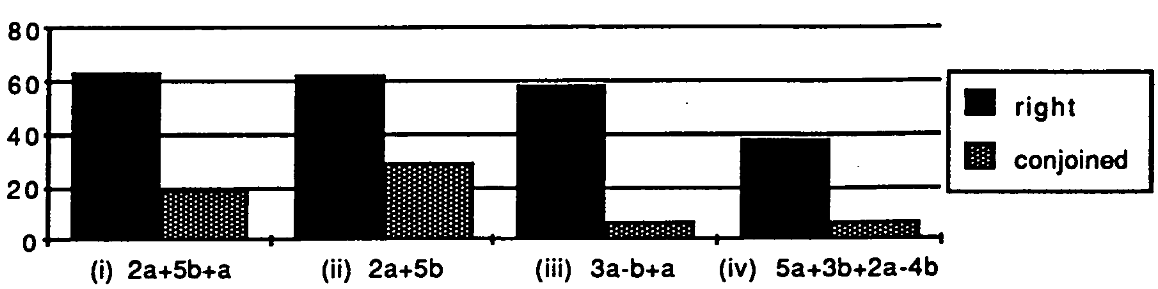


Figure 3D-5. Percentage frequencies on Item 5 (a) Quiz 1/Item 5 Quiz 2
(*N* = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

As Küchemann (1980) pointed out, students could answer questions like these correctly even when they thought of letters as objects to be moved around. It did not contribute to a clarification of students' views about the symbols and so it was not used in the final test, even though it tested skills with some aspects of the conventional uses of symbols. Part (b) of the question read as follows: "Can a and b represent anything here, do they stand for anything? If so, what?" This produced responses which were not really useful in finding out about students' views about symbols. The wording was too vague and it failed to contribute to the objectives of the research. Therefore, that style of item was not used in the final test.

Comments on Question 6. The five parts of this item were retained in the final test. Numerical responses indicated students' interpretation of the given expressions. For instance, students giving the value '23' for '2*y*' when '*y*' equalled 3 indicated that they regarded the coefficient '2' as determining the place value for the '*y*' which led to the conclusion that '2*y*' equalled '20 + 3'. Hence, the item was able to provide data for the objectives to investigate difficulties students experienced with understanding the use of symbols in a conventional way. The outcomes were far more informative than those obtained from items asking students to select equivalent expressions from a given list, such as Items 3 and 11 from Quiz 1. Figure 3D-6 presents the rates of response types for each part of Question 6.

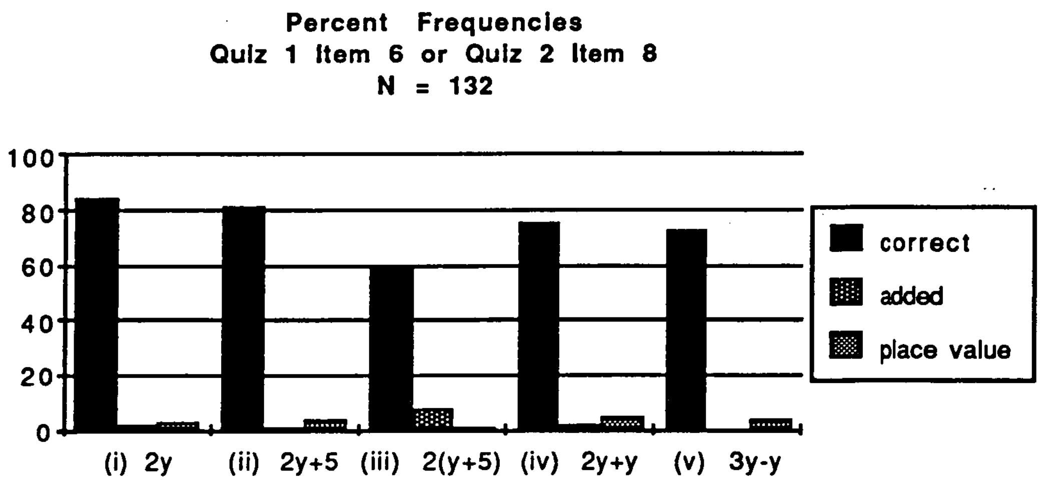


Figure 3D-6. Percentage frequencies for Item 6 Quiz 1/Item 8 Quiz 2
(*N* = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Comments on Question 7. Item 7 was a measure of students' acceptance of positive integers, fractions and negative numbers as possible meanings for the letter '*d*' in an arithmetical context. The success rate on this question depended on the type of number required for '*d*', as can be seen in Figure 3D-7. It was, thus, a question which revealed much about the range of number types that students were ready to incorporate in their understanding of the possible meanings for algebraic symbols. Although this question was not included in the final test, Item 6 (a) was included to obtain similar data in the context of the equation '$c + d = 10$'.

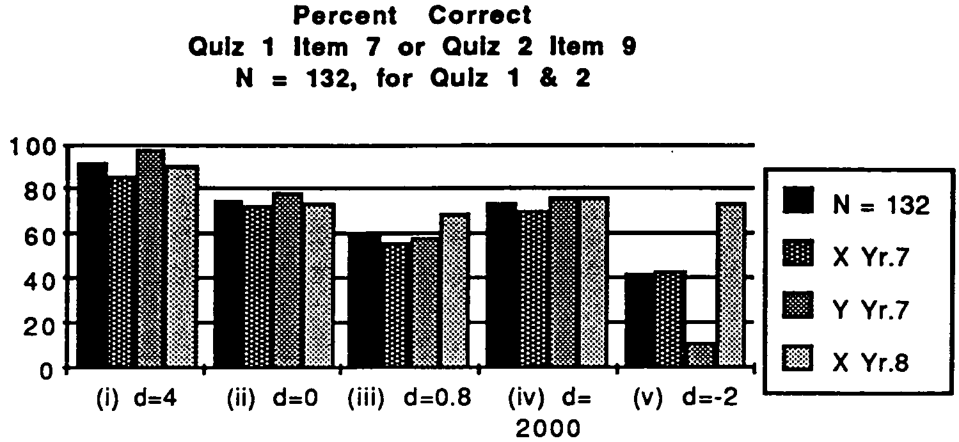


Figure 3D·7. Percentage correct on Item 7 Quiz 1/Item 9 Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Comments on Question 8. This question was included in the final test, with a re-writing of part (i). In Quiz 1, part (i) read "What does $3f$ tell us?". This was clarified for Quiz 2 to read "What does the value of $3f$ tell us?" Response frequencies are summarized in Figure 3D·8.

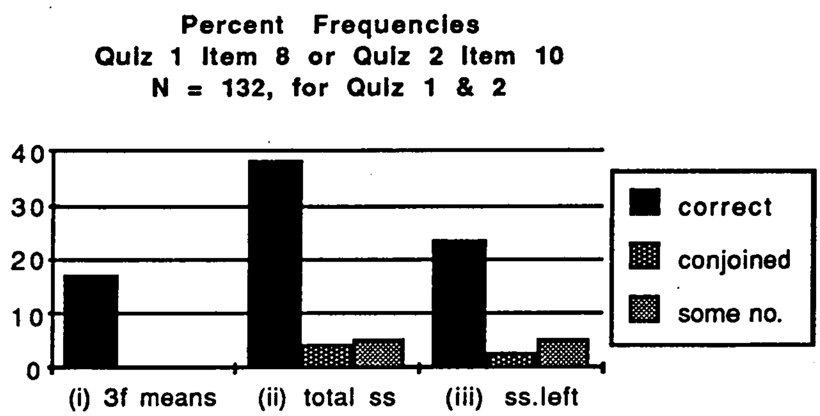


Figure 3D·8. Percentage responses for Item 8 Quiz 1/Item 10 Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Some of the answers obtained from these open-ended questions were used to format the question in a multiple-choice style for the final test as a way to simplify the processing of responses. The item was then included, becoming Item 14, as it achieved the objectives of determining students' ability to interpret the meanings of algebraic symbols in a real-life context and to carry out operations on numerical variables without knowing their values.

Comments on Question 9. This item was included unchanged in the final test as Item 9 because it measured student ability to operate with unknowns at various levels of difficulty. It was based on questions used by Küchemann in 1980. Rates for different responses to each part are graphed in Figure 3D·9

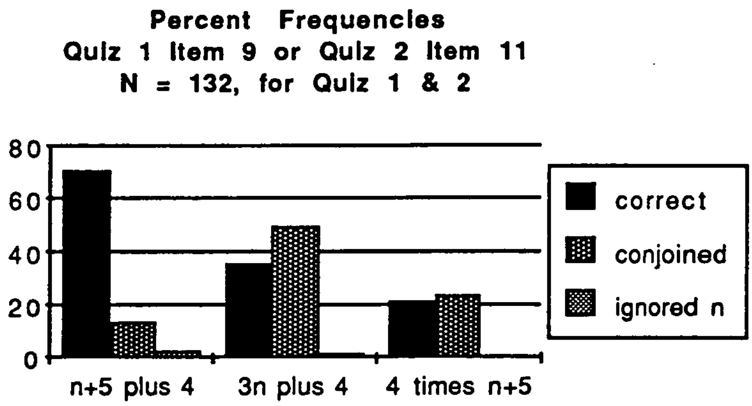


Figure 3D·9. Percentage responses for Item 9 Quiz 1/Item 11 Quiz 2 (N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42] Percentages are for responses when students did the test for the first time)

Comments on Question 10. Percentage frequencies of correct responses and of those showing the reversal error are depicted in Figure 3D·10.

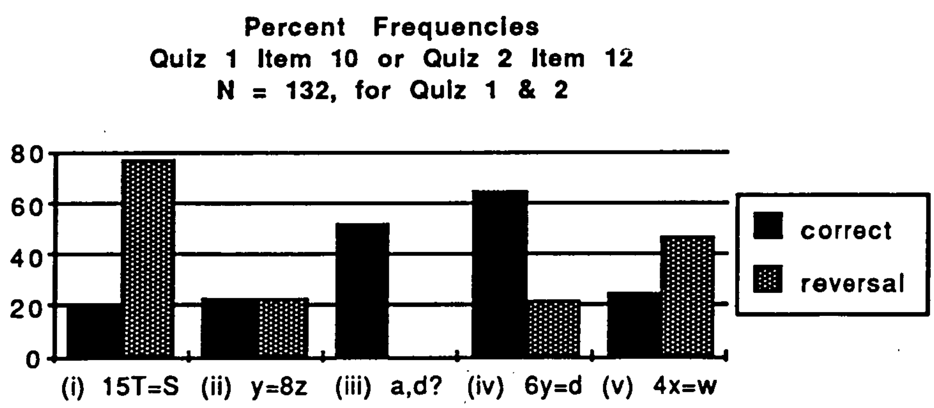


Figure 3D·10. Percentage responses for Item 10 Quiz 1/Item 12 Quiz 2 (N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42] Percentages are for responses when students did the test for the first time)

The final test included only parts (iii) and (iv) of this question, as Item 2, but the other parts were replaced with a multiple-choice version of the well-documented students-and-professors problem (Rosnick, 1981). All these questions measured degrees of ability for interpreting the meanings of letters as generalized numbers and for understanding relationships between two variables. Part (iii) tested students' understanding that the letters could stand for any numerical values, yet, as Figure 3D·10 reports, only about half of the trial candidates saw this. Part (iv), based on work by MacGregor (1989), was kept as a case set in an abstract context, while the professors-and-students problem gave a similar question in a real-life context. The contrast in performances shown in Figure 3D·10 between parts (i) and (iv) indicated that the contexts had a significant influence on the outcomes. It was judged that the reversal error could related to one's understanding of what the symbols represented in an equation. Therefore, these questions were relevant to the research objectives.

Comments on Question 11. Neither Question 11 from Quiz 1 nor Question 14 from Quiz 2 was included in the final test as neither contributed much to revealing students' understanding of the meanings of symbols, as was the case with Question 4. The questions presented students with similar sets of algebraic expressions in the same two variables, '*s*' and '*t*' and asked them to identify the expressions which correctly described some aspect of an accompanying diagram. The basic difference was in the mappings between algebraic symbols and "real-world" referents in the diagrams. In the case of Quiz 1, the mapping was between symbols and *objects*, viz., squares or triangles, whereas the referent for the symbols in the Quiz 2 question was *numbers* of sweets. The comparison of the percentage frequencies for success on the items given in Figure 3D·11 shows that the performances were similar for the two items, indicating that the type of referent in the given diagrams did not make a significant difference.

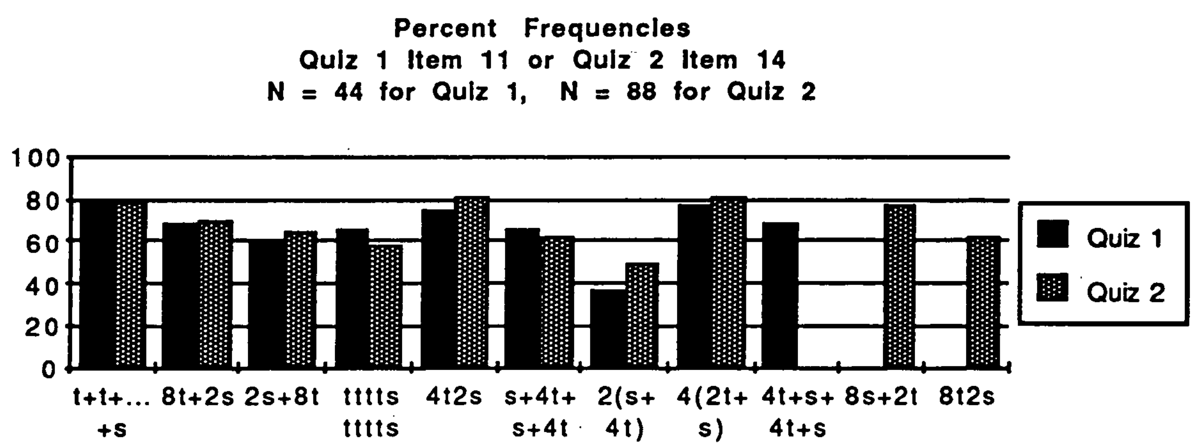


Figure 3D·11. Percentage responses for Item 11 Quiz 1/Item 14 Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Comments on Question 12. This item was not included in the final test. The two parts of the item were based on questions analyzed by Küchemann (1980). Both examined students' ability to interpret and use symbols in a geometric context. The first part, about the perimeter of a pentagon, left as ambiguous whether or not the students were thinking of the symbols in terms of numbers (of centimetres) or objects (labels for the sides). The wording of the problem in the second part, about the perimeter of an incomplete polygon, appeared difficult for many of the students and this distracted from the efficiency of the question for testing algebraic ability. Figure 3D·12 records the percentage frequencies of responses.

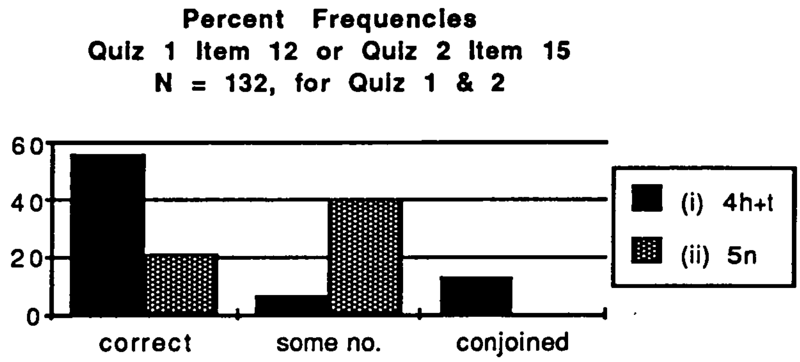


Figure 3D·12. Percentage responses for Item 12 Quiz 1/Item 15 Quiz 2 (N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42] Percentages are for responses when students did the test for the first time)

Comments on Question 13. This item was adapted into a Harper-style question for use as Item 12 in the final test. The question, as used in Quiz 1 and Quiz 2, was based on an item used by Küchemann (1980). The success rate was very low in this format, as is shown in Figure 3D·13, but when it was divided into stages similar to questions used by Harper (1979), a useful gradation of levels of success resulted.

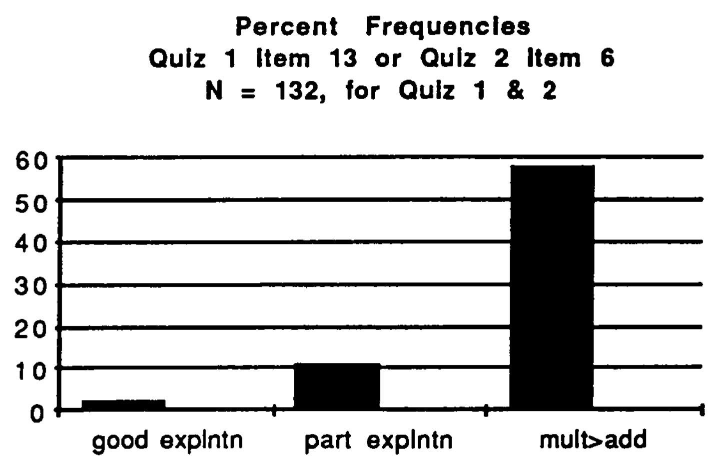


Figure 3D·13. Percentage responses for Item 13 Quiz 1/Item 6 Quiz 2 (N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42] Percentages are for responses when students did the test for the first time)

Comments on Question 14a, b. This item was included in the final test, as Item 6, together with some amendments in part (a). It was based on an item from

Küchemann (1980) which was simply part (c). This was expanded by asking the students, in part (a), to record what they thought were possible meanings for '*c*' and then, in part (b), to record their thoughts about the relationship between the values of '*c*' and '*d*' given that '$c + d = 10$'. It was part (a) that replaced Items 1 and 2 to avoid duplication. For this reason some additional options were added in the final version, namely, "the number of apples in a box", "an object like a cabbage", and "an object like a pear". Percentage responses are given in Figure 3D·14a.

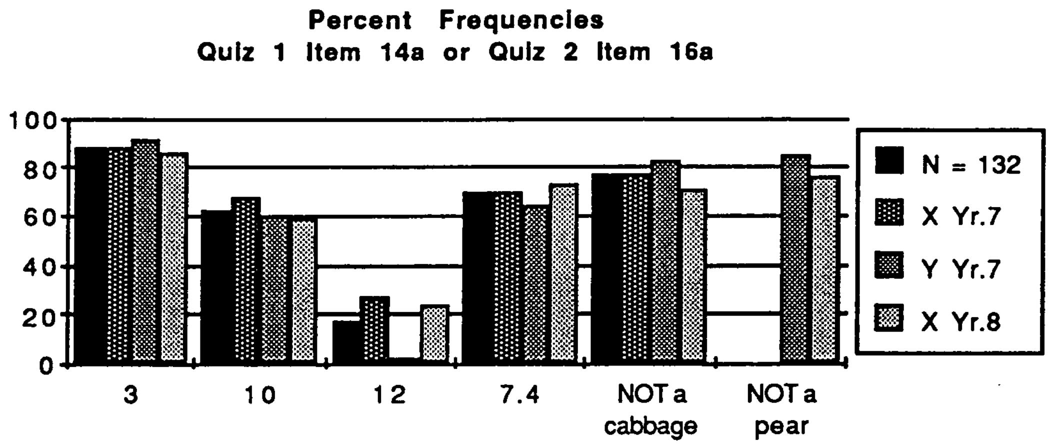


Figure 3D·14a. Percentage responses for Item 14a Quiz 1/Item 16a Quiz 2 (N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42] Percentages are for responses when students did the test for the first time)

Comments on Question 14c. Küchemann (1980) found that 30% of his subjects responded satisfactorily to this question when it was used without the introductory parts (a) and (b). It seems that these introductory parts helped the students to form their ideas about the problem, resulting in about 70% of the 1990 students giving satisfactory responses, as is shown in Figure 3D·14c.

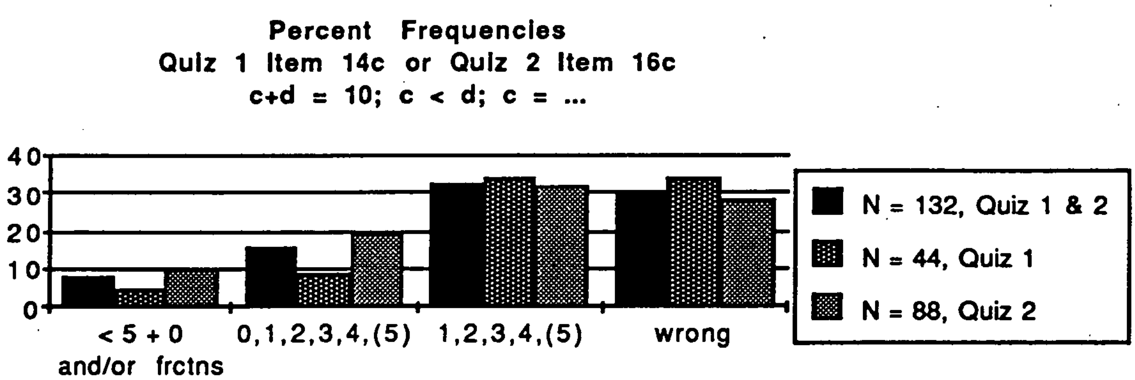


Figure 3D·14c. Percentage responses for Item 14c Quiz 1/Item 16c Quiz 2 (N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42] Percentages are for responses when students did the test for the first time. 'fractns' = 'fractions'.)

Comments on Question 15i. Despite the poor success rates on parts (b) and (c), as shown in Figure 3D·15a, this question was preserved in the final test as Item 4 using the wording as revised for Quiz 2. The second part used an open-ended style question to find out how students interpreted the meaning for the symbol '*g*' in the real-life context of bunches of flowers. Testing ability in this aspect of algebraic understanding was judged to be important in terms of the objectives of the study.

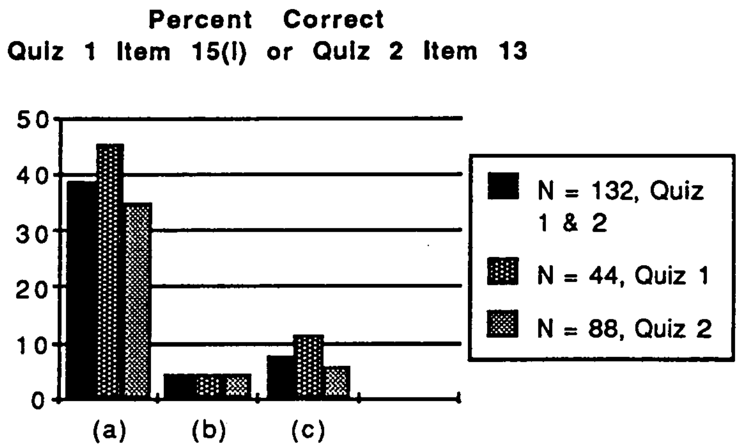


Figure 3D·15a. Percentage correct for Item 15i Quiz 1/Item 13 Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Comments on Question 15ii. This question was used only in Quiz 1 and was not kept for the final test. It was seen as duplicating the outcomes from part (i) of the item and, although a searching question, it was not required. Figure 3D·15b reports the rates of different responses.

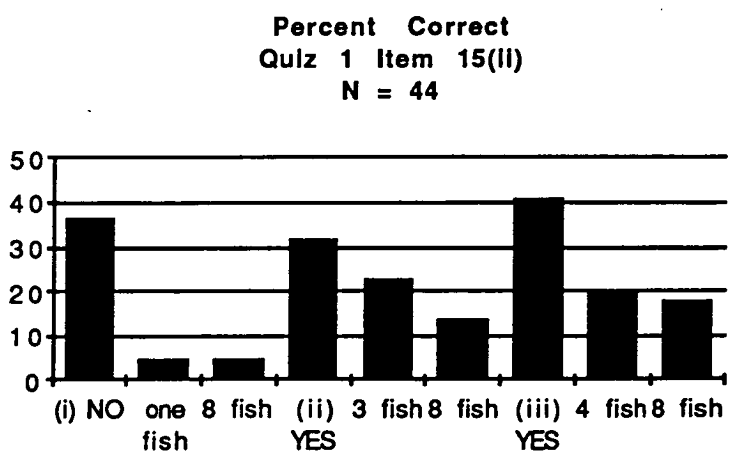


Figure 3D·15b. Percentage responses for Item 15ii Quiz 1
(N = 44: Yr.7 School X
Percentages are for responses when students did the test for the first time)

Comments on Question 16. The whole of Item 16 was used in the final test, as Item 15. Three parts of the item were questions used by Collis (1975a) and the fourth (part iii) was modelled on these. These questions were considered to be valid tests of students' ability to work with the concept of a numerical variable. The percentage frequencies of responses to parts (i) and (ii) are presented in Figure 3D·16a and the remaining parts in 3D·16b.

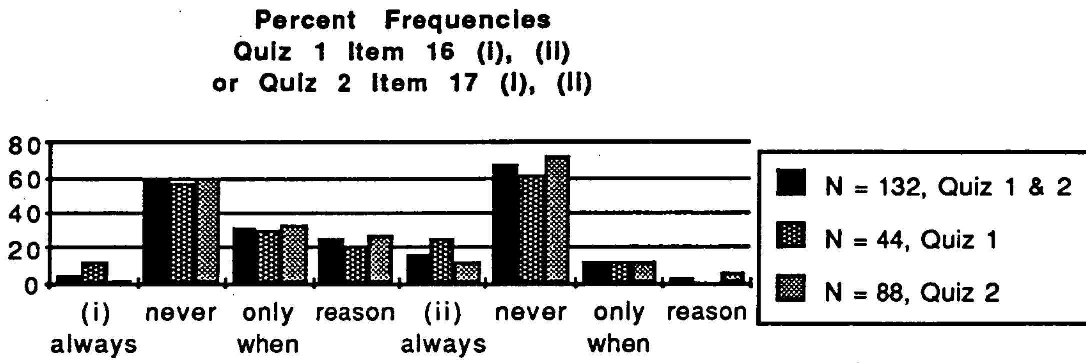


Figure 3D·16a. Percentage responses for Item 16i,ii Quiz 1/Item 17i,ii Quiz 2 (N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42] Percentages are for responses when students did the test for the first time)

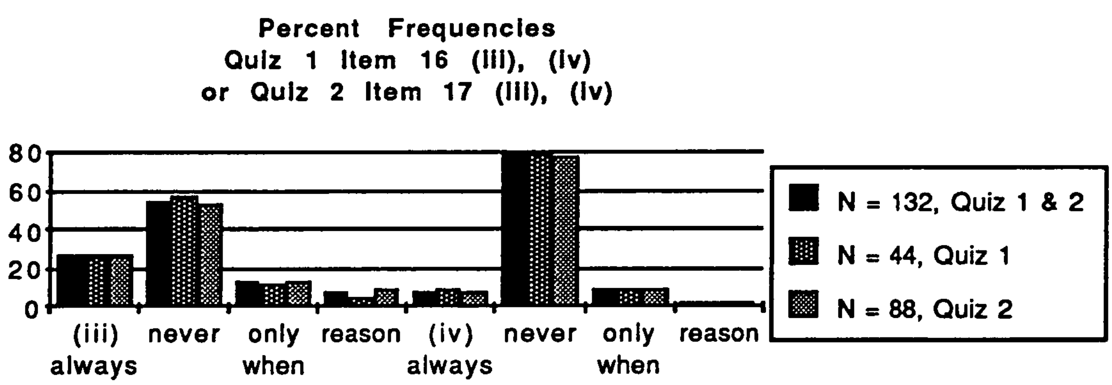


Figure 3D·16b. Percentage responses for Item 16iii,iv Quiz 1/Item 17iii,iv Quiz 2 (N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42] Percentages are for responses when students did the test for the first time)

Performance Outcomes on Quiz 2 Item 7 - not in Brain-Box Quiz No.1 Pretest, but included in the Posttest

Comments on Question 7a, Quiz 2. This question was included in the final test as Item 8 (a). The answers accepted as correct were "one apple" or "an apple". Response rates are shown in Figure 3D·17a.

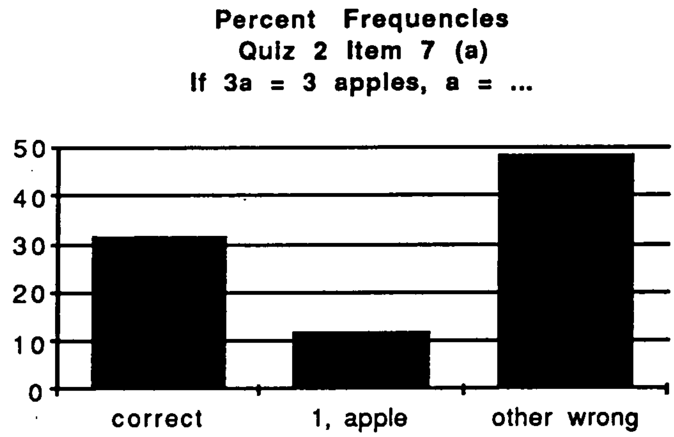


Figure 3D·17a. Percentage frequencies on Item 7 (a) Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Comments on Question 7b, Quiz 2. The question was retained in the final test (as Item 8 b) because it incisively tested the degree of commitment students had to the belief that conjoining in algebra meant multiplication. As can be seen in Figure 3D·17b, over 10% of the students tested in the trialling chose '$a = 6$', indicating that they read '3a' in place value terms as $30 + a$.

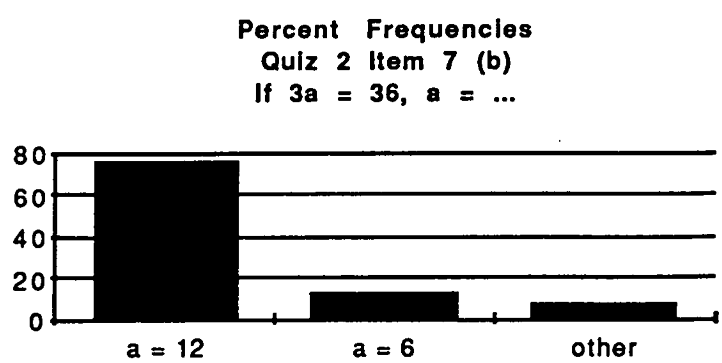


Figure 3D·17b. Percentage frequencies for Item 7 (b) Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time)

Comments on Question 7c, Quiz 2. This question was included in the final test as Item 8 (c) since it measured students' willingness to view algebraic symbols in a real-life context as representing numbers with unstated values. Figure 3D·17c records frequencies of responses.

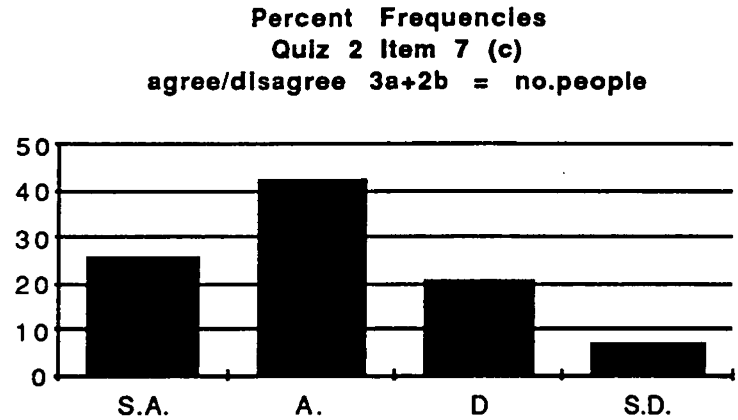


Figure 3D·17c. Percentage frequencies on Item 7 (c) Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time.
S.A.=strongly agree, A.=agree, D.=disagree, S.D.=strongly disagree)

Comments on Questions 7d and 7e, Quiz 2. The last two parts of the item were combined to form Item 8 (d) for the final test. As Figures 3D·17d and 3D·17e imply, a considerable number of students agreed with the statements given in both parts (d) and (e) and so these two questions did not result in the anticipated dichotomy. In the combined format, students were asked to choose only one of the reasons for not adding $3a$ and $2b$, and so they placed themselves in either the category who favoured the objects argument (in terms of apples and bananas) or the numbers argument.

