The evaluation of an intelligent learning system as a learning tool for the solution of numerical problems

Teresa Farran

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy

April 2008

Degree awarded by Liverpool John Moores University

Abstract

It has been widely acknowledged for a long time that acquisition of mathematical understanding and skills is an important attribute for life. In the context of interest in raising academic standards nationally, both employers and post compulsory and higher education establishments have shown concern over the level of mathematical ability of school leavers, even those who have achieved a pass grade at GCSE level. Furthermore some organisations and teaching profession bodies such as the Teacher Training Agency have shown enough concern to introduce their own assessment tools to judge mathematical ability. Following the introduction of the national grid for learning and various national schemes for the development of the use of computer systems and information technology to support learning it has been broadly recognised that there is an immense potential for "e-learning" to impact upon the understanding and skills of learners. The study focused on the performance including errors made and GCSE grade previously attained of two sample groups of learners, in the pilot study undergraduates during their induction into higher education and in the final trial two matched groups of learners who had progressed from school to a sixth form college. In addition the opinion of practicing teachers regarding common errors and misconceptions was investigated. Comparison of the performance of the learners and the opinion of the teachers with the national findings of the Key Stage 3 Mathematics SATS reports and the views of a range of researchers enabled triangulation of findings. Furthermore in the final trial through the use of matched pair groups the effectiveness of a computer system with feedback focusing on the specific nature of the error made is compared with a computer system without feedback. The result of this enquiry indicates that:

1. In some instances there is an extension of areas of common misconception from compulsory education to subsequent levels.
2. GCSE grade is a reliable measure but the level it is at may not be considered satisfactory as a pass grade maybe achieved without understanding of some basic concepts.
3. In addition it was suggested that there was a relationship between the ability to understand brackets and negatives signs and values and the GCSE Mathematics grade achieved.
4. The computer systems used showed progression in results for the sample group, with significantly greater improvement related to the system with feedback providing guidance relating to the error made. Those learners who were empowered by directive feedback which could guide to enable the users to amend erroneous responses were more willing to preserve and subsequently improve their ability to answer questions in the trial and post test.

Preface

In the 1990s as a tutor of post-16 students undertaking re-sit GCSE Mathematics I recognised that many common errors and misconceptions were of a fundamental nature. In fact these misconceptions related to the Key Stage 2 (junior level) Mathematics programme of study. Furthermore it seem that as a consequence of these problems in learning confidence, motivation and self-concept of the learner as well as the understanding of subsequent learning that called upon these concepts was adversely affected.

Years later when I was a university lecturer involved in key skill development and in particular Numeracy the same errors and misconceptions were identified amongst some non-mathematical undergraduates. From this tutoring experience it was clear to me that these learners required support which was individualised being based upon the nature of the error and that the learning be presented by a different approach. Within the education sector the requirement to be economically feasible predominates. As a specialist within ICT Education I started to consider how ICT could effectively support this learning and develop long term deep understanding for a range of undergraduates from various backgrounds and prior experiences. This initiated the focus of this study.

Chapter 1 Background

1.1 Context

Numeracy and mathematics are areas of low achievement for many school leavers and the Dearing Report highlighted the need for higher education programmes to enhance students' basic skills including numeracy. Similarly the concern raised by the London Mathematical society (1995) reported that "There is unprecedented concern amongst mathematicians, scientists and engineers in higher education about the mathematical preparedness of new undergraduates." In addition Moser(2002) indicated that the percentage of adults having some numeracy problems ranges from 30% to 50%.

- one in three adults in this country cannot calculate the area of a room that is 21 by 14 feet, even with the aid of a calculator;
- one in four adults cannot calculate the change they should get out of £2 when they buy goods.

Furthermore according to the DfES(2002) GCSE Mathematics results for 2002 as shown in table 1.1 of the 591.3 who attempted GCSE only 312.6 (52%) achieved a grade A-C.

A*	A	B	C	D	E	F	G	U	X	Entries
22.9	55.4	108.7	125.6	96.1	86.9	52.5	24.0	15.7	3.1	591.3

Table 1.1:GCSE attempts and achievements 2002 (Figures in thousands)

Similarly a concern relating to the poor mathematical ability of school children was identified by Cockroft(1992) yet no significant improvement was reported over 20 years later on by the Smith Inquiry (2004).

The Government 's White Paper, *The Future of Higher Education* and their response to the *Dearing Inquiry HE for the 21st century* included a clear commitment to widen participation in higher education. Ramsden (2000) suggests that "today's lecturers are expected to deal with an unprecedented broad spectrum of student ability and background" and that, "courses and teaching methods must be amended to deal with classes that are now not only larger, but also more mixed in their attainment". However the London Mathematical Society (1995) suggest, "it is not just the case that some students are less well-prepared, but that many 'high-attaining' students are seriously lacking in fundamental notions of the subject."

Employers require workers who are numerate and literate, and complain about the inadequacies of graduates.

> "As a kid I struggled with both numeracy and literacy. Even today many of my colleagues joke that I find it hard to read company accounts and grapple with the difference between gross and net... The more confidence we can give young people to get the skills to survive in the 21st Century, the better British companies will do around the world, which in turn, will result in a better quality of life for all of us."
>
> Sir Richard Branson, Chairman of the Virgin Group of Companies

This concern is reflected in the Teacher Training Agency (TTA) extending requirements for all trainee teachers regardless of curriculum specialism to include the ability and competence to pass specific Qualified Teacher Status (QTS) skills tests including numeracy.

Consequently individual and institutional issues remain. For some, following years with limited success at school there are inevitably problems of anxiety and hence motivation among prospective learners. Lalonde et al (1993) suggest "this anxiety most likely stems from the individual's history of performance and affective reactions in learning mathematics, and is present when an individual enters a university program." Programmes of study may not be accessible enough to encourage potential learners to take part or generally seem inviting.

The GCSE qualification does not provide the institution with sufficient information for designing self-study individualised materials that are useful to the majority of students, focused and able to provide adequate feedback without incurring enormous economic and environmental resource costs. The problem is compounded by the variation in prior learning. The London Mathematical Society (1995) report that "since 1994 students can routinely obtain a Grade B taking only the intermediate tier GCSE papers, which assess students on a reduced syllabus (requiring, for example, very little algebra)." Furthermore GCSE is only one of a variety of Key Stage 4 qualifications. Qualifications such as Basic skills or Key skills Numeracy do not include algebra. The importance of this aspect of mathematics was explained by Kramarski & Hirsch (2003) in that "both arithmetic and algebra are useful for describing important relationships in the world" but they also report that various researchers have shown that learning algebra is complex.

Many educationalists including Papert (1993) have emphasised the need for software, which can enable learners to develop personal knowledge and understanding that is meaningful and transferable. So that understanding of a concept in a specific context can then be applied more generally. Individual tutoring in Mathematics has been shown to produce marked improvements such as that reported by Bloom (1984). However financially this would not be feasible. Recent research referred to by Manoucherhri(1999) provides strong evidence of the usefulness of computers in mathematical learning. For instance in using an interactive learning system for enhancing understanding in calculus, Monson and Judd (2001) reported that 78% agree or strongly agree with the statement that the system was good for learning. The government claims that "E-learning exploits interactive technologies and communication systems to improve the learning experience. It has the potential to transform the way we teach and learn across the board. It can raise standards, and widen participation in lifelong learning". It is "recognised in all the Department's recently published age-related strategies – primary, secondary, and further and higher education" the potential opportunities that e-learning can provide. DfES(2003a) this endorses the recommendation in the Dearing Report to utilise IT in teaching and learning within Higher and Post Compulsory education.

Brown and McIntyre (1981, p245) write:

"The research questions arise from an analysis of the problems of the practitioners in the situation and the immediate aim then becomes that of understanding those problems. The researcher, at an early stage, formulates speculative, tentative, general principles in relation to the problems that have been identified; from these principles, hypotheses may then be generated about what action is likely to lead to the desired improvements in practice."

1.2 Statement of the Problem

The purpose of this study is to evaluate the development of an ILS capable of responding to the individual learning needs of a range of learners. "Can an Intelligent Learning System (ILS) effectively tutor learners beyond compulsory education in numeracy and algebra". These issues were addressed by asking the following questions

Q1. What common errors and misconceptions should be anticipated? Are these transferred and maintained from secondary school mathematics?

Q2. Which factors could enable software to be effective for learning?

Q3. Is the GCSE Mathematics grade achieved appropriate for base line assessment of learner ability?

Q4. To what extent will those who were low attainers at school become anxious when required to use mathematical skills and understanding in post compulsory and higher education?

Furthermore these findings could be used to consider the fundamental question as to whether the system can 'transform' learning thereby changing the use of teaching staff time?

Following the literature review presented in Chapter 2 in which a study of

areas of mathematical weakness in secondary school learners,
learning beyond compulsory education and
the effective use of software in learning

these questions were rephrased as measurable hypotheses.

The first question directly links to hypothesis 1

Hypothesis 1: difficulty in understanding specific areas of Numeracy and algebra as identified by the QCA at KS3 and KS4 will continue into post-compulsory and higher education

Question 4 was revised to

Hypothesis 2: Learners who experience persistent failure experience anxiety and have a negative attitude towards mathematics.

Question 3 became

Hypothesis 3: the GCSE grade or equivalent attained is an adequate indicator of mathematical understanding related to Numeracy and basic algebra

Question 2 was reconstructed to become

Hypothesis 4: there is a difference in the effectiveness for learning between computer systems with and without feedback which focuses on common errors

1.3 Experiments undertaken and outcomes

The following methodologies and stages as detailed in Chapter 3 were used to collect data relating to these two aspects

- Electronic Statement Questionnaire survey — Stage 1
- Pilot study
  - Pre-Test — Stage 2
  - Test — Stage 3
  - Software Evaluative Questionnaire — Stage 4
  - Group interview — Stage 5
- Final trial
  - Matched Pairs Trial and Post Test — Stage 6

This research included a Questionnaire Survey, Pilot Study and Final Trial. The Questionnaire survey gathered the opinion of teachers of specific mathematical strengths and weaknesses of pupils. The Stage 1 findings presented in Chapter 4 indicated that the teachers' opinion did not fully concur with that of the published government findings. The teachers opinion of mathematical concepts identified in government reports gathered in

Stage 1 were used in framing specific questions in Stages 2, 3 and 6. Within the literature review in Chapter 2 the six areas of weakness repeatedly reported with secondary school learners were identified for use in investigating the progression of common errors and misconceptions beyond compulsory education.

The Pilot study Stages 2 and 3 presented in Chapter 5 aimed to gain information regarding common errors and misconceptions of learners in higher education. In addition Stage 2 provided the opportunity to compare the GCSE Mathematics grade attained with mathematical performance during the trial. The Stage 2 findings indicated that Weakness 1 (Division), Weakness 2 (Brackets) and Weakness 3 (Indices) were common areas of weakness. Stage 3 findings show that weaknesses 1 (Division), weakness 2 Brackets, weakness 3 (Indices) and weakness 5 (Negative Signs) were common areas of weakness in learners in higher education. Stage 2 findings also showed that there was a positive correlation between GCSE Grade and Pre Test score. From the log files common errors were identified for use in the feedback in Stage 6.

In Chapter 6 details regarding Stages 4 and 5 of the study were presented. The objective of these stages was to gather the learners' opinions of characteristics of software to support learning from users of the software. Feedback on errors made and guidance regarding a method of solution were emphasised by the learners as being useful in software. In Stage 5 signs of mathematical anxiety were revealed by learners of a range of abilities.

Stage 6, the final study as presented in Chapter 7 focused on testing the effectiveness of software using feedback which focuses on the error made in supporting learning of Numeracy and Basic Algebra being used by learners in post compulsory education. Furthermore this study provided an additional opportunity to compare GCSE Mathematics grade attained and performance in the Pre Test. The findings indicated that the system using feedback which focuses on the error made a significantly difference in the performance of learners and was more effective than the system without feedback. Moreover the log files indicated that the users of the system with feedback attempted more questions despite making more errors and proceeded to gain more correct answers to questions. In addition the results indicated that Weakness 5 Negative signs and values and Weakness 2 Brackets are areas of common weakness and misconception in learners in post compulsory education. Furthermore a correlation between GCSE Mathematical grade and Pre Test score was found.

Subsequently in Chapter 8 the findings of each of these stages of investigation are examined to test each of the aforementioned hypotheses. Finally in Chapter 9 the implications for further research as a consequence of this study are outlined.

Chapter 2 Literature Review

2.1 Introduction

Ramsden (2000) suggests that good practice within higher education will involve an understanding of

> " What do I want my students to learn? ...
> How should I arrange teaching and learning so that students have the greatest chance of learning what I want them to learn? This is the problem of teaching strategies.
> How can I find out whether they have learned what I hoped they would learn? This is the problem of assessment.
> How can I estimate the effectiveness of my teaching?.. This is the problem of evaluation." Ramsden (2000 p123)

Fundamental to this study is the nature, complexity and format of the questions, the 'what', presented to learners. Focused, carefully worded questions for all the intended content can enable understanding to be developed. To inform the learning requirements of students beyond compulsory education, where limited investigation has been undertaken, it has been necessary to review the findings relating to learner performance at secondary school level, where extensive research has been conducted. This approach has then enabled the investigation of the progression of learning.

Crook (1994) described how the rapid evolution of the microcomputer encouraged by government support such as the Computers in Teaching Initiative for Higher Education has enabled widespread use of ICT in learning. However pressure politically and educationally to develop efficient and flexible learning packages which make innovative use of technology persists. The relevance of a computer algebra package is dependent upon the range of methods of solution and associated common errors and misconceptions to be anticipated within and supported by the system. Hence these factors relating to teaching strategies are crucial to the design of such a system.

Underlying the learning process is the interaction derived from the style of presentation of teaching strategies including feedback to learners. Assessment provides the learner and teacher with the necessary information to improve and develop learning and teaching.

Evaluation measuring the effect on learning is necessary to maximise the possible impact of all teaching strategies including those employed by software applications. Clear identification of strengths and areas for development can enable targeted development.

2.2 Identifying areas of mathematical focus

2.2.1 General Overview

It has long been apparent that the skills and understanding required to accurately solve numerical equations and to manipulate algebraic expressions have been unattainable for many secondary pupils. The Third International Mathematics and Science Study (TIMSS) (1996) and (2003) showed that British performance in Mathematics was poor in comparison with other countries whilst that of Science at ages 9 and 14 was good by comparison as indicated in the 1996 survey. Indeed Cockroft (1982) commented that "algebra is a source of considerable confusion and negative attitudes among pupils". If, on leaving school these misconceptions have not been corrected then they can only continue. A suggestion which is supported by the subsequent findings reported by The London Mathematical Society (1995) that

> "The serious problems perceived by those in higher education are:
> (i) a serious lack of essential technical facility —the ability to undertake numerical and algebraic calculation with fluency and accuracy;
> (ii) a marked decline in analytical powers when faced with simple problems requiring more than one step;"

Furthermore Dearing (1996) conducted the review which led to Curriculum 2000 for 16-19 education. These proposals included that all 16-19 year olds in education or training should have opportunities to develop the three keys skills as specified by the QCA, namely Numeracy, Communication and ICT. Consequently UCAS now awards 10 points per skill at level 2 (equivalent to GCSE) and 20 points per skill at level 3 (equivalent to A level). However for those on a non-mathematical higher education programme there is not a specific curriculum for Numeracy, although Gillespie (2003) reports that "Some HEIs are using the QCA key skill unit specifications, particularly at levels 3 and 4 including the Open University and UCLAN". More recently ACME (2002) reported that "The government recognises that there is an urgent need to improve the mathematical skills of the general population" noting that there is "poor uptake of school pupils continuing mathematics through to the age of 19 and beyond" and "the reduced numbers of students qualifying for Higher education in numerate disciplines".

This study focused on areas of common misconception within basic algebra and numeracy. The need to direct learning towards areas of misconception is a widely accepted principle of good practice within education. The Diagnostic Teaching Project, Nottingham University Shell Centre (Bell 1993) reported improvements in achievement and long term retention of mathematical skills as a result of using teaching packages that were designed to elicit and address misconceptions. The importance of identifying and correcting misconceptions has also been given prominence in official UK mathematics education documents. Spooner (2003) explains that the report of the Numeracy Task Force (DfEE 1998b), the Framework for teaching Mathematics (DfEE 1999) and the National Curriculum for mathematics for initial teacher training (DfEE 1998c) all place recognition and remediation of misconceptions at the centre of effective practice.

It is necessary to describe the nature of and distinction between an error and a misconception. Spooner (2003 p3) advises that "an error can be the result of a misconception but could also be caused by a number of other factors, including carelessness, problems in reading or interpreting a question and lack of number knowledge" whilst "a misconception is the product of a lack of understanding".

Considerable reporting and research of learner performance in basic mathematics at secondary school level has been undertaken by academics and government organisations. However despite a range of sources investigating the nature of undergraduates' writing skills it is evident that there is a lack of research regarding basic mathematics learning within higher education. Similarly within post compulsory education little research relating to basic mathematics and numeracy has been undertaken. The identification of areas for investigation was informed by analysis of national surveys of mathematical performance in secondary schools, the areas of mathematics required for teacher professional competence and associated research findings. The hierarchy of learning indicated within National Qualifications Framework suggests that the research findings relating to Key Stage 3 and 4 National curriculum Programme of Study for Mathematics are relevant sources. However it cannot be assumed that the misconceptions identified during secondary school mathematics are always the same as those in higher education. In fact Riall and Burghes(1996) reported that most of the employers do not test mathematics at interview or require GCSE at grade C "since it was said that neither the test nor the GCSE qualification gave an accurate assessment of the competencies that the employer required" and proceeded to state that "Percentages was highlighted by employers as being a real problem area". Furthermore Drake (2002 p208) suggests that there are signals that there is "a lack of confidence in the GCSE certificates".

2.2.2 Mathematics at Key Stage 3

The school curriculum comprises all learning and other experiences that each school plans for its pupils in each phase of education. The national curriculum sets out the requirements to be taught in each subject, at each key stage. Within secondary school education there are two key stages, Key Stage 3 (KS 3) phase representing years 7, 8 and 9 with Key Stage 4 (KS 4) representing the final years of 10 and 11. Assessment provides information to teachers, parents and pupils about how a child is progressing at school. Schools are required to keep records on every child including the results of the end of phase national tests. Each pupil is reported as having attained specific levels of competence. Normally KS 3 relates to levels 5 to 8, details of these levels are included in appendix 1.

Nationally the standard assessment tests (SATS) in Mathematics are compulsory for pupils at ages 7, 11 and 14, near the end of Key stages 1, 2 and 3 respectively. The KS 3 tests in mathematics are taken by some five hundred thousand 14-year olds. As reported by Clausen-May (1998) "the questions go through a complex cycle of development with trials, pre-tests and frequent amendments". Participating schools indicated their opinion of the validity of this assessment tool within the QCA Mathematics KS3 SATS Survey. Over 75% of respondents rated their satisfaction with the papers for each of the pupil (ability) groups as high or good. Fewer than 5% rated their satisfaction as low for any of the pupil groups. Consequently the results of these tests are considered for the purpose of this research to be a reliable indicator of national mathematical performance.

The balance of marks within the KS 3 Mathematics SATS is

Number and algebra	Approximately 60 marks
Space, shape and measures	Approximately 30 marks
Handling data	Approximately 30 marks

Hence approximately half of the marks are allocated to the assessment of skills and understanding associated with the solution of numerical and algebraic problems or the manipulation of algebraic expressions.

Table 2.2 Mathematics national results, key stage 3, 1996–2002

	(Percentage of cohort at each level)							
Levels	Below 3	3	4	5	6	7	8	EP
1996	3	11	23	23	22	10	1	0
1997	2	10	22	23	25	11	1	0
1998	2	11	22	24	23	11	2	0
1999	3	9	21	24	24	12	2	0
2000	2	9	20	24	23	16	3	0
2001	2	8	18	24	23	17	3	0
2002	3	8	18	22	25	16	4	0

Source: QCA(2000) Report to QCA on The evaluation of statutory assessment in 2000 Key stage 3 and QCA(2002) Standards at Key Stage 3 Mathematics (Details relating to these levels are stated in Appendix 1.)

The results in Table 2.2 show that there has been a progressive increase in the proportions of pupils achieving higher levels with 56% achieving level 5 or more in 1996, 60% in 1997, 60% in 1998, 62% in 1999, 66% in 2000, 67% in 2001 and 2002. The QCA(2000) stated that "The number of pupils gaining levels 7 or 8 shows a significant increase over the last four years" and "The number of pupils gaining levels 5 or 6 has changed little over the last four years but fewer pupils now gain the lower levels." However QCA(1999) commented "The stability of test results at key stage 3 over the past four years suggests that the improvements in teaching and learning reflected in the significant improvements in key stage 2 test results since 1995 are not being carried through to key stage 3." These statements support the notion that resources that can support and enhance the learning of KS 3 and KS 4 Mathematics are needed.

Resources which enable targeted additional learning opportunities can provide effective support for learners. Those in the middle band of each cohort could achieve a low grade C rather than a grade D or E.

The QCA has responsibility to keep the school curriculum under review. This is undertaken through a range of research, evaluation and monitoring activities. As part of this, each year the QCA produces for each of the phases 1, 2 and 3 an Annual Mathematics Report summarizing the findings of the national Statutory Assessment Tests (SATS). A comparison of the findings relating to KS 3 from 1996-99 has been undertaken to deduce areas of commonality. A full report is given in Appendix 2.

The following key findings were reported

1. Pupils achieving level 5 showed a good understanding of algebra as generalised arithmetic.
2. Pupils achieving level 5 and above successfully used substitutions in algebraic expressions, performed inverse substitutions, constructed algebraic expressions to represent patterns, constructed patterns to fit algebraic expressions and interpreted algebraic expressions in context.
3. At the higher levels (level 6 and above), when the substitutions, manipulations and work with equations became more complex, errors appeared in pupils' answers, particularly when **negative numbers** were involved.
4. Pupils are generally unable to handle **indices** well in both number and algebra. This applies across the ability range.
5. Pupils could perform inverse operations and procedures well, in a variety of contexts. Pupils achieving level 3, however, did not understand the nature of number equations. In this context of **arithmetic pre-algebra**, they were unable to use **inverse operations** to find missing numbers.

From analysis of the algebraic components of the QCA 2000 KS 3 Mathematics SATS Report the following findings were deduced.

Areas of weakness and common misconception

- At level 5 the majority of pupils made **arithmetic errors** when solving a Linear Equation
- At level 5 37% were unaware of a standard approach for collecting together like terms when solving a **Linear Equation**
- Algebraic manipulation errors increased to 57% at level 6 when **handling negative signs and negative numbers**
- At level 6, 56% had difficulty **writing an algebraic expression**

- At level 7, 66% were unable to **write** the required expression
- At level 6, 78% were not able to **substitute** accurately
- At level 7, 59% were unable to **solve the equation**
- At levels 3 & 4, comparing algebraic expressions is more problematic than comparing numerical values
- At levels 4-7, solving an equation involving **two separate operations** proved more difficult than those which require only one type of operation.
- At level 3, 82%, at level 4, 58%, at level 5, 45% and at level 6, 42% were unable to **explain** adequately the **algebraic term** within an algebraic representation of a tessellation sequence (or series of patterns such as crosses being represented by (number of tiles = 4n+1) see Numbers and Sequences in Appendix 2)
- 78% at level 4, 54% at level 5 and 28% at level 6 were unable to accurately solve **by substitution** when given one equation from which the value of the variable can be deduced and then used within another
- 100% at level 3, 95% at level 5, 84% at level 6, 59% at level 7 and 31% at level 8 were unable to solve by substitution
- 95% of pupils at level 3, 74% at level 4, 36% at level 5 and 19% at level 6 could not accurately **manipulate equations** which involve **negative signs and numbers**
- 81% of level 3 pupils, 59% of level 4 pupils, 25% at level 5 and 15% at level 6 unable to accurately **substitute a numerical value into a formula** without any other manipulations..
- Weak performance in **solving simultaneous equations** of whatever format

These findings indicate that areas of common misconception include the following labeled KS 3 a-f for ease of reference

KS 3 a	Negative signs and values
KS 3 b	Substituting values (in formula, expressions and equations)
KS 3 c	Indices
KS 3 d	Solving equations (linear)
KS 3 e	Writing an algebraic expression
KS 3 f	Brackets (expanding and removing)

A full analysis of the KS3 Mathematics SATS: QCA Evaluation 96 – 99 is situated in Appendix 2a. Subsequently in the QCA(2002) KS3 Mathematics SATS reports, common errors relating to each of the weaknesses KS3 a-f as identified above were reported. A full mapping of these weaknesses and errors is given in Appendix 2b.

2.2.3 Common Errors and Misconceptions identified at GCSE

Prior to commencing the GCSE Programmes of Study the pupils will have been assessed by the KS 3 (KS 3) SATS measuring of Mathematical attainment at QCA(2000b) National Curriculum levels 3-8 (details of the levels are given in Appendix 1). At the end of compulsory education and completion of the Key Stage 4 phase most pupils undertake a GCSE Mathematics qualification. The Joint Council for General Qualifications reported that 673,056 pupils sat a GCSE in Mathematics in the year 2000 with the following classification results

	A*	A	B	C	D	E	F	G	U
Cumulative %	2.8	10.7	27.3	49.2	67.1	82.7	91.8	96.1	100.0
%	2.8	7.9	16.6	21.9	17.9	15.6	8.9	4.3	3.9

Source: Joint Council for General Qualifications

The results of the previous three 1997 to 1999 years illustrate a similar pattern of performance.

	Year	A*	A	B	C	D	E	F	G	U	Number A*-U
Cumulative	97	2.1	9.6	24.2	47.4	63.8	79.4	91.6	97.9	100.00	686982
%	97	2.1	7.5	14.6	23.2	16.5	15.6	12.2	6.3	2.1	
Cumulative	98	2.2	9.8	25.3	46.5	63.7	79.9	89.1	94.3	100.00	682143
%	98	2.2	7.6	15.5	21.2	17.2	16.2	9.2	5.2	5.7	
Cumulative	99	2.3	10.2	26.4	47.7	65.9	81.2	90.6	95.6	100.00	691826
%	99	2.3	7.9	16.2	21.3	18.2	15.3	9.4	5.0	4.4	

Source: Joint Council for General Qualifications

From these results we can deduce that at least 50 % of candidates failed to achieve a grade C or higher result and in fact 33.5% achieved either a grade D or E. Grades A* to C are widely considered by employers, Further Education and Higher Education institutions to be the only acceptable standard at GCSE level.

QCA (2000) states that GCSE Mathematics syllabus consists of three strands: number and algebra; shape, space and measures; and handling data. The Number and Algebra aspect consists of

- Using and applying number and algebra
- Numbers and the number system
- Calculations
- Solving numerical problems
- Equations, formulae and identities
- Sequences, functions and graphs

The QCA requirements for the GCSE Mathematics Curriculum demands that candidates at the higher level

> "use proportional reasoning with fluency and develop skills of algebraic manipulation and simplification... extend their knowledge of functions ...

and solve a range of equations, including those with non-integer coefficients. They use short chains of deductive reasoning, develop their own proofs, and begin to understand the importance of proof in mathematics. Pupils develop the confidence and flexibility to solve unfamiliar problems and to use ICT appropriately. By seeing the importance of mathematics as an analytical tool for solving problems, they learn to appreciate its unique power"

2.2.4 Analysis of algebraic components within Edexcel GCSE Mathematics 2000

Edexcel (2000) identified problematic areas of the curriculum. These included manipulative algebra. Algebra is reported to be an area of weakness with misconception of simplifying and combining terms, expressing a general form of a pattern such as n+3 (KS 3 b) and writing an appropriate formula (KS 3 e) evident. Expanding brackets (KS 3 f) was identified as a problem for "a significant number of candidates", and where there was "a combination of the multiplication sign the fraction and the bracket" greater difficulties were observed. Various arithmetic errors are outlined including careless use of signs (KS 3 a), evaluating expressions which involve the multiplication or division of decimals, with several references to difficulties connected with the concepts of percentage, standard form, rounding and non-commutativity of division. The full details of reported errors are included in appendix 3.

2.2.5 Other findings of common errors and misconceptions in basic algebra and numeracy

Research such as that of Matz(1982) has shown the widespread uniformity of both correct and incorrect answers. The following erroneous solutions to questions illustrate common errors identified by Matz (1982). A coding of M1 to M32 has been applied to enable ease of reference.

Ref	Finding	Ref	Finding
M1	x/x=0, a* 1/a = 0	M17	(AX+BY)/(X+Y) = A+B
M2	Evaluate 4X when X=6, then 4X=46	M18	X/(2X+Y)=1/(2+Y)
M3	Evaluate XY when X=-3 and Y=-5, XY=-8	M19	(X+3Z)/(2X+Y)=3Z/(2+Y)
M4	Evaluate 2(-3) as –1 (NB bracket implying multiplication)	M20	(X-3)/2X=-3/2
M5	$(-1)^3$ as –3	M21	$(X^2+2XY+Y^2)/(X^2-Y^2)=2XY$
M6	$3r^2$ as $3+r^2$	M22	2(X+3)=2X+3
M7	$3r^2$ as $(3r)^2$	M23	-(3X-W)=-3X-W
M8	Simplify 3+23(s-4) to 26(s-4)	M24	$(AX+B)(CX+D)=ACX^2+BD$
M9	Simplify 3XY+ 4XZ to 7XYZ	M25	Solve for x (X+1)/(X+4)=5/6 X=4,2
M10	0*a=a	M26	Solve for x 2X+5=11, X+5 = 11/2
M11	$(A+B)^2=A^2+B^2$	M27	Solve for x 3X+5=Y+3, X+5=Y
M12	A(BC) = AB*AC	M28	Solve for R 1/R = 1/R1+1/R2+1/R3, R=R1+R2+R3
M13	a/(b+c) = a/b +a/c	M29	Solve for x $1/X+1/X^2=3/X^2+6X^2$
M14	(a+b)/(c+d) = a/c +b/d	M30	FACTORING X^2+5/6X + 1/6 AS X(X+5/6) + 1/6
M15	$2^{a+b} = 2^a + 2^b$	M31	Solve for x (X-5)(X-7)=3, X-5=3 OR X-7=3, X=8 OR X=10
M16	$2^{ab} = 2^a 2^b$	M32	Solve for x 5/(2-X)+5/(2+X)=4, 5(2+X)+5(2-X) = 4

Nature of misconceptions

From examination of these questions and erroneous solutions the nature of specific misconceptions can be inferred.

M1 suggests that there is a common misconception related to **division** by the same quantity.

M2 and M3 indicate a lack of understanding in **substituting values** into an algebraic representation of a given expression (KS 3 b). M3 also suggests a difficulty with the use of negative values and signs (KS 3 a).

M4 and M6 suggest a lack of knowledge regarding **notational conventions.**

M5, M7 and M16 show a misunderstanding of the concept of **indices.** (KS 3 c)

M8, M11, M12, M22 and M24 indicate confusion with the understanding of the representation of **brackets** which is also shown in M23 although this also encompasses the concept of a negative quantity. (KS 3 f)

M13, M14, M17, M18, M19, M20, M21, M25, M28, M29, M32 outline misconceptions relating to **algebraic division.** With M25 and M26 indicating difficulties with solving equations. (KS 3 d)

2.2.6 Areas of commonality

In comparing the findings from the reports of KS 3 SATs, GCSE Examiners and Matz we can identify the following recurrent areas for misconception and weakness.

W1 Division
W2 Brackets
W3 Indices
W4 Substituting values
W5 Negative signs and values
W6 Solving equations (linear)

Writing an algebraic expression (W7) however is not encompassed by the research of Matz(1982). A full mapping of the specific QCA findings as stated in the report are mapped against those detailed by Edexcel(2000) and Matz(1982) is given in appendix 4.

2.2.7 Mathematical hierarchy of questions

Sangwin(2002) adapted Smith's mathematical question taxonomy which is based on the well known general scheme of Bloom's Taxonomy to classify mathematic skills.

Group A	Group B	Group B
1 Recall factual knowledge	4 Information transfer	6 Justifying and interpreting
2 Comprehension	5 Application in new situations	7 Implications, conjectures and comparisons
3 Routine use of procedures		8 Evaluation

Bloom's Taxonomy

1. Factual recall
2. Carry out a routine calculation or algorithm
3. Classify some mathematical object
4. Interpret situation or answer
5. Prove, show, justify – general argument
6. Extend a concept
7. Construct instance
8. Criticise a fallacy

Sangwin's Alternative Mathematical Question Taxonomy

In Sangwin's Alternative Taxonomy Levels 1 - 4 are defined as adoptive learning which are essentially based on reproducing processes hence requiring the application of well-understood knowledge in bounded situations. This investigation into an aspect of Basic Mathematics and Numeracy focuses on Level 2 'Carry out a routine calculation or algorithm' which necessarily requires the ability to use Level 1. The higher levels in this alternative taxonomy relate to the proposition raised by Aczel (1998) "are there "deeper" insights that have to be obtained in learning mathematics than the gaining of merely operational knowledge".

2.3 Learning in Higher Education

2.3.1 Hierarchy and how students learn

Cotton (1995 p17) outlined the knowledge steps in the affective domain suggested by Bloom and co-workers through which higher education aims to facilitate and nurture the development of understanding within learners

- Receiving or attending
  - passive stage, no real commitment except a willingness to attend.
- Responding
  - first signs of interest and attention. More emotional commitment. Compliance with work and task undertaken.
- Valuing
  - beginning of commitment. Learners start to defend the work they are undertaking.
- Organization
  - personal commitment and emotional involvment. Learners begin to rationalize and judge their work by professional standards. Learners are able to apply knowledge analysing specific examples and begin to evaluate.
- Characterisation
  - professional commitment is internalised. Learners believe in their work and the profession is now part of their own self-image and self-esteem.

This hierarchy outlines the progression of responsiblity and motivation required for effective learning. Furthermore within the climate of lifelong learning denoted by the DfEE (1998 a) it is desirable to encourage and enable the development of autonomous learners who assume ownership of their personal progression. Vygotsky's constructivist learning theory advocates active learning and the opportunity for learners to negotiate their learning activity. Theorists such as Gagne consider motivation to be the first step in learning. Motivation has been described as being either intrinsic or extrinsic, that is, internally or externally produced. Armitage et al (2003 p69) suggest that within post compulsory education "for external motivation to be sustained students must have attainable goals to work towards, be given immediate feedback on their performance". Cotton (1995 p 137) outlines the benefits to learning in higher education of the move towards self assessment. Jerome Bruner saw a theory of instruction derive from the knowledge of cognitive developmental processes and stages which led to a movement in curriculum development and teaching focusing on the design of experience-based educational programs enabling subject matter to be taught respectively to learners at any age or stage of cognitive development. The challenge became how to translate abstract symbolic principles into concrete modes of representation. This freed learners from memorising principles and rules. Learning became the process of discovering knowledge, not just the content. Hence effective learning should enable the development of an awareness of process, of 'how to learn'. A computer aided system can provide immediate feedback and self assessment opportunities. Harding (2001) reports that the use of ICT in assessment will encourage the use of ICT in learning yet is still not widely utilised despite the potential of the use of new technologies in education having been recognised since 1960s including the benefits of self assessment and feedback.

Additionally as Cunningham (2001) suggests in learning there is 'the need to emphasise motivation whereby people feel better about themselves' (p. 38). Achieving success could promote self esteem through progress and attainment. To embark into post compulsory education students have made a choice which generally suggests that they have intrinsic motivation; however this must be sustained.

2.3.2 Changes in the nature of learning

The nature of learning within higher education is changing to include more use of ICT. Many researchers including Saunders et al (2003) report that throughout the sector there is a notable change in the size of classes and range of educational backgrounds which coupled with developments in ICT have led to different and more flexible approaches to learning. In introducing the Information Communication Technology National Curriculum 2000 QCA(2000a) Charles Clarke, Education Minister claims that "E-learning has the potential to revolutionise the way we teach and how we learn" and within DfES(2003) there is encouragement to build a better e-learning market which encourages innovation and exploitation of the technology. Within the White paper The future of higher education a vision for a higher education system characterised by inclusion, excellence and flexibility in which e-learning plays a significant role is detailed. Furthermore within DFES(2003) the government indicates that e-learning is important because of the contribution towards challenging objectives including

"achievement – by providing new and creative ways of motivating and engaging pupils and learners of all abilities ...
Reducing the number of adults without level 2 qualifications – by offering private and individualised feedback ...
Ensuring wider participation and fairer access to higher education – by creating the opportunity to start learning and choose courses and support according to the learners' needs"

Armitage et al (2003) suggest that technology in learning has no intrinsic value but presents an overview of advantages of e-learning as identified by Shepherd(2002) and Seale (2002) to include widening student access, supporting large groups, motivation, flexible support, enabling active and independent learning.

An additional complexity is that "higher education has expanded to accommodate a much more diverse student population" (Armitage et al (2003) p67). In the context of the financial burdens of tuition fees and the pressure to earn a living while studying the need for learners to be strategic is inevitable. Consequently learning needs to be flexible so that it can be accessed when the learner is available and content should be individualised focusing on personal requirements. DfES(2003) identifies that "E-learning can help to provide individualised feedback to help learners progress" and "offers individual empowerment with greater control over their own learning".

Most higher education institutions recognise the need for 'writing support'. There are numerous books on 'how to write' and 'how to read' and texts for professional development of teachers of undergraduates and post graduates. Mackenzie(2002) comments that these texts "rarely mention either emotional barriers to learning, or mathematical skills required in HE.". Yet it has long been recognised that for many learners there is considerable anxiety and a poor attitude towards learning mathematics.

For instance Mackenzie (2002) stated that

"Over the last 20 years there has been considerable interest in negative attitudes and beliefs relating to maths learning. The Dearing report (1997) highlighted the need for HE programmes to enhance students' basic skills and to deliver numerate graduates. Therefore many HE institutions are now looking seriously at how basic skills are enhanced within degree programmes. One of these is Numeracy".

Mackenzie's survey found considerable negative attitudes (25%) and low confidence and even higher level of anxiety (over 40%) in students taking a Humanities programme. This research suggests that "students arrive with a two year memory of working with maths/number skills in order to pass Maths at GCSE rather than recently practised skills, or a love of maths." and suggests that these learners may be unprepared for the post-Dearing emphasis on key skills. In fact at York University 25% admitted to some concern and avoidance of mathematics learning which is the same proportion found in research relating to US undergraduates over the past 20-30 years. Furthermore she suggests that confidence levels in basic skills may decrease and anxiety increase if they are not being practised in the last years of full-time school. This research identified the following aspects of mathematics as a concern: Converting units (79%), changing scales (52%) and fractions (51%) the skills with the lowest confidence levels. Furthermore Percentages with a 38% confidence rating, interpreting figures 39% and use of computers 31% were also recognised as concerns as these skills are frequently used both in degree studies and in the workplace. However attitudes to Mathematics will be formed during earlier experiences in compulsory school education. This view is suggested by the Smith Inquiry concern over attitudes to Mathematics as a consequence of the 14-19 Curriculum and its inability to motivate learners. Indeed Smith(2004) reported that "the influence of the teacher is clearly important; in particular, poor teaching is likely to turn students off mathematics".

Joffe and Foxman (1988) reported that strong usually negative feelings are often proved by the mere mention of the word 'Maths' and these have a detrimental affect on learners attitude and achievement. Ho et al (2000) supported the notion that mathematics anxiety has a negative relationship with mathematics performance and achievement caused by negative attitudes towards mathematics. Merttens (1997) advised that those who get answers wrong may then fear or become anxious answering questions.

Furthermore Githua et al (2003) conducted a study of 649 students from 32 randomly selected schools that indicated that student mathematical self concept, that is their self-perceptions of their perceived personal mathematical skills, ability, and enjoyment as defined by Marsh (1990,1996), is critical to their motivation to learn mathematics. The dimensions of motivation are highlighted as being relevance, interest, likelihood of success and satisfaction. Hence this suggests that enhanced mathematical self concept could lead to improved motivation, probability of success and incentive to learn.

Boyd et al(1998 p108) questions whether there is really a significant Numeracy problem amongst undergraduates. Their survey of 200 students recorded that the performance of 10% of the group suggested that they had weak Numeracy skills. They reported that "around 10% of the class showed weakness in conversions and percentages, and 15% had problems with order of operators. Two students scored 100%." The evaluation by Boyd et al (p115) showed how student responses "indicated a degree of student surprise at how well these students had performed in the survey. The confidence boost was very evident". However Drake(2002) suggests that tests are not the most valid assessor of numerical understanding as within real life students use numerical skills in context.

Ramsden (1992) sees effective teaching at HE level as that which promotes high motivation, interest in the subject matter and a perception of the relevance of the content. Hence it should be recognised that many students use the number and mathematics skills required in everyday life and that the needs of students on different programmes should be considered rather than merely focusing on those with problems with Numeracy. Learning which motivates whilst limiting anxiety should be active, owned and focussed. The content should not be perceived as being either solely repetitive or only focusing on difficulties. In considering the role of e-learning in facilitating the development of cognitive understanding McDougall (2002 p84) identifies that "the challenge for teachers and for educational software designers, is thus to present to learners a situation provoking an optimal amount of cognitive conflict, while appreciating that the optimal amount may vary widely among individual learners". This objective would enable 'self actualization' according to the widely know Maslow's hierarchy of need.

2.3.3 Numeracy within Initial Teacher Education in Higher Education

All prospective teachers are required to comply with the statutory numerical requirements for Initial Teacher Education (ITE). To gain Qualified Teacher Status each trainee must pass a statutory numeracy test. These tests are designed to be relevant to the teacher's professional role, and the majority of questions will be written in the context of data commonly available within schools which a newly qualified teacher (NQT) could be expected to use. The questions require candidates to interpret statistical information or perform calculations which may be related to:

- Performance, e.g. national test data, improvement data, target setting
- Schools and/or individual pupils e.g. progress, attainment, free school meals, special educational needs (SEN), gender, destinations, absences, subject choices (e.g. GCSE)
- Finance, e.g. costings, budgets
- Teacher support, e.g. mark schemes, timetabling, planning, school trips

The test is designed to assess those aspects of numeracy that are required by teachers to carry out their professional role effectively. It will be presented in two parts, the first part containing mental arithmetic questions for which calculators are not allowed, and the second part containing on-screen questions for which candidates will be able to use the on-screen calculator.