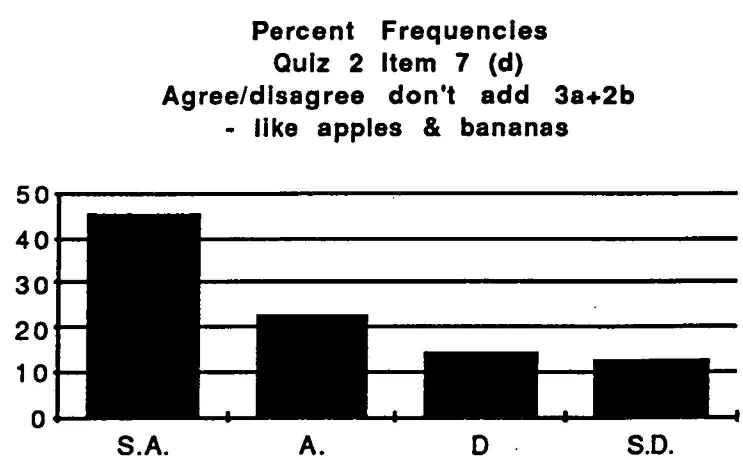


Figure 3D·17d. Percentage frequencies on Item 7 (d) Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time.
S.A.=strongly agree, A.=agree, D.=disagree, S.D.=strongly disagree)

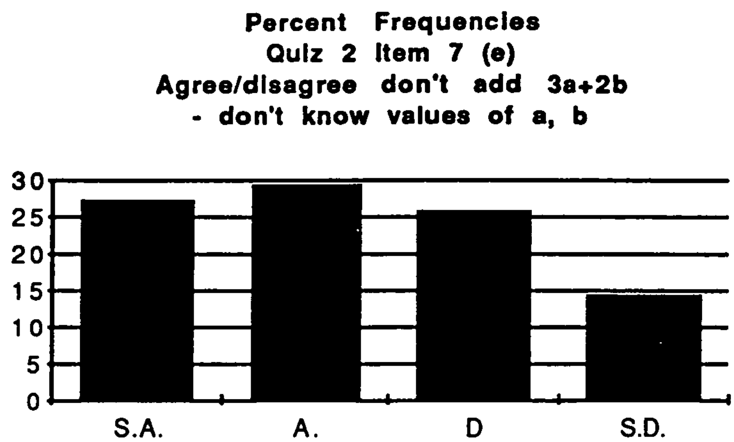


Figure 3D-17e. Percentage frequencies on Item 7 (e) Quiz 2
(N = 132: Yr.7 School X [44],Yr.7 School Y [46], Yr.8 School X [42]
Percentages are for responses when students did the test for the first time.
S.A.=strongly agree, A.=agree, D.=disagree, S.D.=strongly disagree)

ALGEBRA PROJECT 1990

NAME: .... M/F DATE: .... CLASS: ....

DATE OF BIRTH: ....(month).....(year) SCHOOL: ....

1. At a certain university there are six times as many students as there are professors.
This fact is represented by the equation $S = 6P$.

CIRCLE YOUR CHOICES IN THE FOLLOWING QUESTIONS:

(a) In this equation, what does the letter P stand for?
(i) Professors
(ii) Professor
(iii) Number of professors
(iv) Students
(v) Student
(vi) Number of students
(vii) None of the above
(viii) More than one of the above (if so, indicate which ones)
(ix) Don't know.

(b) In this equation, what does the letter S stand for?
(i) Professors
(ii) Professor
(iii) Number of professors
(iv) Students
(v) Student
(vi) Number of students
(vii) None of the above
(viii) More than one of the above (if so, indicate which ones)
(ix) Don't know.

[ANSWERS: a iii; b vi.]

Comments on Question 1. This multiple-choice version of the professors-and-students problem was used as Item 7 in the final test even though only 10 percent (2 students out of 20) of the secondary students tested succeeded with both parts. Figure 3E·1 summarizes the responses, including those of Group VI, students from the Australian Catholic University, and Group VII, students from the University of Tasmania. More the 60% of the tertiary students specializing in the teaching of secondary mathematics were successful, whereas less than 30% of tertiary students studying for primary teaching answered correctly. The reversal error was in evidence mainly from those who had chosen options which indicated that they thought the symbols represented *people* rather than *numbers of people*. The question was valuable in obtaining data contributing to the objectives of the research by asking students in a detailed way about their views of algebraic symbols in a real-life setting.

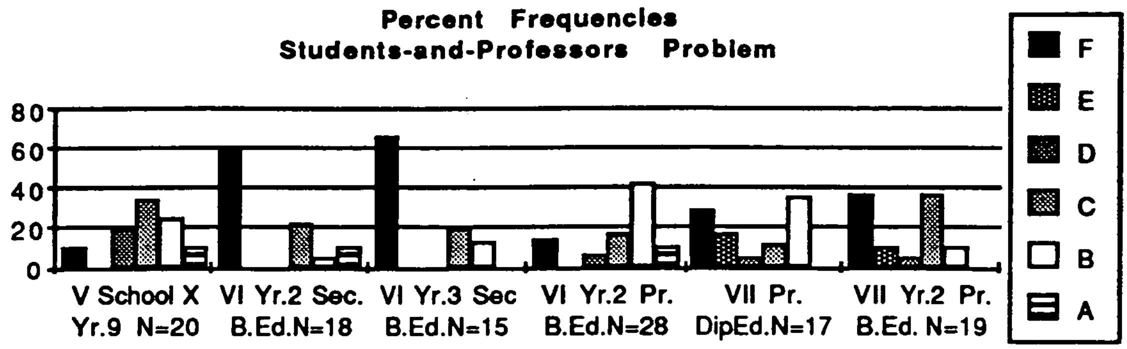


Figure 3E·1. Percentage frequencies of responses to students-and-professors problem when using trial test "Algebra Project 1990", Question 1
(F - correct [*numbers of people*, both parts], E - reversal using *nos. of people*, D- mixed choices of *people* and *nos. of people*, C - *people* [no reversal], B- reversal using *people*, A - other [e.g. "don't know"])

1990 Algebra Project

NAME: .... M/F DATE: .... SCHOOL: ....

MONTH OF BIRTH: .... YEAR OF BIRTH: .... CLASS: ....

1. This question is about x and y in the equation $x + y = 10$.

(a) If the equation is true,
is the value of x always, sometimes or never
greater than the value of y? WHY?

- -

- -

- -

(b) When is the value of x greater than the value of y?

- -

(c) When is the value of x equal to the value of y?

- -

(d) When is the value of x less than the value of y?

- -

(Harper, 1979)

2. This question is about $t + t$ and $t + 4$.

(a) Which is larger, $t + t$ or $t + 4$? WHY?

- -

- -

- -

(b) When is $t + t$ larger?

- -

(c) When is $t + 4$ larger?

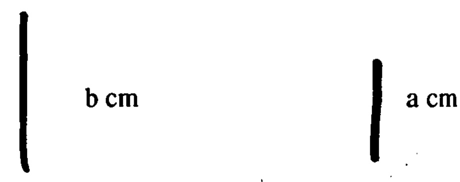
- -

(d) When are they equal?

- -

(Harper, 1979)

3. This question is about the two lines shown in the sketch.



(a) Is the red line longer than the green line,
the green line longer than the red line,
are they equal in length, or
could any of these be possible? WHY?

(b) When is the green line longer than the red line?

(c) When is the red line longer than the green line?

(d) When are they equal in length?

(Harper, 1979)

4. Show all your working as you try to solve this problem:

There are two piles of stones.
The first pile has 13 more stones in it then the second.
There are 53 stones altogether.
How many stones in each pile?

(Quinlan et al., 1989, Unit 4)

5. If 708 times 263 is 186 204

$$\boxed{708 \times 263 = 186\,204}$$

are the following True (T) or False (F) ?

Put T or F in the box and write in your own words how you tried to work out the answers to these questions.

(a) 708 = 263 x 186 204 ☐

(b) 263 = $\frac{186\,204}{708}$ ☐

(c) 263 x 708 = 186 204 ☐

----

6. Look at this:

$$\boxed{3 * 4 = 6 * y}$$

Tick the correct answers below.

(a) * could be ADD (+): ○ yes ○ no ○ can't tell

If you ticked "yes", then y must be ....

(b) * could be TIMES (x): ○ yes ○ no ○ can't tell

If you ticked "yes", then y must be ....

(Collis, 1975a)

ACCEPTABLE ANSWERS (See below for further details about responses)

1a Sometimes, depending on the possible values of x and y; **b** when x > 3; **c** when x = y; **d** when x < 3. **2a** The answer depends on the value of t; **b** when t > 4; **c** when t < 4; **d** when t = 4. **3a** Any could be possible as the values of a and b could change from the values shown in the sketch; **b** when a > b; **c** when b > a; **d** when a = b. **4** Let x = no.of stones in second pile. x + 13 = no.in first. Hence, (x + 13) + x = 53. So, 2x = 40 and x = 20. There are 33 stones in the first pile and 20 in the second. **5a** F: 708 too small; **b** T: same as box; **c** T: same as box. **6a** Yes, 1; **b** Yes, 2.

Comments on Question 1. This question was trialled under written test conditions in March 1990 with Group V, 30 Year 9 boys in Hobart, in order to find out whether written responses would be spread across the categories that Harper had found through interviews. Figure 3F·1 reports both the responses obtained by the Group V boys and those obtained by Harper when he interviewed 48 subjects of similar ages. There is no intention here of seeking to find a similar distribution of response frequencies between the two groups, but simply of reporting that the written format resulted in a spread of responses across the categories identified by Harper. The last Response Type shown in the figure, namely, responses that consisted of simply restating the question, was found in the written responses but was not reported in Harper's interview analyses. As the question successfully probed types of views that students held about the meanings of symbols, this question was found to be most suitable in furthering the objectives of the study. However, it was not included in the research test as preference was given to a similar task used by Harper and involving a more complex equation, namely, '$2x + y = 9$'. The latter was trialled in "New Test 2 1990" and, as was expected, gave even more insights into the students' understanding of the use and meaning of symbols. (See Appendix 3G.)

Table 3F·1
Frequencies of responses to Question 1

RESPONSE TYPE	NUMBER (Percentage)		DESCRIPTIONS of Response Types
	GROUP V (N = 30)	HARPER'S SUBJECTS (N = 48) [a]	
A	7 (23.3)	5 (10.4)	"Fictitious measure" - wrong idea
B	4 (13.3)	18 (37.5)	"Placeholder" - lists one or more examples
C	9 (30.0)	9 (18.8)	"Border-line algebraic" e.g., (b) x=6,7,8,9,10 or $x \geq 6$
D	4 (13.3)	16 (33.3)	"Algebraic" e.g., (b) $x > 5$
Repeats question	6 (20.0)		e.g., (b) when $x > y$

[a] From figures for Yrs. 2 and 3 in Tables 18a and 18b, Harper, 1979, pp. 339-340.

Comments on Question 2. As Harper had used the question in an interview situation, it was trialled under written test conditions with Group V. The written responses were spread across the categories that Harper had found through interviews and were distributed as shown in Table 3F·2. These outcomes compared favourably with Harper's results for 48 students who were in Years 2 and 3 and, hence, were of about the same age, as are given also in Table 3F-2. The question was considered to be suitable for use under written test conditions and was included in the final test instrument as it met the research objective of producing data which shed light on students' levels of thinking about algebraic symbols.

Table 3F-2

Frequencies of Responses to Question 2

RESPONSE TYPE	NUMBER (Percentage)		DESCRIPTIONS of Response Types
	GROUP V (N = 30)	HARPER'S SUBJECTS (N = 48) [a]	
A & B	8 (26.7)	13 (27.0)	"False ordering"
C	2 (6.7)	5 (10.4)	Numerical replacements
D	17 (56.7)	30 (62.5)	Algebraic e.g., (b) $t > 4$
Repeats question	2 (6.7)		
Omit	1 (3.3)		

[a] From figures for Years 2 and 3 in Tables 20a and 20b, Harper, 1979, pp. 351-2.

Comments on Question 3. This question was another used by Harper in an interview mode and was trialled in a written mode to check whether or not students would respond in the different categories found by Harper. Table 3F-3 reports in the affirmative.

Table 3F-3

Frequencies of Responses to Question 3

RESPONSE TYPE	NUMBER (Percentage)			DESCRIPTIONS of Response Types
	GROUP V (N = 30)	GROUP VII (N = 36)	HARPER'S SUBJECTS (N = 49) [a]	
A	13 (43.3)	28 (77.8)	37 (75.5)	"Fictitious measure" - focus on geometry
B	15 (50.0)	7 (19.4)	12 (24.5)	Algebraic and correct
Repeats Question	1 (3.3)	1 (2.8)	-	e.g., "when it is longer"
Wrong	1 (3.3)	-	-	no idea

[a] From figures for Years 2 and 3 in Tables 16a and 16b, Harper, 1979, p. 326.

The trialling reported here includes the responses obtained from 30 Year 9 boys from School X (part of Group V) and from 36 University of Tasmania students (Group VII). All of the latter group were preparing to become Primary teachers, 17 were following a Diploma of Education Course, and the remaining 19 were in the second year of a Bachelor of Education Course. That the Group V boys were more successful than both the University students and the Harper subjects is not the important issue. Rather, the important outcome was the fact that Harper's two categories of responses actually were reported. It was noted that 7 (or 19.4%) of the University students wrote in terms of the perspective for viewing the two lines in the sketch. The question was included in the final test instrument.

Comments on Question 4. Harper used what he labelled the "Zetetic Task", a problem about the sum and product of any two numbers, which could be answered with or without the use of algebraic symbols. He related the response type to the historical stages in the development of algebra (See discussion in Chapter 2). Question 4 was trialled to investigate whether or not a spread of responses would be obtained in written form to a problem about the numbers of stones in two piles of stones. As Table 3F·4 records, this resulted in a spread of responses.

Table 3F·4
Frequencies of Responses Students to Question 4

RESPONSE TYPE	NUMBER (Percentage)		DESCRIPTIONS of Response Types
	GROUP V ($N = 30$) on Piles of Stones Task	HARPER'S SUBJECTS ($N = 48$) [a] on Zetetic Task	
A	12 (40.0)	35 (72.9)	Wrong
B	12 (40.0)	8 (16.7)	"Rhetorical" Right but no algebra
C	Not applicable	4 (8.3)	"Diophantine"
D	4 (13.3)	1 (2.1)	"Viètan" Right; used right algebra
Omit	2 (6.7)		

[a] From figures for Years 2 and 3 in Tables 22a and 22b, Harper, 1979, p. 360.

It was decided, however, not to include a question of this type because it was found that students could give appropriate solutions without calling upon a symbols approach. Thus, it did not contribute directly to the objective of delineating levels of understanding of the use and meaning of symbols.

Comments on Question 5. This question was used by the N.S.W. Algebra Research Group during a four-year action research project which led to the publication of four booklets on the teaching of algerba (Quinlan et al., 1989). It tested ability with arithmetic operations. The use of large numbers was intended to lead students to respond in terms of their understanding of the processes involved, but it was found that some students, nevertheless, carried out arithmetic calculations to make their decisions. Explanations about how they worked out their answers were not always given and were often not easy to follow. The question was not kept in the final test. The success rates in each part were quite high, as Table 3F·5 shows.

Table 3F·5

Percentage Frequencies for Responses by 30 Year 9 Students on Question 5

Q.5	(a)	(b)	(c)
TRUE	20.0	70.0 (correct)	86.7 (correct)
FALSE	80.0 (correct)	30.0	13.3

Comments on Question 6. This was another question to test ability with arithmetic operations. To succeed, the student had to be sufficiently flexible to try different operations as possible meanings for the symbol '*' and then the solve a corresponding equation to give the value of the unknown 'y'. The values given for 'y' revealed precisely how they had thought about the problems and the question was preserved in the final test as Item 1. Table 3F·6 summarizes the responses in a percentage frequency format.

Table 3F·6

Percentage Frequencies for Responses by 30 Year 9 Students on Question 6

Q.6	(a)	(b)
YES	90.0	93.3
NO	3.3	3.3
Can't tell	6.7	3.3
y = 0	3.3	-
y = 1	76.7 (correct)	-
y = 2	-	73.3 (correct)
y = 6	10.0	13.3
y = 8	-	3.3

ALGEBRA PROJECT 1990

NAME: .... M/F DATE: ....SCHOOL: ....

MONTH OF BIRTH: .... YEAR OF BIRTH: ...... CLASS: ....

1. At a certain university there are six times as many students as there are professors.
   This fact is represented by the equation $S = 6P$.

   CIRCLE YOUR CHOICES IN THE FOLLOWING QUESTIONS:

(a) In this equation, what does the letter P stand for?
- (i) Professors
- (ii) Professor
- (iii) Number of professors
- (iv) Students
- (v) Student
- (vi) Number of students
- (vii) None of the above
- (viii) More than one of the above (if so, indicate which ones)
- (ix) Don't know.

(b) In this equation, what does the letter S stand for?
- (i) Professors
- (ii) Professor
- (iii) Number of professors
- (iv) Students
- (v) Student
- (vi) Number of students
- (vii) None of the above
- (viii) More than one of the above (if so, indicate which ones)
- (ix) Don't know.

2. This question is about x and y in the equation $2x + y = 9$.
   (a) If the equation is true,
   is the value of x always, sometimes or never
   greater than the value of y? WHY?

   (b) When is the value of x greater than the value of y?

   (c) When is the value of x equal to the value of y?

   (d) When is the value of x less than the value of y?

P.T.O. (Harper, 1979)

3. If the expression **4g + 8** represents a number of flowers, could **4g** represent the number of flowers in **4** same-sized bunches of flowers ? **YES / NO.**

   If YES, the **g** represents ....

   If YES, the **8** represents ....

4. Jack and Jill say you must not add **3a** and **2b.**
   CIRCLE ONE of the following as what you consider to be the BETTER reason:

   (i) because it would be like trying to add 3 apples to 2 bananas.

   (ii) because **a** and **b** stand for numbers but you do not know what the numbers are.

5. If y = 3, what is the value of

   (i) 2y + 5 ? ....

   (ii) 2(y + 5) ? ....

   (iii) 2(5y) ? ....

6. Decide whether the following statements are TRUE always, never or sometimes.
   Tick the correct answer.
   If you tick **"true only when .."**, write **when it is true.**
   All the letters stand for whole numbers or zero (0, 1, 2, 3, 4, ...)

   (i) **a + b + c = a + x + c**
   O true always
   O never true
   O true only when ....

   (ii) **2a + 3b + 7 = 5a + 7**
   O true always
   O never true
   O true only when ....

   (iii) **2a = a + 2**
   O true always
   O never true
   O true only when ....

   (iv) **a + 2b + 2c = a + 2b + 4c**
   O true always
   O never true
   O true only when ....

(Collis 1975a)

ACCEPTABLE ANSWERS (See below for further details about responses)

1a iii; **b** vi. **2a** Sometimes, depending on the possible values of x and y; **b** when x > 3; **c** when x = y; **d** when x < 3. **3** Yes; no.of flowers in one bunch; 8 extra flowers. **4** ii. **5i** 11; **ii** 16; **iii** 30. **6i** true only when b = x; **ii** true only when a = b; **iii** true only when a = 2; **iv** true only when c = 0.

Comments on Question 1. On trialling the professors-and-students problem with 18 Year 9 students, the difficulty of succeeding was again highlighted. Only two students (11.1%) correctly chose the "number of professors" option for the meaning of '*P*' and the "number of students" option as the meaning for '*S*'. The most popular choices were both in terms of people rather than numbers of people, as selected by eight students (44.4%). One-third of the students responded with the reversal error, and it was noted that all but one of these had chosen options in terms of people. Despite its difficulty, the question was kept in the research instrument as it gave quite detailed information about how students thought about symbols in a real-life context.

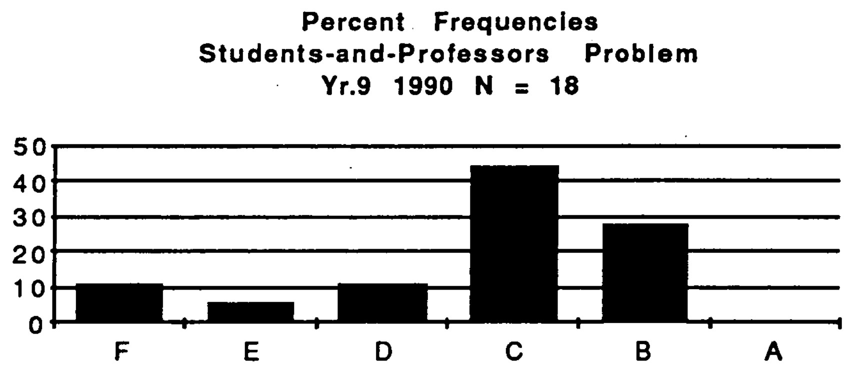


Figure 3G·1. Percentage frequencies of responses to students-and-professors problem when using trial test 'Year 9 Test 1990'
(F - correct [*numbers of people*, both parts], E - reversal using *nos. of people*, D- mixed choices of *people* and *nos. of people*, C - *people* [no reversal], B- reversal using *people*, A - other [e.g. "don't know"])

Comments on Question 2. This was the first trial for this question, the second subtask of Harper's Equations Tasks. The equation '2x + y = 9' was more complex than the equation he used in his first subtask, namely, 'x + y = 10', which was trialled with some of Group V in "1990 Algebra Project" and reported in Appendix 3F. The added difficulty was the use of the coefficient '2'. Two of the students expressed the view that the presence of the coefficient decided which of the variables was the greater, showing that the question could reveal whether or not students understood the use of coefficients. Using the more difficult equation gave an item which was capable of producing more information about students' thinking than the item trialled previously. Table 3G·1 shows that the written format of the question produced student responses across the variety of types that Harper had obtained through interviews. The question was included in the final test instrument.

Table 3G·1
Frequencies of Responses to Question 2

RESPONSE TYPE	NUMBER (Percentage)		DESCRIPTIONS of Response Types
	GROUP V ($N = 18$)	HARPER'S SUBJECTS ($N = 48$) [a]	
A	5 (27.7)	11 (22.9)	"Fictitious measure" - wrong idea
B	2 (11.1)	16 (33.3)	"Placeholder" - lists one or more examples
C	4 (22.2)	11 (22.9)	"Border-line algebraic" e.g., (b) $x \geq 4$
D	4 (22.2)	10 (20.8)	"Algebraic" e.g., (b) $x > 3$
Omit	3 (16.7)		

[a] From data for Yrs. 2 & 3 in Tables 18a and 18b, Harper, 1979, pp. 339 - 340.

Comments on Question 3. When this question had been trialled in Quiz 2 with all of Group V, the success rate was less than 10% on both parts (b) and (c). It was retrialled with part of the group in the following school year and two students (11.2%) had part (b) correct while 5 students (27.7%) had part (c) correct. Although students found it difficult, it was kept in the final test as it tested students' understanding of the meanings of symbols in a real-life context and preserved an open-ended style question.

Comments on Question 4. This question was written as a result of the findings from a pair of questions in Quiz 2, namely Questions 7(d) and 7 (e). (See Figures 3D·17d and 3D·17e in Appendix 3D.) That pair of questions showed that some students thought in terms of symbols as representing either numbers or objects. The new question forced students to declare which view they favoured. The outcome was that 11 students (61.1%) chose objects, 6 students (33.3%) chose numbers and 1 student did not respond. As the new format was found to be more informative on this important aspect than the previous pair of questions, it was used in the final test.

Comments on Question 5. The first two parts of this question had been trialled previously but the last part was added as a consequence of a lengthy discussion with one Year 9 student. He argued that in an expression such as '2(5*y*)' you applied 'the rule' that you multiplied the '5' by the '2' and you multiplied the '*y*' by the '2', so that if '*y*' equalled '3', '2(5*y*)' would equal '60'. He persisted with this argument even though fellow-students argued against his use of "the rule" and even though the correct procedure was verified by using the objects-and-containers model and, later, by using a calculator with a brackets facility. A lengthy extract from the discussion is included as the last few pages of this appendix. In the trial with 18 students from Group V, when they were in Year 9, part (iii) was found to be the most difficult of the three parts. The percentages who answered correctly are as follows: part (i) 83.3, part (ii) 77.8 and part (iii) 55.6. All three parts were included in the final test.

Comments on Question 6. This question had already been accepted into the final test from previous trialling. It was given a second time to 18 students from Group V after they had more experience with algebra and it was found that, as a group, they had improved in their understanding of the variable notion. They found the first two parts easier than the last two parts. The percentage success rates for giving correct reasons for the statements to be true were: part (i) 38.9, part (ii) 22.2, part (iii) 11.1 and part (iv) 5.6.

Comments on Test Score. The subgroup of Group V who completed this trial test attained an average score of 6.94 out of a possible 21 with a range of from 2 to 17. Although many scores were low, the trial was successful in that it showed that all of the items used were able to obtain various measures of students' understanding of the meaning and use of algebraic symbols.

- -

Transcript of a Discussion. This discussion is an example of the way the use of manipulatives can assist students to make explicit the ways they are thinking when discussing mathematical concepts basic to early algebra. It also explains the inclusion of '2(5*y*)' in Question 5 of the "Year 9 Test 1990".

Explanatory Notes:

1. This transcript is from a discussion held with Year 9 volunteers (from Group V) who worked with the researcher after school twice a week for several weeks.
2. Both students and the researcher were trying to assist those students who were having difficulties with brackets, especially when they enclosed a product.
3. The term "blobs" means any small similar objects, such as centimetre cubes.
4. 'E' denotes the experimenter and 'J.C.', 'J.S.', 'J.T.', 'J.P.', 'M.F.' were students.
5. The transcript is given in full.

[In responding to Q.3, J.C. had a LOT of trouble with 2(3*y*):
He picked up two cups containing 2 blobs each, for *y* = 2.]

J.C. I've got 2 cups and that's 4.
E. Yes. How many *y* is that?
J.C. Two.

E. 2*y*. O.K. Now you've got to build 2 times 3*y*

J.C. No! It's in a bracket. So, 2 times 3 is 6 and then you've got to times that by that and that by that so you've got to have 2*y*. Because there's a times between there that's 6 times 2*y*. 2*y* = 4. 6 times 4 = 24.

[He had 2 cups containing 2 blobs each and 6 more blobs next to them.]

E. Let's go back to the previous one. How do you build that?

J.C. 2 times *y* is 2*y* plus 2 times 3 is 6.

E. Alright... See the one with the brackets? Can you all work out two different ways so that you can build that.

J.T. The first one's got a plus in between the *y* aand the 3, and the second one's got a times, so they'll have different answer.

E. All watch Justin build the first one.

J.S. One cup and 3. That's *y* + 3 and then you've got another *y* plus 3.

E. And another *y* + 3. Why is that another *y* + 3?

J.S. Because it's two *y* + 3 and when you do the brackets - 2*y*'s and then plus 6, so you've got 6 blobs. See there.

E. You built *y* + 3 and then you built it again. Is that the idea?

J.S. Yes. And then seeing they're not together, you've got to gather them.

E. Well you can. Can you show me how you gather them on the desk?

J.S. You get two of these and then there's 6.

E. You see that? He built *y* + 3 twice. And you've drawn it like that haven't you? *y* + 3 and *y* + 3. Can you see there's two different ways to do it? I'd like you all to build that for me because it's a very important point. Also a lot of kids get mixed up in it. Build it as two lots of *y* + 3. So build *y* + 3 and then build it again.

[They built a cup with their *y* blobs, 3 more blobs and repeated this.]

E. Now let's do the next one. I think we've got a problem with 2(3y)... Josh's go. Can you show us how you build it? That's what I want to see.

J.C. Um ... 2 times 3. So you have 2 times it by .. 2 times 3 is 6. So have 6

[He put down 6 blobs]

and then you have 2*y* which is that

[He put down 2 cups each with 2 blobs in them]

and then just have times between them that's 6 times 2*y* and 2*y* equals 4 and 6 times 4 is 24.

E. Now, have you got 24 blobs there?

J.C. Yes! .... No!

E. There's 6 there and 2 more and 2 more. You've only got 10 blobs. See, what I want with the model is actually model what your answer is.

J.C. Hang on. If I add 14 it'll be right.

E. (Laugh) It might be. Let's see how someone else goes about it. Who's next? What about James? Let's see what you'd do with that one.

J.P. You want ... What is that? A cup? .. you want 2 of those ...

E. I think these people here are a bit mixed up, so we've got to sort this one out. O.K.?

J.P. Times my *y* number which is 3 ...

J.C. Times is not like a plus or minus. It's not obvious until you've done the sum.

E. It could be obvious with the model.

J.C. It might look right but it's not right. The way you did it is not right.

E. Well let's have a little think about it. How many blobs did you get then?

J.P. 12

E. What's your *y* value?

J.P. 3

[He built 2 cups each with 3 in them, and placed 6 blobs nearby.]

E. 3. Well let's see if there's another way to think about it.

J.S. You've got 2 lots of *y* in the cups and then you've got 2 *y*'s that aren't in the cup, so you've got 7.

[He built a cup with 7 in it, then 14 blobs, then another cup with 7 in it.]

E. How can you have *y*'s that aren't in the cup? 'Cause *y* is defined as the number in the cup.

J.P. y is the cup.
E. What did you say?
J.P. y is the amount in the cup.
E. Yes. y is the amount in the cup. It's not y is the cup. Rightoh. Another way to think about it?
M.F. 3y equals that.
[He showed 3 cups each with 4 blobs in them]
So you do it twice. You get 2 times 3y. That's 3y and that's 3y. 2 times 3y.
[He now had 6 cups with 4 blobs in each - in 2 groups of 3.]
J.C. That's wrong. If you've got a y ... if you got ... if your sum ...
E. It's 2 lots ... It's 2 brackets 3y. That's the sum. Right.
J.C. 2 times 3y. Well you have brackets. You have to times the outside by each number. So you have to expand so its 2 times 3 and 2 times y.
E. No, no, no. You're just quoting a rule. I want you to build what it says. Just explain what you did again, Matthew.
M.F. It's got 3y inside the bracket and you've got to multiply inside the bracket by 2. So you've got 2 lots of 3y.
E. Are you happy with that?
A.J. That's what I've got.
E. You did that over there. I could see that over there. So just explain it again - how you thought that out.
A.J. Umm Well there's 3 lots of y. There's one y.
E. That's 3y in the brackets. You built that first, eh?
A.J. Yes.
E. And then?
A.J. Then there's 2 of them so you need another lot of the same amount.
[He now had 6 cups with 2 blobs in each - in 2 groups of 3.]
E. What was your y value?
A.J. 2
E. And how many blobs have you got?
A.J. 12.
E. Has he got 12 blobs there?.... That matches with the model.
J.C. I used 2 as well, but I ended up with 24.
E. Do you understand this way to do it? You build 3y. And then it's got 2 outside the bracket so it means double what's in the bracket.
J.C. Yes, but I don't think it's right, though. It's not a right way to do it.
M.F. 3y is one number. 3y is one number. It's not two separate numbers.
J.C. But the way you have to do it is times that by that and that by that and add them.
M.F. That's one number.
J.C. When you expand the brackets it's 6 times 2y which is equal ...
E. How do you know that's true?
J.C. It has to be true. I've been doing it for years.
E. All I'm saying is let's look at the model ...
M.F. 3y equals a number..
E. And the model says it is not true.
M.F. 'Cause if y is 4 ... times 3 ... 12 is one number. It's not two ...
E. There's 12. He's got 12 things. And it says get twice that. Which gives you 24.
M.F. It's like saying two 12's
E. Now let me build the one before it ... You build this one for me 2 lots of y + 3.... y + 3 in brackets multiplied by 2.
J.C. Alright, you've got 2 times y and 2 times 3 ... 6.
[He built 2 cups each with 4 in them and 6 more blobs beside them.]
E. Alright. That works in that case. Let's build it another way. Build the brackets first.
J.C. $y + 3$
E. Build me $y + 3$. Now it says ...
J.C. Add another one.
[He built a cup with 4 blobs in it and 3 blobs beside it, and did that again.]

E. Two lots of $y + 3$. That's O.K.?
J.C. Yes. But ...
E. That works out O.K.
J.C. It's a long way of doing it, isn't it?
A.J. No. I did that.
E. Build $y + 3$. Then build it again.
A.J. It's 2 times everything inside the bracket.
E. Now in that case - you've got a plus sign there - you do get $2y$ and 2 times 3. But in the next question it's quite different. Build yours again. You build $3y$.
M.F. You've got $3y$ there.
E. (Laugh) He's built them on top of each other. However ...
M.F. And you've got to that twice so you double it.
E. And that's the answer. I mean that is what it means.
M.F. You do what's inside the brackets twice. That's inside the brackets and then you double it.
E. Now everyone build the last one $2(1 + 3y)$.

J.P.repeatedly built an addition result instead of mutiplication, e.g., for $2(y+3)$ he would build $y+3$ and add 2 - similarly for $2(1+3y)$. He explained that you imagine the brackets are there with the 2 outside. I explained that with the model we want to get the right number of blobs in the answer. With the help of J.S. and M.F. we tried to convince him that just putting 2 more blobs down was really adding 2 more blobs and not multipying. J.C. probably went away not convinced by the model that $2(3y)$ was equivalent to $6y$ - he insisted that it meant 2x3x2xy, giving 24 for y=2. And for $3(2y)$ he said

J.C. I did it the other way. I said 3 times 2 which is 6, and 3 times y which is $3y$ and 'cause you've got to times: 6 times ... $3y$ equals 6 so 6 times 6 equals 36.
E. Where are your 36 blobs? ... See you're still in the land of make-believe. We've been working for two weeks on reality. We've got to use the model that shows what you're thinking. [Pause]
J.C. I've got another way: 2 times y .... $2y$... you want 3 of them. There's another one, and there's another one.
Chorus: Hurray!!
[Clap ... He had built 3 lots of $2y$ but said]
J.C. Yeah, I knew that, but it's not the right way.
A.J. It has to ... It is! It works out properly that way
E. .Go on! .... Someone's brain-washed you.
J.C. Yeah! 12 years ... 9 years of school.
A.J. When you did like this though it was all wrong wasn't it? But when you did it like that it's all right?
J.C. I was using them as numerals.
Building it as 3 lots of $2y$ was not the real thing to him:
J.C. We got that marked wrong in school today [A doubtful proposition!]

NOTE: Clearly $2(3y)$ was harder than $2(1+3y)$, which J.C. did with ease as $2+6y$, without arguing for $2 + 12y$.
In the test some eight weeks after this discussion, for Question 5 (ii), the value of '$2(y + 5)$' when 'y' equalled '3', all of them had the correct answer, '16', and for Question 5 (iii), the value of '$2(5y)$' when 'y' equalled '3',
J.P. wrote '11' which was a combination of multiplying '2' by 'y' and adding '5', and
J.T. wrote '16' which was a combination of adding '5' and 'y', then multiplying by '2',
but the others, including J.C., wrote '30' for '$2(5y)$' when 'y' had the value '3', and so were correct.

NEW TEST 2 1990 **1**

NAME: .... **M/F** **DATE:** .... **SCHOOL:** ....

MONTH OF BIRTH: .... **YEAR OF BIRTH:** .... **CLASS:** ....

1. Look at this: **3 * 4 = 6 * y**

   Tick the correct answers below.

   (a) * could be ADD (+): .... yes .... no .... can't tell

   If you ticked "yes", then y must be ....

   (b) * could be TIMES (X): .... yes .... no .... can't tell

   If you ticked "yes", then y must be ....

(Collis 1975)

2. This question is about x and y in the equation $2x + y = 9$.

   (a) If the equation is true,
   is the value of x always, sometimes or never
   greater than the value of y? WHY?

   -

   -

   (b) When is the value of x greater than the value of y?

   -

   (c) When is the value of x equal to the value of y?

   -

   (d) When is the value of x less than the value of y?

   -

(Harper, 1979)

3. This question is about $t + t$ and $t + 4$.

   (a) Which is larger, $t + t$ or $t + 4$? WHY?

   -

   -

   (b) When is $t + t$ larger?

   -

   (c) When is $t + 4$ larger?

   -

   (d) When are they equal?

   -

P.T.O. (Harper, 1979)

NEW TEST 2 1990 **2**

4. This question is about **2n** and **n + 2**.

(a) Which is larger, **2n** or **n + 2** ? WHY?

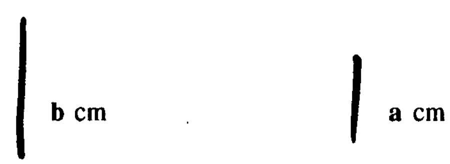
(b) When is **2n** larger?

(c) When is **n + 2** larger?

(d) When are they equal?

(Küchemann, 1980, adapted to Harper 1979 style)

5. This question is about the two lines shown in the sketch.



(a) Is the red line longer than the green line,
the green line longer than the red line,
are they equal in length, or
could any of these be possible? WHY?

(b) When is the green line longer than the red line?

(c) When is the red line longer than the green line?

(d) When are they equal in length?

P.T.O. (Harper, 1979)

NEW TEST 2 1990 3

6. At a certain university there are six times as many students as there are professors.
This fact is represented by the equation $S = 6P$.

CIRCLE YOUR CHOICES IN THE FOLLOWING QUESTIONS:

(a) In this equation, what does the letter **P** stand for?
(i) Professors
(ii) Professor
(iii) Number of professors
(iv) Students
(v) Student
(vi) Number of students
(v) None of the above
(vi) More than one of the above (if so, indicate which ones)
(vii) Don't know.

(b) In this equation, what does the letter S stand for?
(i) Professors
(ii) Professor
(iii) Number of professors
(iv) Students
(v) Student
(vi) Number of students
(v) None of the above
(vi) More than one of the above (if so, indicate which ones)
(vii) Don't know.

(Rosnick 1981, adapted)

7. If y = 3, what is the value of
(i) $2y + 5$? ....
(ii) $2(y + 5)$? ....
(iii) $2(5y)$? ....

8. Decide whether the following statements are TRUE always, never or sometimes.
Tick the correct answer. If you tick "**true only when ..**", write **when it is true.**
All the letters stand for whole numbers or zero (0, 1, 2, 3, 4, ...)

(i) $a + b + c = a + x + c$
O true always
O never true
O true only when ....

(ii) $2a + 3b + 7 = 5a + 7$
O true always
O never true
O true only when ....

(iii) $2a = a + 2$
O true always
O never true
O true only when ....

(iv) $a + 2b + 2c = a + 2b + 4c$
O true always
O never true
O true only when ....

P.T.O. (Collis 1975a)

NEW TEST 2 1990 4

9. If y and d are two positive numbers and $6y = d$,
which is the bigger number, y or d? ....
(MacGregor 1988, adapted)

10. If the expression $4g + 8$ represents a number of flowers,
could $4g$ represent the number of flowers in 4 same-sized bunches of flowers ? **YES / NO.**

If YES, the g represents ....

If YES, the 8 represents ....

11. For a school excursion, 3 buses take f students each and
4 cars take g students each.

(i) CIRCLE the ONE which best says what the value of $3f$ tells us:

(a) 3 buses X f students
(b) How many students took buses
(c) That there are the same number of students on each bus
(d) Three buses, f students
(e) The number of buses which take the children

(ii) Give the total number of students taken by these buses and cars. ....

(iii) One car leaves early with g students. How many students remain? ....

ACCEPTABLE ANSWERS (See below for further details about responses)

1a Yes, 1; **b** Yes, 2. **2a** Sometimes, depending on the possible values of x and y; **b** when $x > 3$; **c** when $x = y$; **d** when $x < 3$. **3a** The answer depends on the value of t; **b** when $t > 4$; **c** when $t < 4$; **d** when $t = 4$.**4a** Depends - for $n > 2$, $2n$ is larger, but for $n < 2$, $n + 2$ is larger; **b** when $n > 2$; **c** when $n < 2$; **d** when $n = 2$. **5a** Any could be possible as the values of a and b could change from the values shown in the sketch; **b** when $a > b$; **c** when $b > a$; **d** when $a = b$. **6a** iii; **b** vi. **7i** 11; **ii** 16; **iii** 30. **8i** true only when $b = x$; **ii** true only when $a = b$; **iii** true only when $a = 2$; **iv** true only when $c = 0.9$ d. **10** Yes; no.of flowers in one bunch; 8 extra flowers. **11i** b **ii** $3f + 4g$; **iii** $3f + 3g$.

General Comments on New Test 2. The items in New Test 2 were trialled with 41 students from Groups III and IV on 5th April, 1990, when they were in Year 8. These groups formed the 1989 Year 7 Manipulatives and Textbook Groups from School Y. In April, interest focused on their overall performance on these items to help the decision-making process regarding the format of the final research test instrument, rather than searching for significant differences between the two groups. The input to the Manipulatives Group had been only two lessons taught by the researcher the previous school year, so little, if any, residue of differences could have been expected. Hence, the response frequencies are reported below for the combined cohort and not separately for Groups III and IV.

Comments on Question 1. As Figure 3H·1 records, about three-quarters of the students were able to correctly answer this item. It was kept in the final test as a measure of ability with arithmetic processes.

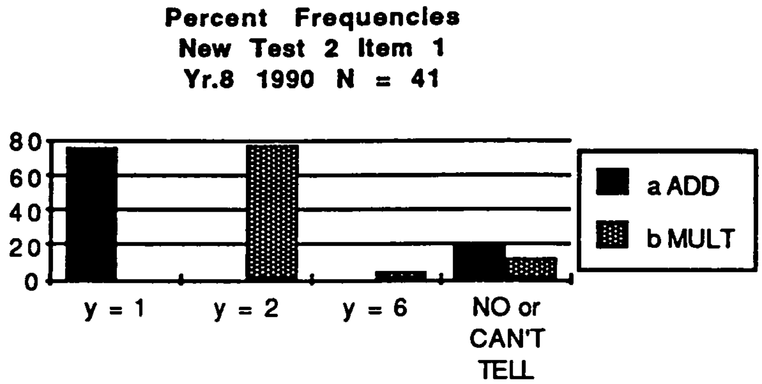


Figure 3H-1. Percentage frequencies of responses to Item 1 New Test 2 ($y = 1$ was correct in part (a); $y = 2$ was correct in part (b).)

Comments on Question 2. This question made use of Harper's Subtask 2 from his Equations Task (Harper, 1979). Figure 3H-2 reports that student responses ranged across most of the categories designated by Harper. The inclusion of a coefficient in the equation '2x + y = 9' provided added difficulty, as was expected, and the question was kept in the final test rather than a corresponding item (trialled as Item 1 in 1990 Algebra Project and reported in Appendix 3F) using Harper's Subtask 1 which built upon the simpler equation 'x + y = 10'

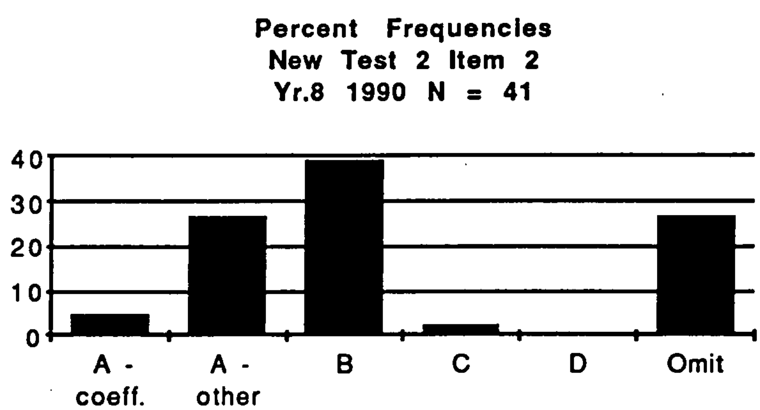


Figure 3H-2. Percentage frequencies of responses to Item 2 New Test 2
(A - Fictitious level: wrong regarding coefficient, or other aspects;
B - placeholder: lists one or more example values;
C - border-line algebraic: e.g., in (b), x > 4;
D - algebraic: e.g., in (b), x > 3)

Comments on Question 3. This question was based on Harper's Subtask One from his Literal Number Task (Harper, 1989) and trialling again showed that using it in written form produced similar categories of responses as were obtained by Harper using an interview form. The response rates are summarized in Figure 3H-3.

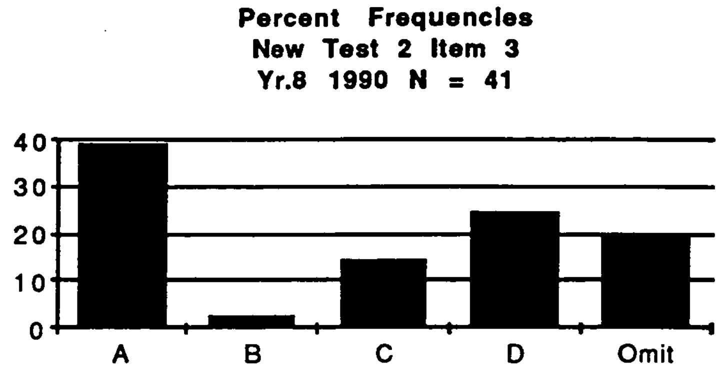


Figure 3H·3. Percentage frequencies of responses to Item 3 New Test 2
(A - Fictitious level: wrong regarding the basic meaning;
B - placeholder: lists one or more example values;
C - border-line algebraic: e.g., in (b), $t > 5$;
D - algebraic: e.g., in (b), $t > 4$)

Comments on Question 4. This item was based on Küchemann's (1980) question "Which is larger, $2n$ or $n + 2$? Explain". A Harper-style approach was used in the adaptation and, as Figure 3H·4 shows, the responses spread across the categories identified by Harper (1979) in his research using other questions. The Year 8 students tested found the question difficult with just under 10% giving correct algebraic responses.

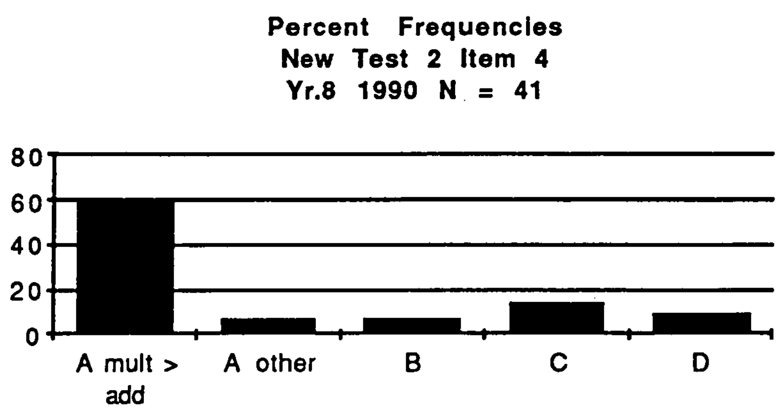


Figure 3H·4. Percentage frequencies of responses to Item 4 New Test 2
(A - Fictitious level: wrong by regarding multiplication as larger than addition, or other aspects; B - placeholder: lists one or more example values;
C - border-line algebraic: e.g., in (b), $n > 3$;
D - algebraic: e.g., in (b), $n > 2$)

Comments on Question 5. This item made use of Harper's Subtask Three of his Parallel Lines Task (Harper, 1979). The majority (nearly 60%) of the Year 8 students responded in terms of geometry and did not see the significance of the '*a*' and '*b*' as variable lengths for the two given lines. However, Figure 3H·5 reports that slightly over one-fifth of them were successful. The powerful geometric distracter made this a question of value for the final test as it served the role of penetrating students' thinking about algebraic symbols.

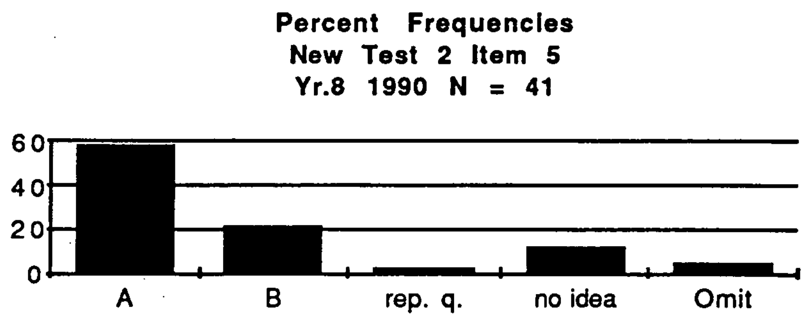


Figure 3H·5. Percentage frequencies of responses to Item 5 New Test 2
(A - fictitious level: geometric rather than algebraic;
B - correct algebraic responses;
"rep.q." - repeats the question, e.g. (b) "when it is longer" or "when the red is shorter")

Comments on Question 6. The professors-and-students problem once again yielded a valuable spread of response types, as is shown in Figure 3H·6.

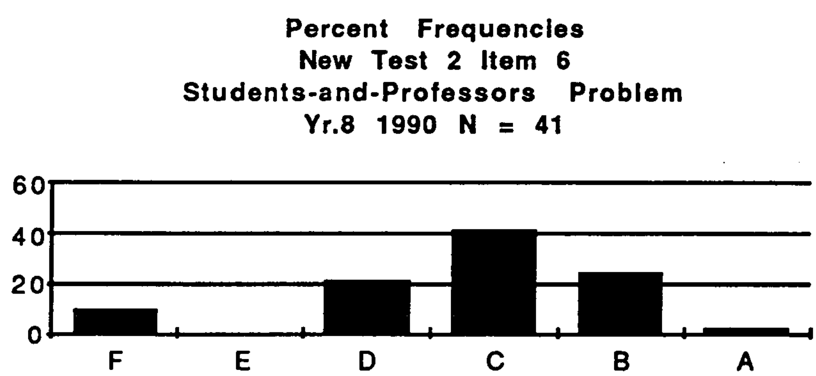


Figure 3H·6. Percentage frequencies of responses to students-and-professors problem when using trial test 'New Test 2 1990'
(F - correct [*numbers of people*, both parts], E - reversal using *nos. of people*, D- mixed choices of *people* and *nos. of people*, C - *people* [no reversal], B- reversal using *people*, A - other [e.g. "don't know"])

Comments on Question 7. This substitution exercise revealed much about students' understanding of some of the conventions for writing first degree algebraic expressions. Success rates are reported in Figure 3H·7. The item was trialled with Year 9 students and is discussed in Appendix 3G (as Question 5).

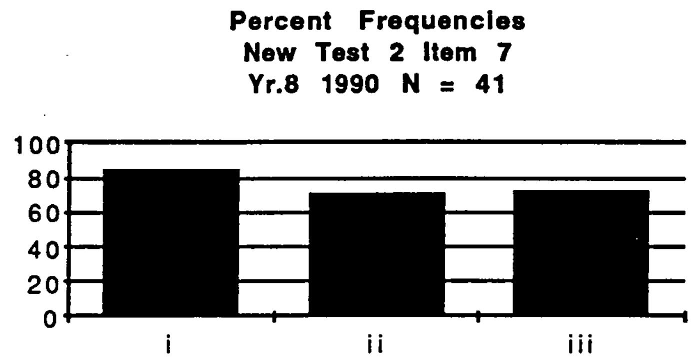


Figure 3H·7. Percentage frequencies of correct responses to parts i, ii, and iii of Item 7 New Test 2

Comments on Question 8. This question on the variable concept was tested in 1989 with the same students when they were in Year 7, as Item 17 of Brain-Box Quiz No.2. They showed some improvement in this trial in 1990 and, as Figure 3H·8 depicts, they clearly found parts (i) and (iii) easier than the other two parts. It was noted that 78% of these students responded in consistent ways to both part (iii) of this item and to Item 4, which asked them to compare values for '2n' and 'n + 2'. These outcomes called for some follow up in the main research.

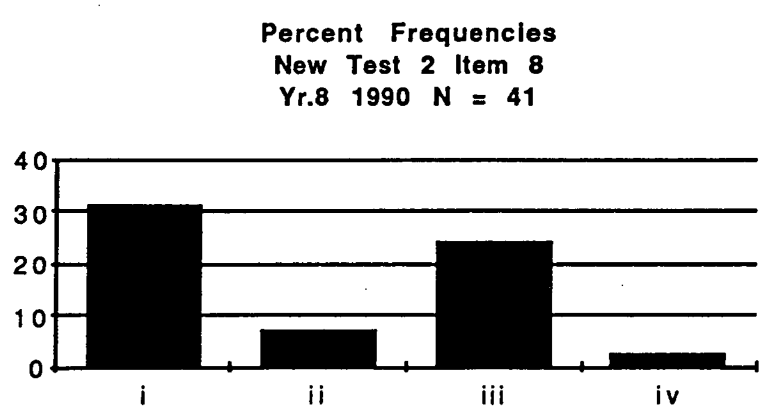


Figure 3H·8. Percentage frequencies of correct responses to Item 8 New Test 2

Comments on Question 9. One-tenth of the students showed the reversal error in the abstract context of this question, whereas nearly one-quarter of them made the corresponding mistake in Question 6 where the context was real life. Such an outcome was worthy of further investigation and this question was kept in the final test. Figure 3H·9 displays the response frequencies.

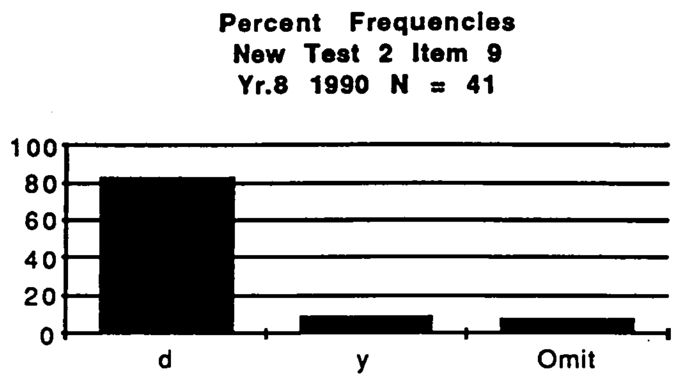


Figure 3H·9. Percentage frequencies of responses to Item 9 New Test 2 ('*d*' is the correct response)

Comments on Question 10. Trialling of this question again showed that students found it difficult to interpret the meanings of the symbols in the given real-life setting. The question was preserved in the final test as it had the potential to discriminate between the levels of understanding of symbols. Figure 3H·10 displays the success rates on the three parts of this question.

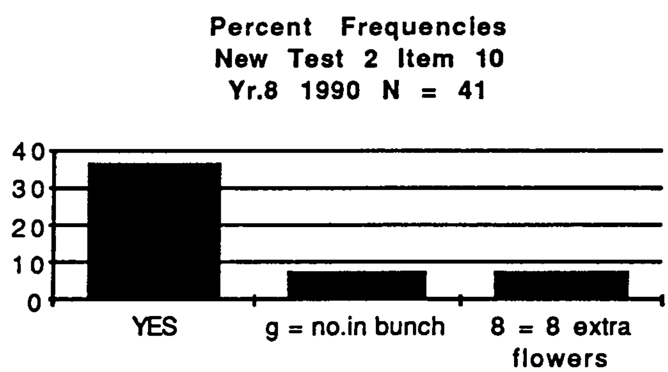


Figure 3H·10. Percentage frequencies of correct responses to Item 10 New Test 2

Comments on Question 11 Part i. Part (i) of this question was a multiple-choice form of what was trialled as an open-ended question in the Brain-Box Quiz trials. All the options were chosen by at least one of the Year 8 students, as is reported in Figure 3H·11a. Although the distribution of choices was uneven, the item was left unchanged in the final test.

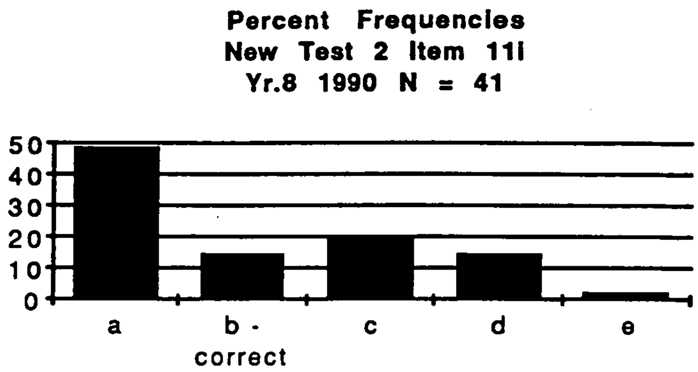


Figure 3H·11a. Percentage frequencies of responses to Item 11i New Test 2

Comments on Question 11 Parts ii and iii. Only two of the Year 8 students gave numerical answers to part (iii) of this question and none gave numerical answers to part (ii). A variety of algebraic answers were recorded and Figure 3H·11b indicates the percentage that wrote correct algebraic responses. The question was kept unchanged for the final test.

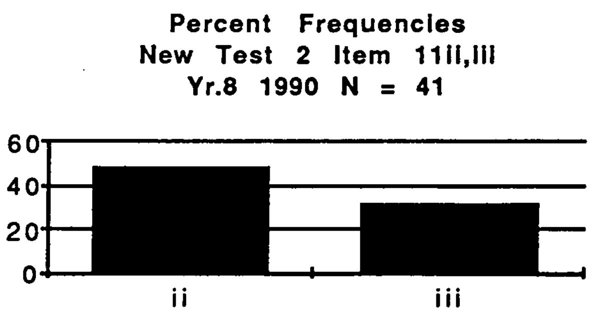


Figure 3H·11b. Percentage frequencies of correct responses to Items 11ii and 11iii New Test 2

Comments on Total Test Scores. After marking the test out of 40, the average mark for these Year 8 students was found to be 12.07 with a standard deviation of 7.28. The range was from 1 to 30. The test items were considered suitable for inclusion in the final research instrument.

NAME: .... M/F DATE: .... SCHOOL: ....

MONTH OF BIRTH: .... YEAR OF BIRTH: .... CLASS: ....

1. Look at this: 3 * 4 = 6 * y

Tick the correct answers below.

(a) * could be ADD (+): .... yes .... no .... can't tell

If you ticked "yes", then y must be ....

(b) * could be TIMES (x): .... yes .... no .... can't tell

If you ticked "yes", then y must be ....

(Collis 1975a, adapted)

2. (i) If a and d are any two numbers, which, if either, is the bigger? Give a reason for your answer.

....

(ii) If y and d are two positive numbers and 6y = d, which is the bigger number, y or d? ....

(MacGregor, 1989, adapted)

3. If y = 3, what is the value of

(i) 2y ? ....

(ii) 2y + 5 ? ....

(iii) 2(y + 5) ? ....

(iv) 2y + y ? ....

(v) 3y - y ? ....

(vi) 2(5y) ? ....

4. If the expression 4g + 8 represents a number of flowers, could 4g represent the number of flowers in 4 same-sized bunches of flowers ? YES / NO.

If YES, the g represents ....

If YES, the 8 represents ....

5. In a football match, one team scored p points and the other scored r points

How many points altogether were scored in the match? ....

(Booth 1983, adapted)

6. (a) If c + d = 10, tick ALL the meanings that c could have:

3 10 12 7.4 the number of apples in a box

an object like a cabbage an object like a pear

(b) If c + d = 10, what happens to d as c gets bigger?

....

(c) If c + d = 10, and c is always less than d, what values may c have?

....

 (Harper, 1979, adapted, & Küchemann, 1980, adapted)

7. At a certain university there are six times as many students as there are professors.
This fact is represented by the equation $S = 6P$.

CIRCLE YOUR CHOICES IN THE FOLLOWING QUESTIONS:

(a) In this equation, what does the letter P stand for?
- (i) Professors
- (ii) Professor
- (iii) Number of professors
- (iv) Students
- (v) Student
- (vi) Number of students
- (vii) None of the above
- (viii) More than one of the above (if so, indicate which ones)
- (ix) Don't know.

(b) In this equation, what does the letter S stand for?
- (i) Professors
- (ii) Professor
- (iii) Number of professors
- (iv) Students
- (v) Student
- (vi) Number of students
- (vii) None of the above
- (viii) More than one of the above (if so, indicate which ones)
- (ix) Don't know.

(Rosnick 1981, adapted)

8. (a) If 3a represented 3 apples, what would a represent? ....

(b) If $3a = 36$, what would be the value of a ? ....

(c) Kay and Ray say that $3a + 2b$ could represent the total number of people seated in a restaurant, some at 3 large tables (the same number at each) and some at 2 smaller tables (the same number at each).
Tick ONE of the following to show how strongly you agree or disagree with Kay and Ray:

I strongly agree.... I agree.... I disagree.... I strongly disagree....

(d) Jack and Jill say you must not add 3a and 2b.
CIRCLE ONE of the following as what you consider to be the BETTER reason:

(i) because it would be like trying to add 3 apples to 2 bananas.

(ii) because a and b stand for numbers but you do not know what the numbers are.

9. (i) Add 4 onto $n + 5$. ....

(ii) Add 4 onto 3n. ....

(iii) Multiply $n + 5$ by 4. ....

(Küchemann, 1980)

P.T.O.

10. This question is about **t + t** and **t + 4.**

(a) Which is larger, t + t or t + 4 ? WHY?

(b) When is t + t larger?

(c) When is t + 4 larger?

(d) When are they equal?

(Harper, 1979)

11. This question is about the two lines shown in the sketch.

(a) Is the red line longer than the green line,
the green line longer than the red line,
are they equal in length, or
could any of these be possible? WHY?

(b) When is the green line longer than the red line?

(c) When is the red line longer than the green line?

(d) When are they equal in length?

(Harper, 1979)

P.T.O.

12. This question is about **2n** and **n + 2**.

(a) Which is larger, **2n** or **n + 2** ? WHY?

(b) When is **2n** larger?

(c) When is **n + 2** larger?

(d) When are they equal?

(Küchemann, 1980, adapted to Harper 1979 style)

13. This question is about x and y in the equation **2x + y = 9**.

(a) If the equation is true,
is the value of x always, sometimes or never
greater than the value of y? WHY?

(b) When is the value of x greater than the value of y?

(c) When is the value of x equal to the value of y?

(d) When is the value of x less than the value of y?

(Harper, 1979)

P.T.O.

14. For a school excursion, 3 buses take **f** students each and 4 cars take **g** students each.

(i) CIRCLE the ONE which best says what the value of **3f** tells us:

(a) 3 buses x f students

(b) How many students took buses

(c) That there are the same number of students on each bus

(d) Three buses, f students

(e) The number of buses which take the children

(ii) Give the total number of students taken by these buses and cars. ....

(iii) One car leaves early with **g** students. How many students remain? ....

15. Decide whether the following statements are TRUE always, never or sometimes.
Tick the correct answer.
If you tick "**true only when ..**", write **when it is true.**
All the letters stand for whole numbers or zero (0, 1, 2, 3, 4, ...)

(i) **a + b + c = a + x + c**

O true always
O never true
O true only when ....

(ii) **2a + 3b + 7 = 5a + 7**

O true always
O never true
O true only when ....

(iii) **2a = a + 2**

O true always
O never true
O true only when ....

(iv) **a + 2b + 2c = a + 2b + 4c**

O true always
O never true
O true only when ....