Candidates will be expected to carry out mental calculations of more than one stage, using for example:

time; amounts of money; proportions, fractions and/or decimals; percentages; measurements (e.g. distance, area); conversions (e.g. from one currency to another, from fractions to decimals or percentages); and combinations of one or more of the following processes: addition, subtraction, multiplication, division.

There are two major aspects of numeracy covered in this part of the test:

1. Interpreting and using statistical information

Candidates will be expected to identify trends correctly; make comparisons in order to draw conclusions; and interpret information accurately.

2. Using and applying general arithmetic

Candidates will be expected to use and apply general arithmetic correctly using time; money, proportion and ratio; percentages, fractions and decimals; measurements (e.g. distance, area); conversions (e.g. from one currency to another, from fractions to decimals or percentages); averages (including mean, median, mode and range where relevant); and simple given formulae.

These are the skills and understanding identified by the Teacher Training Agency (TTA) as being crucial in functioning effectively in the profession. 'Using and applying general arithmetic' will encompass the previously identified difficulties KS 3 a,b,c,d,e,f. McNamara et al (2002) report that "many students perceived the test as degrading other symbols of their proficiency in Mathematics" including Mathematics GCSEs. However the existence of the test could suggest that there is a widespread concern about the Mathematical attainment of Newly Qualified Teachers. For the purpose of this research let us label these areas as

TTA1	Time
TTA2	Money
TTA3	Proportion and ratio
TTA4	Percentages
TTA5	Fractions and decimals
TTA6	Measurements
TTA7	Conversions
TTA8	Averages
TTA9	Simple given formula

From these findings there are two major aspects of mathematics have been identified for investigation.

1. Basic Algebraic Methods

Candidates will be examining how to

a) solve a simple given formulae by substitution
b) express a problem algebraically
c) simplify a given algebraic expression

d) solve linear equations
e) solve simultaneous equations
f) solve quadratic equations

2. Using and applying general arithmetic

Candidates will be considering how to use and apply general arithmetic correctly to given word problems using:

a) time;
b) money;
c) proportion and ratio;
d) percentages, fractions and decimals;
e) measurements (e.g. distance, area);
f) conversions (e.g. from one currency to another, from
g) fractions to decimals or percentages);
h) averages (including mean, median, mode and range)

From a consideration of these basic algebraic methods (1 a – f) and uses of general arithmetic (2 a – h) it is evident that there is a need to ensure an adequate understanding of the identified areas of common weakness (W1 – W6) to ensure successful completion of the Numeracy test for QTS.

An additional requirement of all trainee teachers is the awareness, development and progression of personal knowledge, understanding and skills and experience of a range of uses of ICT. Despite reform initiatives such as the inclusion of the use of IT in the teaching of mathematics within the National Curriculum Programme of Study since 1995 there is strong evidence that the use of computers in mathematical explorations are not widely found. Manouchehri (1999) reported that in the context of various movements and recommendations for school mathematics to utilise computer technologies and appropriate educational software to impact upon students' mathematical learning, numerous local and national surveys showed that the majority of school pupils had never used a computer in their mathematics classes. This research reported that teachers' limited use of computers in classrooms was due to their lack of knowledge about the use of computers within mathematics access to educational software and training. Research by Dooley et al (1999) supports the view that inexperience and inappropriate training are the main concerns of teachers who are low level users of ICT. Thus it is the teacher's own ability to use the software rather than scepticism on the impact of learning. Most teachers had learnt and progressed in their ability to use software by 'doing' rather than attending training courses. Monaghan(1999) supports the view that effective software design should be 'accessible' to teachers. Also research ((Delozanne and Bruillard 1993) for example) underlines the important role an ITS can play in the learning and teaching of mathematics. It is difficult to transform the educational system to effectively utilise electronic resources but it is necessary to prepare all future teachers with the opportunity to experience a range of uses of ICT. Research undertaken in 1999 would be relevant to many current teachers however it must be acknowledged that research from 1999 may not be altogether timely to a current study.

2.4 Learning in Post Compulsory Education

2.4.1 Government initiatives

New government initiatives and policies relating to the post compulsory learner were many as Trowler(2001) identifies:

- 1986 – Working Together: Education and Training – NCVQ
- 1988 – White paper – Employment for the 1990's
- 1989 – CBI Paper – Towards a Skill Revolution
- 1991 – White paper – Education and Training for the 21st Century
- 1992 – Further and Higher Education Act
- 1994 – White Paper – Competitiveness
- 1996 – Dearing Report on Qualifications for 16 – 19 year olds
- 1996 – Student Loans Act
- 1997 – Kennedy Report – Lifelong Learning
- 1998 – The Learning Age – Green Paper on lifelong learning
- 1998 – The Teaching and Higher education Act
- 1999 – White Paper – Learning to Succeed: a new framework for post – 16 education
- 2000 – Excellence Challenge – widening participation of young people
- 2000 – Learning and Skills Act

In addition, since Trowler produced his up-dated information, there have been further governmental developments including:

- 2004 - 14 – 19 Reform– Tomlinson produced an interim report
- 2002/3 - White paper – Success For All
- 2002 /3 – White Paper for Higher education

Recent changes in post compulsory education driven by governmental policy have included

- Incorporation in 1993, where the further education colleges became autonomous and left the jurisdiction of the local education authorities
- Re-focusing on levels of ability of the British workforce leading to the instigation of 'Key Skill' qualifications and frameworks.

- 'widening participation' with regard to education, also linked to 'inclusive education' and 'lifelong learning'.

The successful institutions within the sector have been proactive in their response to change, with flexibility to teach a vast range of students in terms of age and ability. From the mid to late eighties, Previous to incorporation in 1993 colleges received centralised finance from the local education authority. Funding was largely guaranteed irrespective of recruitment, retention and achievement figures for the various curricular activities. This is no longer the case. Ecclestone comments that:

"the full glare of the political spotlight now shines over the whole post – 16 education and training sector. Policy – makers in government, and their many intermediaries in the Learning and Skills Councils, inspectorates, awarding and qualification bodies, are more interested in the processes and outcomes of post – 16 teaching, learning and assessment than ever before."

(Ecclestone. 2003, vii)

Based on this rising profile, and the national quality measures being implemented colleges need to ensure that effective learning and teaching is occurring.

The Incorporation of Colleges came into effect following the Government's 'White Paper' (DfE, 1991) under the direction of the Further Education Funding Council.
One of the main aims of the Funding Council's leadership was to:

"lift England to the top of the international league in the staying on rates for 16 – 19 year olds" (Stubbs and McClure, 1992)

In meeting the Governments National Training and Education Targets for staying on post sixteen (NTETs), colleges were expected to show an increase in number of students of 25% over three years between 1993 and 1996 by the Council, without equally matching funds.

Within the background of the numerous policies and initiatives institutions and teachers are "under other pressures to be efficient and cost-effective in getting more learners through the system more cheaply". (commented by Ecclestone (2003)). The government in its guidance on the establishment of the Further Education Funding Council (FEFC, 1992, p. 5) made explicit the twin aims of expanding the sector and doing it for the least cost:

The Council should strike a balance between securing maximum access to the widest possible range of opportunities in further education (FE) and avoiding a disproportionate charge on public funds.

The proportion of young people undertaking training was to be increased 'in a way which encourages institutional efficiency' (FEFC, 1992, pp. 6-7).

2.4.2 Research in post compulsory education

Use of research relating to compulsory school education and higher education has been necessary, as within post compulsory education and specifically Adult Education Coben (2004) reports that "Evidence on the impact of adult numeracy tuition is sparse and unreliable. Detailed studies are required" and specifically that "Some teachers' inadequate subject knowledge is a continuing concern. Studies with children suggest that: initial and ongoing teacher education increases subject knowledge, facilitates career development and encourages future research and development". The Smith Inquiry was announced, by the Chief Secretary of the Treasury, in July 2002, as part of the Government's response "Investing in Innovation" to Sir Gareth Roberts' UK wide review "Set for Success: The supply of people with science, technology, engineering and mathematics skills" Smith (2004) recommended the creation of "a forum for school, college, FE and HE local links and joint working" and "bring together all major groups and agencies involved in mathematics education, including from England the DfES, National Strategies, QCA, Ofsted, LEAs, HEIs, LSC, SSCs, ACME". Also recommended was the need for "learning and teaching materials, including distance learning materials and materials to enhance the teaching of mathematics through the use of ICT". Smith had highlighted the important link in learning with compulsory school education, post compulsory education and higher education and the important function that ICT could undertake.

2.5 Use of software in teaching and learning

2.5.1 General use

> "The information technology revolution is having and will have profound impacts on the educational process" Goodman (2001)

Potential benefits from using IT in education have long been reported. NCET (1994) stated that that IT can impact positively on teaching and learning in schools for reasons including:

- IT can provide a safe and non-threatening environment
- IT has the flexibility to meet the individual needs and abilities of each student
- Children who have not enjoyed learning can be encouraged by the use of IT
- Computers give children the chance to achieve where once they have failed
- IT can present information in new ways which help children to understand, assimilate and use it more readily
- Difficult ideas are made understandable when IT makes them visible
- IT gives children the power to try out different ideas and take risks

- Computer simulations encourage analytical and divergent thinking

However in the same report the importance of the learning and teaching strategies in assuring effective learning with and through IT is emphasised.

- Computers help children to learn when used in well-designed, meaningful tasks and activities
- Children make more effective use of computers if teachers know how and when to intervene
- It offers the potential for effective group working

More recently it has been shown that the use of ICT in education can help improve memory retention, increase motivation and generally deepen understanding (Dede, 1998). Students will be able to direct their own studies to a greater extent, with the teacher acting as a guide or moderator rather than as a director (Forsyth, 1996: 31). Some studies such as that by Cox(1997) and Passey et al (2004) have demonstrated improved general motivation of learners. However improvements in learning following the introduction of a new medium could be attributable to a change of curriculum or teaching strategy and not specifically to a change in the general motivation of learners. Indeed Passey et al (2004 p3) reported that motivation was linked to subject specific attainment where ICT use supports internal cognitive aspect of learning. The National Council of Mathematics(2000) reported that technology can enhance learning through active engagement but this would require that technology is not the master.

Twining (2002) suggests that when the intended mode of use of computers is to support the curriculum then the content will remain unchanged, the process will be automated but otherwise unchanged but the learning will be more efficient. He proposes that there are three potential focuses of use

- to develop IT skills
- learning tool supporting areas of learning other than ICT
- other reasons determined by practical aspects or external factors

and that in any context all three will apply. Hence when the focus of use is a learning tool the other focus dimensions are necessarily involved. That is learners will be developing their skills, knowledge and understanding of ICT and the use of computers corresponds to the current demands of government, funding and quality assurance bodies. In fact in line with guidance from the QAA a principle enabling aim of the John Moores University's 1999 Learning, Teaching and Assessment (LTA) Strategy was "to exploit the potential of ICT to enhance the quality and range of student learning opportunities". This is in the national context of the government encouraging schools to embrace ICT as a fundamental part of the curriculum. Between 1998 and 2004 the Government supported over £1 billion of expenditure to improve schools' ICT facilities. In 1998-1999, the National Grid for Learning (NGfL), the UK government's funding for ICT development in schools began to have an impact. The funding provided by NGfL has resulted in a growth of connections to the Internet in primary schools. In March 1998 only 17 per cent of primary schools in the UK had Internet access. By March 1999 this had increased to 62 per cent and in the same period there was also an increase of Internet connectivity in secondary schools from 83 to 93 per cent (DfEE, 2000: 18). Many secondary schools and an increasing number of primary schools are now developing websites and announcing their presence in cyberspace. The use of web pages and virtual learning environments to post school news and homework assignments is soon to become common practice, as is the submission of work via e-mail from the child's home to the teacher's mailbox. This culture is already well established in many universities providing flexible 'any place, any time' learning opportunities. Furthermore as Jones and Knezek (1993) reported ICT was originally intended to serve as a means of improving efficiency in the educational process. As an assessment tool ICT can enable on-line testing which can be instant and provide the teacher with a wide range of information associated with the learner's score. Comparisons of previous scores and dates of assessment for example, will indicate progress, and each student can be allocated an individual action plan database stored in electronic format into which each successive test's results can be entered automatically.

We can be certain that the use of ICT technologies will inevitably escalate in education as they have the world of leisure and work. Rapid changes in technology mean that we have moved on a long way from

> 'I forsee no reason for people to have a computer in their home'.
>
> (Ken Olson, President of Digital Equipment Corporation, 1977).

2.5.2 Mathematics and CAL

> "technology is essential in teaching and learning mathematics"
>
> National Council of Teachers of Mathematics (2000)

In recent years personalised instruction and the use of computer technology to facilitate learning within all disciplines at all levels of education and training has received escalating emphasis. The increasing availability and power of electronic technologies such as computers and graphic calculators offer new opportunities for students to communicate and analyse their mathematical thinking, since the objects generated on the screen can act as a common referent for discussion (National Council of Teachers of Mathematics, 2000). Various findings of various research studies including Manourchehri (1999 and 2004) indicate that the use of computers as a learning aid does indeed improve students' ability to problem solve and develop mathematical understanding. Within their evaluation Nicol and Anderson (2000) demonstrated the benefits of computer aided instruction to help teach adults with learning difficulties. From the teacher's perspective these included the capability to vary the levels of difficulty supporting high levels of success and allowing users to work at their own pace. Users were found to enjoy the learning experience and some improvement in performance recorded. Furthermore Abidin and Hartley (1998) in investigating a computer based learning environment evaluated the benefits of problem solving performances and investigatory learning. All users showed substantial improvements between pre and post tests. In addition national guidance for school mathematics teachers from DfES (2003) is that

> "Computers offer powerful opportunities for pupils to explore mathematical ideas .. and to receive fast and reliable, and non-judgmental feedback" p1
>
> "ICT can be used advantageously in most areas of mathematics, but the following areas particularly benefit from the opportunities it offers. Applying mathematics and solving problems .. Equations, formulae and identities" p9
>
> "Computer algebra systems are tools that automate algebraic computation .. although most often used with older pupils ... some useful ways in which computer algebra can be used in mathematics at KS3
>
> Exploring patterns in number and algebra
> Learning about algebraic equivalence
> Developing skills of algebraic manipulation and solving equations"

In response to the issues of anxiety and negative views towards learning mathematics detailed earlier the view stated by the DfES supports the notion that computerised systems could provide an effective approach. In that, they could provide essential support for learners who fear getting answers wrong and then risk becoming reluctant to participate in further learner as suggested by Merttens (1997). Furthermore repeated erroneous answers would exasperate these negative feelings. Also the standardised approach offered by computerised systems can provide an alternative to 'poor teaching' as reported by Smith(2004).

Indeed the development of e-learning within mathematics is more extensive than within the UK. For instance Haspekian (2003) reported that

"Nowadays, there is a firm institutional will to integrate the technologies of information and communication to the practices and curricula in France. In mathematics, teachers are encouraged to use Internet and various software: dynamic geometry and symbolic calculation software, spreadsheets, etc."

2.5.3 Learning models and computer aided learning systems

A progressive and systematic model for a learning environment for algebraic calculus, of NAIADE (Cauzinille-Marmeche et al, 1989; Joab, 1991) can be categorised into three stages:

Learning by example,
Learning by doing and
Learning by reasoning.

This model conforms to Bruner's theory of experience-based learning. Learning by example entails the users observing explained step by step examples. Learning by doing is whereby the user is guided and supported by the system in terms of what action to take and why, but does participate in solving a problem. Learning by reasoning is whereby the user has full control and must make the decisions in order to solve the problem. When errors occur in the user's solution the system can backtrack through these stages and hence the system is required to be intelligent in order to identify which sub problems were correctly solved and those that were not. Help can be provided at three different levels:

Procedural knowledge – skills based
Strategic – how and when to choose a method
Conceptual – applying the method

which reflects the nature of error classification by Donaldson (1963) as being either structural, arbitrary or executive.

A computer aided learning system can provide interactive and supported learning which enables students to extend their tuition whilst enabling individuals to 'learn by doing' (Papert, 1980). Atkins (1993) suggests that when learning with interactive courseware there are two dominant views: behaviourist and cognitivist. The Behaviourist view of learning focuses on the establishment of a behaviour pattern developed from responses given rather than from any mental processes. In consideration of this type of learning the feedback given to a user, as in a drill and practice package, can influence subsequent behaviour. Monson et al (2001) reported that a benefit of an interactive learning system is that students receive immediate feedback enabling them to correct their own misconceptions. However the behaviourist theory does not consider any mental processes that may occur and therefore influence future response to be accountable. The cognitivist view of progress in learning can include constructivism that emphasises the learner's ability to construct their own schema of understanding from experiencing learning opportunities. Many educationalists including Papert (1980) have emphasised the need for software that can enable learners to develop personal knowledge and understanding that is meaningful and transferable, so that understanding a concept in a specific context can then be applied more generally. Hence these software applications can act as 'tools to think with' relating to Bloom's fourth step of knowledge. Schofield and Ashton (2004) suggest that "online assessments can offer students individual and independent learning experiences which can greatly enhance the teaching and learning process. They link the view of Laurillard(2001) that "action without feedback is completely unproductive for a learner" with the concept that "feedback to a submitted answer can take a number of forms" reporting on the benefits of combining immediate or synchronous feedback with post-assessment feedback which supports student reflection.

Jameson (2003) indicates that for learning to be enabled by ICT, putting the pedagogy back into ICT remains a crucial issue. Fundamental to the proposed algebraic ILS is the interactivity, the quality and nature of the feedback. Flexible online tuition will provide guidance on idenfitication of anticipated common pupil errors and misconceptions. BECTa(1999) reported that within the five studies conducted one of the main weaknesses of Integrated Learning Systems was the lack of alignment between the systems content and function with those of the curriculum. In the context of a learning tool designed for a specific purpose the relevance of the subject content is a crucial quality criterion. In addition the software will provide an opportunity for the user, the trainee teacher and school pupil, to experience a tutoring system that is an ICT tool for the advancement of mathematical understanding.

2.5.4 Levels of feedback

Learning technologies include drill and practice and tutorial systems, the main features of which were outlined by Rist and Hewer(1996) as

> **Drill and practice**
> Drill and practice packages offer structured reinforcement of previously learned concepts. They are based on question and answer interactions and should give the student appropriate feedback. Drill and practice packages may use games to increase motivation.
>
> **Tutorials**
> Tutorials are used to teach new concepts and processes. Material is presented to the student in a structured format. Tutorial software usually includes worked examples and gives the learner the opportunity to assess their understanding with questions, answers and feedback. Intelligent Tutoring systems are capable of corrective feedback and adapt their presentations to suit the learner, based on the actions of the learner.

A computer based resource which provides Boolean yes/no feedback may be perceived as being for assessment rather than learning. Mason and Bruning(1999) have identified eight commonly used levels of feedback. Within their hierarchy "Knowledge of response" is the simplest form of feedback, whereby users are only informed whether their answers are correct or incorrect. This style of feedback directly relates to that provided within drill and practice software. Various observations have been made regarding the use of assessment to motivate learners. Assessment is considered to be very positive in terms of considering specifically what it is that students are expected to learn. The importance of assessment within the learning process is indicated by Cotton (1995: 1). They claim that "good assessment is a consolidating tool within the learning process." Brown et al. (1997, 7) argue that the style of assessment can have an important influence on student learning. They suggest that students consider assessment to highlight the most important areas of knowledge. The ability to learn through assessment would require errors to result in feedback which is corrective and the users to become aware of their misconceptions. Kulhavy and Stock (1989) reported that effective feedback provides the learner with verification and elaboration. Verification ascertains whether an answer is correct or incorrect, whilst elaboration provides relevant cues to guide the learner. Without elaboration the user could receive repetitive negative feedback which reinforces a poor mathematics self concept and have a negative effect on motivation. Findings relating to the effects of feedback elaboration and types of elaboration are varied. Mason and Bruning(1999) have advised that generalisations of the results of elaboration could not be made but are dependent on many variables including the nature of the computer system, the subject being studied and the feedback as well as the skills being measured.

Gugerty(1997) suggests that if an intelligent tutoring system (ITS) is to provide stand-alone tutoring then it will be required to handle all aspects of tutoring "including expert problem solving, student diagnosis, tailoring instruction to changing student needs, and providing an instructional environment". This type of system makes inferences from the student's behaviour to determine the users conceptual knowledge and procedural skills. These systems are 'glass box' models in that they contain a detailed model of human though processes which enables specific and accurate diagnosis of student knowledge and misconceptions. However these traditional ITS are expensive to develop. A non diagnostic ITS is unable to automatically adapt to the needs of the user but would require the interjection of the teacher. Diagnostic systems rely on tests to assess mathematical ability. However Drake (2002) suggests tests are not the best tool for assessing mathematical ability. Hence a tool which does not require tests to determine ability but adapts to actual performance should be a more effective learning tool.

2.6 Review of learning systems

In reviewing the use of ICT to support literacy and mathematics relating to the Key Stage 2 Strategy the TTA(2004) report that "the gap between the best and weakest applications of ICT continues to widen." The following systems focus on at least some of the basic algebra and numeracy skills required by undergraduates.

1. Sheffield on line key skills (main contributor for HE supported by TLTP3 funding)
2. MathsBytes Numeracy for Trainee teachers (book and CD ROM aimed at students of initial teacher training)
3. Mathtutor (supported by FDTL4)
4. Equation (produced by Aczel)

2.6.1 Sheffield Hallam University Key skills online

Gillespie (2003) and MacKenzie(2002) report that some HEIs are using learning materials for study skills which include materials to support the development of key skills, such as the Sheffield Hallam University Key skills on line system which includes a Numeracy aspect. MacKenzie(2002) reports that this system is "based on assessment of skills, identification of skill shortfalls and building confidence" however "these initiatives come under the heading of 'helping students with a problem' but may not go far enough".

2.6.2 Keybytes

A book and CD ROM aimed at the numeracy requirements of the national test for all newly qualified teachers. Hence the areas of study are interpreting and using statistical information and using and applying general arithmetic encompassing time; money, proportion and ratio; percentages, fractions and decimals; measurements; conversions using formula and between fractions, decimals and percentages; averages and simple given formulae. The CD ROM aspect only provides assessment opportunities based on the final answer.

2.6.3 Mathtutor

This resource has been produced as part of the project "Mathematics Support at the Transition to University" funded by HEFCE FDTL4. One aspect of the resource is algebra which consists of various subcategories as detailed in Table 2.6.1.

2.6.4 Equation

Linear algebra problems are represented graphically as balance problems depicted as a game-like balance model. Aczel (1998) explains that this model introduces algebraic notation as a convenient abbreviation, which enables negative signs and negative answers within the model; and it then promotes algebra as a tool for solving word problems. It also logs what students see on the screen, and what they click and enter. This system provides progressive challenge and feedback on the effects of operations. It is possible for students to create, test and improve strategic theories for a wider range of transformation and representation problems. Evaluation of the use of the system by a group of KS3 pupils presented by Aczel (1998) suggested that the system was an effective learning tool.

To evaluate these packages for learning the following criteria have been selected:

- Style of learning and appearance
- Target audience of intended users
- Curriculum coverage
- Interactivity and feedback for learning and assessment

Table 2.6.1 Algebra Learning Systems

	Style and appearance	Curriculum	Audience	Interactivity
Sheffield Hallam On line	Appearance of a text book on line. Notes are provided for guidance which include example questions and solutions. Limited opportunity to personally undertake solutions to set questions. Answers given without explanation. Web-based clear and easily navigated interface with some dynamic links providing more details if required.	Number skills, Interpreting data, Decimals, Fractions and Percentages, Formula and Algebra, Shape and Space, Calculators & IT	HE undergraduates All learners receive the same guidance and feedback	None. The learner is required to read the guidance as would be presented in a text book, to undertake questions and then mark these with the given answers. Limited set of questions posed regardless of user performance.
Keybytes	Answers are responded to by YES/NO however the correct answer and an explanation of how this is calculated can be discovered on completion of the test. A calculator is provided for on screen use. Final answer only is expected to be input. No assessment of method or partial answers is included. Clear easily navigated interface. .	Mental arithmetic – adding, subtracting, multiplying, dividing, time, fractions, decimals, percentages, ratio Long Questions – formulae, statistics, tables, graphs, charts, cumulative frequency, percentiles, quartiles and interquartiles, box and whisker diagrams	Trainee teachers undertaking QTS Numeracy test All learners receive the same guidance and feedback	Feedback is provided to the assessment tasks but it is standardised whatever the error incurred by the user. Not responsive to the type of error or misconception made. Same questions are posed regardless of user performance. Does indicate time taken and time remaining.

	Style and appearance	Curriculum	Audience	Interactivity
Mathtutor	For each topic there is a diagnostic test consisting of five questions, a video clip providing a visual presentation of the written text that explains methodologies, written text and a set of questions. The questions are assessed on a right/wrong basis. Questions can be re-attempted having been assessed. Clear, easily navigated, web-based interface with a hierarchy of menus.	Mathematical language, Powers or indices, Logarithms, Substitution, Expanding and removing brackets, Factorising quadratics, Transposition of formulae, Linear equations, Completing the square, quadratic and simultaneous equations, solving inequalities, cubic equations, simplifying and partial fractions, polynomial division	Undergraduates on non-mathematical programmes	Feedback is based on right/wrong however if the wrong answer has been given the correct answer can be displayed by a further click. Methodology is presented in text for screen viewing and hard copy and as video clips however these show the solution rather than discussing possible errors or misconceptions. The same questions are set regardless of user performance in the Diagnostic test or exercises.
Equation	A scale balance representing the two sides of an equation. Attractive, graphical interface providing drag and drop and direct input. Quickly accessible to use.	Linear equations	KS3 pupils and beyond	Users can input partial solutions to equations by adding or subtracting amounts from both sides. There is no feedback for incorrect answers or methods. Complexity of the questions posed does increase with success.

Commercial Products

	Style and appearance	Curriculum	Audience	Interactivity
SAM Learning	Number, algebra, shape and space, data handling Revise and examination practice	GCSE Mathematics	KS 4 Pupils and beyond	Will not allow input of an incorrect answer. The answer selected varies the interactivity.
BBC Skillwise	Fact sheets to give guidance as in a text book then worksheets and answers are provided.	Whole numbers, measure, shape and space, fractions, decimals, percentages, handling data	All learners receive the same guidance and feedback	Interactive game is provided to test addition skills against the clock. Feedback provided is whether the response is correct and correct solution shown.

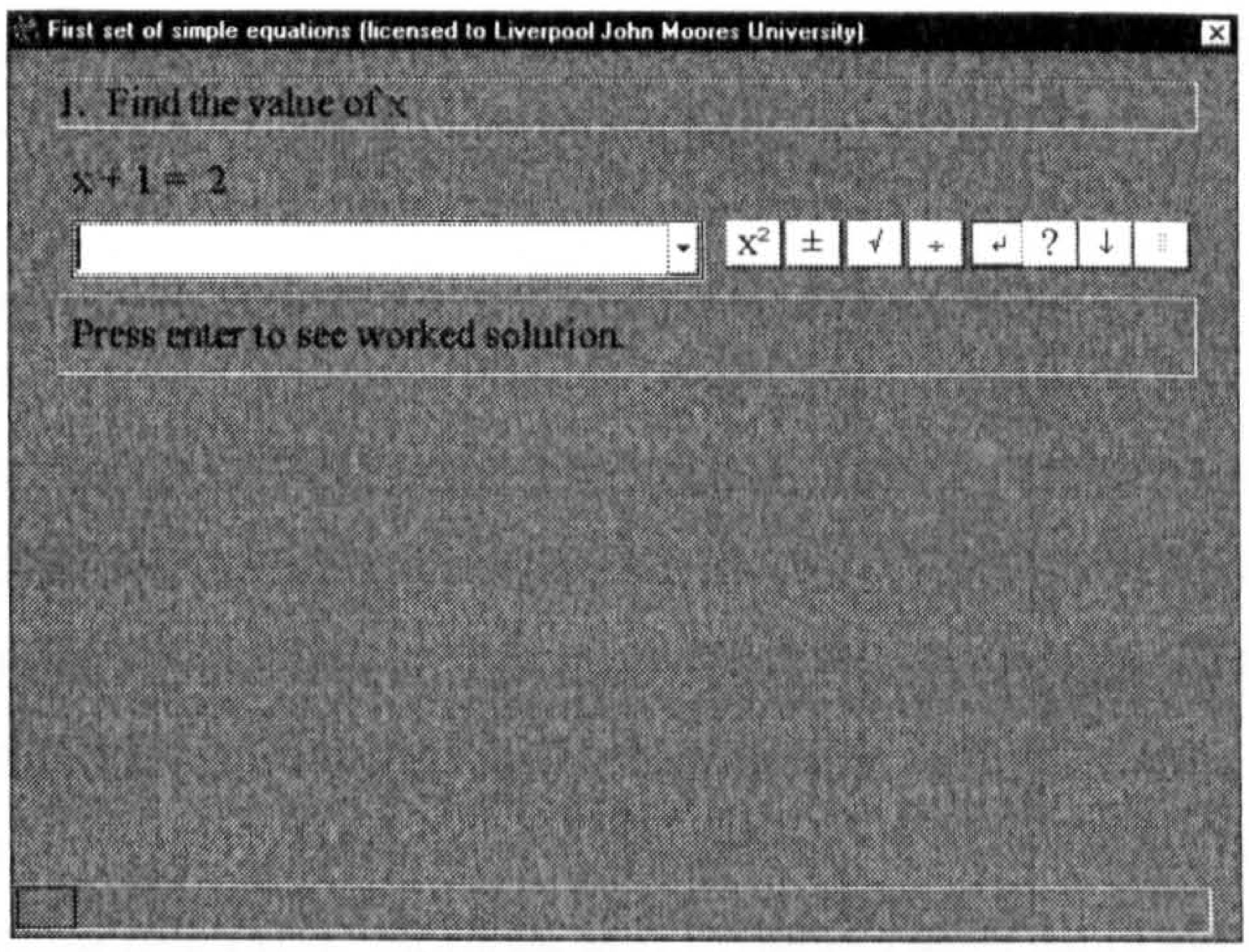
From comparison and analysis of these systems the following common deficiencies for learning systems are evident

- Do not focus on common misconceptions and errors to ensure learners confront these.
- Feedback is not adaptive reacting to the performance of the user but is inflexible and standard for all learners.
- Limited staging of learning such as demonstration opportunities provided for learners

2.6.5 Use of Treefrog

This study is focused on the evaluation of teaching and learning using ICT to enhance and support understanding and achievement. Treefrog, is an ILS which can support the learning and understanding of methods of solution of numerical expressions as well as solving, simplifying or factorising algebraic expressions.

TREEFROG is an algebraic package which can provide hints as well as informing learning with both formative and summative assessment. Educational practitioners in four schools in Liverpool have trailed this software with KS3 pupils. Al-Jumeily & Strickland (1998) have reported and analysed the findings of this focus group by evaluating the learning supported in one secondary school in Liverpool. This study found that there was a significant improvement in the performance of pupils who used TREEFROG as opposed to those whose learning was progressed by use of traditional methods. Questions are determined by the teacher for a group of learners. Feedback provided as hints when the learner is offering an incorrect solution; however this does not vary with the nature of an individual response and consequentially the error made. The teacher can view a log of student performance which details each action taken. Details regarding the format and range of mathematics supported by Treefrog are given in Appendix 9 and in the manual see Treefrog(2000). The software will be adapted to focus on common misconceptions and errors to enable learning to progress and provide feedback which is adaptive reacting to the performance of the user in terms of the error made. The system presented to learners is shown below.



This first question is a worked example, as can be seen from the yellow background in the input area and the advice below that. This means that the pupil can view the supplied solution by hitting the enter key repeatedly. Alternatively, they can type in their own solution a line at a time, hitting the enter key after each line.
The user can obtain help on the function of each item on the screen by moving the mouse pointer over that item. Here is a summary of the actions of the buttons;

- X^2 and ±: insert the special characters for squared and plus or minus.
- √ and ÷ : enable entry of square roots and fractions in conventional mathematical notation, which is then translated into linear form.
- ↵ : is equivalent to hitting the enter key.
- ?: displays the hint; this can also be achieved by hitting the F1 key.
- ↓ : passes on to the next question.
- ⇑ : reviews previous working.

A progress bar at the bottom of the TREEFROG window shows how far through the test the pupil has got; a black-bordered rectangle indicates the position of the present question within the test, and blue blocks indicate successfully attempted questions. Passed questions, and worked examples where the pupil has not tried to enter their own solution, are left in the current background colour (green here). Students are presented with six "tests" or sets of questions to complete. Students could determine the order of the tests selected. Within each test questions were presented in order the user could select to "pass" on a question or review work presented as a solution. Previous questions could not be revisited or work on these reviewed. Performance in all the questions presented contributed towards the trial. For the purpose of this study the only feedback provided by Treefrog is whether the user response is correct and whether it is a valid finishing point for the question posed.

When the test is completed, or the close button on the TREEFROG window is clicked, a summary of performance is presented if applicable.

2.7 Predictive evaluation of software for learning

In the early 1980s educational software was limited, and what little there was consisted mostly of "drill and practice". Teachers often selected software from catalogues and were subsequently disappointed when the software arrived. Twenty years later, ICT is an increasing contributor to the learning process with a diverse range of software and numerous titles available within many curriculum areas. Yet the object of evaluations of software packages in many cases is of a more general 'ticklist' nature and removed from the intended use of the resource. Squires and Preece(1999) reported that many researchers (eg McDougall and Squires) have questioned the suitability of checklists to predict educational issues in all but a naïve and superficial way. They cited many problems which have been identified by evaluators including the following relevant weaknesses:

> Selection amongst educational software of the same type emphasises similarities rather than differences (Squires & McDougall, 1994)
>
> The focus is on technical rather than educational issues (Office of Technology Assessment, 1988)
>
> It is not possible to allow for different teaching strategies (Winship, 1988)

To evaluate the impact of specific computer algebra systems (CAS) on learning we are necessarily concerned with specific educational situation(s). Predictive evaluations focus systematically on 'heuristics' that are justifiably selected pre-determined aspects of 'educational' software rather than the checklist approach. This approach is supported by the increasing dissatisfaction with checklists attempts to develop instruments for use in evaluation. Squires and Preece (1999) developed the usability heuristics published by Nielsen (1994) as follows:

- *Visibility of system status: the system should always keep users informed about what is going on, through appropriate feedback within reasonable time.*
- *Match between the system and the real world: the system should speak the user's language, with words, phrases and concepts familiar to the user, rather than system oriented terms. Follow real world conventions, making information appear in a natural and logical order.*
- *User control and freedom: users often choose system functions by mistake and will need a clearly marked 'emergency exit' to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.*

- *Consistency and standards: Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.*
- *User interface error prevention: even better than a good error message is a careful design which prevents a problem from occurring in the first place.*
- *Recognition rather than recall: Make objects, actions and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for the use of the system should be visible or easily retrievable whenever appropriate.*
- *Flexibility and efficiency of use: Accelerators – unseen by the novice user – may often speed up the interaction for the expert user to such an extent that the system can cater for both inexperienced and experienced users. Allow users to tailor frequent actions.*
- *Aesthetic and minimalist design: Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.*
- *Help users recognise, diagnose, and recover from errors: Error messages should be expressed in plain language (no codes), precisely indicating the problem and constructively suggest a solution.*
- *Help and documentation: Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large."*

Research has shown that the use of these heuristics by five expert evaluators of educational software will typically lead to the identification of about 75% of the design problems associated with a package (Nielsen, 1992). Squires and Preece (1999) related Nielsen's usability heuristics to socio-constructivist criteria for learning in terms of credibility, complexity, curriculum and ownership in order to propose a predictive evaluation tool of 'learning with software heuristics' which considers usability and learning issues. The effect of the usability, the accessibility of software, the ease with which it can be used, can be influential in the actual use of a system as reported by Monaghan (1999) and Cox et al (1999).

Blake et al (2003) suggest that the participation of software developers within the evaluation process is beneficial as they can then appreciate the problems that end users have and how they use the resources. From their experience they also recommend that "it is more appropriate to focus on broader learning and educational advantages once the usability issues are resolved". Accordingly it could be considered that the application of these heuristics by software users in the development of subsequent educational software can be beneficial in determining effective design attributes.

2.8 Summary

There is considerable evidence of the nature of misconceptions within Numeracy and algebra at secondary school levels. Despite the widespread recognition of the problem amongst adults there is a lack of specific detail in identification of the errors and misconceptions applied to learners within the post-compulsory and higher education sectors.

In the context of lifelong learning and a widening of participation in education economical, flexible tutoring is sought. The use of computer technology in supporting learning is developing. The notion of "effectiveness" has come to be recognised as fundamental within this strategy. Investigations have shown that there are potential benefits in using ICT within mathematics. However evaluation criteria illustrate how these benefits may not be associated with every software application.

Roscoe (2003) states

> "if e-learning is to become a key and effective component in higher education then there are many issues that need to be addressed ... issues around tailoring learning to the learner".

He proceeds to suggest that

> "perhaps e-learning packages could include a front-end which as well as testing for learning style could include exercises for strengthening some that are weak in some learners".

However could this be adapted to consider whether the package should be able to recognise the solution method of learners and help understanding of alternatives?

And

> "Quality is an important issue in higher education today ... in an e-learning context we need to be concerned with the quality of the materials, the quality of the teaching and learning and the quality of the assessment and feedback and record keeping of progress"

Computerised algebra systems are able to provide consistent learning opportunities. Treefrog currently monitors every action the learner undertakes however this data is currently only used to establish the score of the associated performance.

Hence there could be a beneficial role for ICT in learning if

- Flexible and adapts to the learners needs

Research suggests that an efficient and effective life long learning tool should

- match the curriculum
- Motivate learners and develop self esteem

- Provide feedback in learning

Hence to develop such a tool we would need to identify the needs of

- the curriculum
- learners

Chapter 3 Methodology

3.0 Philosophical Approach to Methodology

In an attempt to rationalise identity, professional standpoint and ideology in relation to the research, this researcher examined her strategic approach to learning. It was anticipated that such a reflexive process would enable further exploration of my learning strengths and gaps. Kolb (1984) has linked theory to practice and developed a 'cycle' to describe these phenomena. The Kolb cycle which acknowledges the early work on experiential learning by others in the 1900s including Rogers, Jung and Piaget is now widely acknowledged by academics, teachers, managers and trainers as fundamental concepts towards our understanding and explaining human learning behaviour. Kolb's learning theory infers that learning processes undergo transitions involving Concrete Experience (feelings), Reflective Observation (watching) Abstract conceptualisation (thinking) and Active Experimentation (doing) as a cycle. Furthermore these four distinct learning styles on which the cycle is based are preferences and hence provide a method to understanding individual people's learning styles. Hence in utilising this model to understand the researcher's learning style it was anticipated that an established or preferred way of learning would indicate this researchers preferred reasoning strategies (inductive or deductive) and hence the research style. This knowledge would assist in the development of a methodological approach to this research. Any gaps identified in the researchers learning style which would be appropriate to the problem investigated could then be modified and capitalised upon. However, there is no single style that identifies the learning style of the individual. The learning style of each individual is a composition of these four basic styles, which are 'Accommodator', 'Assimilator', 'Diverger' and 'Converger' (Askar and Akkoyunlu 1993).

Honey and Mumford (1992) developed the learning styles inventory. This researcher was subjected to a styles assessment, administered at John Moores University, Liverpool, England. This researcher was scored and the test determined a learning style of accommodator. Thus this researcher, according to the adaptation of Kolb's Learning Styles inventory, combines 'feeling' of Concrete experience (CE) and 'doing' Active Experimentation (AE) attributes as a preferred style. The combination of CE and AE placed this researcher into a quadrant of the model titled Accommodating (reflecting and doing). Perry and Ball (2004) has explored this style identifying "their greatest strength as doing things, carrying out plans and experiments and involving themselves in new experiences ... They like working with others. This group recorded the highest scores found here for Logical-Mathematical Intelligence and for Spatial Intelligence. They had the lowest scores for Linguistic Intelligence. The strength in Logical-Mathematical Intelligence can be described as the enjoyment of ordering, categorising, calculating, experimenting, stating hypotheses and inferring consequences, conducting research, analysing findings and developing logical argument." According to Perry and Ball (2004)

Mathematics-Science educators favoured this quadrant hence the result for this researcher (previously a Mathematics and ICT teacher) is in compliance with this finding.

Through extending the work of Kolb to further examine strategies of learners and researchers, we can begin to better focus on an analysis of choice of methodology. In examining reasoning and links to the cycle McClelland and Yolles (1997) differentiated research approaches broadly into those that are deductive and those that are inductive. In relating this division of research approach to Kolb's learning cycle McClelland and Yolles (1997) have suggested that the deductive approaches can be seen to correspond to those styles on the left hand side of the cycle and the inductive approaches to those on the right hand side. Polit and Hungler (1995 p9) have defined these types of logical reasoning;

> 'inductive reasoning is the process of developing generalisations from specific observations',
>
> And;
>
> 'Deductive reasoning is the process of developing specific predictions from general principles'
>
> (Polit and Hungler 1995 p9)

Deductive reasoning, or deductivism is linked to empiricism, and can be seen as being concerned with rationality and testing theories through hypotheses, a positivist view. The hypothetico–deductive tradition (that scientific knowledge is preceded by the deductive tradition) is intimately related to positivism and causality. Inductivism therefore, as Polit and Hungler (1995) have indicated, is the reverse of deductivism in that it seeks to construct explanation and theories about observations from an empirical world. The theory is the outcome of induction. The models utilised within inductive reasoning processes rely on stimulus, experience, response, interpretation, meaning and action. The researchers learning style and research style as determined within the scored inventory lay on the left hand side of Kolb's cycle, an deductive approach to learning and research, according to McClelland and Yolles (1997).