(Collis, 1975a, adapted)

ACCEPTABLE ANSWERS (See Chapter 4 for further details about responses)

1a Yes, 1; **b** Yes, 2. **2i** Cannot tell as they may be any value; **ii** d. **3i** 6; **ii** 11; **iii** 16; **iv** 9; **v** 6; **vi** 30. **4** Yes; no.of flowers in one bunch; 8 extra flowers. **5** p + r. **6a** 3; 10; 12; 7.4; the number of apples in a box. **b** d gets smaller; **c** c < 5 (e.g., 4.9, 0. -1). **7a** iii; **b** vi. **8a** one apple; **b** 12; **c** I strongly agree; **d** ii. **9i** n + 9; **ii** 3n + 4; **iii** 4n + 20. **10a** The answer depends on the value of t; **b** when t > 4; **c** when t < 4; **d** when t = 4. **11a** Any could be possible as the values of a and b could change from the values shown in the sketch; **b** when a > b; **c** when b > a; **d** when a = b. **12a** Depends - for n > 2, 2n is larger, but for n < 2, n + 2 is larger; **b** when n > 2; **c** when n < 2; **d** when n = 2. **13a** Sometimes, depending on the possible values of x and y; **b** when x > 3; **c** when x = y; **d** when x < 3. **14i** b **ii** 3f + 4g; **iii** 3f + 3g. **15i** true only when b = x; **ii** true only when a = b; **iii** true only when a = 2; **iv** true only when c = 0.

Time Taken. The average time taken on the test which became the final version by the group of Year 7 beginners (Group VIII) was 19.58 minutes, with a range from 15 minutes to 29 minutes. This gave assurance that the test could be conducted within a class period, a practical consideration when collecting data.

Comments on Responses by Year 7 Beginners

Total Scores. When the test was marked out of 65, the average score was 7.95 and the range was from 1 to 20.

As was expected from students who had not begun to study algebra, very few in the class ganswered the test items correctly. The detailed success rates were as follows.

Table 3J·1
Percent correct on Items 1 to 6

ITEM	% CORRECT	ITEM	% CORRECT
1a	31.6	4iii	0
1b	31.6	5	15.8
2i	10.5	6a i	63.2
2ii	31.6	6a ii	42.1
3i	10.5	6a iii	10.5
3ii	10.5	6a iv	47.4
3iii	10.5	6a v	5.3
3iv	10.5	6a vi	47.4
3v	5.3	6a vii	52.6
3vi	5.3	6b	52.6
4i	5.3	6c (c = 1,2,3,4)	31.6
4ii	0	6c (c=0,1,2,3,4)	15.8

Note. Year 7 School X. $N = 19$.

Apart from Questions 1 (i), 1 (ii), and 2 (ii), only one or two students wrote correct answers to these first three items, as Table 3J·1 records. Of the next three items, students succeeded best with Item 6. Over 40% correctly responded to four of the options for possible meanings for '*c*' in part (a), including the rejection of the two options implying that '*c*' could represent an object. Over half of these beginning students were able to understand the covarying nature of the relationship between 'c' and 'd', as they showed by their responses in part (b), and over half of them realised that, given the conditions expressed in part (c), 'c' could take at least four different values. Item 4 was beyond their experience, and Item 5 was correctly answered by only three students.

Table 3J·2 reports that Items 7 and 9 were too difficult for almost all of the beginners. However, Item 8 produced a high success rate (nearly 80%) with respect to the notion that the letters in algebra represented numbers rather than pieces of fruit, as the results for part (d) indicated. Item 14 required a knowledge of the conventions for writing algebraic expressions and an ability to operate on variables without knowing their values. Understandably, the trialling students generally did not succeed with the questions in this item. Likewise, Item 15, with its dependence on an understanding of the notion of a variable, proved to be too difficult for the students.

Table 3J·2

Percent correct on Items 7 to 9 and 14, 15

ITEM	% CORRECT	ITEM	% CORRECT
7a	10.5	9iii	0
7b	5.3	14i	5.3
8a	0	14ii	10.5
8b	5.3	14iii	5.3
8c	31.6	15i	10.5
8d	78.9	15ii	0
9i	0	15iii	0
9ii	5.3	15iv	0

Note. Year 7 School X. $N = 19$.

The set of questions based on Harper's interview style were beyond the reach of most of the Year 7 beginners, as Table 3J·3 records. This could be expected because of the large role played in these items by the concept of a variable.

Table 3J-3

Percent correct on Items 10 to 13 (the 'Harper Items')

ITEM	% CORRECT	ITEM	% CORRECT
10a	0	12a	0
10b	0	12b	0
10c	5.3	12c	0
10d	15.8	12d	0
11a	5.3	13a	5.3
11b	5.3	13b	0
11c	5.3	13c	0
11d	5.3	13d	0

Note. Year 7 School X. $N = 19$.

Comment on Usefulness. Reading the students' responses made it clear that the test items succeeded in revealing preconceptions of students about algebraic ideas before starting to study algebra. It appeared useful to the main research to administer the test to participating students before they began their classroom study of algebra. All questions were accepted for the final format of the research test instrument.

Table 3K·1a

Significant Differences from Pretest to Posttest Using *t*-tests for 1989 Data

ITEM QUIZ 1	ITEM QUIZ 2	Description	GROUPS @			
			I	II	III	IV
1iii	1iv	x or $d = 0$		*		
-	1ii	$d = 1$			**	
2a iii	2a iii	$a = 578$			*	
2a iv	2a iv	$a = 0$			*	*
-	2b i	$b \neq$ apple				*
Q.2 total	Q.2 total				*	
3	3	score =		**		
4iii	4iii	$4n$			**	
4iv	4iv	$n + n + n + n$	*			
4vii	4vii	n x 4			*	
4viii	4ix	$3n+n$			**	
4x	4viii	n^4				*
4ix	4x	$2(n+n)$	*		***	
Q.4 total	Q.4 total		*		***	*
5i	5i	$2a+5b+a$			*	
5ii	5ii	$2a+5b$			*	
5iii	5iii	$3a-b+a$				***
Q.5 total	Q.5 total				*	
6ii	8ii	$y = 3$; $2y+5 =$		*neg		
6iii	8iii	$y = 3$; $2(y+5) =$			*	
Q.6 total	Q.8 total				*	
7iii a	9iii a	$d+6=6.8$; YES		*	*	
7iv a	9iv b	$d+6=2006$; $d=$			*	
7v a	9v a	$d+6=4$ YES		*		

Note. @ I School X Manipulatives, Class 1 ($N = 20$). II School X Textbook, Class 2 ($N = 24$). III School Y Manipulatives, Class 3 ($N = 28$). IV School Y Textbook, Class 4 ($N = 18$). neg = average score decreased.
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$

Table 3K·1b

Significant Differences from Pretest to Posttest Using *t*-tests for 1989 Data

ITEM QUIZ 1	ITEM QUIZ 2	Description	GROUPS @			
			I	II	III	IV
8iii	10iii	$3f+3g$ remain	*			
Q.8 total	Q.10 total		*			
10iii	12iii	a ,d bigger?			*	
11ii	-	$4t+s+4t+s$	*			
11viii	14viii	$2(s+4t)$			***	
Q.11 total	Q.14 total				*	
12b	15b	Perimeter = $5n$		*		
13	6	$2n$, $n+2$ bigger?			*	
14a ii	16a ii	$c+d=10$; $c=10$	*			
14a iii	16a iii	$c+d=10$; $c=12$	**			
14c	16c	$c < 5$		*		
Q.14 total	Q.16 total		*		*	
15i c	13c	8 flowers			*	
Q.15 total	Q.13 total				*	
-	7c	$3a+2b$ = no.of people			***	
-	7e	number argument			* neg	* neg
TOTAL SCORE	TOTAL SCORE		**	**	***	

Note. @ I School X Manipulatives, Class 1 ($N = 20$). II School X Textbook, Class 2 ($N = 24$). III School Y Manipulatives, Class 3 ($N = 28$). IV School Y Textbook, Class 4 ($N = 18$). neg = average score decreased.
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Table 3K-2a

Significant Differences between Groups Using *t*-tests for 1989 Data

ITEM QUIZ 1	ITEM QUIZ 2	Description	GROUPS			
			Class 1 v 2 School X	Class 3 v 4 School Y	Year 7 School X v Year 8 School X	Year 7 School Y v Year 8 School Y
-	1iii	$d \neq$ a cow		3* pre		
1v	1vi	$d \neq -11$				7*
-	2a vi	$a = -6$				8*
-	2a ix	$a \neq$ a football		3* pre		
-	2b i	$b \neq$ an apple		3* pre		
-	2b iii	$b \neq$ a banana		3* pre		
2c ii	2b ii	a, b = no.in a box			7*	7*
4i	4ii	$4 + n$			8***	
4ii	4i	$4 \times n$			8**	
4vi	4vi	$n + 4$			8***	
4ix	4x	$2(n + n)$			8***	8**
5i	5i	$2a+5b+a$			8***	8*
5ii	5ii	$2a+5b$		3** post	8*	
5iv	5iv	$5a+3b+2a-4b$			8**	8***
6iii	8iii	$y=3;\ 2(y+5) =$			8***	8*
6iv	8iv	$y=3;\ 2y+y =$			8**	
6v	8v	$y=3;\ 3y - y =$			8*	
7iv a	9iv a	$d+6 = 2006$; YES	1* pre			
7v a	9v a	$d+6 = 4$; YES			8***	8***

Note. pre = Pretest, post = Posttest.
Numbers under the heading 'GROUPS' indicate which group was significantly better, e.g. in the first row, '3* pre' means that Class 3 was significantly better than Class 4 on the Pretest and the level of significance was $.010 < p \leq .050$.
Pretest responses from Year 7 were used for comparisons with Year 8.
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Table 3K·2b

Significant Differences between Groups Using *t*-tests for 1989 Data

ITEM QUIZ 1	ITEM QUIZ 2	Description	GROUPS			
			Class 1 v 2 School X	Class 3 v 4 School Y	Year 7 School X v Year 8 School X	Year 7 School Y v Year 8 School Y
8ii	10ii	$3f+4g$ students			8***	
9ii	11ii	Add 4 onto $3n$			8**	8*
9iii	11iii	Mult.$n+5$ by 4				8**
10i	12i	$15T = S$		3* post		
10ii	12ii	$y = 8z$		3* post		
11ii	-	$4t+s+4t+s$	1* pre&post			
11viii	14viii	$2(s+4t)$			8**	8**
12a	15a	perim=$4h+t$			8*	
13	6	$2n$, $n+2$ bigger?				8*
14a iii	16a iii	$c+d=10$; $c=12$				8**
14b	16b	d smaller		4* pre		
15i b	13b	g = no.in bunch		3* post		
15i c	13c	8 flowers			7*	7*
Q.15 total	Q.13 total			3* post		
16ii b	17ii b	when $a = b$				8**
16iii b	17iii b	when $a = 2$				8*
-	7b	$3a=36$; $a=12$		3* pre		
-	7c	$3a+2b$=no.people		4* pre		
-	7d	$3a+2b$ not fruit				7**

Note. pre = Pretest, post = Posttest.
Numbers under the heading 'GROUPS' indicate which group was significantly better.
Pretest responses from Year 7 were used for comparisons with Year 8.
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Timetable of Events for Main Research Project

Year 7 Beginners: Teaching, Testing, Interviewing

For the Year 7 beginners, the timing of the intervention teaching to introduce algebra, and the associated testing and interviews was as follows:

School A. For about three weeks in April 1990, the researcher taught a group of eight Year 7 students in School A. This group (Class 1) and another Year 7 group (Class 2) from the same school were tested three times during this period (before they started algebra, about midway in the period, and at the end of it) and students in Class 1 were each interviewed twice. A week's Easter holiday break occurred during the teaching intervention.

School B. From late April until well into May 1990, three Year 7 classes (Classes 3, 4, and 5) in School B were monitored by the researcher as they began their study of algebra. Interviews were conducted with samples of students from each of the classes, some being interviewed twice. The three classes were tested before they started algebra and twice more over the period.

School C. During May 1990, the first three weeks of lessons in algebra were monitored by the researcher for two Year 7 classes (Classes 6 and 7) in School C. Samples of students from each class were interviewed, some of them twice, during this period. Testing was carried out at the start, half-way through, and again at the end of the three-week session.

School D. For the first three weeks of June 1990, the researcher monitored lessons for three Year 7 classes (Classes 8, 9, and 10) in School D as they were introduced to algebra. The test instrument was administered three times in each of the classes. Interviews were conducted with samples of students from each class. Some of these interviews followed soon after the first test so that students' preconceptions could be discussed. During the three weeks some students were interviewed twice.

Testing of Other Classes.

In mid-June 1990, the test was taken by a total of 13 older classes in three schools. In School B, four classes participated: one Year 9 Advanced class (Class 33) and three Year 12 classes of different ability levels (Classes 61, 62, and 63). Five classes in School C responded to the test: one Year 8 Advanced class (Class 24), one Year 9 Advanced class (Class 34), and two Year 11 classes of diffierent ability rankings (Classes 51 and 54). Participation in testing by four classes in School D was also organized: one Year 9 Advanced class (Class 35), two Year 10 groups of differing ability levels (Classes 45 and 46), and one Advanced Year 11 class (Class 55).

In October 1990, test data were obtained in School A from two Year 9 classes of different ability rankings (Classes 31 and 32).

In December 1990, four more classes were also tested: two slow-learner Year 7 classes, one from School C (Class 14) and one from School D (Class 15), and two slow-learner Year 8 classes, both from School D (Classes 25 and 26).

Delayed Posttests and Associated Interviews

During November 1990, delayed posttests were administered to the previously-tested Year 7 classes in Schools A and B, and samples of students were interviewed.

In early December 1990, delayed posttests and sample interviews were completed for the previously-tested Year 7 classes in Schools C and D.

Follow-up Tests and Interviews in 1991

In July 1991, five classes of students at advanced levels were retested on three of the test items they had completed in 1990. Three Advanced Year 10 classes were retested, one from each of Schools B, C, and D. These three goups had previously been tested as Year 9 students in Classes 33, 34, and 35. Two Advanced Year 12 classes were retested, one from School C (formerly tested in Year 11 as Class 54), and one from School D (formerly tested as Class 55 in Year 11). Of the 115 students retested, 52 were interviewed to elucidate a paradox that had arisen from the responses given in 1990 by these able mathematics students.

UNIT ONE WORKSHEET ONE Part B

NAME: .... **DATE:** ....

1. LINE OF TRIANGLES

This pattern of a line of triangles may be built as long or as short as you like.

R---- F---- wrote, on 28th Feb. that:

"Times the number of triangles by 3, tells you the number of matches."

(a) RE-WRITE R----'s statement about *his* way of building the pattern using a letter, like y, in place of "the number of triangles":

. . . .

. . . .

(b) STATE clearly what the letter stands for:

. . . .

. . . .

2. LINE OF SQUARES - FIRST METHOD:

A---- B---- wrote, on 28 Feb., about *his* method of building the squares:

"The number of squares x 3 + 1 will give the number of matches."

(a) RE-WRITE A----'s statement using a letter, like y, in place of "the number of squares":

. . . .

. . . .

(b) STATE clearly what the letter stands for:

. . . .

. . . .

3. LINE OF SQUARES - SECOND METHOD:

J---- F---- wrote, on 28 Feb., about *his* method of building the squares:

"Use 4 matches for the first square and then add 3 for each other square."

(a) RE-WRITE J----'s statement using the expressions "times" and "the number of squares" as part of it: [Discuss your ideas with a friend.]

. . . .

. . . .

(b) RE-WRITE J----'s statement using a letter, like k, to stand for "the number of squares":

. . . .

. . . .

(c) STATE clearly what the letter stands for:

. . . .

4. LINE OF SQUARES - THIRD METHOD:

J---- S---- has been working on this, *his*, method of building the pattern. He uses four matches for each square and then takes out the unwanted matches.

(a) See if you can finish his rule for the number of matches needed:
"You times the number of squares by four and minus ....

. . . .

. . . .

(b) RE-WRITE your rule by replacing the words "number of squares" by some letter:

. . . .

. . . .

(c) STATE clearly what the letter stands for:

. . . .

5. TRY TO WRITE RULES FOR OTHER WAYS OF BUILDING THE PATTERN OF SQUARES. Write the rule(s) in words, and then use a letter.

. . . .

. . . .

PART I

Area Model for Non-integers

Figure 3N·1 shows possible representations for the algebraic symbol 'y' in the area model. These would be element mappings in Halford's (1987) scheme.

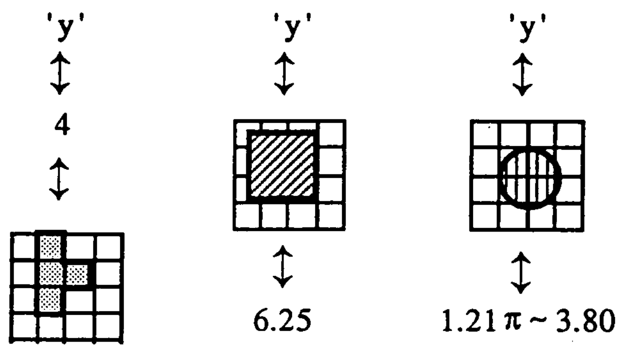


Figure 3N·1. Examples of Element Mappings using the Area Model

In the first case, the value of 'y' as 4 could be chosen before building an area of 4 square centimetres but, in the other cases, the area shape would probably be selected first, as either a plastic square (2.5 cm x 2.5 cm) or a circle (radius 1.1 cm), and the numerical value is consequent on the selection. Different students could choose different-sized shapes to give different representations of a variable 'y'. Validity, therefore, applies because the algebraic variables were being represented as numbers which can vary. The model is suitable for taking into account Dienes' recommendation of "mathematical generalization" (1963, p. 98): as is shown in Figure 3N·1, the classes of number that can be modelled range from integral, to fractional and even irrational values, allowing for the fact that these are restricted to positive values.

PART II: Additional Questions for Area Model

UNIT ONE WORKSHEET TWO

5. (i) This time choose a flat shape with a known area which is NOT a whole number of square centimetres, e.g., a circular plastic counter, or a cut-out shape. You will need THREE of the same size.

(ii) Letting **m** represent the number of square centimetres in the area of one of these shapes, and build and shade and calculate the total area for:
$m + 2$, $3m + 2$, and $3(m + 2)$.

(iii) Using a calculator to work out these areas, explain to a student or teacher what appears on the calculator display after EACH key is pressed.

Additional Exercise:

USING A CALCULATOR WITH BRACKETS

1. For 2(C + 3), record what is on the calculator display after EVERY keystroke, and explain each display. Do this for three different values of C.

KEYSTROKE	DISPLAY first C	EXPLANATION	DISPLAY second C	DISPLAY third C
2				
x				
(				
C value				
+				
3				
)				
=				

2. For 2(3C), record what is on the calculator display after EVERY keystroke, and explain each display. Do this for three different values of C.

KEYSTROKE	DISPLAY first C	EXPLANATION	DISPLAY second C	DISPLAY third C
2				
x				
(				
3				
x				
c value				
)				
=				

6. For this question you will need:
   a sheet of centimetre grid paper,
   six flat shapes each of the same area, each to be called k cm^2, and
   twelve shapes each with unit area (or 1 square centimetre of area).

NOTE that k equals the number of square centimetres covered by each of your six flat shapes, BUT the NUMERICAL VALUE of k is NOT to be used in this question!
ALL ANSWERS ARE TO BE GIVEN IN TERMS OF k, and NOT as simple numbers!
YOUR ANSWERS NEED TO BE TRUE NO MATTER WHAT THE VALUE OF k HAPPENS TO BE - EVERY GROUP SHOULD HAVE THE SAME ANSWERS!
YOU MUST BUILD AREAS FOR QUESTIONS a, b, c, BUT THERE IS NO NEED TO SHADE YOUR AREAS.

(a) Build k +3.
WRITE, in terms of k, the number of square centimetres of the grid that you have now covered: ___________
Add 4 more square centimetres.
WRITE, in terms of k, the number of square centimetres of the grid that you have now covered: ___________

(b) Build 4k
WRITE, in terms of k, the number of square centimetres of the grid that you have now covered: ___________
Add 6 more square centimetres.
WRITE, in terms of k, the number of square centimetres of the grid that you have now covered: ___________

(c) Build 2k + 4
WRITE, in terms of k, the number of square centimetres of the grid that you have now covered: ___________
Triple what you have built, that is, show three times 2k + 4.
WRITE, in terms of k, the number of square centimetres of the grid that you have now covered: ___________

(d) Complete without building: Twice 3k + 5 = ___________
6k + 2 plus 4 = ___________
3(k + 1) plus 2k = ___________

7. (i) Take a sheet of grid paper.
Choose another shape and trace around it on the grid paper.
Let g be the number of square centimetres covered.
Write on the grid paper, the your value of g:
(a) g = ...

(ii) Build 3g + 6, shade it in, and write on your grid paper the value of 3g + 6:
(b) 3g + 6 = ...

(iii) Now take away an area of 2 square centimetres, and write your result:
(c) 3g + 6 - 2 = ... (in symbols)
= .... (in value)

(iv) Now take away an area of g square centimetres from what you had left in (iii), and write your result:
(d) .... - g = ... (in symbols)
= .... (in value)

(v) Compare your results with those of someone else who used a different value for g.

8. Use cardboard cut-outs, or counters, or other shapes that do not fit the grid lines exactly, to build, shade and give values of expressions such as:
(i) y and then 3(y + 1) + y
(ii) x and then 7 + 2x - x
(iii) w and then 3(w + 2) - 2w - 1.
Write these expressions in simpler ways (and give their values).

NOTE: JUST SHADE YOUR OWN SHAPES if you have NOT got any objects that do not fit the grid lines.

9. Discuss and model cases like
(i) 3d + 2 , when d = 0;
(ii) 2(4t + 3), when t = 0;
(iii) 2(4t + 3) - 3(t + 1), when t = 0.

10. Use any letters you wish, and draw and shade areas to represent any expressions you want to look at. Label each drawing.

11. Without building or drawing any areas, write the following expressions down and work out their values:
(i) 3c + 4, when c = 4.5;
(ii) 3h + 4, when h = 0;
(iii) 2(3S + 4) - 3S, when S = 1.2.

12. Invent your own questions to match the following ANSWERS:
(i) 5 w
(ii) 3x + 7
(iii) 8
(iv) 0
(v) b - 6.

PART I: COMMUTATIVITY AND ISOMORPHISM

Commutativity

The following diagram (in Figure 3P·1) shows that a person's interpretation or representation of the environment is consistent when the given algebraic relationship is modelled with the objects and containers.

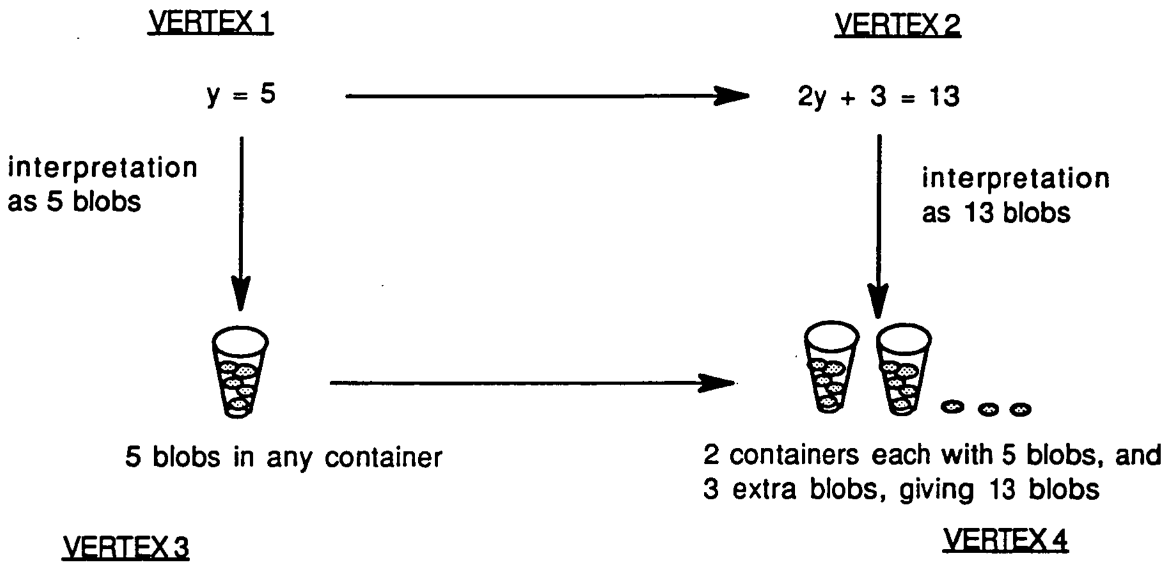


Figure 3P·1. Commutativity using objects-and-containers model

You can proceed from Vertex 1 to Vertex 4 either clockwise or anti-clockwise and deduce the same result from either path. Hence, we can say that this cognitive system is "commutative" in the sense used by Halford and Wilson (1980, p.373) .

Isomorphisms

1. Single algebraic expression modelled in terms of area. Figure 3P·2 shows two specific instances of modelling '$2y + 3$', breaking each of them down into the two operations of doubling and adding 3. The same structure is common to both examples and to the algebraic generalization of the procedure in terms of 'y'. There is isomorphism between example and example, as well as between example and the algebraic system. The manipulatives visually display the structure (double and add 3). Modelling '$2(y + 3)$' produces quite a different visual result, helping to clarify the difference between '$2(y + 3)$' and '$2y + 3$'. Mapping from arithmetical examples to algebra does not have the advantage of a clear visual display of the isomorphisms present.

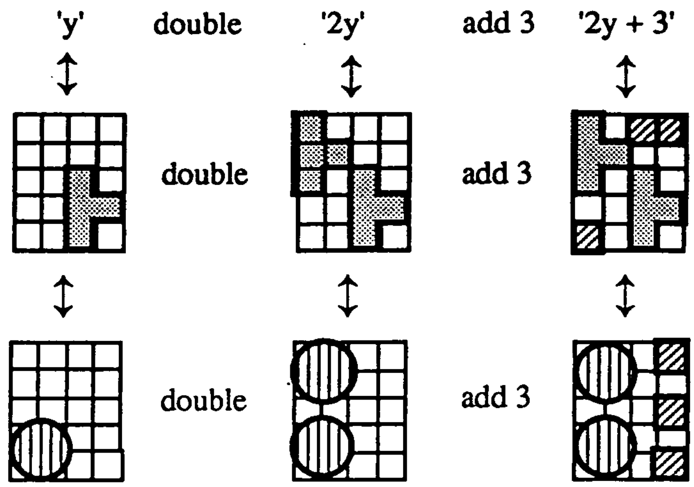


Figure 3P·2. Multiple-System Mappings from example to example and from example to algebraic generalization or vice versa

2. Equality of two algebraic expressions. Figure 3P·3 shows how the objects-and-containers model and the area model could be used to represent visually the truth of the algebraic identity

$$2(y = 3) = 2y + 6.$$

This visual form of cummunication is possible because, in the case illustrated, both the abstract algebraic system and the concrete objects-and-containers model have a similarity of structure.

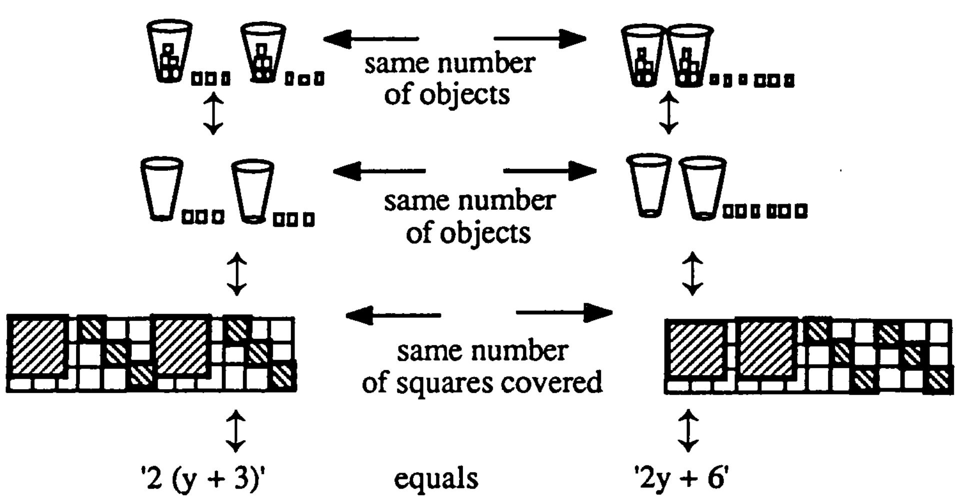


Figure 3P·3. Isomorphisms between concrete systems and algebraic system

11. For this question, let m = the number of blobs in one cup.
You will be told what values of m to use in parts (i), (ii) and (iii).
Again, build each using the containers and blobs,
draw what you build,
write how many blobs there are.

(i)

m = 4	m + 5 =	2m + 3=	3(m + 1)=	2(2m)=

(ii)

m = 3	m + 5 =	2m + 3=	3(m + 1)=	2(2m)=

(iii)

m = 2	m + 5 =	2m + 3=	3(m + 1)=	2(2m)=

(iv) **Discuss** the following two problems and use the results you have just obtained in order to solve them. Write your answers.

a. If m + 5 = 2m + 3,
what is the value of m ?

. . . .

b. If m + 5 is 1 less than 2m + 3,
what is the value of m ?

. . . .

(v) **Discuss** the following, using the results already obtained, and write your answer with a good explanation in the space below:
Which is bigger: 3(m + 1) or 2(2m) ?

12. Work with a partner for this question.
To make it easier for giving directions, one partner will be referred to as **Alpha** and the other as **Beta**.
Let w = the number of blobs in one cup.

(a) (i) Alpha decides on a value for w and builds 4 + 3w.
(ii) **Beta** takes 3 blobs from what **Alpha** has built, without changing the value of w.
(iii) What is Alpha left with? . . . .
(iv) Copy and complete: (4 + 3w) - 3 = ...

(b) (i) Beta decides on a value for w and builds 4w.
(ii) **Alpha** takes w blobs from what **Alpha** has built, without changing the value of w.
(iii) What is Beta left with? . . . .
(iv) Copy and complete: 4w - w = ...

(c) (i) Alpha decides on a value for w and builds 4 + 5w.
(ii) **Beta** takes 3 + 2w blobs from what **Alpha** has built, without changing the value of w.
(iii) What is Alpha left with? . . . .
(iv) Copy and complete: (4 + 3w) - (3 + 2w) = ...

(d) (i) Beta decides on a value for w and builds 2(3w + 4).
(ii) **Alpha** takes 3(w + 2) blobs from what **Alpha** has built, without changing the value of w.
(iii) What is Beta left with? . . . .
(iv) Copy and complete: 2(3w + 4) - 3(w + 2) = ...

(e) Make up some more subtraction examples that may be shown with the blobs in containers model, and try them out on your partner.

(f) Copy and complete:

(i) 3w - w = ... (ii) 3 + 3w - 3 = ...

(iii) 3w + 3 - 3w = ... (iv) (5w + 2) - (1 + 2w) = ...

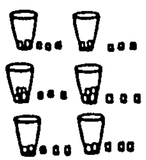
(v) 2(4 + y) - 2y = ... (v) 5(2K + 3) - (7K + 8) = ...

(vi) 3(4 + 3d) - 3(3 + d) = ... (A short way for this one!)

13. For each part of this question, imagine that you have been sent on a message during your mathematics class.
While you were away, the teacher wrote on the board an expression in algebra and asked the class groups to build a model for that expression using objects and containers.
When you came back you found that different groups had built models as shown in the diagrams.

[i] What algebraic expression had the teacher written on the board if three of the groups had built the following? ....