Popper (1981) writing of progress in science has suggested that; 'we do not discover new facts or new effects by copying them, or by inferring them inductively from observation... we use, rather, the method of trial and elimination of error' (Popper 1981 p 90). This was suggesting that man must continually try out their hypotheses through trial and error elimination. In generating the resultant theory, Lakatos (1981) in reference to this work of Popper has suggested that to qualify as scientific such theory must; 'predict facts which are novel, that is unexpected in the light of previous knowledge' (Lakatos 1981 p 13). The predominant styles of educational research are positivist expecting and proving as opposed to inductivist.

Brown and McIntyre (1981, p245) advise that

"The research questions arise from an analysis of the problems of the practitioners in the situation and the immediate aim then becomes that of understanding those problems. The researcher, at an early stage, formulates speculative, tentative, general principles in relation to the problems that have been identified; from these principles, hypotheses may then be generated about what action is likely to lead to the desired improvements in practice."

Within this chapter there is a discussion of the theoretical underpinning of quantitative and qualitative research methods, selection of groups and ethical considerations, an examination of the rationale for the chosen methodological approaches, and a personal reflection of the systematic research design selected. This chapter proceeds to an overview of the research design and an extensive discussion of the six-stage methodological approach that was finally undertaken. In addition it outlines the various research tools utilised to examine the main and sub research questions within this study.

3.1 Quantitative methods

> "Quantitative Research is the function that allows us to obtain information and data about activities, events and occurrences in order that we can identify, define, monitor and better understand issues, problems and processes, through quantitative evaluation. The research specifies the data required to address the issues, problems or processes, designates the information and data gathering design and methodology, provides interpretation of the results and presents arguments, discussion and critical evaluation of the findings"
>
> (McClelland, R, 2002)

Many researchers including Polit and Hungler (1995) maintain that scientific research is the most sophisticated method of acquiring knowledge. The use of checks and balances minimise both bias and the effect of the researcher on conclusions made. Robson (1994) describes the methodical process of investigation which enables control and order. The search for an identifiable truth is clearly the focus of such an approach and would obviously influence the chosen methodology. The world view of a researcher is also heavily influenced by such an approach, often recognised to be emergent from a realist ontology, one which Guba and Lincoln (1994) have suggested is often called;

> 'naïve realism. An apprehendable reality is assumed to exist, driven by immutable natural laws and mechanisms. Knowledge of the 'way things are' is summarised in the form of time and context free generalizations, some of which take the form of cause effect laws...the basic posture of the paradigm is argued to be both reductionist and deterministic' (Guba and Lincoln 1994 p109)

Some authors infer that the falsification of theory is the only way to achieve scientific revolution. Popper (1981) writing of progress in science has suggested that;

'we do not discover new facts or new effects by copying them, or by inferring them inductively from observation... we use, rather, the method of trial and elimination of error'

(Popper 1981 p 90)

This was suggesting that man must continually try out their hypotheses through trial and error elimination. In generating the resultant theory, Lakatos (1981) in reference to this work of Popper has suggested that to qualify as scientific such theory must; 'predict facts which are novel, that is unexpected in the light of previous knowledge'

(Lakatos 1981 p 13)

Assessment of learning is widely considered by schools, external examination boards universities and employers to determine student understanding. Interim assessment results which track progress by means of pre and post testing can indicate learning progress. In Mathematics education the awarding of numerical marks for answers to questions is a widely adopted practice. This type of activity will necessarily produce numerical data. Mathematics education research extensively uses quantitative approaches to data analysis. For instance as outlined in Chapter 2 Section 6.5 Strickland and Al-Jumeily (1999) and (2001) evaluated the ability of a computer algebra system in improving learners algebraic manipulative skills by means of a matched pairs groups whose attainment was measured quantitatively and compared by means of statistical methods. Cooper and Harries (2003) analysed the effectiveness of teaching method by comparison of quantitative results from a pre and post test. Ricketts and Wilks (2002) selected the use of quantitative statistical methods on results of achievement in analysing the effectiveness of computer based assessment in improving student performance in numeracy. Aczel (1998) evaluated the effectiveness of a computerised tutoring system for learning linear equations by means of comparison the results of two groups of students who undertook a pre test and a post test as well as interviews with the students. The system was described in Chapter 2 Section 6.4. One group used the computer environment whereas the other did not undertake any tuition in algebra. In the research of Nicol and Anderson (2000) as discussed in Chapter 2 Section 5.2 a pre test and post test are used for comparison of two groups to compare effectiveness of two different methods of instruction specifically in this case one being teacher implemented and the other computer assisted. However other factors could have contributed to a variation in results such as motivational factors of one group using computer technology as investigated by many researchers including Cox (1997) and Passey et al (2004) as outlined in Chapter 2 Section 5.1. This effect of this variation could be eliminated by the two groups using different computer technology. The total mean scores of the two groups were compared by means of quantitative statistical methods. Abidin and Hartley(1998) as discussed in Chapter 2 Section 5.2 evaluated the use of a computer based learning environment for developing problem solving skills in algebra by means of a paper based pre test and post test whereby the answers were tabulated for right or wrong answers and the progression mapped by comparison. In addition through analysis of the wrong answers types of error were identified. This approach was selected in the Undergraduate survey to identify common errors and misconceptions.

3.2 Qualitative methods

> "Qualitative researchers study things in their natural settings attempting to make sense of, or interpret, phenomenon in terms of the meanings people bring to them"
> (Denzin and Lincoln, 1998 The Landscape of Qualitative Research)

Guba and Lincoln (1994) suggest that qualitative data are useful for revealing and 'uncovering' views. In the same way as they described the general principles of a quantitative approach to research Polit and Hungler (1995) also described the underlying principles of a qualitative approach which generally;

- Attempts to understand the entirety of a phenomenon rather than focus on specific concepts
- Has few preconceived hunches, stresses the importance of peoples interpretation of events and circumstances rather than the researchers interpretation
- Collects information without formal structured instruments
- Does not attempt to control the context of the research but, rather, attempts to capture it, in its entirety
- Attempts to capitalize on the subjective as a means for understanding and interpreting human experiences
- Analyses narrative information in an organized, but intuitive fashion.'

(Polit and Hungler 1995 p16)

This clearly illustrates the key differences between the quantitative and qualitative approaches.

Bailey (1997) cites Schwandt (1994 p18) in describing the goal of the qualitative research paradigm as;

> 'to provide a research methodology for understanding the complex world of lived experience from the point of view of those who live in it '

(Schwandt 1994 p18)

3.3 Statistical Methods of analysis

Tuckman (1999 p282) advises that "statistical tests are major tools for data interpretation. By statistical testing, a researcher can compare groups of data to determine the probability that difference between them are based on chance, providing evidence for judging the validity of a hypothesis or inference." The type and nature of the data determines the appropriateness of statistical tests. Parametric tests make assumptions about the distribution and spread of the data whereas non parametric tests do not require normal distribution or equal group variances but are based on ordinal or nominal data rather than more precise ratio data. Nominal data is where numbers are used to classify different groups such as those who use "Webfrog" and those who use "Webfrog with Feedback". Where this classification has an order such as in a Likeart scale the data is ordinal. For instance grades achieved in a GCSE examination or gradings in a questionnaire are ordinal. Ratio data is precise data such as scores achieved in tests where the interval between marks is continuous and the marks are awarded on a comparative basis. Interval data informs on order of data and the intervals or distances between. With ratio data all arithmetical operations are possible and cross scale comparisons as they have absolute zeros and the ratio of stated intervals is the same. Furthermore precise data may be considered as being parametric and hence statistical tests applied which are more powerful than nonparametric statistical tests. In selecting the appropriate statistical test the number of variables must be determined and the types of data. Tuckman (1999 p290) recommends that

> "when you are dealing with two interval variables, use a parametric correlation called Pearson product-moment correlation. When dealing with two ordinal variables, most researchers use a Spearman rank-order correlation. With two nominal variables, they use the chi-square statistic. For a study with a nominal independent variable and an interval dependent variable with only two conditions or levels use a t-test; use analysis of variance to evaluate more than two conditions or more than one independent variable. Finally the combination of a nominal independent variable and an ordinal dependent variable requires a Mann-Whitney U-test (a non-parametric version of the t-test)."

Where groups of nominal data is non parametric and related, Wilcoxon can measure differences in rank and confidence in the result.

Independent variables operate within a study to affect behaviour. An example of an independent variable within this study is the use of two different types of software to support learning. The dependent variable is a response to a variable that is the factor that is observed or measured to determine the effect of the independent variable such as the post-test score after the use of different types of software to support learning.

The following rationales based on the details Wikipedia(2005) last updated in 2005. These principles will be applied to the findings from each of the stages of this investigation to determine which test is appropriate to use. The specifics of selection will be outlined in the following Sections (3.11 - 3.14) in which the methodology of each stage of this study is detailed.

Selection of Statistical methods

Cronbach's Alpha Coefficient (α)

Is a measure of reliability of data, it indicates reliability of treating a set of test items as measuring one variable. α can take values between minus infinity and 1 (although only positive values make sense). As a rule of thumb, a proposed psychometric instrument should only be used if an α value of 0.70 or higher is obtained on a substantial sample. Cognitive tests (tests of intelligence or achievement) tend to be more reliable than tests of attitudes or personality.

Spearman's Rho Rank Correlation

is a non-parametric measure of correlation – that is, it assesses how well an arbitrary monotonic function could describe the relationship between two variables, without making any assumptions about the frequency distribution of the variables. Unlike the Pearson product-moment correlation coefficient, it does not require the assumption that the relationship between the variables is linear, nor does it require the variables to be measured on interval scales; it can be used for variables measured at the ordinal level.

Pearson's (PMCC) product-moment correlation coefficient (r)

is a measure of how well a linear equation describes the relation between two variables X and Y measured on the same object or organism. It is defined as the sum of the products of the standard scores of the two measures divided by the degrees of freedom: If X and Y jointly normally distributed, this can be used to "predict" the value of one measurement from knowledge of the other. That is, for each value of X the equation calculates a value which is the best estimate of the values of Y corresponding the specific value of X.

Significance testing

Wilcoxon Signed rank test

involves comparisons of differences between measurements, so it requires that the data are measured at an interval level of measurement. However it does not require assumptions about the form of the distribution of the measurements. It should therefore be used whenever the distributional assumptions that underlie the *t*-test cannot be satisfied. The Wilcoxon signed-ranks method tests the null hypothesis that two related medians are the same. This test allows you to compare a single median against a known value or paired medians from the same (or matched) sample

Mann-Whitney U-test

Similar to the Wilcoxon signed rank test in that medians are compared and data does not have to be normal and variances do not have to be equal. However data sets are independent from each other. Sample sizes should be as equal as possible but some differences are allowed

Paired T Test

A ***t* test** is any statistical hypothesis test in which the test statistic has a Student's *t* distribution if the null hypothesis is true. The validity of use of this test is based on the assumptions of normal distribution of data and equality of variances.

Charts can be effective and efficient in determining distributions and relationships between variables. Outlier data is that which is inconsistent with the rest of the data set. Many statistical procedures are sensitive to the inclusion of extreme data and hence the rejection of these values has been proposed by many researchers. Although Clegg (1999 p135) suggests that when excluding outliers this must be reported in the discussion of findings as this will effect the value of the coefficient of correlation. In deciding whether to omit extreme values can be clearly identified by means of box plots. A scattergram is a graphical representation of the relationship between two variables. Robson (1999 p336) advises that "scattergrams are a powerful pictorial device, giving a clear picture of the nature and strength of the relationship between the variables".

Testing the nature of data

To ascertain validity of statistical method the linearity of relationships will be investigated by means of a scatter graph, the normal distribution by use of Tukey PP Plot and identification of outliers and extreme values by Box and Whisker Plots.

Data sets from each stage of the investigation will initially be investigated to determine whether parametric tests are appropriate for selection.

3.4 Survey approach

Tuckman (1999 p11) advises that "survey research" is a common approach in education research in which "variables are studied using a simple counting procedure" but that the interpretations of the answers may be misleading without a basis for comparison. But by including a control or comparison group of students who have not had the experience being evaluated, the researcher can discover whether the interpretations of data correspond to the real situation". Hence this would support the view that the research has internal validity. Research methods that draw on ethnographic techniques include participant observation, interviews, survey instruments, (Burns, 1997) are suitable for gathering data regarding opinion and attitude towards techniques and methods and are broadly used in educational research.

Many researchers including Goos et al (2003) have investigated attitudes to technology in learning by means of a questionnaire survey. Hibberd, Litton, Chambers and Rowlett (2004) evaluate the use of online formative assessment mechanisms undertaken by analysis of qualitative results relating to student performance and feedback on the environment. Similarly investigations into the effectiveness of approaches to teaching of mathematics can adopt a qualitative approach such as that by Reynolds and Muijs (1999) whereas Mackenzie (2002) quantitatively surveyed opinions and attitudes towards mathematics. The approach proposed by Squires and Preece (1999) combined the two aspects of learning and use of technology in their evaluation tool of qualitative heuristics relating to both. Carr-Hill (1997) reminds us of the three most persistent criticisms of qualitative research;

- Researchers are subjective and data are biased
- Data collection is uncontrolled and cases have been selected non randomly
- Generalisations are not possible'

(Carr- Hill 1997 p186)

Some researchers translated qualitative data to be quantitative such as Manouchehri(2004) in examination of the impact on learner attitude of using interactive computer algebra software by means of qualitative observation data which was converted methodically to quantitative data. Ricketts and Wilks (2002) investigated by means of quantitative analysis of a questionnaire based survey of users opinion of a computer environment. This researcher selected to adopt a survey approach which combined qualitative and quantitative data based on the use of heuristics which focused on the learning as proposed by Squires and Preece and the software followed by an interview of some participants. This combination approach of analytical approaches enables comparison of findings.

3.5 Mixed methodologies

The difficulty in utilising one methodological paradigm over another is illustrated well by Popper (1981) who despite making reference to the natural sciences clearly illustrates the importance of change and how it occurs;

'very similar organisms may sometimes respond in very different ways to some new environmental challenge'

(Popper 1981 p82)

Use of a variety of qualitative and quantitative methods enables the collection of complementary data from a variety of sources and contexts and the "reduction of inappropriate certainty" Robson (p290) since the use of a variety of qualitative and quantitative methods can support the researcher "rather than focusing on a single, specific research question, they may be different but complementary questions within a study; the Complementary purposes model."

By means of mixed methodologies Lilley, Barker and Britton(2004) evaluated software by means of comparison of performance between Computer Adaptive Tests and Computer Based Tests using statistical tests applied to quantitative data collected by means of a questionnaire and qualitative data from a focus group session after using the software. Data collected was used to investigate perceptions of the adaptive nature of the software and the usability of the interface. They suggested that with the combined methods approach "one of the main advantages of a focus group is the possibility of gathering information about complex or sensitive issues that were likely to be overlooked in quantitative methods employed earlier". The combined approach of quantitative methods with qualitative investigation was selected by the researcher to encompass both the quantitative evaluation of results and the qualitative nature of attitude and 'feelings' toward the research environment. Kramarski and Hirsch (2003) investigated the effectiveness of two different computer systems in mathematical classrooms by means of quantitative comparison of attainment and questionnaire responses and qualitative comparison of behaviours display the appropriateness of data and tools for different purposes.

3.6 Selection of groups

3.6.0 Randomised Controlled Trials

The main features of a randomised controlled trial are the randomisation process and the presence of a control group. The rationale for selecting this approach is outlined below.

3.6.1 Randomisation

The use of this technique should aim to produce a representative sample, typical of the population under investigation such as learners of a programme of study. Therefore, care must be taken in the method of randomisation employed, with 'pseudo' methods of randomisation avoided, such as students who are all from one subject area, because bias can arise if such an approach to deciding which group of a trial participants are allocated is used. Biased sampling occurs when the sample chosen fails to represent the true make up of the overall population of interest. For example, if only IT students users are asked to volunteer in the study they may be more enthusiastic and motivated to use a computer based system than those from other disciplines, consequently, unrepresentative. A biased sample relates to those in which some individuals had a greater or lesser chance of being included. Random sampling should overcome such bias, by ensuring that every individual within the target population has the same chance of being chosen to take part. Through this randomisation all extraneous variables of ability, gender, main programme of study will be evenly distributed across the control and experimental groups. Therefore, the rationale behind randomisation is to attempt to produce equal groups in terms of participants is adhered to.

3.6.2 Use of a control group

Individuals in control groups will either not receive the intervention under investigation, or will receive something else, often 'standard care' applied to the specific problem or situation. The control group used will be as closely matched as possible to the experimental group to avoid any confounding factors interfering with results produced, such as ability, age, gender, social class. This allows for a comparison between control and experimental group. Any differences exhibited between the two groups should relate to the effect of the independent variable, as long as the groups are of a similar makeup at the outset.

3.6.3 Allocation to Treatment Group

There are three principle means by which an individual may be assigned to an experimental or control group during a trial: independent sample design, matched pair design, or repeated measures design.

3.6.4 Independent sample design

"For this design a group of participants is obtained for the experiment as a whole, and then individuals are allocated randomly to one or other of the experimental conditions" (Robson, 1994: 17). Such a design involves individuals being exposed to different conditions of an investigation. However variables such as gender, subject of study, and ability may influence outcomes. As these factors are randomly distributed the results may not be comparable.

3.6.5 Repeated measures (or within subjects) design

Participants in this approach undergo both sets of conditions under investigation, i.e. all participants are involved in all conditions. The problem with this design is that of 'order effects', i.e., the results are influenced by the sequence in which conditions are administered. A learning effect may ensue when a participant is involved with two conditions, so that whatever is carried out second tends to receive a higher score, which can not be directly attributed to the independent variable. Alternatively, a 'fatigue effect' may occur, giving a negative overall result, so that whatever is carried out second will tend to receive lower marks. "In either case we cannot make unambiguous statements about the experimental effect, i.e. as to whether condition A or condition B produces better results, because what we actually measure is the combination of the experimental effect and the order effect" (Robson, 1994: 20). To overcome this difficulty, 'counterbalancing' may be adopted, so that half of the participants undergo condition A followed by condition B, and the other half are involved with condition B followed by condition A

3.6.6 Matched pairs design

When using this design, participants are usually matched into pairs and then allocated randomly to either the control or experimental condition. A researcher may decide to carry out this form of design if they feel there is a third variable that they suspect could affect the dependent variable. For example, if gender was thought to affect outcomes measured, the research would match people according to their gender and then randomise participants to an experimental or control group in relation to these paired groups (one of the pair going to the experimental group, the other to the control). Such an approach helps to ensure that this potentially moderating variable is controlled, by being equally distributed within each group. It is therefore less likely to interfere with the relationship between the independent and dependent variables.
To enable use of this design the participants will be required to undertake a pre-test and post-test as well as undertake the investigation.

Achieving validity in research is essential but not an easy task. Selection of method by means of philosophical intention can support this aim. A study must have internal and external validity. Internal validity is achieved if the outcome achieved is a result of the approach being tested rather than of other causes. Ideally there would be only one variable and hence all change could be attributed to the change in this variable. It is of course in some learning situations impossible to fully apply the rules of internal validity. By utilising matched pair groups the researcher was limiting the affect of the variation in the subjects. External validity is within a study if the findings have generality that is they would be obtained if the results would be achieved again in similar circumstances hence the selection of 'typical' learners with no particular bias.

3.7 Ethical considerations

Ethics is important for educational researchers as the focus of their studies is learning. The experience of participation in the research should not cause anxiety or deter the subject from learning. The subjects should participate freely and hence informed consent must be sought. This researcher followed the ethical research guidelines of John Moores University. All the students involved in the pre pilot, Undergraduate and final trials were advised verbally in detail and this was then confirmed in writing about the intended use of the research data, their right to withdraw at any time, their right to privacy and their right to anonymity before deciding whether to participate at any stage in the research. Students were reassured that much of the data would be analysed as group data and individuals would not be identifiable other than to the researcher. Signed agreements verifying the participants understanding of their rights were collected from participants at all stages of the investigation see Appendix 6.

3.8 Rationale of the selected methodology

This investigation considers issues related to the establishment of the research questions, whilst examining the challenges, justifications and rationales that underpin the selected methodical approaches undertaken within this study. The studies main question of "Can an Intelligent Learning System (ILS) effectively tutor learners beyond compulsory education in numeracy and algebra?" is subdivided in two parts (question 1 and 2 in Chapter 1 Section 1.2)

1. What common errors and misconceptions should be anticipated? Are these transferred and maintained from secondary school mathematics?

2. Which factors could enable software to be effective for learning?

The first sub-question is approached through a triangulation of three areas of concern. These are the 'published findings, 'teacher opinion', and the 'learners', each containing a further sub question, that relates back to the main research question. According to Robson (1999) triangulation "is an indispensable tool in real world enquiry. It is particularly valuable in the analysis of qualitative data where the trustworthiness of the data is always a worry. It provides a means of testing one source of information against other sources" (1999, p 382).