Group A built
Group B built
Group C built



For each group, state their value of 'y' (or whatever letter was used).

[ii] What algebraic expression had the teacher written on the board if three of the groups had built the following? ....

Group A built
Group B built
Group C built

For each group, state their value of 'y' (or whatever letter was used).

[iii] What algebraic expression had the teacher written on the board if three of the groups had built the following? ....

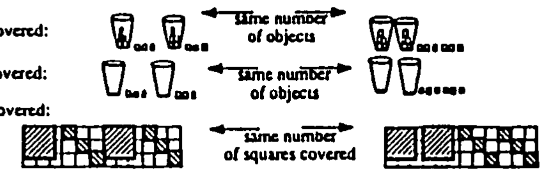
Group A built
Group B built
Group C built

For each group, state their value of 'y' (or whatever letter was used).

14. In each part of this question, you are to do two things:
(a) Write one equation in algebra which would be true for the discoveries made by all three groups; and
(b) Explain why you claim that your equation is true.

[i]

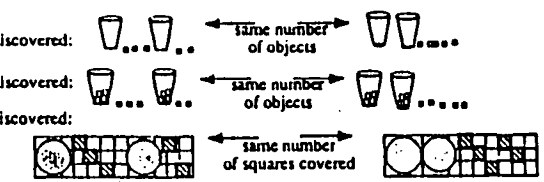
Group A discovered:
Group B discovered:
Group C discovered:



(a) My equation: ....
(b) My explanation: ....
....

[ii]

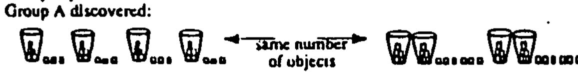
Group A discovered:
Group B discovered:
Group C discovered:



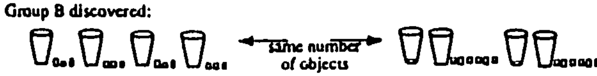
(a) My equation: ....
(b) My explanation: ....
....

[iii]

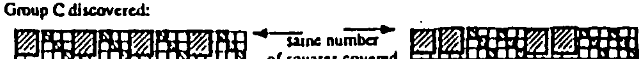
Group A discovered:



Group B discovered:



Group C discovered:



(a) My equation: ....
(b) My explanation: ....
....

PART III: PILOT STUDY ON LONG-TERM EFFECTS OF MANIPULATIVES

Applications of Halford's Structure-Mapping Theory in Interviews after Delayed Post-test

The researcher conducted 50 videotaped interviews of Year 7 students after they had completed the test instrument for the fourth time some six to seven months after the teaching intervention sessions. Dependence on models was not in evidence: Only five interviewees said that they sometimes had diagrams in mind during Test 4. The vast majority of those interviewed from Manipulatives Classes remembered how to use the models correctly. Of 33 students asked to map from algebra to one of the models, 23 used the models very well, 7 were rated "fair" and 3 were considered failures. Those from the Textbook Classes showed that they could very quickly learn how to use the models with a minimum of instruction: 6 out of 7 succeeded in mapping from algebra to model. The models proved 100 percent effective as self-correcting aids: All 17 students who were asked to model cases they had wrong in the test arrived at the correct answers via the models. For instance, six students who had written '$7n$' in answer to the question "Add 4 onto $3n$", when asked to model '$3n$' and then to add 4, recognised '$3n + 4$' as correct and quite different from '$7n$' in the model, as is shown using the objects-and-containers model in Figure 3P-4.

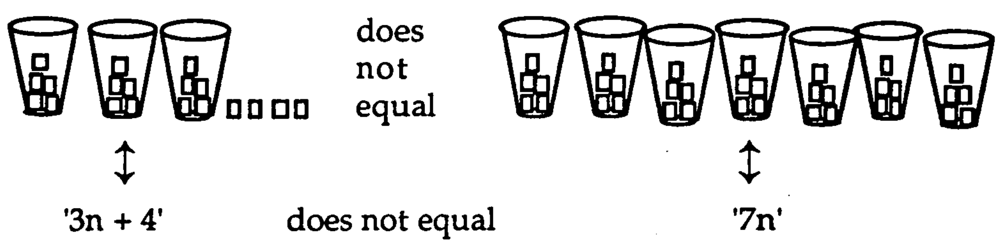


Figure 3P-4. Example of self-correction using structure mapping

Ability to map from model(s) to algebra was tested in 30 of the interviews. The researcher asked the students, for instance, to imagine that they had been sent on a message during a mathematics lesson and while they were out the teacher had set the class the task of building something in algebra. When they came back they found that

one group had built 6 objects near 2 containers each holding 3 objects,
another group had built 6 objects near 2 containers each holding 5 objects, and
another group had built 6 objects near 2 containers each holding zero objects.

The students then were asked what the teacher had set all the groups to build, in this case '2*y* + 6', the groups having chosen values 3, 5, and 0 respectively for '*y*'. Those familiar with the models were highly successful (20 out of 24), and 2 out of 6 from the Textbook Groups managed this.

Four students familiar with the models were tested on their ability to see the equality of two algebraic forms by mapping from models to algebra. All four succeeded. As an example of the generalization '$n(y + c) = ny + nc$', the models just described were re-built with the two containers separated and three objects placed beside each. The students wrote that this showed that '2*y* + 6' equalled '2(*y* + 3)'. Figure 3P·5 shows three of the mappings successfully used in interviews.

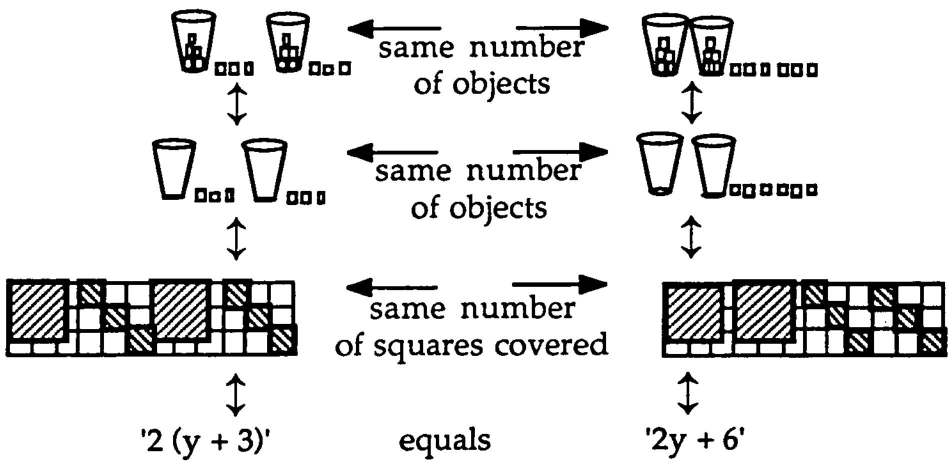


Figure 3P·5. Relational and Multiple-System Mappings

Brief reports on these posttest findings are given in Quinlan (1990, 1991).

NEW ITEM FOR
UNIT TWO WORKSHEET ONE

10. [i] Choose a new value for x. Build 3x, draw it , label the drawing, and write down the number of blobs used.

[ii] Now decide on the value of y so that $y = 3x$. Build y, draw it , label the drawing, and write down the number of blobs used.

[iii] Compare results with those of someone who used a different x value.

[iv] Which is bigger, x or y, if $y = 3x$? Discuss.

[v] Solve and discuss this problem:
In a certain class the number of boys (z) and the number of girls (w) are related by the equation $2z = w$.
What does this equation tell you about the numbers of boys and girls in the class?

NAME: .... M/F DATE: .... SCHOOL: ....

MONTH OF BIRTH: .... YEAR OF BIRTH: .... CLASS: ....

6. (a) If c + d = 10, tick ALL the meanings that c could have:

3 10 12 7.4 the number of apples in a box an object like a cabbage an object like a pear

7. At a certain university there are six times as many students as there are professors. This fact is represented by the equation S = 6 P.
CIRCLE YOUR CHOICES IN THE FOLLOWING QUESTIONS:

(a) In this equation, what does the letter P stand for?
- (i) Professors
- (ii) Professor
- (iii) Number of professors
- (iv) Students
- (v) Student
- (vi) Number of students
- (v) None of the above
- (vi) More than one of the above (if so, indicate which ones)
- (vii) Don't know.

(b) In this equation, what does the letter S stand for?
- (i) Professors
- (ii) Professor
- (iii) Number of professors
- (iv) Students
- (v) Student
- (vi) Number of students
- (v) None of the above
- (vi) More than one of the above (if so, indicate which ones)
- (vii) Don't know. (Rosnick 1981, adapted)

15. Decide whether the following statements are TRUE always, never or sometimes. Tick the correct answer.
If you tick "true only when ..", write when it is true.
All the letters stand for whole numbers or zero (0, 1, 2, 3, 4, ...)

(i) a + b + c = a + x + c
O true always
O never true
O true only when ....

(ii) 2a + 3b + 7 = 5a + 7
O true always
O never true
O true only when ....

(iii) 2a = a + 2
O true always
O never true
O true only when ....

(iv) a + 2b + 2c = a + 2b + 4c
O true always
O never true
O true only when ....

(Collis, 1975a)

Last year, did you answer questions like these?
...... Yes ..... No ..... Not sure [TICK ONE]

Description of Years 7 to 12 Students Who Avoided "Missing Values" in Their Test Responses, using Test 3 Responses for Year 7 Students Who Did the Test More Than Once

Table 5A·1

% Distribution Across Quartiles of Students Who Avoided "Missing Values"

CATEGORY by Quartile Divisions	NUMBER of Students	PERCENTAGE of Students
Quartile 4 (top)	109	45.2
Quartile 3	64	26.6
Quartile 2	47	19.5
Quartile 1 (low)	21	8.7
TOTALS	241	100

Table 5A·2

Distribution Across Year Groups of Students Who Avoided "Missing Values"

YEAR GROUP	NUMBER for ALL STUDENTS	NUMBER for STUDENTS INCLUDED#	PERCENT of YEAR GROUP INCLUDED#
7	232	96	41.4
8	54	17	31.5
9	99	57	57.6
10	36	12	33.3
11	65	46	70.8
12	31	13	41.9
TOTALS	517	241	46.6

"included" in group who avoided "missing values"

Factor Loadings After Varimax Rotation [$n = 241$; loadings less than 0.1 are omitted; 74.5% of variance explained]

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16
% Variance:		24.9	6.3	4.9	4.7	4.3	3.7	3.5	3.4	3.2	2.8	2.6	2.3	2.3	2.0	1.8	1.8
A	10c	.81	.20	.15	.13									.24			.12
	12b	.81	.21	.13	.24	.11											
	12c	.80	.22	.10	.26						.10						
	12d	.79	.22	.15	.17	.11											
	10b	.78	.18	.18										.23			.12
	10d	.75	.21	.14		.11	.13							.19			
	15iiib	.72	.15	.11	.11	.16	.12		.38				.11	-.13			-.17
	15iiia	·.70		.11		.17	.12		.39			.16	.10	-.13			-.20
	12a	.67	.11	.12	.19	.22	.10		.21	.11			.11				
	10a	.62	.21		.11	.17			-.11				.17	.16	.12		.34
	6aiii	.41	.18		.27				.14	.11	.19			-.13		.24	
	8a	.33	.22		.11	.17			.22	.18	.21				-.11		.14
B	9ii	.31	.82		.11	.14				.14					.10	-.11	.10
	9iii	.28	.81		.15	.11				.12							
	9i	.22	.80		.11					.12		.12	.10		.14	-.17	.12
	5	.27	.77	.11	.11	.13					.10		.12				
	14ii	.27	.69	.13	.19	.12			.11			-.15				.27	-.14
	14iii	.34	.64	.15	.19	.13			.15			-.13				.27	-.18
C	3ii	.17		.88													.13
	3iv		.13	.86													
	3i	.15		.81													.11
	3vi	.16		.76			.17										-.11
	3v			.76		.12											
	3iii	.16		.73										.15			
	8b	.26	.23	.38	.15					.23	-.10	.19		.15	-.12	-.11	

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16
% Variance:		24.9	6.3	4.9	4.7	4.3	3.7	3.5	3.4	3.2	2.8	2.6	2.3	2.3	2.0	1.8	1.8
D	13d	.26	.16		.86	.12											
	13b	.31			.81	.16											.11
	13a	.28	.17	.11	.78	.17											
	13c	.22	.23		.76		.11		.11								
E	11b	.18	.12		.12	.92											
	11c	.21	.10		.12	.91											
	11d	.25	.14		.10	.83											
	11a		.11		.10	.77											
F	1aii	.11					.89	.10								-.12	
	1ai						.86										.17
	1bii	.35	.21		.17		.61						.14		.26	.24	-.17
	1bi	.26	.20		.14		.60						.11		.24	.35	
G	4b							.90									
	4a							.89					-.11				
	4c					.11	.12	.88									
H	15iia	.22							.82			.16		.22			
	15iib	.25	.12						.80			.12		.26			
I	6avi									-.94							
	6avii		-.11							-.92		.11					
	6av	.17	.16		.24				.20	.53	.12		.13			.18	
J	7b										.91						
	7a										.90	.12					
	8d		-.18							-.20	.31		-.15	.27	.14	.16	.26

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16
% Variance:		24.9	6.3	4.9	4.7	4.3	3.7	3.5	3.4	3.2	2.8	2.6	2.3	2.3	2.0	1.8	1.8
K	15iva											.93					
	15ivb								.14	-.10	.13	.91					
L	2ib	.12	.12										.89				
	2ia	.13	.18				.11						.88				
M	15ia	.22	.12	.12	.16	.13			.34		.11		.14	.74			
	15ib	.25	.19	.10	.15	.12			.32				.14	.74			
	6c	.26	.19		.19		.14		.28		.13			-.31	.17		
N	6ai											.14			.83		
	6aiv	.19	.13	.14			.12		.11	.11		-.15	.13		.72		
	2ii	.18	.10	.28	.23		.30		.16			-.15			.36	-.25	-.14
Q	6b	.13	.24	.18			.12		.17		.13		.17		.15	-.56	
	6aii	.25	.12	.22			.12		.22	.19		.12	.12	-.12		.54	.19
R	8c	.11		.14	.19				.22		.13					.13	.64
	14i	.18	.25		.25				.14		.30						-.44
Eigenvalues		15.41	3.91	3.05	2.93	2.64	2.29	2.15	2.10	1.99	1.74	1.58	1.45	1.42	1.26	1.14	1.11

Factor Loadings After Varimax Rotation [$n = 241$; loadings less than 0.1 are omitted; 72.7% of variance explained]

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17
% Variance:		24.2	5.4	4.9	4.4	3.8	3.6	3.3	3.2	2.9	2.5	2.4	2.4	2.2	2.0	1.9	1.9	1.7
A	10c	**.82**	.15	.10		.10									.23	-.13	.18	.11
	10b	**.80**	.15	.12											.21	-.11	.18	.12
	12b	**.79**	.19			.23									.10	.13		
	12c	**.76**	.22			.24					.11		.11			.19		
	12d	**.70**	.30	.16	.13	.15			.12					.10		.17		
	10d	**.70**	.22	.11	.18		.18		.15	.11		.14		.14				
	15iiib	**.66**	.21	.13	.15	.14	.12		.33			.12				.27	-.15	
	15iiia	**.63**		.12	.17		.16		.36			.13	.13			.29	-.27	
	12a	**.61**	.15	.14	.26	.19	.11	.14	.24	.15						.13		
	10a	**.49**	.23	.16	.20	.13	.18	.15	.11	.24		.13		.16			.19	
B	9iii	.31	**.75**	.11	.15	.19		.12								.10		
	9ii	.34	**.75**	.11	.17			.13	.13							.10	.11	
	9i	.20	**.72**	.11		.11								.13	.13			-.15
	5	.29	**.66**	.13	.18	.11				.13	.11							
	14ii	.26	**.65**		.13	.16				.25		.13						.38
	14iii	..33	**.50**	.12	.11	.18			.16	.22								.48
C	3ii	.17		**.85**												-.10		
	3i	.12		**.84**														
	3iv		.16	**.79**						.11						.13		
	3v			**.78**	.15				.11									
	3vi	.15		**.70**			.15											
	3iii	.14		**.55**	.12	-.11					.10	.13		-.11	.33	.19		
	8b	.20	.31	**.41**		.15	.10	.21					.17			-.16		

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17
% Variance:		24.2	5.4	4.9	4.4	3.8	3.6	3.3	3.2	2.9	2.5	2.4	2.4	2.2	2.0	1.9	1.9	1.7
E	11c	.20	.15		.93													
	11b	.19	.14		.93													
	11d	.18	.12		.86													.11
	11a				.83	.13								.11	.12			
D	13d	.24			.12	.84									.10	.13	.10	
	13b	.24			.10	.81				.11	.13						.20	
	13a	.25	.28	.14	.18	.68			.16			.13				-.12		
	13c	.16	.29	.15		.68	.16		.20								-.13	
F	1aii		.10				.87											
	1ai						.84										.11	-.14
	1bi	.24	.14			.11	.61			.14	-.10			.23		.12		.34
	1bii	.33	.16			.16	.59			.18				.20				.30
I	6avi							-.94										
	6avii							-.94					.10					
	6av	.18	.17			.24		.53	.11	.11						.21		
H	15iia	.20		.12					.78				.17		.21	.10		
	15iib	.24	.15			.12			.77		.11				.19			
L	2ib	.12	.15				.11			.89								
	2ia	.15	.15				.15			.89								
J	7a										.90		.10					
	7b					.12					.90							

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17
% Variance:		24.2	5.4	4.9	4.4	3.8	3.6	3.3	3.2	2.9	2.5	2.4	2.4	2.2	2.0	1.9	1.9	1.7
G	4iii	.14					.10					.83						
	4ii					.10				.20		.76	.11				.19	
	4i								-.11	-.22		.71	-.11			.13	-.14	
K	15iva										.10		.89					
	15ivb	.19							.17				.82				.11	
N	6ai												.11	.83	.11	.13		
	6aiv	.20	.13	.10	.13		.11	.13		.16				.73				
	2ii	.21	.18	.26		.20	.30		.11				-.11	.34		-.19	-.13	-.20
M	15ia	.20	.15	.10	.17	.14			.33	.12					.74			
	15ib	.27	.25		.15	.14	.14		.31	.13				.11	.69			
S	6aii	.16		.23			.14	.16	.16					.14		.63	.24	.14
	6aiii	.36	.20			.35					.14				.14	.51		
	6c	.18	.22			.15	.18	-.13	.24					.18	-.27	.35		-.16
T	8c	.12		.13		.11				-.11						.13	.72	
	8a	.20	.32		.10	.11		.17	.32								.38	
	8d		-.18				.10	-.18	.17	-.10	.35			.15		-.13	.35	.23
U	6b		.34	.16			.10		.10	.21	.10			.14			-.14	-.45
	14i	.17	.18			.21					.23		.24	-.14	.18	.26	-.32	.37
Eigenvalues		15.02	3.38	3.07	2.71	2.36	2.25	2.06	1.97	1.79	1.57	1.49	1.46	1.36	1.23	1.16	1.15	1.05

Factor Loadings After Varimax Rotation [$N = 571$; loadings less than 0.1 are omitted; 69.7% of variance explained]

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14
% Variance:		28.3	6.7	5.2	4.2	3.6	3.5	3.0	2.8	2.5	2.3	2.2	2.0	1.8	1.7
A	10c	.79	.17		.19	.11	.12							-.16	.22
	12b	.78	.11		.18		.11	.25							
	10b	.76	.15		.18	.13	.12			.14				-.20	.23
	12c	.76	.11		.16			.26							
	12d	.71	.20		.24	.17	.16	.20	.14					.14	-.11
	10d	.70	.14		.20	.22	.18		.21			.12			
	12a	.67	.16		.21	.23	.19	.22	.13	.12				.14	-.10
	15iiib	.65	.14		.21	.13	.24	.25	.16			.13	.12	.21	-.16
	15iiia	.60	.14		.13	.15	.26	.21	.18			.14	.20	.16	-.26
	10a	.55	.27		.21	.15	.22		.12	.26		.11			
	6c	.32	.11	.19	.12				.30		.12		.10	.19	
C	3i	.13	.89		.10										
	3ii	.15	.89		.10										
	3iv		.86		.11										
	3v	.12	.81		.15	.10			.11	.11					
	3vi	.19	.76	.10	.13				.13						
	3iii	.17	.73	.13	.12		.17								
	8b	.17	.47		.35		.11	.11							
P	6avi			.86			-.12							-.32	-.11
	6avii			.85			-.10							-.34	-.13
	6ai		.22	.84					.16						
	6aiv	.14	.22	.74	.14	.13			.18	.16				.17	
	6aii	.13		.73			.18							.19	
	6aiii	.21		.65	.19		.16	.28						.14	
	6av		.17	.53	.12		.18	.12						.45	.16

PC Analysis on Dichotomous Data While Including Missing Values

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14
% Variance:		28.3	6.7	5.2	4.2	3.6	3.5	3.0	2.8	2.5	2.3	2.2	2.0	1.8	1.7
B'	14ii	.19	.18		.75	.16		.15		.13					
	9ii	.34	.19		.67	.12	.18		.10					.15	
	9iii	.32	.23		.67	.13	.19	.18	.15					.11	
	14iii	.27	.11		.67		.13	.16		.16					
	9i	.18	.17		.62		.12		.14					.15	
	5	.24	.19		.58	.17			.14						
	14i				.38		.12	.28		-.16			.27	-.16	-.17
E	11c	.23	.13		.14	.91		.11							
	11b	.23	.12		.14	.90		.11							
	11d	.19	.12		.16	.86	.13	.14							
	11a	.11	.11		.11	.81		.12		.12					
H&M	15iia	.20					.74						.24		
	15iib	.22			.17	.10	.72	.10					.15		
	15ia	.30	.18		.14	.15	.68		.12	.15				-.13	
	15ib	.36	.16		.23	.10	.66		.16	.11					
D	13d	.28			.13	.14		.84							
	13b	.29				.10		.81			.14				
	13c	.21	.15		.31	.18	.17	.65	.17						
	13a	.27	.15		.29	.21		.61	.10	.12					
F	1aii	.17	.16						.83						
	1ai			.14					.82						
	1bii	.33	.21		.28	.12	.14	.14	.61						
	1bi	.22	.15		.23		.12	.10	.60						.11
	6b		.24	.14	.22	.20			.34		.18		.11		

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14

APPENDIX 5D **PC Analysis on Dichotomous Data While Including Missing Values**

% Variance:		28.3	6.7	5.2	4.2	3.6	3.5	3.0	2.8	2.5	2.3	2.2	2.0	1.8	1.7
L	2ia	.18	.18		.13		.13		.13	.88					
	2ib	.16	.18		.14		.12			.88					
	2ii	.20	.12		.17	.15			.15	.24	.12			.23	.17
J	7a										.92				
	7b				.11			.13			.90				
G	4iii	.11										.81			
	4i											.74			.12
	4ii							.11	.12			.71			
K	15iva						.16						.85		
	15ivb	.20			.11		.12				.11		.79		.13
V	8d		.10		-.21		.20			.21				-.47	.24
W	8a											.12	.11		.74
	8c	.27			.24		.33	.10		.12				.19	.34
		------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Eigenvalues		17.53	4.13	3.22	2.59	2.22	2.17	1.85	1.75	1.54	1.41	1.39	1.26	1.12	1.08

APPENDIX 5E ALPHA Factor Analysis on Ordinal Data While Rejecting Missing Values

Factor Loadings After Varimax Rotation [$n = 241$; loadings less than 0.1 are omitted; 66.4% of variance explained]

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16
% Variance:		24.3	65.7	4.4	4.3	3.7	3.2	3.0	3.0	2.8	2.4	2.1	1.7	1.9	1.5	1.3	1.1
A	12b	.80	.21	.14	.12	.23											
	12c	.79	.21	.10		.26						.11				.10	
	10c	.79	.19	.16		.13								.25			.17
	12d	.76	.21	.16	.12	.16											
	10b	.75	.17	.19										.21			.15
	15iiib	.73	.15	.11	.17	.11	.11		.33				.11	-.13			-.20
	10d	.71	.22	.15	.12		.13							.16			.11
	15iiia	.69		.12	.17		.12		.33		.16			-.12			-.23
	12a	.65	.12	.13	.22	.19	.10		.17	.10			.10				
	10a	.57	.21		.18	.11							.16	.14	.12		.32
	6aiii	.40	.18			.24			.11			.16				.23	
	8a	.31	.22		.15	.13			.19	.15		.17					
	6c	.27	.16			.16	.12		.14					-.16	.11	.11	
B	9ii	.31	.83		.14	.12			.11	.14					.11		.12
	9i	.23	.78			.12				.12	.11		.11		.14	-.12	.11
	9iii	.29	.78		.12	.16				.12							
	5	.28	.73	.11	.14	.12						.11	.12				
	14ii	.29	.65	.12	.13	.19					-.14					.30	-.11
	14iii	.36	.59	.14	.14	.19			.11		-.12					.30	-.14
C	3ii	.16		.89													.13
	3iv		.12	.85													
	3i	.14		.77													
	3vi	.16		.72			.15										
	3v			.70	.11												
	3iii	.15		.68										.13			
	8b	.25	.20	.34		.15				.17	.12			.11			

APPENDIX 5E ALPHA Factor Analysis on Ordinal Data While Rejecting Missing Values

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16
% Variance:		24.3	65.7	4.4	4.3	3.7	3.2	3.0	3.0	2.8	2.4	2.1	1.7	1.9	1.5	1.3	1.1
E	11b	.18	.12		.93	.12											
	11c	.21			.91	.12											
	11d	.25	.14		.78	.10											
	11a		.12		.67	.11											
D	13d	.26	.15		.12	.87											
	13b	.31			.17	.76											
	13a	.29	.18	.11	.17	.73											.13
	13c	.24	.23			.70	.10										
F	1aii	.11					.90										
	1ai						.76										.14
	1bii	.36	.19			.17	.56						.13	.12	.24	.26	-.14
	1bi	.28	.18			.13	.54					-.12		.12	.21	.34	
G	4c			-.10	.11		.12	.86									
	4b							.85									
	4a							.84					-.11				-.12
H	15iib	.27	.11						.80		.12	.10		.20			
	15iia	.24							.77		.16			.15			
I	6avi									-.94							
	6avii		-.11							-.90	.12						
	6av	.19	.16			.23			.16	.42						.17	
K	15iva										.93						
	15ivb								.14	-.10	.83	.14					

APPENDIX 5E ALPHA Factor Analysis on Ordinal Data While Rejecting Missing Values

Cluster	Quest.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16
% Variance:		24.3	65.7	4.4	4.3	3.7	3.2	3.0	3.0	2.8	2.4	2.1	1.7	1.9	1.5	1.3	1.1
J	7b											.90					
	7a										.12	.80					
	14i	.21	.23			.22			.12			.25					-.21
	8d		-.15							-.13		.19		.17			.19
L	2ia	.13	.18				.11						.89		.10		
	2ib	.12		.13									.79				
M	15ib	.25	.19	.11	.13	.15			.38				.16	.64			
	15ia	.22	.12	.13	.14	.16			.40			.12	.16	.63			
N	6ai										.10				.66		
	6aiv	.19	.13	.14			.13			.11	-.13		.12		.64		
	2ii	.19	.13	.27		.22	.26		.14		-.11				.29	-.14	
Q	6aii	.27		.20					.16	.17						.44	.17
	6b	.13	.24	.18			.11		.14			.12	.15		.14	-.27	
X	8c	.13		.15		.15			.19			.12					.41
Eigenvalues		15.07	3.56	2.73	2.69	2.29	1.99	1.84	1.85	1.73	1.47	1.31	1.03	1.21	0.92	0.78	0.71

Response Categories Used in Scales

Table 5F·1

Q.1 Response Categories Used in the AR Arithmetic Scale

Parts	Response Type	AR
ai, bi	Yes	✔
aii	$y = 1$	✔
bii	$y = 2$	✔

Table 5F·2

Q.2 (i) Response Categories Used in Scales

Part	Code	Response Type	AD	GNV	AL1	NV	PRE
a	1	*a*				✔	✔
a	2	*d*				✔	✔
a	3	neither	✔	✔			
a	4	same				✔	✔
a	5	not sure		✔			
a	6	other error				✔	✔
b	1	irrelevant "reason"					✔
b	2	order in alphabet			✔		✔
b	3	neither; can't tell	✔	✔			
b	4	they aren't numbers					✔

Table 5F-3

Q.3 Response Categories Used in Scales

Part	Expression	Response	SUB	SUBS	CON	IG	PV
i	$2y$	6	✔	✔			
i	$2y$	5			✔		
i	$2y$	2				✔	
i	$2y$	23					✔
ii	$2y + 5$	11	✔	✔			
ii	$2y + 5$	10			✔		
ii	$2y + 5$	7				✔	
ii	$2y + 5$	28					✔
iii	$2(y + 5)$	16	✔	✔			
iii	$2(y + 5)$	10			✔		
iii	$2(y + 5)$	7				✔	
iii	$2(y + 5)$	28					✔
iv	$2y + y$	9	✔	✔			
iv	$2y + y$	8			✔		
iv	$2y + y$	26					✔
v	$3y - y$	6	✔	✔			
v	$3y - y$	3			✔		
v	$3y - y$	30					✔
vi	$2(5y)$	30	✔	✔			
vi	$2(5y)$	10			✔		
vi	$2(5y)$	28; 106; 253; 630					✔

Table 5F-4
Q.4 Response Categories Used in Scales

Part	Code	Response Type	FL	NFL	JFL
i	1	No			
i	2	Yes	✔		
ii	2	'g' = flowers/bunches			✔
ii	4	no.of flowers in a bunch	✔	✔	
ii	5	no.of bunches		✔	
ii	6	any no.of flowers		✔	
ii	7	flowers in bunches			✔
ii	8	no.of grams		✔	
ii	9	grams			✔
iii	3	'8' = 8 extra flowers	✔	✔	
iii	4	no.in a bunch		✔	
iii	5	(the) bunches			✔
iii	6	8 bunches		✔	
iii	7	flowers			✔
iii	9	any no./no./the no.		✔	

Table 5F-5
Q.5 Response Categories Used in Scales

Code	Response Type	SYM	ALC	SC2
1	uncertainty			✔
2	34			✔
3	no.other than 34			✔
4	*pr*		✔	
6	$p + r$	✔	✔	
7	$p + r = pr$		✔	

Table 5F-6a

Q.6(a) Response Categories Used in Scales

Part	Response Type	FZN	GN#	GNV#	C2	NBR	JCP	NJCP
a i	Accept $c = 3$		✔	✔	✔			
a ii	Accept $c = 10$	✔	✔	✔				
a iii	Accept $c = 12$	✔	✔	✔				
a iv	Accept $c = 7.4$	✔	✔	✔	✔			
a v	Accept $c =$ no.of apples in a box					✔		
a vi	Accept $c =$ an object like an apple						✔	
a vi	Reject $c =$ an object like an apple							✔
a vii	Accept $c =$ an object like a pear						✔	
a vii	Reject $c =$ an object like a pear							✔

a score of "1" was merited on the GN and GNV Scales if two or more of the options from parts i to iv were accepted.