The need for such an approach was considered vital to this study, due to the three levels of interpretation being closely interrelated. For example, the initial basis for the research stems from the 'published findings' (government agencies and academics) viewed as a means of establishing the overarching background and context for the investigation. This involved clarification of the reported common errors and misconceptions of secondary school learners. In examining this position, a further factor is dependant upon gauging the teachers' opinion of learners' knowledge, understanding and skills of learners within secondary education (the second level of interpretation – 'teacher opinion'). As a result, the third level of investigation addresses the actual knowledge, understanding and skills of learners in post compulsory and higher education (the third level, the performance of 'learners') having progressed from secondary. Consequently, this final phase acts as the ultimate benchmark against which published findings and professional opinion can be judged in relation to progression and development of learners. The quantitative nature of some of this data enables the factual study of the relationship of one set of qualitative facts against another.

The second sub-question is evaluated qualitatively and quantitatively by means of a questionnaire based on learner experience the criteria of which were informed by the findings of Squires and Preece (1999) as outlined in Chapter 2 Section 2.7 and further probed within a group interview context. This approach was selected due to the nature of the data required as Bell (1992) rationalizes "researchers adopting a qualitative perspective are more concerned to understand individuals' perceptions". In addition these findings could be tested quantitatively by comparison of performance of learners having used a system with the proposed features with the performance of a control group. The design of the studies will ensure that the additional questions (stated in Section 1.2)

> Q3. Is the GCSE Mathematics grade achieved appropriate for base line assessment of learner ability?
>
> Q4. To what extent will those who were low attainers at school become anxious when required to use mathematical skills and understanding in post compulsory and higher education?

can be tested.

Research methodology processes and procedures need to demonstrate that they have both a systematic structure, (Robson 1999, Moore 2000, Lloyd 2000) and be able to satisfy objectives of 'transferability, credibility, dependability and conformability' (Robson 1999). The methodological processes within this study were constructed to ensure that the key objectives set out by Robson (1999) became an integral component of the research design. It is argued the theoretical framework within this study can be readily reassigned to other educational settings (transferability), whilst at the same time taking note of the real world settings, themes and dilemmas associated with this area of research. Thus, the triangulation of the three areas of concern (published findings, professional opinion and practice, and the learners) recognises the inter-relationship and complexity of mathematical understanding beyond compulsory school education. As a result, each of the six stages of the study outlined later in this chapter reflect different 'real world' aspects of the current position of learning in post compulsory and higher education. In undertaking such an approach, the study offered opportunities to interpret both isolated, and progressive perspectives on a wide and varied range of issues concerned with e-learning in post compulsory and higher education.

In relation to 'credibility' of research design, this was achieved through the involvement of learners, teachers, examiners and academics, triangulation of the three levels of concern, evidence of briefing, informed consent, and confidentiality of subjects. Findings from the initial examination of published literature, through to the analysis of secondary school teachers' opinions, and finally performance of learners in a Undergraduate and final study were presented. Equally from the adaptation of published criteria for software evaluation through to the canvassing of learner opinion through questionnaires and a group interview is clearly evidenced. As a result, it is argued that the 'dependability' and accuracy of the studies research processes are, systematic, clearly evidenced and sufficiently rigorous.

In addition, consideration of bias and trustworthiness of the chosen methodologies has been addressed at each of the six stages of the research design. This has been examined with particular reference to recognition of the potential for both researcher and subject bias. In regard to Robsons fourth key objective of 'conformability', the research design examined within this chapter enables researchers and professionals external to the study the opportunity to follow and replicate the chosen methodology, within a systematic and cohesive design structure. This would enable the findings to be checked against the different levels of analysis of published findings, professional opinion and practice, and learners' performance, as well as providing a level of transparency and justification of the studies outcomes within the wider context of e-learning. It would be expected that another researcher following these methods would get similar results.

According to Robson (1999), entering into any kind of investigation involving other people is necessarily a complex, and sensitive undertaking and to do this you need to know what you are doing. In exploring the methodological approaches within this study, it is contended that the research design stands the test of Robson's (1999) four key themes of transferability, credibility, dependability and conformability. Hence internal and external validity outlined in Section 3.6 is assured. Within this context, the next part of this chapter proceeds to a personal, and critical reflection of the rationale behind each of the aspects of the chosen research methodology. It is envisaged this will offer a helpful insight into the many decisions that were taken in arriving at the final six stage methodological approach.

3.9 A personal reflection on the research design

In designing this study, there were many issues that had to be considered prior to arriving at an appropriate methodological approach.

In attempting to arrive at a strategy that would enable analysis of the research questions it was important to initially identify a baseline position of post compulsory learners misconceptions and errors within Numeracy and basic algebra. In order to address the main research question of "Can an Intelligent Learning System (ILS) effectively tutor learners beyond compulsory education in numeracy and algebra?" a first step was to identify the difficulties, the common errors and misconceptions. Following consideration of this question various sources were identified as worthy reference for difficulties related to secondary school learners however during the literature review process a lack of sources relevant to post compulsory and undergraduate learners became apparent. In order to determine these difficulties firstly the findings relating to school learners were clarified and the question to investigate the nature of these errors and misconceptions amongst undergraduates adapted as to how these compared with those reported at compulsory school level to enable an investigation of progression of understanding.

In addition the consideration of presenting challenge at an appropriate level as outlined in Chapter 2 section 2.7 and section 3.2 was fundamental to the study. The second aspect of this study was essential to inform upon the requirements of software to effectively support learners in post compulsory education and undergraduates who displayed these misconceptions and errors. Through the experience of using software to solve questions encompassing anticipated common errors and misconceptions actual performance could be recorded, measured and perspectives developed. Hence this would enable examination of the main research question through all the sub-questions.

3.10 The methodological approach selected

This research consisted of a survey, an undergraduate study and an investigation with learners in post compulsory education. The undergraduate enquiry consisted of two aspects

- Identification of common errors, misconceptions and methods of solution
- Clarification of characteristics of software to support learning

The final exploration focused on the effectiveness of the software in supporting learning as well as providing additional opportunities for investigating common errors and misconceptions.

The following methodologies and stages were used to collect data relating to these two aspects

- o Electronic Statement Questionnaire survey — Stage 1

- o Undergraduate and Preliminary pilot study
    - o Pre-Test — Stage 2
    - o Test — Stage 3
    - o Software Evaluative questionnaire — Stage 4

- o Group interview — Stage 5

- o Final trial
    - o Matched Pairs Trial and Post Test — Stage 6

Design of experiments overview

Stage 1 was an electronic survey of statements from QCA reports which was circulated to a web-based community of Mathematics teachers to enable comparison of opinion between practicing teachers and the QCA. Teacher responses were provided using a Likeart scale indicating level of agreement.

Through Stages 2, 3, 4 and 5 the mathematical ability and performance of a sample group of undergraduates as well as their opinion on criteria for software to be "effective" for learning was gathered. The software criteria findings from stages 4 and 5 and the errors made in Stage 3 informed the feedback included and the design of the software used in Stage 6. In addition the performance of participants in Stage 3 related to specific weaknesses enabled the comparison of common errors in learners beyond compulsory education with those reported for Key Stage 3 learners. Furthermore the results of the Pre Test in Stage 2 would be used to compare with the GCSE grade attained to enable investigation of the effectiveness of this as an indicator of mathematical ability.

In Stage 6, the Final trial through use of two computer systems "Webfrog" and "Webfrog with Feedback" by a group of learners in post compulsory education the effectiveness of feedback based on common errors and misconceptions was investigated. This trial required the use of a Pre Test to establish matched pair groups. The inclusion of the Post Test provided the findings to compare the performance of both groups before and after use of the software as well as comparison of the two groups after use of the two different computer systems. In addition the performance in the Pre Test would be used to compare with the GCSE grade attained and the specific weaknesses reported for Key Stage 3 learners.

The processes for each of these aspects can be described by diagrams 3.10.1, 3.10.2 and 3.10.3.

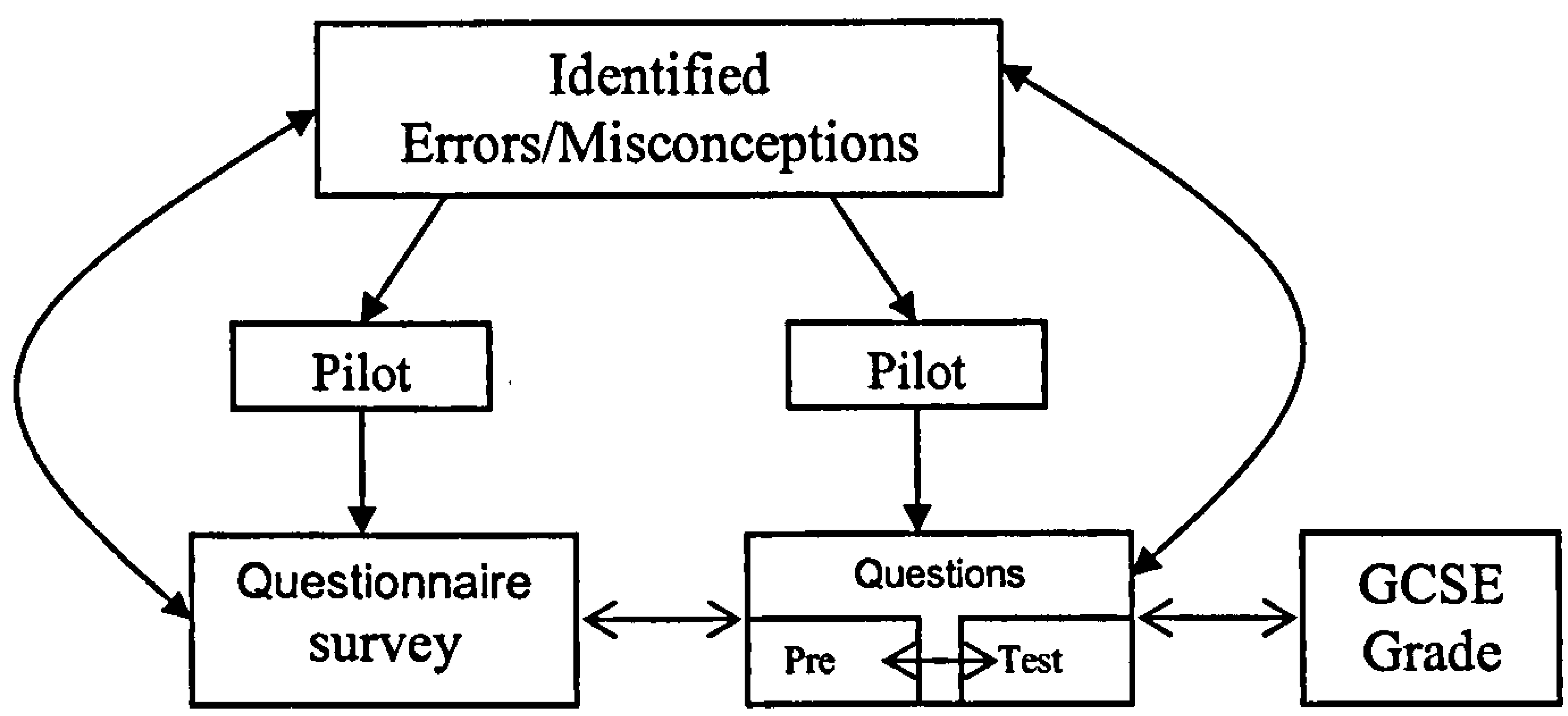


Diagram 3.10.1: Stages 1-4 Process of identification of common errors, misconceptions and methods of solution

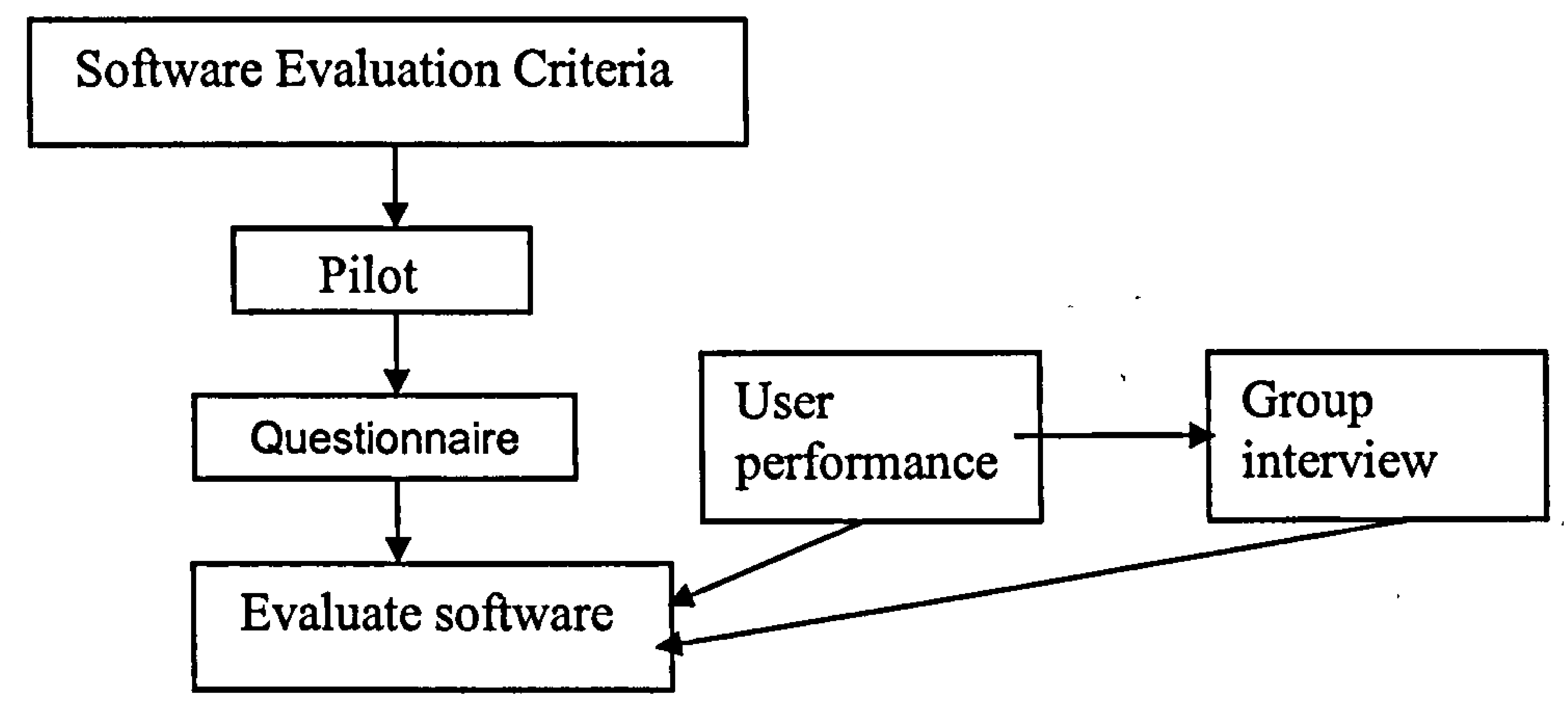


Diagram 3.10.2: Stages 4 & 5 Process for clarification of characteristics of software to support learning

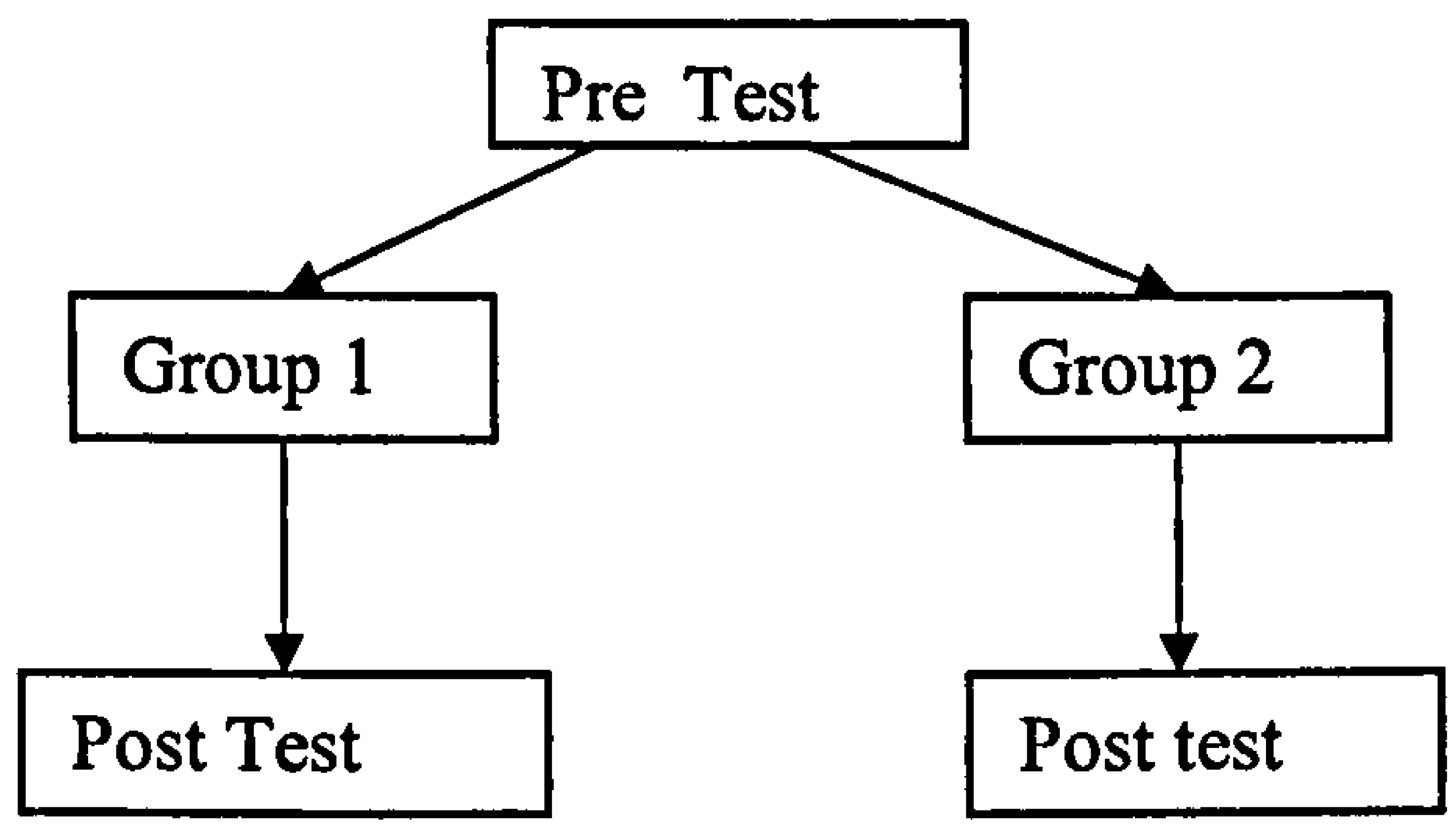


Diagram 3.10.3: Final Trial (Stage 6) Investigation of effectiveness of software

The explorations shown in diagrams 3.10.1 and 3.10.2 provide triangulation of the findings. Delamont (1992) highlights three different types of triangulation, however this study adopted the 'Between Method' as it involves drawing on evidence from more than one source based on the Quality Curriculum Agency (QCA) Examiners Reports; statement questionnaire (Stage 1), and responses to mathematical questions (Stages 2 and 3) and a software evaluation questionnaire by participants (Stage 4) and a group interview with a target group (Stage 5). The triangulation being depicted in diagrams 3.10.2. Comparing the findings from these different sources enabled a broader and more varied investigation. The electronic questionnaire was distributed to a community of secondary school mathematics teachers through the use of an e-learning mathematics support network. The aim of this enquiry was to investigate the findings of the QCA Statutory Assessment Test (SAT) Examiners by canvassing opinion relating to teachers' experience of learners abilities in areas of perceived strength and weakness. Responses to mathematical questions provided numerical data. This type of quantitative research produces data that is 'objective' in that results attained are independent of the researcher and are not the result of undue influence on the part of the researcher (Brown & Dowling, 1998). The group interview and software evaluation of Treefrog resulted in qualitative data representing participants opinions. Furthermore quantitative data relating to evaluative heuristics will be collated to provide evidence related to opinion on software requirements.

The objectives of these initial investigations (Stages 1-5) were to

- Compare areas of Numeracy and algebra which are problematic for many learners within post compulsory and higher education with those reported for secondary school pupils
- Explore learners attitudes towards interaction with mathematics learning software
- Appraise the accuracy and usability of the GCSE grade in determining individual understanding

Within chapter 4 the findings related to Stage 1 the investigation of school teachers' opinions relating to QCA statements are presented. Chapter 5 is concerned with Stages 2 and 3 and Chapter 6 with Stages 4 and 5.