Table 5F-6b

Q.6(c) Response Categories Used in Scales or as Individual Item

Code	Response Type	VBL	GNV	FZN	PRE	Single Item GN6c	Single Item NV6c	Single Item INT6c
1	Other error				✔			
2	1, 2, 3, 4 (5)		✔			✔		✔
3	0, 1, 2, 3, 4 (5)		✔	✔		✔		
4	< 5	✔	✔	✔				
5	< 5, specifically mentioning fractions, zero and/or negatives	✔	✔	✔				
6	"3"		✔		✔		✔	
7	"4"				✔		✔	
9	any value		✔					

Table 5F·7

Q.7 Response Categories Used in Scales

Part	Code	Response Type	PS	NBR	OBJ	NRPS	RPS
a	1	i			✔	✔	
a	2	ii			✔	✔	
a	3	iii	✔	✔		✔	
a	4	iv			✔		✔
a	5	v			✔		✔
a	6	vi		✔			✔
a	8	2 or 3 from i,ii,iii			✔	✔	
a	9	2 or 3 from iv,v,vi			✔		✔
a	0	Omit (a) not (b)					
b	1	i			✔		✔
b	2	ii			✔		✔
b	3	iii		✔			✔
b	4	iv			✔	✔	
b	5	v			✔	✔	
b	6	vi	✔	✔		✔	
b	8	2 or 3 from i,ii,iii			✔		✔
b	9	2 or 3 from iv,v,vi			✔	✔	
b	0	Omit (b) not (a)					

Table 5F·8

Q.8 Response Categories Used in Scales or for Individual Item

Part	Code	Response Type	OBJ	PV	EQL	SUBS	Ind.Item
a	2	apples	✔				
a	3	an apple or one apple					✔
b	12	correct: $a = 12$			✔	✔	
b	6	$3a = 30 + a$: so $a = 6$		✔			
c	4	strongly agree = "4"					✔
c	3	agree = "3"					✔
c	2	disagree = "2"					✔
c	1	strongly disagree = "1"					✔
d	1	i (fruit) = "1"					✔
d	2	ii (numbers) = "2"					✔

Table 5F·9

Q.9 Response Categories Used in Scales or for Individual Item

Part	Code	Response Type	SYM	ALC	CON	SC2	IG	AL1
i	1	$n+9$	✔	✔				
ii	1	$3n+4$	✔	✔				
iii	1	$4n+20$ or $4(n+5)$	✔	✔				
i	2	$9n$, $n9$		✔	✔			
ii	2	$7n$		✔	✔			
iii	2	$20n$		✔	✔			
i	3	other algebra, e.g., $4n + 5$, $4 + n + 5$		✔				
ii	3	other algebra, e.g., $3n4$		✔				
iii	3	other algebra, e.g., $n + 20$, $n + 5 \times 4$		✔				
i	9	"9"				✔	✔	
ii	4, 7	"4" or "7"				✔	✔	
i	nos.	"19","23","74"				✔		✔
ii	nos.	"21","46","318"				✔		✔
iii	76	"76"				✔		✔
i,ii,iii	nos	other nos.				✔		

Table 5F·10a

Q.10 (a) Response Categories Used in Scales

Code	Response Type	VBL	NV	PRE
1	Wrong idea, e.g., $t + 4 = 4t$			✔
2	No sense of variable, e.g. $t = 20$		✔	✔
3	Variable notion: Correct answer	✔		
4	can't proceed without value of t (no closure)			✔
5	t not regarded as a number			✔
6	can't add t and 4			✔
8	$t + t$ always greater			✔
9	$t + 4$ always greater			✔
0	Omit this part but answered other part(s) of Q.10			

Table 5F·10b

Q.10 (b), (c) Response Categories Used in Scales

Code	Response Type	VBL	GN	GNV	12REP	1REP	NV	PRE	INT
1	Wrong idea or repeats question							✔	
2	False ordering e.g., t's not equal							✔	
3	More than one replacement value		✔	✔	✔				✔
4	Algebra not quite correct e.g., (b) If $t \geq 5$		✔	✔					✔
5	Correct answer e.g., (b) If $t > 4$	✔		✔					
6	One replacement value				✔	✔	✔		✔
7	"Always" or "Now"						✔	✔	
8	"Never"						✔	✔	
9	Literal comparison e.g. (c) If second t is 4							✔	
0	Omit this part but answered other part(s) of Q.10								

Table 5F·10c

Q.10 (d) Response Categories Used in Scales

Code	Response Type	VBL	PRE	EQL
1	Wrong idea or repeats question		✔	
2	False ordering e.g., t's not equal		✔	
5	Correct answer: $t = 4$	✔		✔
7	"Always" or "Now"		✔	
8	"Never"		✔	
9	Literal comparison e.g. (c) If second t is 4		✔	
0	Omit this part but answered other part(s) of Q.10			

Table 5F·11

Q.11 Response Categories Used In Scales

Parts	Code	Response Type	PL	EQL	GNV	PRE
a - d	1	Repeats question, e.g., (c) "When green is shorter"				✔
a - d	2	No idea, e.g., 'a' > 'b' always				✔
a - d	3	Focus on geometry, e.g., perspective, double the length, bend line				✔
a - d	4	Algebra correct, e.g,. (b) when $a > b$	✔		✔	
a - d	5	"always" or "now"				✔
a - d	6	"never"				✔
a - d	7	vaguely variable notion, e.g., either could be longer			✔	
d	4	Algebra correct, i.e.,. (d) when $a = b$		✔		

Table 5F·12a

Q.12 (a) Response Categories Used in Scales

Code	Response Type	VBL	GNV	PRE	SC1
1	Wrong idea, e.g., "n + 2 is like p"			✔	
2	"Same"			✔	
5	More than one replacement value		✔		
6	Algebra nearly correct, e.g,. (b) $n \geq 3$		✔		
7	Need value of n (no closure)			✔	✔
9	Correct: e.g. (b) $n > 2$	✔	✔		
0	Omit this part but answered other part(s) of Q.12				

Table 5F-12b

Q.12 (b), (c) Response Categories Used in Scales

Code	Response Type	VBL	GN	GNV	12REP	1REP	NV	PRE	INT	SC1
1	Wrong idea e.g., $2n$ is a 2-digit number							✔		
2	Repeats question							✔		
3	"Always" or "Now"						✔			
4	"Never"						✔			
5	More than one replacement value		✔	✔	✔				✔	
6	Algebra nearly correct e.g. (b) $n \geq 3$		✔	✔					✔	
7	Need value of n (no closure)							✔		✔
8	One replacement value				✔	✔	✔		✔	
9	Correct answer e.g. (b) When $n > 2$	✔		✔						
0	Omit this part but answered other part(s)									

Table 5F-12c

Q.12 (d) Response Categories Used in Scales

Code	Response Type	VBL	PRE	EQL	SC1
1	Wrong idea e.g., $2n$ is already added together		✔		
2	Repeats question		✔		
7	Need value of n (no closure)		✔		✔
9	Correct answer When $n = 2$	✔		✔	
0	Omit this part but answered other part(s) of Q.10				

Table 5F·13a

Q.13 (a) Response Categories for Scales used

Code	Response Type	EQN	GNV	PRE	12REP	2REP	SC1
1	Repeats question, or wrong idea, e.g., "never because y is further in the alphabet"			✔			
2	One replacement pair of values				✔		
3	False ordering by coefficients			✔			
4	Fictitious measures e.g. x, y independent variables		✔				
5	Two or more replacement pairs		✔		✔	✔	
6	Algebra nearly correct e.g., integers only		✔				
7	Correct answer	✔	✔				
8	"Never", "Always" or "Now"			✔			
9	Need values (no closure)			✔			✔
0	Omit this part but answered other part(s) of Q.10						

Table 5F·13b

Q.13 (c) Response Categories for Scales used

Code	Response Type	EQN	NV	PRE	CF	EQL	SC1
1	Repeats question, or wrong idea, e.g., "When $2x = y$."			✔			
2	One replacement pair of values		✔				
3	False ordering by coefficients			✔	✔		
7	Correct answer: $x = 3$ ($= y$)	✔				✔	
8	"Never", "Always" or "Now"		✔	✔			
9	Need values (no closure)			✔			✔
0	Omit this part but answered other part(s) of Q.10						

Table 5F·13c

Q.13 (b), (d) Response Categories for Scales used

Code	Response Type	EQN	GN	GNV	NV	PRE	12REP	1REP	INT	CF	SC1
1	Repeats question, or wrong idea, e.g., "When it is used to represent a number."					✔					
2	One replacement pair of values				✔		✔	✔	✔		
3	False ordering by coefficients					✔				✔	
4	Fictitious measures, e.g. x, y independent variables		✔	✔							
5	Two or more replacement pairs		✔	✔			✔		✔		
6	Algebra nearly correct, e.g., (b) $x = 4, 5, 6, \ldots$		✔	✔					✔		
7	Correct answer e.g., (b) $x > 3$	✔		✔							
8	Never, Always or Now				✔	✔					
9	Need values (no closure)					✔					✔
0	Omit this part but answered other part(s) of Q.10										

Table 5F·14a

Q.14 (i) Response Categories for Scales used

Code	Response Type	NBR
2	How many students took buses	✔
5	The number of buses which take the students	✔

Table 5F·14b

Q.14 (ii) (iii) Response Categories for Scales used

Code	Response Type	SYM	ALC	SC2	AL2
1	Correct answer: (ii) $3f + 4g$ (iii) $3f + 3g$	✔	✔		
2	Other algebraic answers e.g. (ii) $3f4g$		✔		
3	Using $f = 6$, $g = 7$			✔	✔
4	Other numerical answers			✔	

Table 5F·15

Q.15 Response Categories for Scales used

Part	Code	Response Type	BXBA	VBL	CZ	EQL	FZN
i a	3	true only when ...	✔				
ii a	3	true only when ...	✔				
iii a	3	true only when ...		✔			
iv a	3	true only when ...			✔		
i b	2	correct ($b = x$)	✔			✔	
ii b	2	correct ($b = a$)	✔			✔	
iii b	2	correct ($a = 2$)		✔		✔	
iv b	2	correct ($c = 0$)			✔	✔	✔

Caution. Some of the correlation coefficients in Table 6·9 were fairly low in size although highly significant statistically (at the 1% level). The corollary was to expect exceptions to the general conclusions about relationships between the variables, without discrediting the significance of the correlations.

Tables 6A·1 and 6A·2 display cross-tabulations of scores on the BXBA Scale (one of the Variables View scales) against scores on the NBR Numbers View Scale and the OBJ Objects View Scale respectively. The scores have been grouped to avoid having too large a table and small cell frequencies.

Table 6A·1

Cross-tabulation Scale BXBA by Scale NBR

		Scale NBR 0	1 or 2	3 or 4	Row Totals
Scale BXBA	0	79 (18.2)	72 (16.6)	21 (4.8)	172 (39.6)
	1 or 2	39 (9.0)	73 (16.8)	22 (5.1)	134 (30.9)
	3 or 4	18 (4.2)	65 (15.0)	45 (10.4)	128 (29.5)
	Column Totals	136 (31.3)	210 (48.4)	88 (20.3)	434 (100.0)

Note. Frequencies are given for each cell, with percentages of N in brackets. N =434. $\chi^2 = 47.3$, $df = 4$, $p \leq .000$; $r = .315$, $p \leq .000$.

The chi-square (χ^2) and correlation (r) statistics both indicate a strong relationship between scores on the NBR and BXBA Scales. The cell frequencies for the extremes of equal score categories on each scale (the top left and the bottom right) accounted for nearly 30% of the responses (18.2 + 10.4 = 28.6), while the most discordant frequencies (in the top right and the bottom left cells) tallied only 9.0% of students with a high score on one scale and a zero score on the other scale. These features of the distribution of scores underlined the strength of the positive interaction between the Numbers View of symbols and ability to work with the notion of variable, while acknowledging that not all cases belonged to the categories which supported this claim.

Table 6A·2
Cross-tabulation Scale BXBA by Scale OBJ

		Scale OBJ 0	1	2 or 3	Row Totals
Scale BXBA	0	42 (9.5)	66 (15.0)	69 (15.6)	177 (40.1)
	1 or 2	51 (11.6)	42 (9.5)	43 (10.0)	136 (30.8)
	3 or 4	59 (13.4)	48 (10.9)	21 (4.8)	128 (29.0)
	Column Totals	152 (34.5)	156 (35.4)	133 (30.2)	441 (100.0)

Note. Frequencies are given for each cell, with percentages of N in brackets.
N =434. $\chi^2 = 25.1$, $df = 4$, $p \leq .000$; $r = -.230$, $p \leq .000$.

The chi-square (χ^2) and correlation (r) values in Table 6A·2 strongly indicate that there was a negative interaction between the Object View of symbols and the development of the variable concept. Nearly 30% of cases (15.6 + 13.4 = 29.0) were located in the cells at the top right or bottom left, showing the negative nature of this strong relationship. The major cases showing the opposite tendency were the 14.3% who had either high scores on both scales or zero score on both (bottom right and top left cells). The negative interaction was strong despite the existence of cases which showed some form of positive interaction.

Table 6B·1 gives correlations with two scale measures derived from Question 4, the question about applying algebra to a number of flowers. These figures complement those given in Table 6·9 in the main text.

Table 6B·1

Correlations Between Variables-, Numbers-, and Objects- Views and Test Scores

VARIABLES View	NUMBERS View	OBJECTS View	Corrected TEST TOTAL
	NFL Scale	JFL Scale	
VBL Scale	N.S.	N.S.	.762 (442) ***
EQN Scale	N.S.	- .163 (395) ***	.590 (441) ***
PL Scale	.080 (431) *	- .112 (431) **	.472 (491) ***
BXBA Scale	.106 (413) *	- .085 (413) *	.619 (468) ***
CZ Scale	N.S.	N.S.	.228 (470) ***
Corrected TEST TOTAL	.085 (448) *	- .109 (448) **	-

Note. Year 7 to 12, using Test 3 for Year 7 students who did the test more than once. Scales formed from dichotomous data. Figures in brackets give the number of cases. Corrected TEST TOTAL = TEST TOTAL *minus* marks gained from items in scale.
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$, N.S. not significant.

In the following crosstabulations, responses to item 8 (d) were scored "1" for choosing option (ii), the Numbers View, and "0" for choosing option (i), the Objects View.

Table 7A-1
Cross-tabulation of Scores on Item 8 (d) and the NFL Scale

NFL (Q.4) / Q. 8 (d)	0	1	2	Row Totals
0	109	50	27	186
1	128	64	62	254
Column Totals	237	114	89	440

Note that in Table 7A-1, just over half of the 254 students who chose a Numbers View in Item 8 (d) rejected a similar view in Question 4. (128 of them scored "0" on the NFL Scale.) The correlation between these measures was not statistically significant, showing that students tended to be inconsistent in responding to these questions.

Table 7A-2
Cross-tabulation of Scores on Item 8 (d) and the NPS Scale

NPS (Q.7) / Q. 8 (d)	0	1	2	Row Totals
0	147	14	45	206
1	165	32	85	282
Column Totals	312	46	130	488

Note. The NPS Scale tallied Numbers View responses on Question 7, the Professors-and-Students problem.

Table 7A-2 records that, of the 282 students who favoured the Numbers View in Item 8 (d), 165 rejected that view in Question 7 (as shown by their score of "0" on the NPS Scale). The correlation ($r = .119$) between the two measures was, however, statistically significant ($p = .004$), mainly because most of those who scored "1" or "2" on the NPS Scale had scored "1" on the Item 8 (d), and that 147 rejected the Numbers View in both Question 7 and Item 8 (d).

Table 7A·3

Cross-tabulation of Scores on Items 8 (d) and 6 (a) (v)

Q. 6 (a) (v) / Q. 8 (d)	0	1	Row Totals
0	108	93	201
1	153	113	266
Column Totals	261	206	467

Note. "1" was scored on Item 6 (a) (v) for choosing "the number of apples in a box" as a possible meaning for '*c*'.

Responses to these two items were not significantly correlated, as could be expected from the distribution of frequencies in Table 7A·3: Those who scored "1" on Item 6 (a) (v) were spread almost evenly across scores of "1" and "0" on item 8 (d), showing the inconsistency of student views.

Table 7A·4

Cross-tabulation of Scores on Items 8 (d) and 14 (i)

Q. 14 (i) / Q. 8 (d)	0	1	Row Totals
0	150	52	202
1	211	64	275
Column Totals	361	116	477

Note. "1" was scored on Item 14 (i) for choosing a Numbers View option, either (b) or (e).

The frequencies in Table 7A·4 record inconsistencies in students' views of symbols in the two items. The correlation was not statistically significant.

Manipulatives and Textbook Classes

The research project was not organized along the lines of a classical experiment (e.g., Randomized Design With Two or More Treatments, as in Kirk, 1982) to compare and contrast two different teaching approaches. However, it did provide some interesting reflections about the effects that different classroom approaches might have on the rates of development of different aspects of learning to understand algebraic symbols, which might form the basis in the future for a fully-controlled research project based on intervention in classroom practices. As the data were available, a comparison between the groups of the rates of development of the understanding of algebraic symbols was undertaken to see if trends could be discerned. The three classes in the Textbook Group were matched with three of the Manipulatives classes according to the schools' ratings of classes on mathematics ability. The matching is summarized in Table 7B·1, where it can be noted that each group was composed of one Mixed Ability class of girls, one Mixed Ability class of boys and one Advanced class of girls, selected in pairs from three schools.

Table 7B·1
Matched Manipulatives and Textbook Groups

CLASS (Code)	SCHOOL	Average AGE* in years	TYPE**	No.	ABILITY rating	GENDER
1	A	12.79	Manip.	8	mixed	GIRLS
2	A	12.43	Text.	15	mixed	GIRLS
4	B	12.50	Text.	21	mixed	BOYS
5	B	12.56	Manip.	23	mixed	BOYS
6	C	12.40	Manip.	28	advanced	GIRLS
7	C	12.43	Text.	29	advanced	GIRLS

Note. * AGE when given their first test.
** TYPE: Manip. = Manipulatives, Text. = Textbook

Differential Rates of Development for Manipulatives and Textbook Groups

Several statistically significant differences were identified between the rates of development of the understanding of algebraic symbols when *t*-tests were carried out between the mean scale scores for the matched Manipulatives and Textbook groups. These are summarized in Tables 7B·2 and 7B·3.

Table 7B·2

Significant Differences Between Manipulatives & Textbook Groups on Tests 1 & 2

Test	Scale	Comment	Max	Mean Manip.	Mean Text.	*t* value	*df* ✻	*p*	Favours
1	BXBA	$b = x, b = a$	4	0.24	0.60	2.14	79.5	*	Text.
1	AL2	Alphabetic Code Q.14	2	0.18	0.017	2.32	60.3	*	Text.
2	Total	Test Total	65	19.9	16.0	2.55	122	*	Manip.
2	SUBS	Substitute & Solve	7	5.02	2.00	6.32	116	***	Manip.
2	SYM	Symbols in answers	6	0.47	1.08	- 2.68	95	**	Text.
2	Q.8d	No.>object	1	0.60	0.38	2.32	106	*	Manip.
2	CZ	c = zero Q.15 iv	2	0.20	0.05	2.47	79.6	*	Manip.
2	EQL	Equality	9	1.90	1.05	2.17	76	*	Manip.
2	NRPS	No Reversal Prof-Student	2	0.43	0.17	2.06	97.6	*	Text.

Note. Matched groups. Manip. = Manipulatives ($n = 59$); Text. = Textbook ($n = 65$).
✻ *df*: decimal point if using separate variances; otherwise, pooled variance.
Max. = maximum possible score (= number of items for scales).
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

There were five main indications of differences that are possibly related to the teaching approach used. Those introduced to algebra with the aid of concrete representations had developed significantly more success in four general areas:

1. overall competence in algebra (Totals for Tests 2 & 3),
2. substitution and solving-a-simple-equation tasks (SUBS Scale in Tests 2 and 3, and SUB Scale in Test 4),
3. completing tasks which required the concept of a variable (CZ and EQL Scales in Test 3, and BXBA Scale in Test 3), and
4. avoiding the view that algebraic symbols stood for objects (OBJ Scale in Test 3, and Q.8d in Test 2, 3, and 4).

Also, those using a text-book approach were more inclined to write symbols correctly in answers, when appropriate (SYM Scale in Test 2, and SC2 Scale in Tests 3 and 4). This fifth difference favoured the Textbook classes and arose mainly from the pair of classes in School A (Table 7B·4).

Table 7B·3

Significant Differences Between Manipulatives & Textbook Groups on Tests 3 &4

Test	Scale	Comment	Max	Mean Manip.	Mean Text.	t value	df ∗	p	Favours
3	BXBA	$b = x, b = a$	4	1.56	0.78	3.32	97.1	***	Manip.
3	SUBS	Substitute & Solve Qq.3,7B	7	6.41	4.94	4.02	83.2	***	Manip.
3	AD	a, d any values	2	1.49	1.02	2.71	105	**	Manip.
3	SC2	Seek closure Qq.5,9,14	6	2.69	1.56	2.68	101	**	Text.
3	OBJ	Objects view Qq.7,7B	3	1.93	2.49	- 3.22	115	**	Manip.
3	CON	Conjoining Qq.3,9	7	0.54	1.25	- 2.78	93.3	**	Manip.
3	Total	Test total	65	26.05	22.62	2.05	122	*	Manip.
3	Q.10	Compare $t + t$, $t + 4$	4	1.75	1.10	2.19	114	*	Manip.
3	PS	Prof-Student Q.7	2	0.53	0.26	2.01	122	*	Manip.
3	Q.8d	No.>object	1	0.77	0.59	2.11	119	*	Manip.
4	Q.8d	No.>object	1	0.60	0.38	2.32	106	*	Manip.
4	SUB	Substitute Q.3	6	5.47	4.90	2.01	99.7	*	Manip.
4	REP	2 or more replacements Qq.10,12,13	6	0.25	0.06	2.20	57.5	*	Manip.
4	SC2	Seek closure Qq.5,9,14	6	0.85	0.32	2.34	66.6	*	Text.

Note. Matched groups. Manip. = Manipulatives ($n = 59$); Text. = Textbook ($n = 65$).
∗ df: decimal point if using separate variances; otherwise, pooled variance.
Max. = maximum possible score (= number of items for scales).
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

The group using concrete materials seem to have gained an advantage over their counterparts by attaining a greater understanding of algebraic expressions, as was shown in the outcome numbered 2 above by their higher success in the substitution items of Question 3. This advantage may have been derived from exercises, undertaken by manipulative classes only, in which first degree algebraic expressions were modelled and discussed. With each model used, teachers of manipulative classes emphasized that the letter-symbols stood for numbers rather than objects: number of square centimetres of area, or number of objects inside a container, or number of centimetres of length. This heightened the awareness of their students to the distinction between using letters to stand for numbers of objects rather than the objects themselves, and could have contributed to Outcome 4 above.

Individual Schools. To examine more closely the differential rates of growth, t-tests were carried out between the Manipulatives and Textbook Groups in each school separately. The results of these tests are assembled Tables 7B·4 to 7B·6.

Table 7B·4
Significant differences School A Matched Manipulatives and Textbook Groups

Test	Scale	Comment	Max	Mean Manip.	Mean Text.	t value	df ∗	p	Favours
2	SYM	Symbols in answers Qq.5,9,14	6	0.20	1.92	3.71	14.5	**	Text.
2	SUBS	Substitute & Solve	7	5.14	2.54	2.17	18	*	Manip.
2	CZ	c = zero Q.15 iv	2	0.57	0	2.83	6.0	*	Manip.
3	BXBA	$b = x$, $b = a$	4	2.00	0.13	4.00	5.9	**	Manip.
3	SYM	Symbols in answers	6	1.43	3.43	- 2.69	19	*	Text.
3	PS	Prof-Student Q.7	2	0.88	0.20	2.10	21	*	Manip.
3	OBJ	Objects view Qq.7,7B	3	1.75	2.62	- 2.20	19	*	Manip.
4	SUB	Substitute Q.3	6	5.50	4.15	2.41	17.5	*	Manip.

Note. Matched groups. Manip. = Manipulatives (n = 8); Text. = Textbook (n = 15).
∗ df: decimal point if using separate variances; otherwise, pooled variance.
Max. = maximum possible score (= number of items for scales).
*** $p \le .001$, ** $.001 < p \le .010$, * $.010 < p \le .050$.

Only the pair of classes in School A showed a significant difference on the SYM Symbols Scale measure. This result is recorded in Table 7B·4 for both Test 2 and Test 3. The Textbook Group in this case had followed a different textbook from that used in the other schools: The text introduced algebra by exercises that asked students to write answers in algebraic symbols, such as, "Express ... as simply as possible: The sum of p and q " (McLeod, et al., 1988, p. 72). Classroom practice of this form of exercise was probably a contributing factor to the greater success by these students on questions requiring symbolic answers.

That students were capable of writing symbolic expressions to answer specific types of algebra questions did not necessarily indicate that they had a sound grasp of the meaning of algebraic expressions. This issue is treated in the discussion of Proposition 9 in Chapter 8. It was found, in this case, that students from School A who followed the concrete approach to introductory algebra were significantly better, by Test 2, at substituting and solving (SUBS Scale) and better at using the concept of a variable in cases such as realizing that equality could be established for the pairs of algebraic expressions given in Question 15 part (iii) (CZ Scale), by Test 2, and Question 15 parts (i) and (ii) (BXBA Scale), by Test 3.

Table 7B·5

Significant differences School B Matched Manipulative and Textbook Groups

Test	Scale	Comment	Max	Mean Manip.	Mean Text.	*t* value	*df* ✱	*p*	Favours
2	SUBS	Substitute & Solve Qq.3,7B	7	4.67	2.43	2.59	40	*	Manip.
3	SUBS	Substitute & Solve	7	6.47	3.42	4.46	22.3	***	Manip.
3	CON	Conjoining Qq.3,9	7	0.44	1.52	- 2.19	26.5	*	Manip.
4	SUBS	Substitute & Solve	7	6.50	4.95	2.44	21.4	*	Manip.
4	IG	Ignore symbol Qq.3,9	5	0.35	0.05	2.14	24.4	*	Text.

Note. Matched groups. Manip. = Manipulatives ($n = 23$); Text. = Textbook ($n = 21$).
✱ *df*: decimal point if using separate variances; otherwise, pooled variance.
Max. = maximum possible score (= number of items for scales).
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

The main difference to emerge between the two classes in School B was that those taught with the assistance of concrete materials consistently evinced a better grasp of the conventions for first degree expressions, as shown by their skill in substitution. The appropriate statistics are assembled in Table 7B·5. As mentioned earlier, an influence on this outcome could well have been the time spent by manipulative classes on modelling first degree expressions and discussing their meaning.

The two classes from School C were Advanced level, with the Textbook Class rated by the school mathematics department as better at mathematics than the

Manipulatives Class. Significant differences for these classes are listed in Table 7B-6. The manipulatives group advanced faster with the concept of a variable and viewing symbols as representing numerical variables. After six months, however, only one significant difference remained.

Table 7B-6

Significant differences School C Matched Manipulative and Textbook Groups

Test	Scale	Comment	Max	Mean Manip.	Mean Text.	*t* value	*df* ∗	*p*	Favours
1	BXBA	$b = x, b = a$	4	0.18	0.71	- 2.30	41	*	Text.
1	FZN	Fraction,zero Negative	5	0.55	1.36	- 2.28	23	*	Text.
1	JCP	c = cabbage, pear	2	0	0.29	- 2.30	27.0	*	Manip.
2	Total	Test Total	65	20.57	14.38	3.42	55	***	Manip.
2	SUBS	Substitute& Solve Qq.3,7B	7	5.26	1.45	5.81	54	***	Manip.
2	NRPS	No reversal Prof-Student	2	0.96	1.76	- 3.48	43.0	***	Text.
2	AD	a, d any values	2	1.25	0.54	2.71	44	**	Manip.
2	EQL	Equality	9	1.86	0.89	2.45	38	*	Manip.
2	NV	Non-variable Qq.10 to 13	8	0.39	1.53	- 2.21	24.1	*	Manip.
3	OBJ	Objects view Qq.7,7B	3	1.67	2.59	- 3.52	52	***	Manip.
3	BXBA	$b = x, b = a$	4	2.15	0.86	4.33	52	***	Manip.
3	AD	a, d any values	2	1.74	1.00	3.22	43.5	**	Manip.
3	NVall	Non-variable Qq.2, 10 to 13	12	0.87	2.45	- 3.17	31.5	**	Manip.
3	Total	Test Total	65	29.43	24.83	2.38	55	*	Manip.
3	VBL	Variable Qq.6c 10,12,15iii	11	3.81	2.20	2.13	41.3	*	Manip.
4	BXBA	$b = x, b = a$	4	2.05	1.08	2.60	45	**	Manip.

Note. Matched groups. Manip. = Manipulatives ($n = 28$); Text. = Textbook ($n = 29$).
∗ *df*: decimal point if using separate variances; otherwise, pooled variance.
Max. = maximum possible score (= number of items for scales).
*** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

The following extracts are from interviews in 1990 with Year 7 students who had selected the options "an object like a cabbage" and/or "an object like a pear", but gave these selections a number interpretation in the interviews.

Extract 1. (May 1990, after Test 2; student 'D' not too sure but hints at regarding a pear as the number 1)

E D, what did you have in mind when you ticked an object like a pear?
D I don't know. You know how they say ... cabbage, I thought that was just one. Then I just looked at this, an object like a pear, you know how he [the teacher] said like my number is your number, so I just put a pear.
E A piece of fruit?
D Yes, something like that.

Extract 2. (Dec.1990, after Test 4 - student 'L' a bit confused but speaks of "however many" and "a set amount" as at least part of her interpretation of "an object like a pear")

L ... and then '*c*' could be the number of apples in a box, because then you could have ... '*d*' could be another sort of thing, ... and an object like a cabbage because I just felt '*c*' could be an object like a cabbage and then ... I don't know ... I don't know why I picked that one.
E Or an object like a pear?
L Yes, um ... I think it might have been like '*d*' could be an object like a pear and '*c*' might be however many, and there could be a set amount of them.
E Oh yes, not just the object ... You're thinking of the number of them ? ... All right.

Extract 3. (Dec.1990, after Test 4 - student 'M' is not really clear about his view; he mentions '*c*' as "a couple of pears" which implies a number, but vaguely says it could stand for "any objects")

E Um ... one of the things you ticked off here - you said it could be an object like a pear. What were you thinking about there?
M Like '*c*' could stand for a couple of pears, ... or whatever. It could stand for any objects.
E Or a number like 3 or anything?
M Yes

Extract 4. (Dec.1990, after Test 4 - student 'K' is unsure but is inclined towards interpreting '*d*' as some number of cabbages)

K Tell me what I put.
E What you wrote down a couple of days ago was a lot of different numbers and an object like a cabbage or a pear. So could you just tell me what you're thinking about?
K Well I thought it could be 3 or 10, because they're numbers that can go into 10 if you add something else onto them, ... and 7.4. I knew it couldn't be 12. An object like a cabbage is something though - I was thinking like, yes? [Pause] Probably wrong, but I did.
E What would '*d*' be if '*c*' was an object like a cabbage?
K It's probably wrong then.
E No. I'm just asking what you're thinking about.
K Another cabbage ... a number of cabbages or something.

Table 7D-1

Frequency Distribution of 1991 Interviewees

Category		Yr.12 3U School C	Yr.12 3U School D	Yr.10 Adv. School C	Yr.10 Adv. School D	Yr.10 Adv. School B	TOTAL
Number Tested		25	11	31	25	23	115
Number Interviewed		13	9	14	11	5	52
Chose cabbage & pear	Total (%)	9 (36.0%)	2 (18.2%)	7 (22.6%)	11 (44.0%)	3 (13.0%)	32 (27.8%)
	Interviewed	8	2	7	8	3	28
Rejected cabbage & pear	Total (%)	16 (64.0%)	9 (81.8%)	24 (77.4%)	14 (56.0%)	20 (87.0%)	83 (72.2%)
	Interviewed	5	7	7	3	2	24

Note. $N = 115$; Percentages are of numbers tested in each class.