The findings from Stages 1-5 relating to the objectives given above were then used to inform the final trial (Stage 6) as stated in Chapter 7. The purpose of the final trial was to evaluate the effectiveness of an intelligent learning system "Webfrog with feedback" by comparison with a computer assisted learning system "Webfrog" in terms of attainment and user perceptions of using the software. Webfrog being an online rather than server based version of Treefrog. The adaptive nature of the feedback of "Webfrog with feedback" being based upon the common errors and methods of solution identified in the earlier stages of the investigation. Further enquiry regarding the accuracy and usability of the GCSE grade achieved in compulsory education was also enabled. Subsequently in Chapter 8 there is a comparison of findings relating to the Undergraduate and Final Trials to test the hypotheses stated in Chapter 1 Section 1.2.

Chapter 4 Stage 1: Electronic Questionnaire Survey

4.0 Introduction

This electronic survey based on statements relating to aspects of Numeracy and algebra from QCA reports was circulated to a web-based community of Mathematics teachers. The objective of this enquiry was to establish areas of corroboration, disagreement and indecision regarding aspects of Numeracy and algebra identified by the QCA Examiners reports as being a strength or weakness with the opinions of Key Stage 3 mathematics teachers.

4.1 Sample Group

The questionnaire was responded to by a sample group of eighteen. Some of the participants represented individuals and others whole departments within an institution. The questionnaire was circulated to a web-based community of Mathematics teachers. No further details other than the uniqueness of each questionnaire received were obtained about each of the respondents in the sample group. Responses were collected electronically as online replies. This method of distribution and collection was efficient.

4.2 Participant Bias

As teachers of mathematics electing to respond to a blind request it was possible that there may be a participants' bias to disagree with the findings of the QCA in an attempt to discredit this government appointed body. Alternatively it could be possible that teachers unconditionally agreed considering the original source to be infallible. This possibility was minimised by evaluating responses for an extreme approach.

4.3 Selection

This was a randomly selected group whose involvement derived from blind canvassing. None of the participants were known by the researcher. As members of the online community these teachers may not be wholly representative of the teaching body. However being IT competent or not should not indicate a difference in awareness of pupils strengths and weaknesses in algebra and Numeracy.

4.4 Expectancy

Responses to the questions are coded using a ordinal scale to enable quantitative analysis and hence minimise researcher bias and to standardise answers. The use of this scale distinguishes order but not the extent of the difference between for instance "Agree" and "Strongly Agree". The respondents were asked to indicate their opinions and views derived from their experience of teaching and learning mathematics using the following grading system using a Likeart scale

Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree	No Opinion
1	2	3	4	5	X

4.5 Ethics

The identity of all participants has been concealed throughout the study so no school or individual could be associated with any specific responses.

4.6 Piloting the questionnaire

As Bell (1992) indicates "care has to be taken in selecting question type, in question writing, in the design, piloting, distribution and return of the questionnaires." The initial draft of the questionnaire was distributed to a small group of five prospective mathematics teachers and initial teacher training mathematics tutors for evaluation of the questions posed and their format. Respondents were asked to consider as suggested by Bell (1992)

- How long did it take you to complete?
- Were the instructions clear?
- Were any of the questions unclear or ambiguous? If so, will you say which and why?
- Was the layout of the questionnaire clear/attractive?
- Any comments?

Time taken to complete the form was not considered an issue by any of the pilot participants. The instructions were considered to be clear. The file format was converted to HTML to aid visual layout. The wording of the questions was already determined and an essential aspect of the survey as outlined below.

4.7 The questionnaire statements

Each year the QCA produces an Annual Mathematics Report summarising the findings of the national SATS. The construction of this questionnaire to survey teachers' opinion regarding KS3 Mathematics was based on this report citing 23 findings relating to general pupil performance when tackling algebraic and numerical manipulations. The questionnaire is situated in Appendix 5.

4.8 Phrasing of the Statements

The statements used within the questionnaire are taken from the QCA examiners annual report. The mapping of these is given in Appendix 2. The wording of these statements generally indicates the National Curriculum Attainment Level (see Appendix 1). However some phrases are more general. Statements 19, 20 and 21 refer to "all levels". Statements 9, 22 & 23 relate to "higher levels" and statement 12 states "lower levels". Rather than indicating actions that pupils are considered to be unable to do, the phrases in statements 5, 7 and 14 are positively constructed. Only statement 12 involves a comparison of two different operations.

4.9 Choice of statements

The statements included directly relate to the findings cited by the QCA and span the associated attainment levels (3 to 8). There are three pairs of related statements, 13 and 14, 15 and 16 and 17 and 18.

Statement 13

Equations which involve only one variable (3k=18, 5k=?) are problematic for the majority of pupils at levels 4 and 5

Statement 14

Equations which involve only one variable (3k=18, 5k=?) are not problematic for the majority of pupils at level 6 and higher

Statement 13 referring to the difficulties anticipated to be encountered by pupils at 'levels 4 and 5' whereas statement 14 refers to the lack of difficulties anticipated by 'level 6 and higher pupils.

Statement 15

A value n being deduced from an equation of the form 26-2n=8 is extremely difficult to solve for pupils at level 6 and lower.

Statement 16

A value n being deduced from an equation of the form 26-2n=8 is problematic to solve for pupils at level 6 and lower

The different responses to these two statements above should indicate the level of difficulty associated with specific types of solution by lower achieving pupils. Whereas the following pair of statements relate to variations in learning and understanding of pupils of differing ability groups.

Statement 17

The solution of simultaneous equations is weak at levels 7 & 8

Statement 18

The solution of simultaneous equations is very weak at levels 5 & 6.

4.10 Statistical Analysis of the Electronic Questionnaire Survey

The findings were analysed by means of absolute and percentage figures to examine the level of agreement shared between the examiners and teachers. A respondent could indicate 'No Opinion' due to a lack of opportunity to experience the basis of a specific statement because of the nature of their particular school environment. For instance there may be restrictive access to either lower or higher attainment levels. Hence for each statement the responses that indicate No Opinion have been removed from the sample data. However a response of Unsure indicates that the respondent has some experience of the statement and hence their response is valid and remains within the sample group. The results will be sorted into Sample Size. To ease the process of analysing levels of agreement 'Strongly Agree' and 'Agree' will be combined as will 'Disagree' and 'Strongly Disagree' However this means the loss of some detail. This broader range of opinion will be used in assessing reliability by analysis of individual submissions to all questions. Data will be inspected for extreme values which could indicate a bias and potentially skew findings by means of modal score per participant.

As the data sample size varies depending on how many respondents returned "No Opinion" the data will be converted to percentages of the sample. Next this data will be sorted into agreement order to enable calculation of levels of agreement to determine consensus of opinion.

Reliability of the data set will not be tested by means of the Cronbach's alpha coefficient as this data relates to personal opinion rather than cognitive tests

4.11 Findings Overview

Table 25.1.1 in Appendix 25 summarises the results for each question by scale category sorted by rate of agreement. A full table of responses for each question from each participant is given in Appendix 14. The statements included in the questionnaire survey are detailed in Appendix 5.

4.12 Statements by level of agreement with QCA findings

The rates of agreement given in Table 4.1.1 elected from the survey results given in Appendix 14 and the full set in Appendix 25 tables 25.1.1, 25.1.2 and 25.1.3 represent those statements whose results indicate the strongest agreement or disagreement.

Table 4.1.1 Rate of agreement in questionnaire statements

			Percentages		
	Statement	Sample Size	Agree	Disagree	Unsure / No
14	Equations which involve only one variable (3k=18, 5k=?) are **not problematic** for the **majority** of pupils at level 6 and higher	18	100	0	0
23	At higher levels pupils need **more opportunity** to structure and organise solutions to problems.	18	100	0	0
5	At level 5 **most** pupils do know a standard approach to collecting together like terms	18	94	6	0
7	At level 6 in re-writing an expression **most** are **successful** if	17	88	0	12

	positive terms only such as 3b+1 = b+ .... and 5a-4 = 2a+2 +......				
22	At higher levels pupils need **more opportunities** to handle questions where little initial mathematical orientation is given	16	88	0	13
2	At level 3: **most** are unable to complete equations of the form ... / 24 = 16	16	88	6	6
6	At level 5 when negative numbers or signs are involved **many** pupils make arithmetic errors such as $4\text{-}2y=10\text{-}6y \Rightarrow 4y=14 \Rightarrow y=3.5$	18	78	17	6
19	Pupils at all levels **need help** with **handling** indices	17	76	12	12
1	At level 3: **most** are unable to complete equations of the form 962 - ... = 476	16	69	25	6
9	At higher levels, when the substitutions, manipulations and work with equations is more complex, **errors** appear in pupils' answers	15	67	13	20
18	The solution of simultaneous equations is **very weak** at levels 5 & 6	18	67	22	11
20	Pupils at all levels **need help understanding** negative quantities in algebra	18	67	28	6
10	At levels 6 & 7, **most** pupils are **unable** to write the required expression from a sequence eg. 1 hut needs 6 matches, 2 huts need 11 matches	18	11	67	22
15	A value n being deduced from an equation of the form 26-2n=8 is **extremely difficult** to solve for pupils at level 6 and lower	18	11	72	17
13	Equations which involve only one variable (3k=18, 5k=?) are **problematic** for the **majority** of pupils at level 4 & 5	18	11	83	6
11	At level 6 **most** were **not able** to substitute accurately a numerical value into an expression	18	6	83	11

These results indicate that 100% of the group agreed with two statements, 94% with another. All of these statements presented in Appendix 4 are positively constructed and refer to the higher ability levels. Three other statements were widely confirmed with 88% of the sample group in agreement. At least three quarters of the sample who replied with certainty agreed with two further statements. Whilst a slightly smaller compliance of at least 67% for statements 1, 9, 18, 20.

Four statements were disputed by the majority of the full sample group. From the full sample of eighteen only one respondent agreed with statement 11 and two respondents agreed with the other three statements. Seven statements caused a variable response from the sample group.

4.13 Analysis of results

From the findings presented in Table 4.1.1 it is evident that the participants of the survey confirmed only twelve of the twenty three statements with more that 65% of the sample in agreement and with six of these being agreed with by at least 88% of the sample group. However eleven statements were not confirmed by at least 65% of the sample group. In fact four statements were refuted by at least 89% of the sample group. The results from the remaining seven statements were inconclusive.

Analysis of survey results

Table 4.1.2 presents the mapping the statements supported and refuted by the sample group as presented in Table 4.1.1 to the relevant weaknesses of division, brackets, indices, substitution, negative signs and values and solving linear equations as identified in Chapter 2 Section 2.2.6 will provide evidence of the sample group opinion related to each of these weaknesses.

Table 4.1.2 Agreement with weaknesses in questionnaire statements

	Statement	% Agree	Related to weaknesses
14	Equations which involve only one variable (3k=18, 5k=?) are **not problematic** for the **majority** of pupils at level 6 and higher	100	Division, Substitution and Linear Equations refuted
23	At higher levels pupils need **more opportunity** to structure and organise solutions to problems.	100	
5	At level 5 **most** pupils do know a standard approach to collecting together like terms	94	
7	At level 6 in re-writing an expression **most** are **successful** if positive terms only such as 3b+1 = b+ .... and 5a-4 = 2a+2 +......	88	Positive terms i.e. not negatives supported
22	At higher levels pupils need **more opportunities** to handle questions where little initial mathematical orientation is given	88	
2	At level 3: **most** are unable to complete equations of the form ... / 24 = 16	88	Division supported
6	At level 5 when negative numbers or signs are involved **many** pupils make arithmetic errors such as 4-2y=10-6y ⇒ 4y=14 ⇒ y=3.5	78	Negatives and Linear Equations supported
19	Pupils at all levels **need help** with **handling** indices	76	Indices supported
1	At level 3: **most** are unable to complete equations of the form 962 - ... = 476	69	Negatives supported
9	At higher levels, when the substitutions, manipulations and work with equations is more complex, **errors** appear in pupils' answers	67	Substitutions supported
18	The solution of simultaneous equations is **very weak** at levels 5 & 6	67	
20	Pupils at all levels **need help understanding** negative quantities in algebra	67	Negatives supported
10	At levels 6 & 7, **most** pupils are **unable** to write the required expression from a sequence eg. 1 hut needs 6 matches, 2 huts need 11 matches	11	Not in study
15	A value n being deduced from an equation of the form 26-2n=8 is **extremely difficult** to solve for pupils at level 6 and lower	11	Negatives and Linear equations refuted
13	Equations which involve only one variable (3k=18, 5k=?) are **problematic** for the **majority** of pupils at level 4 & 5	11	Division, substitution and linear equations refuted
11	At level 6 **most** were **not able** to substitute accurately a numerical value into an expression	6	Substitution refuted

From this mapping results of statements 14, 2 and 13 suggests that the teachers consider division to be a problem at level 3 but not at higher levels. The result for statement 19 indicates the teachers agree that indices are a problem for pupils. Where the process of substitution is inferred in three statements the teachers disagree that this is an area of weakness conversely in statement 9 where substitution is stated but in 'more complex' equations it is supported. However the complexity of the equations could account for this variation in opinion. The view that Negative signs and values as an area of weakness is supported in statements 6, 1 and 20 however the result for statement 15 indicates that the majority of the sample group disagree that expression given including the negative sign is 'extremely difficult' to solve. The results for Statement 16 which directly relates to statement 15 other than the use of 'problematic' rather than 'extremely difficult' given in Appendix 25 Table 25.1.1 are indecisive with 8 in agreement, 6 disagreeing and 4 unsure. This could signify that it is the use of the wording 'extremely' in statement 15 which has led to this result. Hence these findings could indicate that the sample group do support the view that negative signs and values are an area of weakness. Solving linear equations is refuted as an area of weakness in the results of statements 14 and 13. Whilst the outcome of statement 6 could indicate that solving linear equations is problematic. However statement 6 also includes the use of negative quantities whereas statements 14 and 13 do not. In addition statements 13 and 14 as presented in section 4.9 were selected as an opposing pair and the results of these show consistency of opinion. These findings therefore suggest that the sample group do not consider solving linear equations to be an area of weakness.

To summarise these results could indicate that there is a disagreement of opinion between the teacher sample group and the range of opinion given in the literature review in three aspects in that solving linear equations, division and substitution are not areas of weakness. Whereas there is agreement from the sample group with the view that indices and negative signs and values are areas of weakness. The statements given in Table 4.1.1 do not provide evidence to test the notion that brackets are an area of weakness.

The survey data was derived from the opinions of a sample group of eighteen teachers. The questions were based on QCA reports. Teachers' responses were derived from their personal opinion which may not be scientifically based or informed by national reports. The QCA findings are from scientific analysis of the results of nationally set SATS. As outlined in QCA(2003) these findings result from four sources including a detailed analysis of pupils' responses to the mathematics tests, a nationally representative statistical analysis, evaluation data from approximately 400 schools as well as the opinions of Local Education Authorities (LEAs), individual schools and teachers. These reports may not necessarily be seen by teachers and hence would not inform their teaching strategies in the classroom. Furthermore the inconsistency in opinion could support the recommendation of the 2004 Smith Report that to improve standards in mathematics teachers require training. It would be expected that teachers who are inclined to participate in research studies are more likely to be informed about QCA findings. It was not expected that so many of the QCA statements would not be supported by practicing mathematics teachers particularly those who are motivated to investigate interactive teaching methods and share good practice. It could be interpreted that these teachers are more likely to be informed than mathematics teachers in general. Hence mathematics teacher generally could be less in agreement with the QCA than this sample group. A lack of awareness could support and enable the progression of errors from Key Stage 3 into post-compulsory and higher education. Consequently failure to address areas of misconception could cause anxiety for the learner.

In subsequent stages of this study the areas of weakness identified in Chapter 2 Section 2.2.2 of division, brackets, indices, substitution, negative signs and values and solving linear equations will be investigated further. The learner performances within stages 2, 3 and 6 relating to each of these weaknesses will be compared with the views of this teacher sample group as indicated in this survey.

CHAPTER 5 Undergraduate Study Stages 2, 3 and 4

5.0 Introduction

In this undergraduate study a group of learners undertook a Pre Test (Stage 2) using Diagnosys Software, Pilot Trial (Stage 3) using Treefrog Software and responded to a paper based Sofware Evaluation Questionnaire (Stage 4). Stages 2 and 3 were designed to provide data regarding mathematical ability and GCSE grade attained previously as well as the identification of common errors and misconceptions. The findings from Stage 4 would suggest the characteristics that the learners considered would contribute to the effectiveness of software for learning.

5.1 The Context of Undergraduate Trial

Participants in the Undergraduate' trial undertook Stages 2, 3 and 4 in a single two hour slot to enable control of exposure to the software and prevent external influences on the outcomes. Hence factors relating to experiment arrangements were necessarily applicable to Stages 2, 3 and 4 of this study.

Experiment Arrangements

This was during their first week as an undergraduate hence providing an opportunity to determine the ability of students commencing higher education. The artificial nature of this singular session for continuous use until all questions are completed may limit the generalisability of the results. It is possible that the results may incur a possible Hawthorn effect within the questionnaire results, in that, participant react more positively to being included than to the treatment itself.

Due to the computerisation of the instrumentation of the trial pre test and test all participants were necessarily provided with a standard interaction hence eliminating the potential for bias.

The group

To evaluate the success of the tutoring system a random group of ninety newly enrolled trainee teachers from the School of Education was recruited. This size of the sample group was preferred as some students may not make any errors and would provide only limited data with no evidence towards types of error and methods of solution. This diverse student body hails from four routes within a three year undergraduate programme whereby prospective teachers can elect to teach either in primary or secondary school education. The teacher education student community is varied in terms of experience and ability in both mathematics and ICT. Some students may have last studied school mathematics many years ago and once at university it is apparent that many of these students appear not to have actually gained a solid long term understanding of the school mathematics curriculum. The students' previous experience of ICT ranged from those who have studied at A level to those who have never used a computer and included some mature learners who have used a computer in a specific work-based environment. Hence this sampling across a range of routes and within the induction period of new undergraduates should overcome bias in terms of mathematical ability or ICT experience of undergraduates. Following the experimentation time all students are likely to be exposed to various uses of ICT as outlined in DfEE (1998) Annex B, 4/98.

Use of the blind technique whereby participants were unaware of the expectations of the trials minimised the expectancy factor.

Ethics

Ethics were considered throughout the duration of this study. Consent was sought from all participants involved. The consent form made participants aware of the aims of the research and the intended use of the data. All participants were aware that they could withdraw from the research process at any point with no fear of any negative effects. Verbal consent was also sought from all participants throughout the study and specifically in conducting the group interview. The purposes of the work in which they were participating was discussed as well as how the information would be used. Confidentiality and anonymity of all participants was of high importance throughout the research project. All names have been removed to protect the identity of individuals.

5.2 The Preliminary Trial

Objective

To inform the determination of the content and format of a larger scale Undergraduate trial a preliminary pilot trial was undertaken. Consequently the findings from this investigation were analysed to identify required adaptations and amendments to the trial procedure.

The procedure and methodologies

This trial was undertaken with nine undergraduate students randomly selected. Within this group the highest mathematics qualification attained ranged from GCSE Grade C or equivalent to A level Grade B. Prior experience of IT spanned the range of 'none' to 'a lot'. The results for these participants were obtained from one of two sessions conducted in an identical manner. Firstly the purpose and intention of the preliminary pilot trial to inform the Undergraduate was shared with the participants.

What	Why	When & where
Consent form	Ethical Requirement	Beginning of session
Stage 2 Pre Test	Diagnose mathematical ability Compare with GCSE grades	Start of session in test conditions
Stage 3 Test questions	Analyse common errors made Analyse common methods of solution	Using floppy disk In one session
Stage 4 Questionnaire	Evaluate the software used for learning and for usability	End of session

5.3 Stage 2 Pre Test Arrangements

5.3.1 Introduction

To determine the distribution of mathematical ability, previous ICT experience, age and gender a pre-test measurement of understanding was obtained by means of Diagnosys, a computerised knowledge-based diagnostic system. This system has been utilised by more than sixty universities and colleges in the UK and other countries for testing background knowledge of basic mathematics (or other technical subjects). The pre use of Diagnosys enabled a quantitative evaluation of individual students who having already gained a GCSE qualification would have previously studied this area of the school mathematical curriculum. However the use of a pre-test may cause reactive effects. As outlined by Tuckman (1999) participants may have been sensitised by the Pre Test. That is that on another occasion without the Stage 2 Pre Test being undertaken participants may respond to the Stage 3 trial differently.

5.3.2 Content and process

As discussed in Chapter 2 Section 2.3.2 questions and topic areas involved should be selected to be at an appropriate level. If levelled too high this could de-motivate the participants and give evidence only of what a learner cannot do. If pitched too low whereby most participants could successfully tackle all questions then minimal findings regarding the limit and level of student understanding may be attained. Ideally what is required is a test from which we can identify both common areas of success and weakness. The Pre Test was performed by Diagnosys. This software enables a basic mathematics test including areas selected from the following:

Area	Content
Numbers	decimals, negative numbers, inequalities, etc.
Powers (indices)	powers and roots, properties of powers, scientific notation
Basic algebra	expanding brackets and collecting terms, formulae, simple factorization, linear equations, algebraic fractions
Algebraic methods	more factorization and expanding with two brackets, more algebraic fractions, completing the square
Equations	quadratic equations and methods, simultaneous equations
Algebra and Calculus	simple differentiation and integration, max and min, complex numbers, geometric progression
Statistics	Range, mean, simple and conditional probability
Graphs	coordinates, gradient, linear graphs, quadratic and reciprocal graphs
Area and Volume	simple shapes and solids, similar shapes
Miscellaneous	simple trigonometry, Pythagoras, percentages, radians, equation of circle

The objectives in including the use of this software was to

- investigate the level of understanding of specific aspects of mathematics among participants
- compare these results with the previously attained GCSE grade or equivalent
- consider the validity of the findings regarding the ability and understanding from the use of Treefrog, "Webfrog" and "Webfrog with Feedback" the research software.