Table 7D-2

Frequency Table (Part 1) for Cross-tabulation of Responses to Item 6 (a) and Item 15 by 115 Advanced students, July 1991

TOTALS	Number of parts of Item 15 correctly answered				
	4	3	2	1	0
	Subtotals according to choice				
SUBTOTAL DID choose object	7 (21.9%)	7 (21.9%)	10 (31.3%)	6 (18.8%)	2 (6.3%)
SUBTOTAL did NOT choose object	21 (25.3%)	22 (26.5%)	18 (21.7%)	16 (19.3%)	6 (7.2%)
TOTALS	28 (24.3%)	29 (25.2%)	28 (24.3%)	22 (19.1%)	8 (7.0%)

Table 7D-3

Frequency Table (Part 2) for Cross-tabulation of Responses to Item 6 (a) and Item 15 by 115 Advanced students, July 1991

Item 6a	Number of parts of Item 15 correctly answered				
Responses	4	3	2	1	0
	Year 12 students who chose cabbage and pear				
said $c = 1$	2		1		
said $c \geq 1$		2		1	
said c = object	1				
confused	1	1	1		
not interviewed				1	
	Year 12 students who did not chose cabbage or pear				
said c not a thing *or* said object is not a number	9	2			
said object not plural	1				
not interviewed	6	4	2	1	
SUBTOTALS Yr.12	**20 (55.6%)**	**9 (25.0%)**	**4 (11.1%)**	**3 (8.3%)**	**0**
	Year 10 students who chose cabbage and pear				
said $c = 1$	1	3	3	1	
said $c \geq 1$	2	1	3	1	
said c = object			1		
confused				1	1
not interviewed			1	1	1
	Year 10 students who did not chose cabbage or pear				
said c not a thing *or* said object is not a number	2	2	3		
said object not plural	3	2			
not interviewed		12	13	15	6
SUBTOTALS Yr.10	**8 (10.1%)**	**20 (25.3%)**	**24 (30.4%)**	**19 (24.1%)**	**8 (10.1%)**

Summaries of Average Scores on Scales
for Understanding of Symbols as Variables (Level 5)

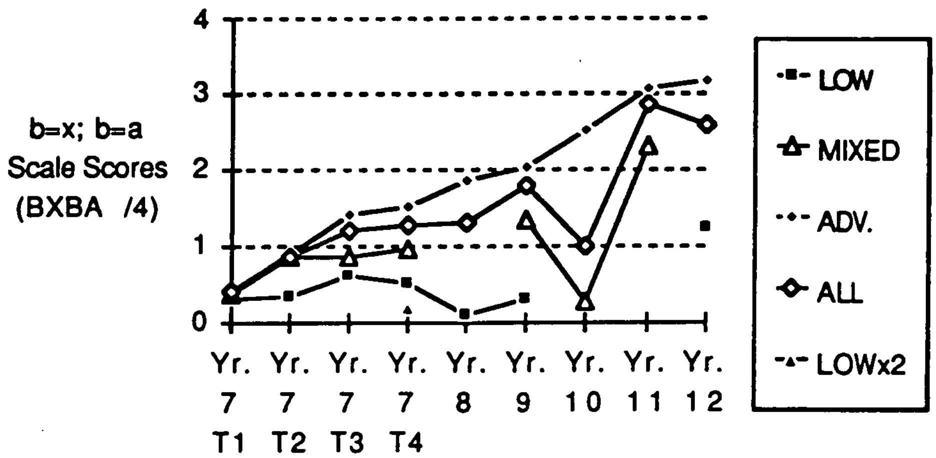


Figure 8A·1. Average scores on BXBA Scale, using responses to Question 15 Parts (i) and (ii)

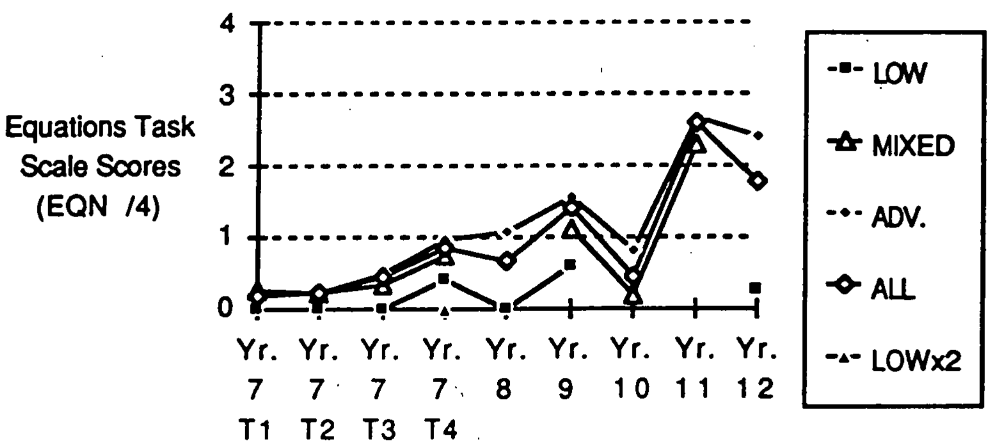


Figure 8A·2. Average scores on EQN Scale, using responses to Question 13

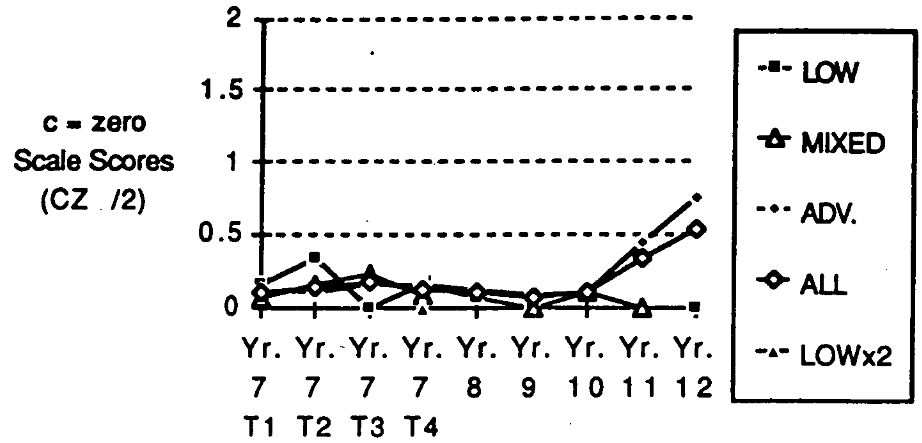


Figure 8A-3. Average scores on CZ Scale, using responses to Question 15 Part (iv)

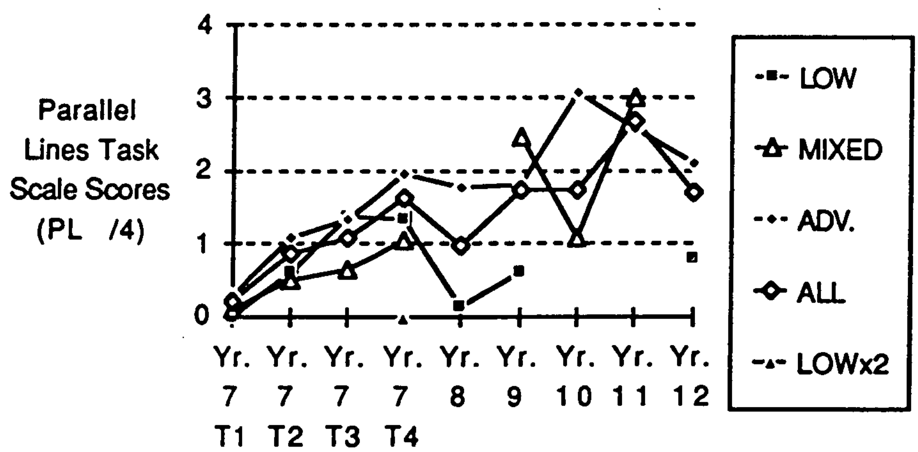


Figure 8A-4. Average scores on PL Scale, using responses to Question 11

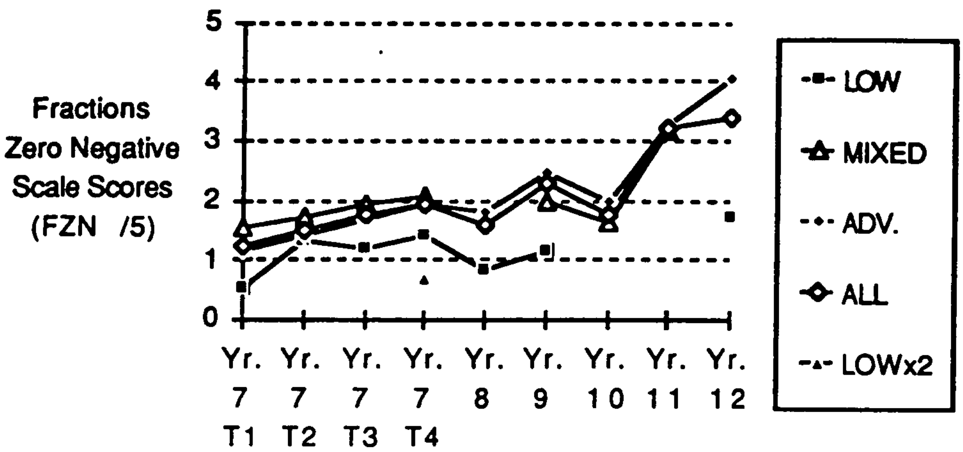


Figure 8B-1. Average scores on FZN Scale, using responses to Questions 6 and 15 Part (iii)

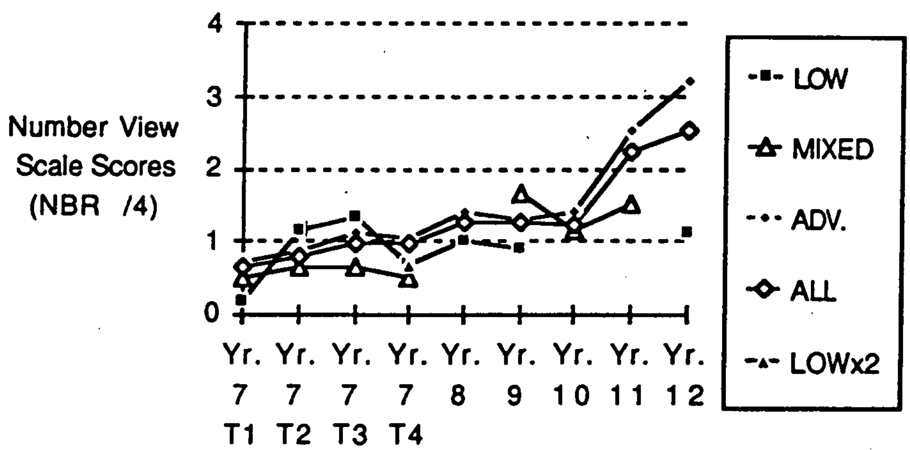


Figure 8B-2. Average scores on NBR Scale, using responses to Questions 6, 7 and 14

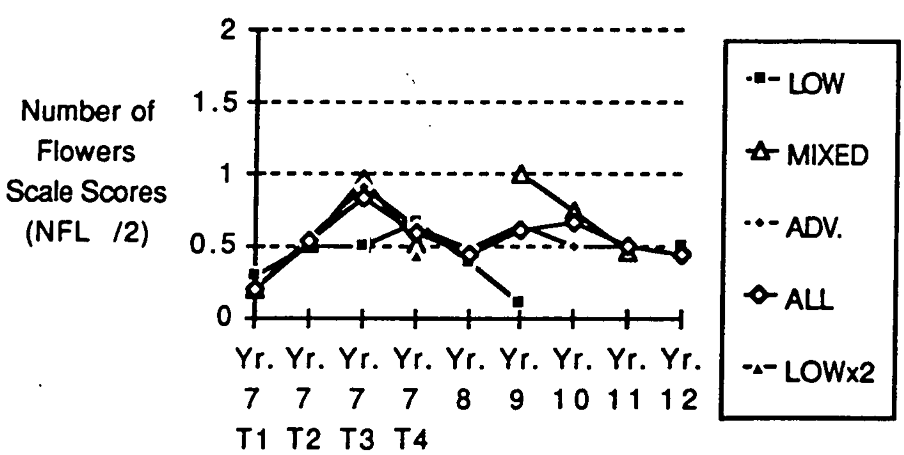


Figure 8B-3. Average scores on NFL Scale, using responses to Question 4

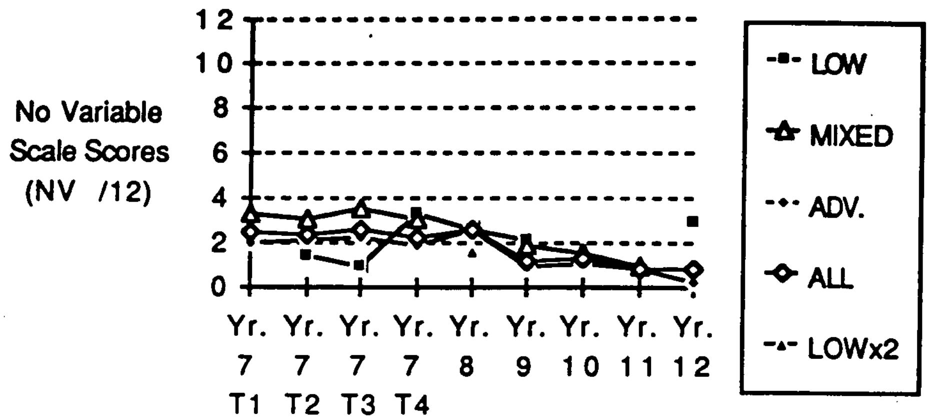


Figure 8C-1. Average scores on the NV Scale, based on Questions 2i, 10b,c, 11b,c,d, 12b,c, and 13b,c,d

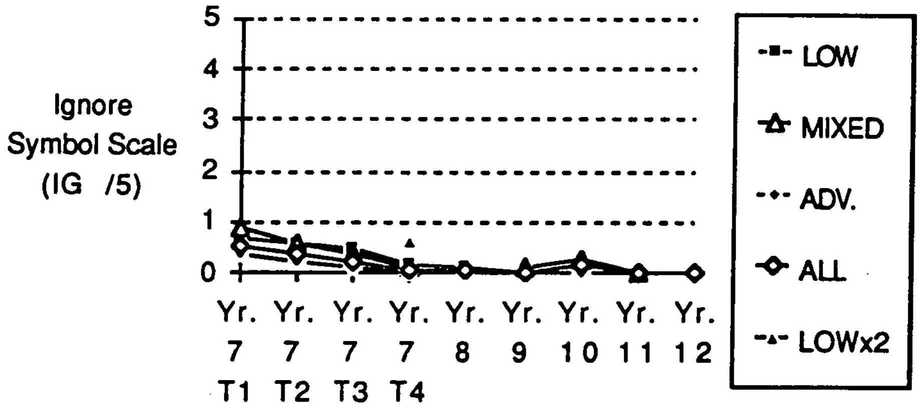


Figure 8C-2. Average scores on the IG Scale, based on Questions 3i,ii,iii, and 9i,ii

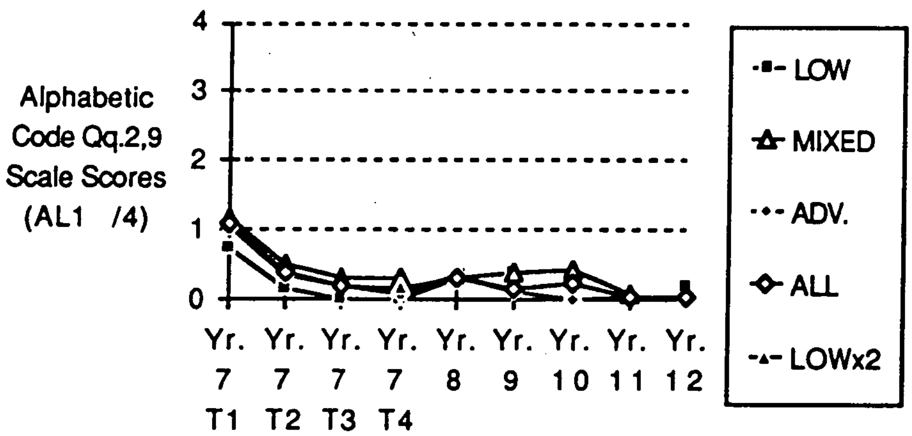


Figure 8C-3. Average scores on the AL1 Scale, based on Questions 2i b, 9

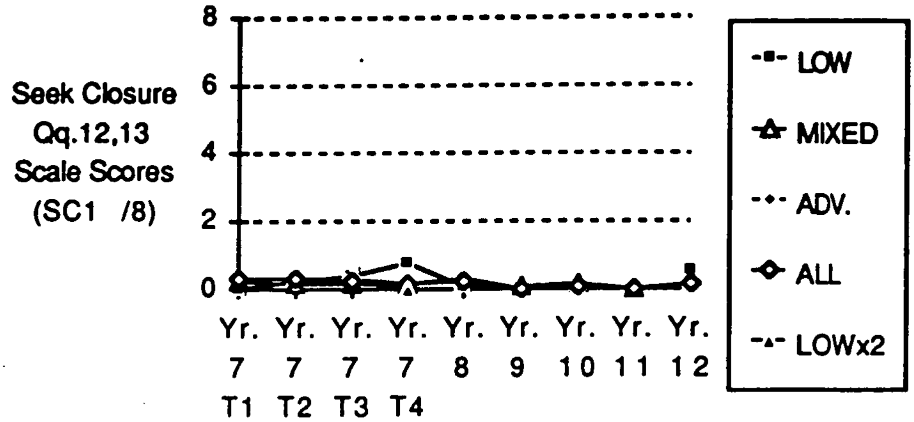


Figure 8C-4. Average scores on the SC1 Scale, based on Questions 12 and 13

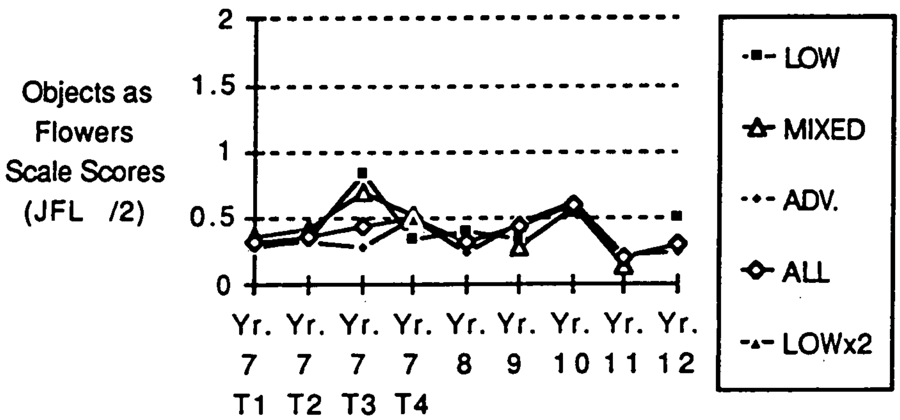


Figure 8C-5. Average scores on the JFL Scale, based on Questions 4b and 4c

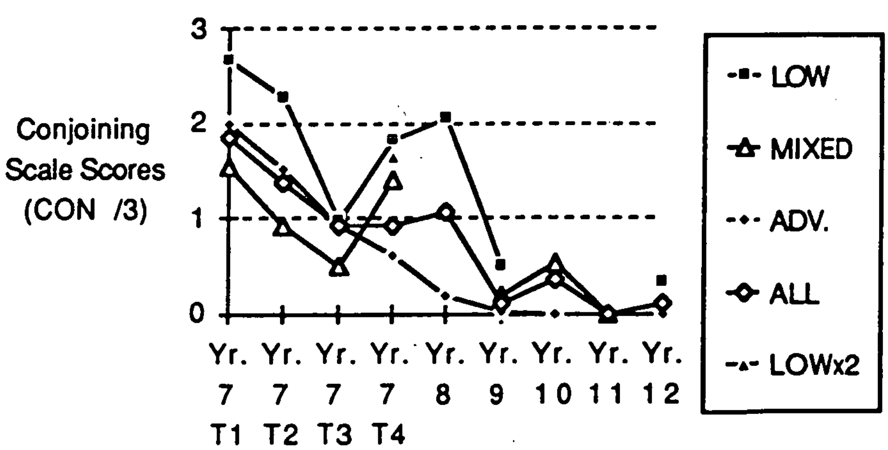


Figure 8C-6. Average scores on the CON Scale, based on Questions 3i,ii, v,vi, and 9

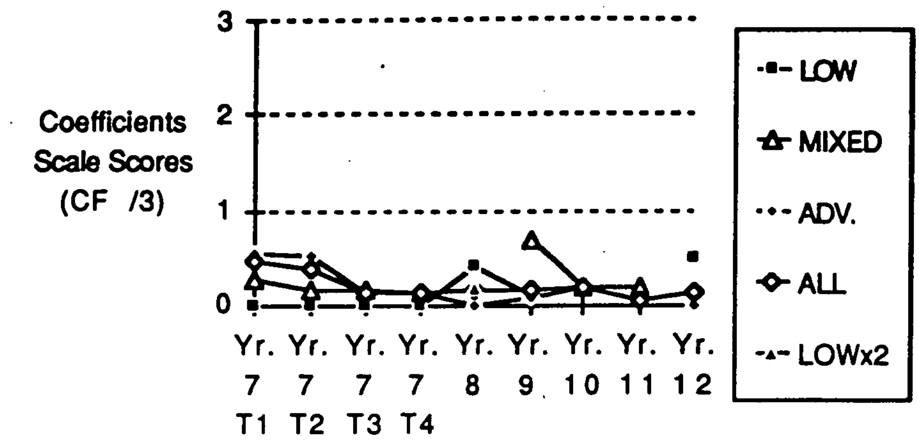


Figure 8C·7. Average scores on the CF Scale, based on Questions 13b,c,d

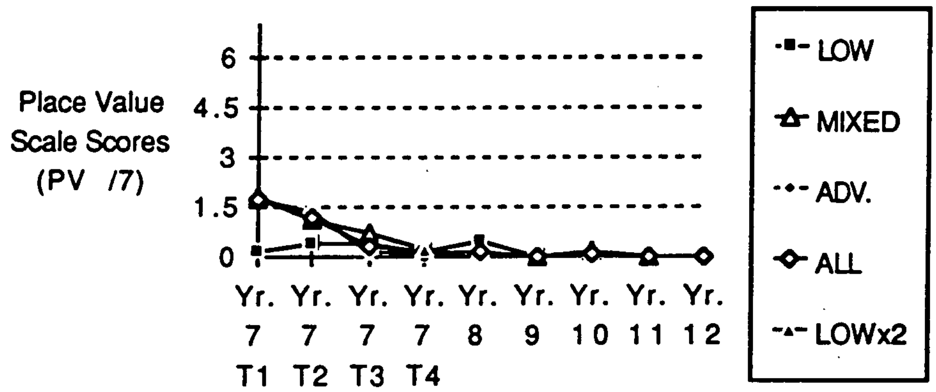


Figure 8C·8. Average scores on the PV Scale, based on Questions 3, 8b

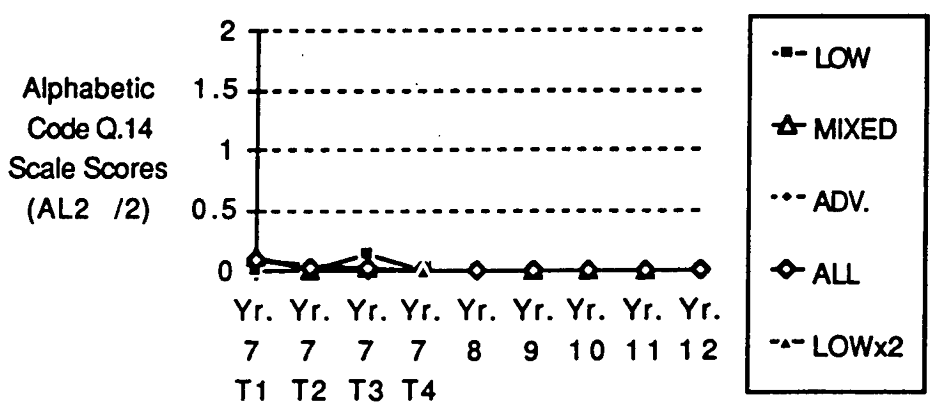


Figure 8C·9. Average scores on the AL2 Scale, based on Questions 14ii, iii

Cross-tabulation Statistics for Hierarchies Within Level 5

Table 8D-1
Hierarchies Within Level 5 for Year 7

Test	Variables		Frequencies			Percentages		Ratio	Correlations	
	A	B	High A & B	High A	High B	%AB	%BA	%AB / %BA	r	p
1	Q.15i	Q.12	1	11	1	100	9.1	11.0	.288	***
1	Q.15iii	Q.10	1	5	1	100	20.0	5.00	.466	***
1	Q.15iii	Q.12	1	5	1	100	20.0	5.00	.610	***
2	Q.15i	EQN	5	41	6	83.3	12.2	6.83	.269	***
2	Q.10	Q.12	5	22	6	83.3	22.7	3.67	.544	***
2	Q.15iii	EQN	5	18	5	100	27.8	3.60	.602	***
2	Q.15iii	Q.12	5	18	6	83.3	27.8	3.00	.708	***
3	Q.15i	CZ	4	59	4	100	6.8	14.8	.197	***
3	Q.15iii	EQN	3	32	3	100	9.4	10.7	.450	***
3	Q.10	Q.12	12	37	14	85.7	32.4	2.65	.602	***
3	Q.15iii	Q.12	11	34	14	78.6	32.4	2.43	.660	***
4	Q.15i	CZ	3	45	4	75.0	6.7	11.3	.136	*
4	PL	EQN	6	56	8	75.0	10.7	7.00	.276	***
4	6c Vbl.	CZ	3	24	4	75.0	12.5	6.00	.281	***
4	Q.10	Q.12	20	58	22	90.9	34.5	2.64	.672	***
4	Q.15iii	Q.12	15	47	21	71.4	31.9	2.24	.588	***
4	Q.15iii	BXBA	17	47	22	77.3	36.2	2.13	.479	***

Note. $N = 208$, less missing values.

$\%AB = \frac{\text{No.high on A \& B}}{\text{No.high on B}} \times \frac{100}{1}$; $\%BA = \frac{\text{No.high on A \& B}}{\text{No.high on A}} \times \frac{100}{1}$;

ordered by ratio %AB to %BA for each test. *** $p \leq .001$, * $.010 < p \leq .050$.

Table 8D-2

Hierarchies Within Level 5 for all Advanced classes

Variables		Frequencies			Percentages		Ratio	Correlations	
A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%AB}{\%BA}$	r	p
Q.15iii	CZ	19	152	19	100	12.5	8.00	.164	**
BXBA	CZ	16	88	19	84.2	18.2	4.63	.250	***
Q.15i	EQN	43	165	53	81.1	26.1	3.11	.298	***
Q.15i	6c Vbl.	40	165	57	70.2	24.2	2.90	.139	**
Q.15iii	EQN	48	151	54	88.9	30.8	2.80	.429	***
Q.15iii	6c Vbl.	47	151	57	82.5	31.1	2.65	.303	***
Q.10	6c Vbl.	41	143	58	70.7	28.7	2.47	.255	***

Note. $N = 316$, less missing values. Ordered by size of ratio %AB to %BA.
*** $p \leq .001$, ** $.001 < p \leq .010$.

Confidence Intervals

When the three criteria are applied to test hierarchies of cognitive difficulty, confidence intervals are not required if the sample is deemed to be the entire population since the sample values are automatically the true population values. However, if the students used in the study are considered as a sample from a larger conceptual population, then the 95% confidence interval for the expected proportion, P, in the population may be calculated from the formula $P \pm 1.96\sqrt{\frac{P(1\text{-}P)}{N}}$, where N is the number of cases (McPherson, 1990, Table 9.2.1).

Taking the example in Table 8·1:

For Crierion 1, the proportion $P = \frac{111}{123} = 0.9024$ (= 90.24% = %AB),

$$1.96\sqrt{\frac{P(1\text{-}P)}{N}} = 1.96\sqrt{\frac{0.902 \times 0.098}{454}} = 0.0273 \ (= 2.73\%).$$

Therefore, the 95% confidence interval is (0.8751, 0.9297) for %AB, or, in percentage form, (87.51, 92.97), which is entirely better than 70%, as required.

For Crierion 2, the proportion $P = \frac{\%BA}{\%AB} = \frac{123}{316} = 0.389$,

$$1.96\sqrt{\frac{P(1\text{-}P)}{N}} = 1.96\sqrt{\frac{0.389 \times 0.611}{454}} = 0.0449.$$

The 95% confidence interval for $\frac{\%BA}{\%AB}$ is (0.344, 0.434), which does not go above 0.5, so that we can be 95% confident that $\frac{\%AB}{\%BA}$ is greater than 2, as required.

For criterion 3, the standard estimate of significance level is applied to the correlation, giving $p \leq .001$, as shown.

The outcomes reported from the application of Criteria 1 and 2 can generally be applied beyond the sample made up of the participating students unless the proportions used are numerically close to the borderline values (of 70% or 2, respectively).

Table 8E-1

Percentage Frequencies for Success on VBL Scale Alone & With EQN Scale

Scale Success (%)	VBL Scale only	VBL Scale & EQN Scale
100%	22.2	11.8
≥ 80%	68.9	29.9
≥ 50%	89.6	59.0

Note. Advanced Classes from Years 9 to 12 (n = 144)

Table 8E-2

% Frequencies for Success on VBL Scale Alone & With BXBA Scale

Scale Success (%)	VBL Scale only	VBL Scale & BXBA Scale
100%	22.7	17.0
≥ 80%	63.1	34.8
≥ 50%	87.9	71.6

Note. Advanced Classes from Years 9 to 12 (n = 141)

Table 8E-3

Percentage Frequencies for Success on VBL Scale Alone & With SYM Scale

Scale Success (%)	VBL Scale only	VBL Scale & SYM Scale
100%	22.1	20.0
≥ 80%	63.4	58.6
≥ 50%	89.0	89.0

Note. Advanced Classes from Years 9 to 12 (n = 145)

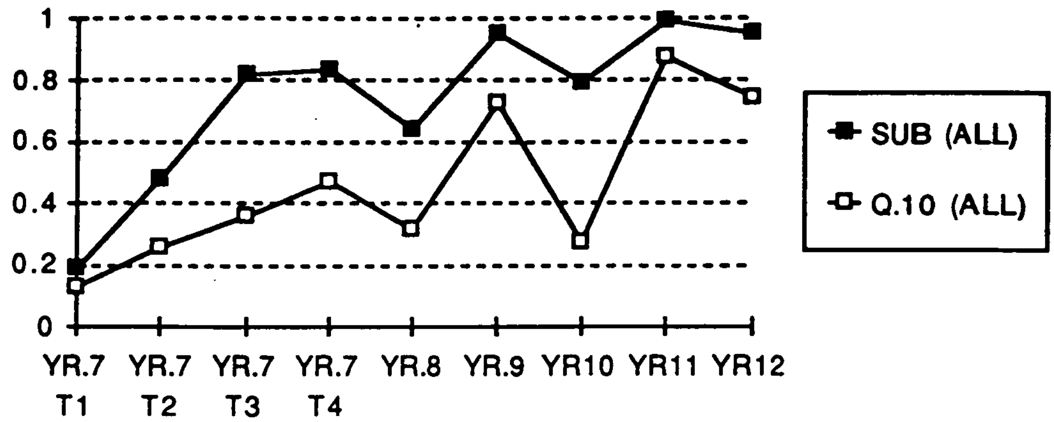


Figure 8F-1. Prop. 2: Av. scores/item by all classes for SUB Scale and Question 10

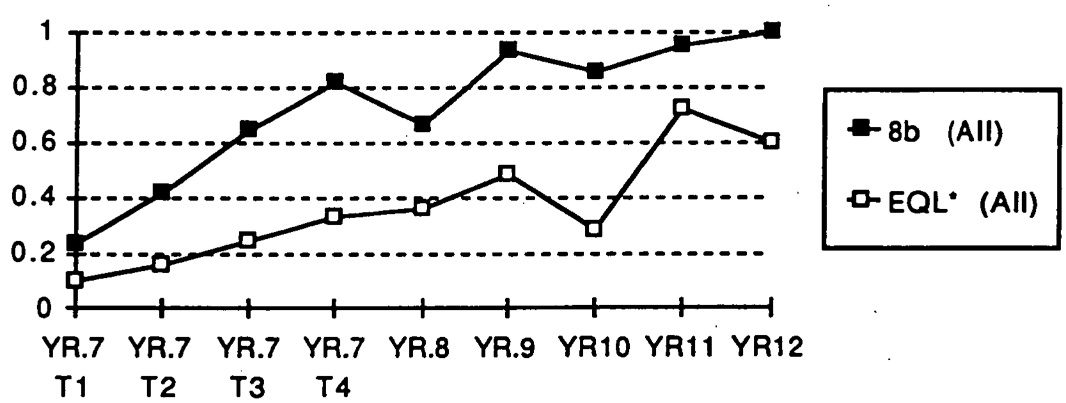


Figure 8F-2. Prop. 3: Av. scores/item by all classes on Q. 8 (b) and EQL* Scale

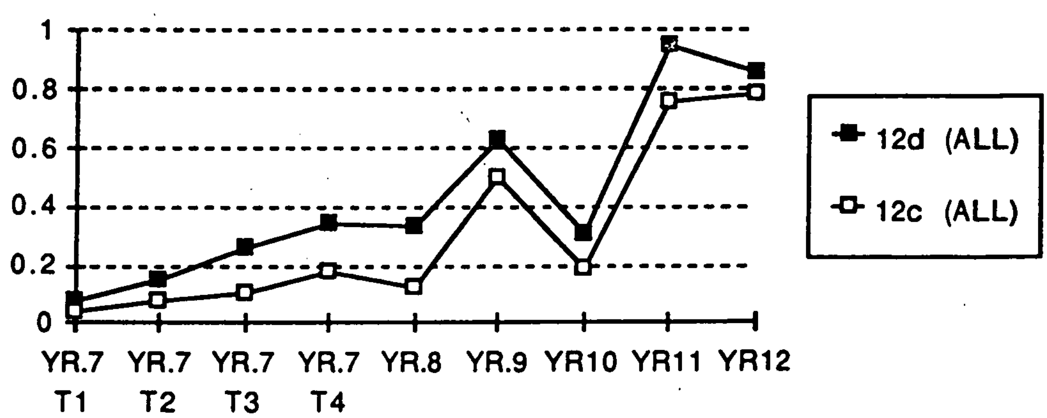


Figure 8F-3. Prop. 4: Average scores by all classes for Questions 12 (d) & 12 (c)

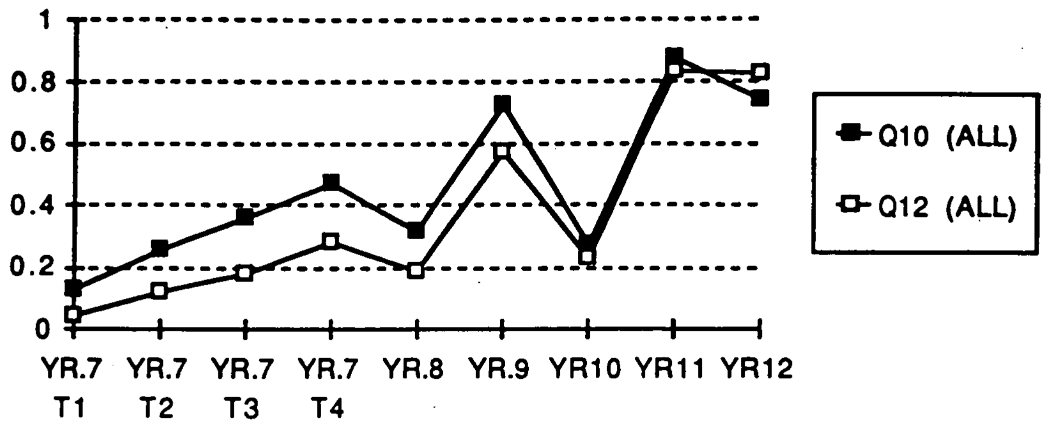


Figure 8F-4. Prop. 5: Average scores/item by all classes for Questions 10 and 12

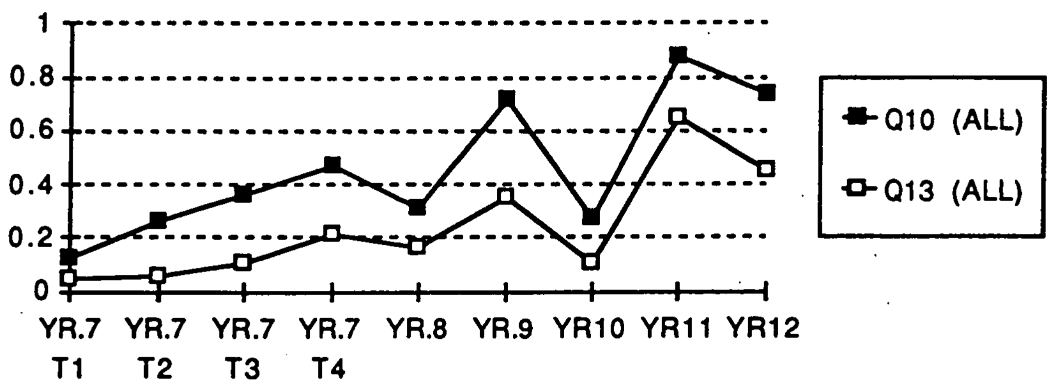


Figure 8F-5. Prop. 6: Average scores/item by all classes for Questions 10 and 13

Table 8F·1
Evidence for Proposition 6

Variables		Frequencies			Percentages		Ratio	Correlations	
A	B	High A & B	High A	High B	%AB	%BA	$\frac{\%AB}{\%BA}$	*r*	*p*
10b	13b	60	217	71	84.5	27.7	3.06	.305	***
10c	13b	64	210	71	90.1	30.5	2.96	.370	***
10b	13d	62	217	75	82.7	28.6	2.89	.299	***
10c	13d	65	210	75	86.7	31.0	2.80	.350	***
12b	13b	63	168	71	88.7	37.5	2.37	.455	***
12b	13d	65	168	75	86.7	38.7	2.24	.451	***
12c	13b	59	152	71	83.1	38.8	2.14	.446	***
12c	13d	64	152	75	85.3	42.1	2.03	.483	***

Note. Yrs.7 to 12. $N = 517$. 6c = Response to Q.6 (c) at variable level. Ordered by size of ratio %AB to %BA. *** $p \leq .001$.

Table 8G·1

Correlations between SYM Scale Scores and Level 5 Score for Yr.7

SYM by	Test 1	Test 2	Test 3	Test 4
EQL	.331 (48) *	.470 (115) ***	.420 (140) ***	.508 (146) ***
GNV	.288 (22) *	.392 (101) ***	.388 (128) ***	.519 (135) ***
VBL	.278 (52) *	.305 (129) ***	.182 (152) *	.502 (160) ***
EQN	.272 (64) *	.430 (138) ***	.335 (158) ***	.323 (161) ***

Note. Frequencies in brackets. *** $p \leq .001$, * $.010 < p \leq .050$.

Acceptance of Lack of Closure

Proposition 9 was expressed in terms of correctly using algebraic symbols in answers, where appropriate. Table 8G·2 displays the cross-tabulation data for using symbols in answers, regardless of whether or not the usage was accurate. The ALC Acceptance of Lack of Closure Scale kept a tally of the number of times students wrote symbols in answers to Questions 5, 9, and the last two parts of Question 14. For instance, in Question 5, they were allocated "1" if they correctly wrote '$p + r$' or an other algebraic expression, such as 'pr'. The analyses show that it was less cognitively demanding for students to write symbols in answers than to attain a correct understanding of symbols at Levels 5 (variables) or 4 (generalized numbers). Hence, success on the ALC Scale was a prelude to success on the other scales listed in the table.

Table 8G-2

Evidence Related to Proposition 9

Source	Variables		Frequencies			Percentages		Ratio	Correlations	
(ALL)	A	B	High A & B	High A	High B	%A B	%B A	\F(% AB, %B A)	*r*	*p*
ALL	ALC	GN	5	259	5	100	1.93	51.8	.136	***
ALL	ALC	FZN	10	258	10	100	3.88	25.8	.292	***
ALL	ALC	INT	13	268	14	92.9	4.85	19.1	.158	***
ALL	ALC	NBR	25	276	27	92.6	9.06	10.2	.224	***
ALL	ALC	VBL	35	261	37	94.6	13.4	7.05	.377	***
ALL	ALC	GNV	32	218	32	100	14.7	6.81	.421	***
ALL	ALC	EQN	57	271	60	95.0	21.0	4.52	.336	***
ALL	ALC	BXBA	85	275	93	91.4	30.9	2.96	.262	***
ALL	ALC	PL	109	281	126	86.5	38.8	2.23	.305	***

Note. ALL = Data from Years 7 to 12 students, using Test 3 results for Year 7 students.

Ordered by ratios %AB to %BA. *** $p \leq .001$.

Chi-square Tests for Differences Between Classes From School D

Table 8G-3

Cross-tabulation for Classes 8 & 9 on Q.5 in Test 2

Class	Q.5 correct	Q. 5 incorrect	Row Totals
8	9	16	25
9	23	2	25
Column Totals	32	18	50

Note. n =25 for each class. $\chi^2 = 14.7$, $df = 1$, $p \leq .001$.

Table 8G·4

Cross-tabulation for Classes 8 & 9 on Q.5 in Test 3

Class	Q.5 correct	Q. 5 incorrect	Row Totals
8	13	12	25
9	22	3	25
Column Totals	35	15	50

Note. n =25 for each class. $\chi^2 = 6.10$, $df = 1$, $p \leq .050$.

Table 8G·5

Cross-tabulation for Classes 10 & 9 on Q.5 in Test 2

Class	Q.5 correct	Q. 5 incorrect	Row Totals
10	4	22	26
9	23	2	25
Column Totals	27	24	51

Note. n =25 for Class 9 and n =26 for Class 10. $\chi^2 = 26.5$, $df = 1$, $p \leq .001$.

Table 8G·6

Cross-tabulation for Classes 8 & 10 on Q.9 (i) in Test 2

Class	Q.9i correct	Q. 9i incorrect	Row Totals
8	14	11	25
10	3	23	26
Column Totals	17	34	51

Note. n =25 for Class 8 and n =26 for Class 10. $\chi^2 = 9.24$, $df = 1$, $p \leq .010$.

Table 8H·1

Distribution of Sources for Items Considered When Examining Propositions

Proposition	New 1990	Collis 1975	Harper 1976	Küchemann 1980	Rosnick 1981	Booth 1983	MacGregor 1989
1	2i,3,4, 6a,b,8,14	1,15	6c,10,11, 13	6c,9,12	7	5	2ii
2	3	15	10,13	6c,12			
3	8b	15	13	12			
4			13#	12#			
5			10	12			
6			6c,10,13	6c,12			
7		15					
8	3, 14			9		5	
9	2i,6a,14	1		9		5	2ii
10					7		2ii

Relationships within parts of one item rather than between items.

Details of Scales Based on Responses to Parts (b) & (c) of Questions 10 & 12

Table 9A·1

Q.10 (b), (c) Response Categories Used in Scales

Code	Response Type	VBL*	GN*	GNV*	12REP*	1REP*	NV*	PRE*	INT*
1	Wrong idea or repeats question							✔	
2	False ordering e.g., t's not equal							✔	
3	More than one replacement value		✔	✔	✔				✔
4	Algebra not quite correct e.g., (b) If $t \geq 5$		✔	✔					✔
5	Correct e.g., (b) If $t > 4$	✔		✔					
6	One replacement value				✔	✔	✔		✔
7	"Always" or "Now"						✔	✔	
8	"Never"						✔	✔	
9	Literal comparison e.g. (c) If second t is 4							✔	
0	Omit this part but answered other part(s) of Q.10								

Note. 10 (b) "When is $t + t$ larger?" 10 (c) "When is $t + 4$ larger?"
VBL*: Variable. GN*: Generalized Number (but not Variable Level). GNV*: Generalized Number or Variable. REP*: Replacement Value(s). NV*: No Variable Idea. PRE*: Prestructural. INT*: Integers only.

Table 9A-2

Q.12 (b), (c) Response Categories Used in Scales

Code	Response Type	VBL*	GN*	GNV*	12REP*	1REP*	NV*	PRE*	INT*
1	Wrong idea e.g., $2n$ is a 2-digit number							✔	
2	Repeats question							✔	
3	"Always" or "Now"						✔		
4	"Never"						✔		
5	More than one replacement value		✔	✔	✔				✔
6	Algebra nearly correct e.g. (b) $n \geq 3$		✔	✔					✔
7	Need value of n (no closure)							✔	
8	One replacement value				✔	✔	✔		✔
9	Correct answer e.g. (b) When $n > 2$	✔		✔					
0	Omit this part but answered other part(s) of Q.10								

Note. 12 (b) "When is $2n$ larger?" 12 (c) "When is n +2 larger?"
VBL*: Variable. GN*: Generalized Number (but not Variable Level). GNV*: Generalized Number or Variable. REP*: Replacement Value(s). NV*: No Variable Idea. PRE*: Prestructural. INT*: Integers only.

Table 9A·3

Summary of Analyses of Scales Based on Items 10 b,c & 12 b,c

Scale for 10b,c, 12b,c	Cognitive Focus	No.of cases	*Alpha*	Item-item Correlations		Corrected Item-Scale Correlations	
				Mean	Range	Mean	Range
VBL*	Variable	468	.91	.72	.60 - .89	.80	.77 - .83
GNV*	Generalized No.or Variable	468	.91	.72	.60 - .87	.80	.76 - .82
GN*	Generalized Number	468	.75	.42	.23 - .71	.55	.43 - .60
INT*	Integers only	468	.77	.46	.30 - .71	.57	.53 - .62
12REP*	1 or 2 replacements	177✳	.67	.34	.18 - .52	.46	.39 - .52
1REP*	1 replacement	177✳	.68	.40	.20 - .66	.50	.34 - .61
PRE*	Prestructural	468	.81	.52	.38 - .79	.63	.57 - .68
NV*	Not Variable	468	.66	.34	.20 - .59	.46	.37 - .52

Note. Dichotomous Variables. Missing values rejected.
For Yrs. 7 to 12, except for ✳ - based on Test 2 Year 7 responses.

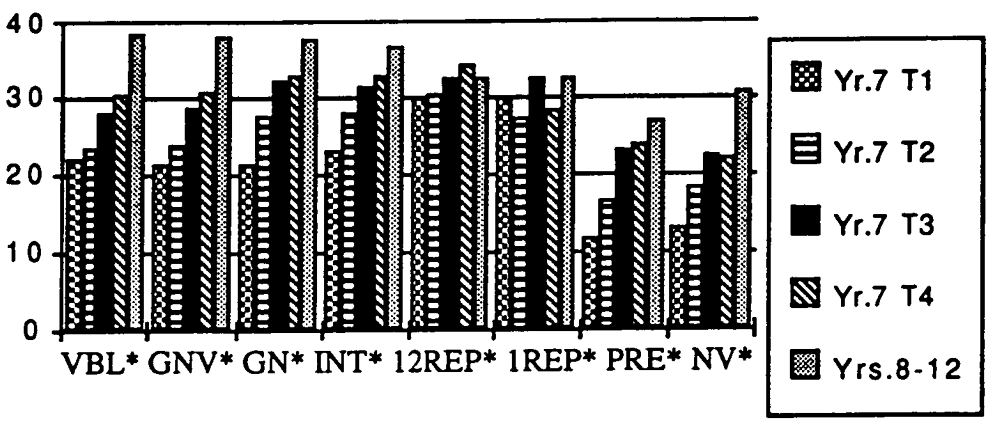


Figure 9A·1. Average test totals corrected for Qq.10 & 12 for Advanced classes according to categories determined by scale scores of "2" or more

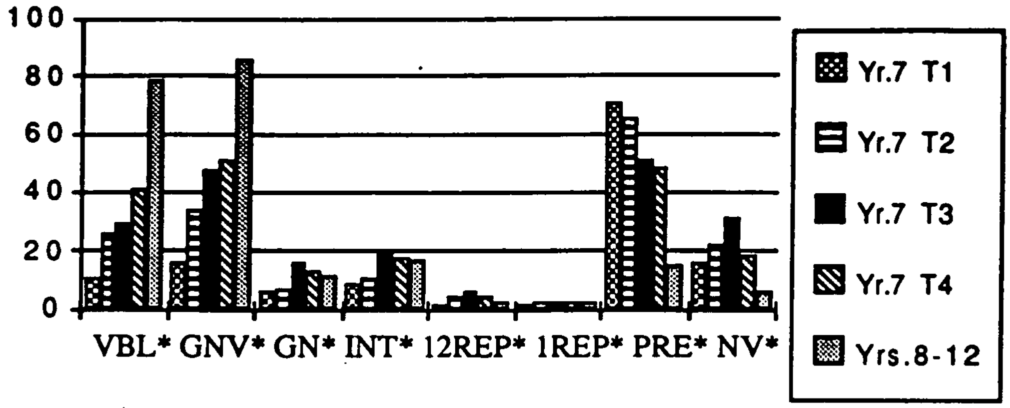


Figure 9A-2. Percentages of valid cases for Advanced classes according to categories determined by scale scores of "2" or more

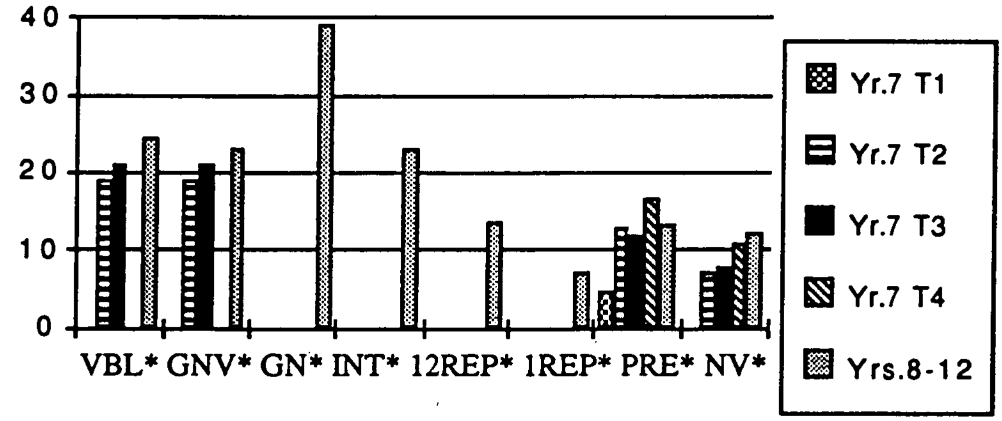


Figure 9A-3. Average test totals corrected for Qq.10 & 12 for Low Ability classes according to categories determined by scale scores of "2" or more

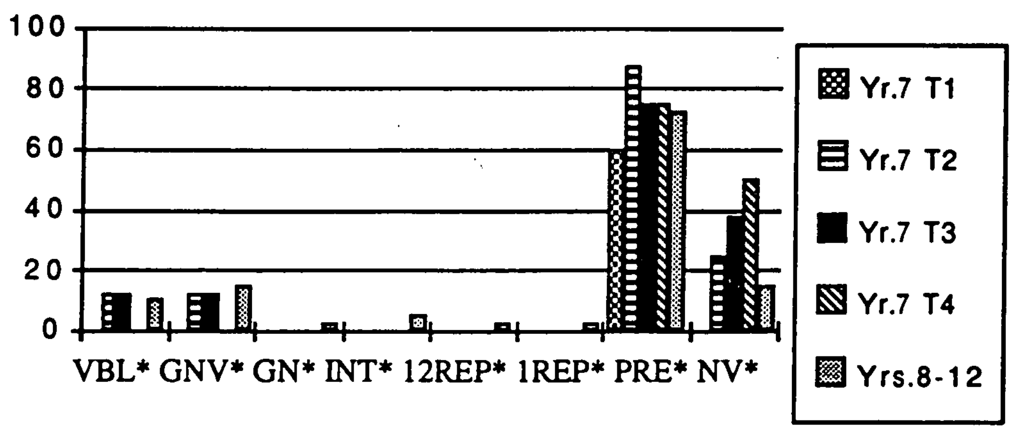


Figure 9A-4. Percentages of valid cases for Low Ability classes according to categories determined by scale scores of "2" or more

Missing data. Those in the SS11 category returned less data than did those in the SS14 category. Differences were not significant where both groups of students omitted answers on early tests but became significant in Tests 3 and 4, as shown in Table 9B·1.

Table 9B·1
Odds Ratio for Missing Data in Terms of Membership of Group SS11 or Group SS14

Test	Scale		Odds Ratio for having missing data $= \frac{\text{odds if in SS11}}{\text{odds if in SS14}}$		Comment	Fisher exact probability	
	Name	Description	Calculation	Conclusion		Value	p
3	VBL	Variable Concept	$\frac{9/8}{0/8}$	complete association	SS14 no missing data	.0119	*
4	VBL	as above	$\frac{6/9}{0/8}$	complete association	as above	.0496	*

* $.010 < p \leq .050$.

When students who had omitted responses were asked in interviews for their reasons for not responding, they gave answers such as:

"I was a bit confused" (Student in Group SS11, regarding Q.3, after Test 3), or

"I didn't understand it" (Student in Group SS14, regarding Q.4, after Test 1).

Membership of Advanced Classes. One could surmise that students in Advanced level classes were more likely to be amongst those who improved the most than those from classes with an average lower ability in mathematics. Table 9B·2 summarizes an odds ratio analysis which confirms this forecast.

Table 9B·2
Odds Ratio for Advanced Students Also Being Members of Group SS14 or SS11

Odds Ratio for coming from Advanced classes $= \frac{\text{odds if in SS14}}{\text{odds if in SS11}}$		Comment	Fisher exact probability	
Calculation	Value		Value	p
$\frac{7/1}{2/15}$	52.5	Advanced students more likely to be in Group SS14	.0005	***

*** $p \leq .001$.

Table 9B-3

Summary of t-tests for Groups SS11 & SS14 on Test 3 Responses

Scale	Comment	Max	Mean SS11	Mean SS14	*t* value	*df* ✶	*p*	Favours
Total	Test Total	65	10.18	39.75	10.68	7.75	***	SS14
EQL	Equality	9	0.25	5.33	9.76	6.10	***	SS14
SUBS	Substitute & Solve	7	1.73	6.38	7.03	18.47	***	SS14
FZN	Fractions, Zero, Negatives	5	1.00	2.86	5.27	18	***	SS14
GNV	Gen.No.and/ or Variable	17	2.20	12.20	5.21	8	***	SS14
NBR	Number view	4	0.46	2.29	4.64	18	***	SS14
PS	Prof-Student	2	0.18	1.50	4.58	23	***	SS14
C2	c=3, c=7.4	2	0.93	2.00	4.37	13.00	***	SS14
VBL	Variable	11	0.25	6.38	3.96	7.50	***	SS14
FL	Q.4 re no.of flowers	3	0.91	2.13	3.87	17	***	SS14
BXBA	$b = x, b = a$	4	0.47	2.71	3.79	7.17	**	SS14
GN	Gen.No.	7	0.33	4.14	3.72	6.33	**	SS14
AR	Arithmetical Processes	4	1.57	3.67	3.66	18	**	SS14
INT	Integers only	6	0	3.43	3.09	6.00	*	SS14
SYM	Symbols in answers	6	1.09	3.20	3.00	14	**	SS14
PL	Parallel Lines	4	0	1.88	2.61	7.00	*	SS14
AD	a, d	2	0.44	1.50	2.60	13	*	SS14
PRE	Prestructural	19	14.2	2.40	- 9.63	8	***	SS14
OBJ	Objects view	3	2.33	0.50	- 3.32	17	**	SS14
NV	Not Variable	8	6.17	0.80	- 3.23	5.53	**	SS14
CON	Conjoining	7	2.13	0.33	- 3.19	17.60	**	SS14
IG	Ignore letter	5	0.40	0	- 2.45	14.00	*	SS14

Note. Sorted in order of t values. Max. = maximum possible score (= no. of items for scales). ✶ *df*: decimal point if using separate variances; otherwise, pooled variance. *** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Table 9B·4

Summary of t-tests for Groups SS11 & SS14 on Test 4 Responses

Scale	Comment	Max	Mean SS11	Mean SS14	t value	df ✶	p	Favours
Total	Test Total	65	9.13	41.38	19.97	21	***	SS14
EQL	Equality	9	0.38	5.88	10.72	9.00	***	SS14
AR	Arithmetical Processes	4	1.23	3.88	7.63	15.27	***	SS14
GNV	Gen.No.and/ or Variable	17	3.00	13.50	6.51	11	**	SS14
SUBS	Substitute & Solve	7	2.69	6.63	6.14	14.04	***	SS14
VBL	Variable	11	0.44	6.50	6.12	15	***	SS14
PL	Parallel Lines	4	0	3.00	4.58	7.00	**	SS14
FZN	Fractions, Zero, Negatives	5	1.20	2.88	4.36	16	***	SS14
SYM	Symbols in answers	7	0.64	4.00	4.29	9.04	**	SS14
EQN	Equation Task	4	0	1.75	4.25	7.00	**	SS14
BXBA	$b=x, b=a$	4	0.33	2.25	3.99	8.77	**	SS14
PS	Prof-Student	2	0.14	1.25	3.68	20	***	SS14
AD	a, d	2	0.80	2.00	3.67	9.00	**	SS14
GN	Gen.No.	7	0.67	2.75	3.57	8.24	**	SS14
C2	c=3, c=7.4	2	1.17	2.00	3.46	11.00	**	SS14
INT	Integers only	6	0	2.50	3.42	7.00	*	SS14
NBR	Number view	4	0.09	1.75	3.32	7.48	*	SS14
12REP	Replacement by 1 or 2 nos.	7	0	1.13	2.83	7.00	*	SS14
1REP	1 replacement	6	0	0.75	2.39	7.00	*	SS14
PRE	Prestructural	19	16.17	2.88	- 8.90	12	***	SS14
OBJ	Objects view	3	2.58	1.00	- 5.22	17	***	SS14
CON	Conjoining	7	2.67	0.50	- 3.70	13.41	**	SS14
AL1	Alphabetic	4	0.44	0	- 2.53	8.00	*	SS14

Note. Sorted in order of t values. Max. = maximum possible score (= no. of items for scales). ✶ df: decimal point if using separate variances; otherwise, pooled variance. *** $p \le .001$, ** $.001 < p \le .010$, * $.010 < p \le .050$.

Table 9B·5

Summary of t-tests for Groups SS22 & SS24 on Tests 2 & 3 Responses

Test	Scale	Comment	Max	Mean SS22	Mean SS24	t value	df ✶	p	Favours
2	GNV	Gen.No.and/ or Variable	17	4.00	10.5	6.45	8	**	SS24
2	VBL	Variable	11	0.50	5.00	4.47	13	***	SS24
2	Total	Test Total	65	20.0	33.8	4.17	15	***	SS24
2	EQL	Equals Scale	9	0.83	3.78	3.67	10.5	**	SS24
2	PL	Parallel Lines	4	0.14	2.56	3.64	8.77	**	SS24
2	RPS	Reversal Prof-Student	2	0	0.80	2.45	9.00	*	SS22
2	PRE	Prestructural View	19	14.8	6.17	- 9.87	8	***	SS24
3	Total	Test Total	65	21.0	37.8	5.02	15	***	SS24
3	SUBS	Substitute & Solve	7	0.38	5.00	4.41	7.22	**	SS24
3	VBL	Variable	11	0.86	5.22	4.00	9.66	**	SS24
3	SYM	Use symbols in answers	6	1.83	4.40	3.70	14	**	SS24
3	EQL	Equality	9	1.17	4.89	3.61	13	**	SS24
3	GNV	Gen.No.and/ or Variable	17	4.25	11.2	3.02	11	88	SS24
3	BXBA	$b = x, b = a$	4	0.29	2.30	2.82	15	*	SS24
3	PL	Parallel lines	4	0.29	2.22	2.55	10.48	*	SS24
3	RPS	Reversal Prof-Student	2	0	0.80	2.45	9.00	*	SS22
3	ALC	Accept lack of closure	11	3.00	5.20	2.22	14	**	SS24
3	PRE	Prestructural View	19	12..4	4.22	- 5.01	12	***	SS24
3	SC2	Seek closure	6	3.00	0.70	- 2.31	14	*	SS24

Note. Sorted in order of t values for each test. Max. = maximum possible score (= no. of items for scales). ✶ df: decimal point if using separate variances; otherwise, pooled variance. *** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Table 9B·6

Summary of t-tests for Groups SS22 & SS24 on Test 4

Scale	Comment	Max	Mean SS22	Mean SS24	t value	df ∗	p	Favours
Total	Test Total	65	23.5	45.5	15.9	14	***	SS24
EQL	Equality	9	2.00	7.30	9.57	14	***	SS24
VBL	Variable	11	1.17	8.89	8.38	13	***	SS24
BXBA	$b = x, b = a$	4	0.50	3.60	7.14	14	***	SS24
GNV	Gen.No.and/ or Variable	17	5.33	14.67	7.08	13	***	SS24
RPS	Reversal Prof-Student	2	0	1.20	5.01	3.67	**	SS22
SYM	Use symbols in answers	6	2.6	5.00	4.05	13	***	SS24
EQN	Equations Task	4	0.50	2.50	3.62	14	**	SS24
GN	Generalized Number	7	1.00	2.30	2.90	9.00	*	SS24
PRE	Prestructural View	19	11.67	0.89	- 5.57	6.31	***	SS24

Note. Sorted in order of t values. Max. = maximum possible score (= no. of items for scales). ∗ df: decimal point if using separate variances; otherwise, pooled variance. *** $p \leq .001$, ** $.001 < p \leq .010$, * $.010 < p \leq .050$.

Table 9B·7

Comparison of SS11 v SS14 and SS22 v SS24 Outcomes

Scales	SS11 v SS14			SS22 v SS24			Test items used
Test	T2	T3	T4	T2	T3	T4	
1REP			✔				10,12,13
12REP			✔				7,10,12,13
AD	✔	✔	✔				2
AL1			✔				2,9
ALC					✔		5,9,14
AR		✔	✔				1
BXBA		✔	✔		✔	✔	15
C2		✔	✔				6
CON	✔	✔	✔				3,9
EQL	✔	✔	✔	✔	✔	✔	8,10,11,12,13,15
EQN			✔			✔	13
FL		✔					4
FZN		✔	✔				6,15
GN		✔	✔			✔	6,10,12,13
GNV	✔	✔	✔	✔	✔	✔	2,6,10,11,12,13
IG		✔					3,9
INT		✔	✔				10,12,13
NBR	✔	✔	✔				6,7,14
NV		✔	✔				2,10,11,12,13
OBJ		✔	✔				8,7
PL		✔	✔	✔	✔		11
PRE	✔	✔	✔	✔	✔	✔	2,6,10,11,12,13
PS		✔	✔				7
RPS				✔	✔	✔	7
SC2					✔		5,9,14
SUBS		✔	✔		✔		3,8
SYM		✔	✔		✔	✔	5,9,14
VBL	✔	✔	✔	✔	✔	✔	6,10,12,15

✔ denotes a statistically significant difference recorded in earlier tables.