These findings could provide details regarding the selection of a representative sample group.

For the purpose of this study focusing on lower levels of Mathematical attainment a common subset of the relevant topic areas was specified. Usually when using Diagnosys the areas on which a user is tested depends on their prior attainment in Mathematics. It was anticipated that participants prior mathematical experiences can be classified using the following qualifications:

- GCSE Mathematics
- Access Course
- A level Mathematics
- AS Level Mathematics
- Vocational Qualification (GNVQ, BTEC)

Despite the fact that for each qualification, a suitable starting level can be given it was more appropriate for comparative analysis to offer all students the same entry point.

To determine the ability of the sample group to use and understand the required numerical processes questions were selected from the following topic areas

- Numbers.
- Basic algebra
- Statistics
- Graphs
- Miscellaneous

The first two areas relate directly to the focus of the study, the third identified as a TTA target and the latter two being aspects of the National Curriculum Programme of Study at KS 3 and 4 providing an opportunity to compare and contrast achievement in various aspects. The topics selected for inclusion compared against the lowest level of relevant topics within Diagnosys can be seen as a diagrammatic representation of the hierarchy structure of the software in Appendix 7.

Table 5.12.2 indicates the order of topics and within each the level the number of questions posed where level 1 is the lowest and level 4 the highest. Across the topic areas there was a variation in degrees of difficulty in the questions posed with algebraic methods and equation being available at the higher levels only. Levels 3 and 4 represent mathematics beyond GCSE. The focus of this study is Numeracy and algebra hence these areas represent more than 50% of the test. To determine the exact content of the test a Preliminary Trial was conducted.

5.3.3 Preliminary study effect on Stage 2

The results of the nine participants of the Preliminary study Pre Test by area of mathematics are given in Appendix 25 Table 25.2.1. The data is sorted in ascending order by the total score attained within the full test. Also shown is the corresponding highest previously attained Mathematics qualification and grade. Questions relating to Algebraic Methods and Equations were at level 3 and hence only offered to candidates whose prior mathematical qualification was higher than GCSE. Following analysis of this group the adaptations to the Pre Test are indicated in Tables 5.3.1 and 5.3.2 below. Full analysis of the preliminary trial and rationale for the Undergraduate trial is presented in Appendix 26.1

Table 5.3.1 Alteration of the number of questions per Mathematical area for the Undergraduate trial

Mathematical area	Preliminary trial	Undergraduate Trial
Numeracy	17	15
Powers and indices	3	3
Basic algebra	11	11
Algebraic Methods	5	0
Equations	5	0
Statistics	4	3
Miscellaneous	2	2
Algebra & Calculus	2	0
TOTAL	48	34
% Difference		74%

Table 5.3.2 Number of questions posed per mathematical area

Level	Numbers	Powers	Basic algebra	Algebra Methods	Equations	Algebra & Calculus	Statistics	Miscellaneous
1	9	1	3	0	0	0	0	0
2	6	2	8	0	0	0	3	1
3	2	3	0	3	5	0	1	1
4	0	0	0	2	0	2	0	0
Total	15	3	11	5	5	2	4	2

In undertaking the test each participant was given five 'lives'. Thus each was allowed a total of five instances of responding incorrectly before the test would be terminated. There was no time limit imposed. The results for each candidate of the proportion of questions correctly answered are mapped against the prior mathematics qualification in Appendix 15. The data shown in the lives used column represents those used and those successfully used. Also indicated is the time used by individuals. This will inform the time limit set within the Undergraduate study.

Procedure for using Diagnosys

1. Give a demonstration of the tutorial questions so that all candidates can pass through these quickly. Also explain other personal details questions and how to answer.
2. Allow time for all users to be able to ask questions about use.
3. During demonstration use all the interface buttons and explain their use.
4. Confirm that all candidates can use pen and paper when answering questions.
5. All candidates will receive an overview of their results on completion of the test. These may be printed out for personal reference. As evidence for the trial a copy of this file is saved for each participant.
6. Inform candidates that there is a time limit set for this test of 30 minutes.

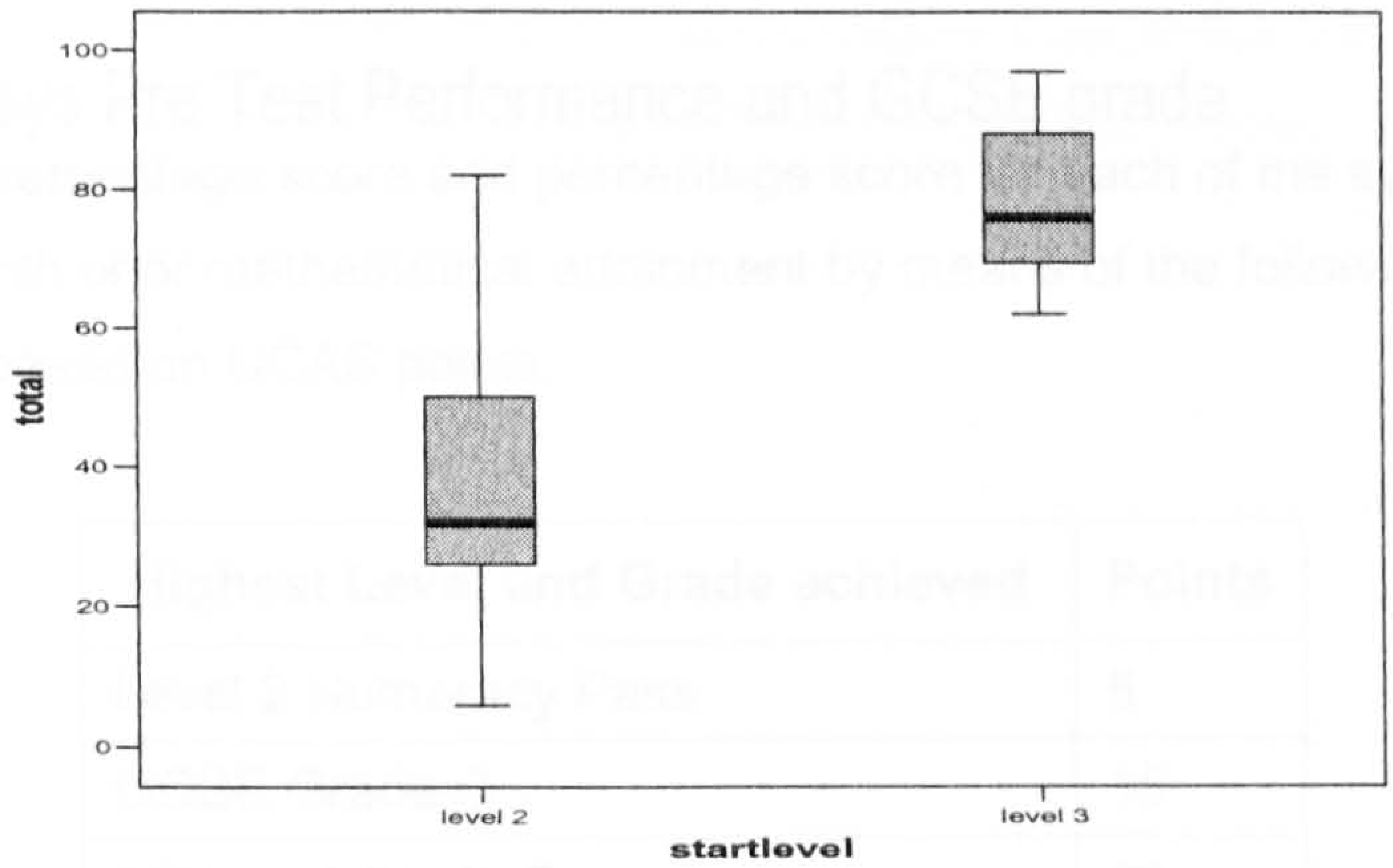
5.4 Stage 2 Analysis and findings

5.4.1 Validity of data

Statistical Method 2.1

To support the validity and accuracy of the test results outliers within each pairing have been identified by means of box plot charts and values removed hence the number of participants may vary between pairings. Box plots comparing total percentage score with the percentage score for each of starting level, numeracy, powers, algebra, statistics and miscellaneous will indicate if there is any outlying data. As the sample size of the Undergraduate trial is seventy three the outliers could be removed for subsequent statistical tests and analysis of the data. This is considered appropriate as the sample group is sufficiently large and can prevent the effect that these extreme values can have on statistical results.

Chart 5.4.1 Box plots of percentile values of Total Diagnosys score and Starting level



These box plots in Chart 5.4.1 indicate that there are no outliers in comparing Starting level with Total Diagnosys score. Outliers are marked with ° these are cases with values between 1.5 and 3 box lengths from the upper or lower edge of the box. These are rogue values that do not seem to be part of the main sample which can skew the coefficient result in a particular direction. Hence where outliers are identified they are removed from the data set to calculate the quartiles, maximum and minimum values.

Statistical Method 2.2

The reliability of the data sets being analysed will be tested by means of the value of the Cronbach's alpha coefficient.

Table 5.4.2 Reliability coefficients of data sets

Cronbach's Alpha	N	Items
.886	4	Percentage score for each of Diagnosys, algebra and number and GCSE grading
.824	9	Starting level, Total number of questions asked, Number of correct questions, Percentage total achieved, Percentage scores for Numeracy, Algebra, Powers, Statistics and Miscellaneous
.871	6	Percentage scores for total achieved, numeracy, algebra, powers, statistics and Miscellaneous
.896	34	Each individual skill
.971	13	Mean scores for total, 6 Weaknesses and 6 Non-Weaknesses

Each of these data sets comprise of mean scores for success rates. All Cronbach's alpha scores are greater than 0.8 indicating that the data is reliable. The results of the Tukey PP Plot indicated that the data sets did not have a normal distribution. Consequently Non parametric tests for correlation and comparison have been undertaken.

5.4.2 Diagnosys Pre Test Performance and GCSE grade

Diagnosys total percentage score and percentage score for each of the subcategories was compared with prior mathematical attainment by means of the following ordinal grading system based on UCAS points.

Highest Level and Grade achieved	Points
Level 2 Numeracy Pass	5
GCSE Grade C	15
AS Level Grade E	20
GCSE Grade B	25
AS Level Grade D	30
GCSE Grade A or A*	35
AS Level Grade C or A Level Grade E	40
AS Level Grade B	50
AS Level Grade A or A Level Grade D	60
A Level Grade C	80
A Level Grade B	100
A level Grade A.	120

This coding based on the UCAS points system will be used in Stage 6 also.

Full details of the individual participant's performance and an overview of the results of the participants of the Undergraduate trial based on their Prior Mathematics Performance are given in Appendix 15. An overview of the distribution of prior attainment of the subgroup of participants is shown in Table 5.4.3.

Table 5.4.3 Summary of Grades for Starting level 2, GCSE and equivalent

A Level	GCSE A	GCSE B	GCSE C	GCSE U	Vocational Pass	Vocational Merit
13	1	15	39	1	1	1

The results relating to each of the prior attainment categories are given in summary Tables in Appendix 15. 56% of the full group had a baseline achievement of Grade C at GCSE or equivalent of those whose highest achievement was GCSE, 73% achieved Grade C

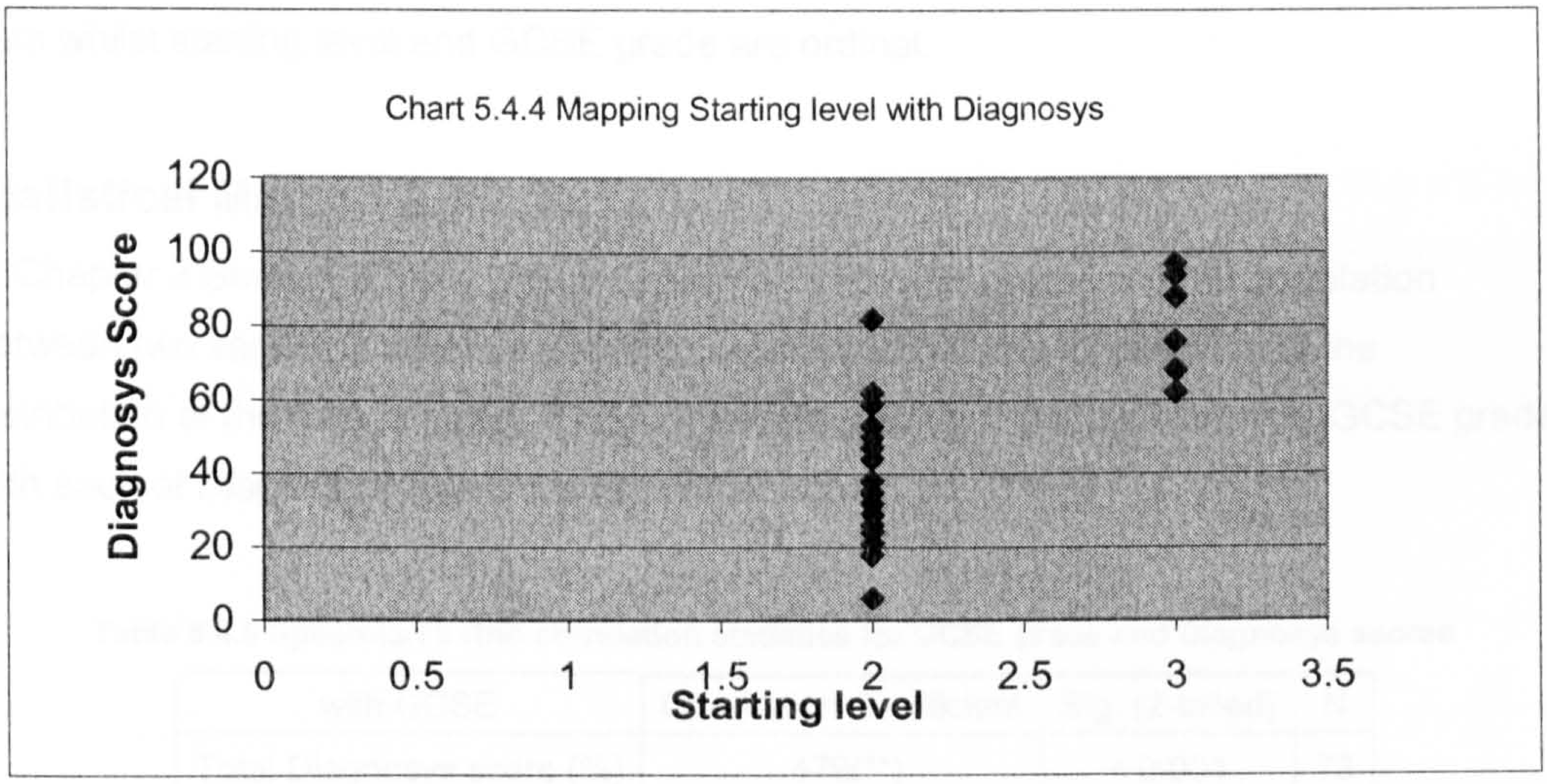
Summary data in Appendix 15b relating to the results of students subdivided by GCSE grade indicates

- o Diagnosys mean score of 34 for students with GCSE Grade C
- o Diagnosys mean score of 49 for students with GCSE Grade B

- the mean scores for grade B students was higher than that of the Grade C students for each of the five subcategories of Numbers, Powers, Statistics, Miscellaneous and Basic Algebra.

This is consistent with the expectation that Grade B students would achieve higher marks than Grade C students. The mean marks of students with an A level in Mathematics was higher than those relating to Grade B and Grade C students in each of the five subcategories and Total.

If GCSE is the highest level of mathematics qualification a participant has been awarded then a 2 is allocated for starting level whereas a 3 is allocated to those who have studied above this. Chart 5.4.4 shows the distribution of Total Diagnosys score clusters for starting levels 2 and 3. Chart 5.4.1 box plots present the descriptives relating to starting value and Diagnosys score.



Level of prior attainment and Diagnosys Score

From the results of every individual given in Appendix 15a a participant with a starting level of 2 who achieved an overall score of 82 in Diagnosys can be identified. Scores of 76, 68 and 62 in Diagnosys were attained by participants whose starting level was 3.

Chart 5.4.4 shows this overlap in Diagnosys achievement between those with a baseline level of 2 and those at level 3.

The box plot in Chart 5.4.1of those with a baseline level 2 and those at level 3 does not indicate any outlying values. This Chart shows a clear difference in the spread of the data by means of median and quartile range. Each percentile value relating to level 3 participants was greater than those of level 2 by at least 50 marks. The range of marks of level 2 students was greater than 75 whilst that of level 3 was less than 40.

Statistical Method 2.3

Starting level which is ordinal and total score is ratio data to test the correlation of these two variables Spearman's Rho will be used, the rationale as outlined in Chapter 3 Section 3.3.

The one tailed Spearman's Rho undertaken on score attained has a Correlation Coefficient of .646 with N=71 and is significant at 0.01 level. This rank correlation determined by score. The results of this test are given in Appendix 25 Table 25.2.3. Hence the higher the start level then the higher the total score. This finding supports the alternative hypothesis at 1% level that there is a significant correlation between starting level and score attained in a Diagnosys test.

Total score attained in Diagnosys and each of the sub categories of Diagnosys is ratio data whilst starting level and GCSE grade are ordinal.

Statistical Method 2.4

In Chapter 3 Section 3.3 the use of Spearman's Rho for measuring the correlation between two variables which are at least ordinal without any assumption of the distribution of the data is given. This will enable testing of the correlation of GCSE grade with each of Diagnosys, algebra and number scores

Table 5.4.5 Spearman's Rho correlation statistics for GCSE grade and Diagnosys scores

with GCSE	Correlation Coefficient	Sig. (2-tailed)	N
Total Diagnosys score (%)	.479(**)	< 0.001	73
Number score (%)	.409(**)	< 0.001	73
Algebra score (%)	.472(**)	< 0.001	73

These results indicate that there is a positive correlation between GCSE and Total Diagnosys, GCSE and Number, and GCSE and Algebra significant at the 0.01 level. Spearman's Rho rank correlation was undertaken on Diagnosys total score basis. Hence the higher the GCSE grade the higher the Diagnosys total, Algebra and number scores. These findings support the alternative hypothesis that there is a correlation between GCSE grade and performance in Diagnosys and in Numeracy and Algebra questions. Although the correlation between GCSE and Number is the weakest suggesting that Numeracy ability is less well predicted by GCSE score than algebraic ability.

Totals and subgroup totals

The percentage scores per participant in answering all questions and questions relating to specified Sections of numbers, powers, basic algebra, statistics and miscellaneous is ratio data.

Box plots given in Appendix 25 Chart 25.2.7 indicate that

- There were no outliers between total percentage score and percentage score for each powers and miscellaneous, that is, all values within the box and whisker range.
- There were three outliers between percentage scores for Total and Algebra and two outliers between Percentage scores for Total and each of Statistics and Numeracy.
- The data subsets Basic Algebra and Numeracy represent the areas of mathematics which are the focus of this study. The Box plot presented in Chart 25.2.7 indicates all values are within the box and whisker range hence there are no outliers in comparing these two subsets.

Statistical Method 2.5

Total and subtotal scores data sets are not normally distributed Spearman's Rho will be used to measure the correlation and Wilcoxon to test for significant differences between two variables. The Spearman's Rho correlation findings for Basic Algebra and Numeracy along with each of the areas of mathematics and the full data set are presented in Table 5.4.6.

Table 5.4.6 Correlation statistics for Diagnosys total and subset scores

	Spearman's Rho Correlation	Sig. (1-tailed)	N
Total Diagnosys (%) and Number (%)	.918(**)	< 0.001	73
Total Diagnosys (%) and Algebra (%)	.871(**)	< 0.001	73
Number (%) and Algebra (%)	.695(**)	< 0.001	73

The correlation coefficient of .695 significant at 0.001 level indicates that there is a positive correlation between scores for algebra and number when rank correlation is based on total score.

Table 5.4.7 Wilcoxon Z scores for Diagnosys total and subset scores

	Z	Sig. (2-tailed)
Number (%) and Total Diagnosys (%)	-6.899(a)	< 0.001
Algebra (%) and Total Diagnosys (%)	-4.404(b)	< 0.001
Algebra (%) and Number (%)	-6.013(b)	< 0.001

a Based on negative ranks. b Based on positive ranks.

Each of the Z scores in Table 5.4.7 is significant at 1% level indicating that there is a significant difference between scores for algebra and number and full data set with each of the subsets of algebra and number. This result indicates that it is significantly more difficult to answer Algebra questions correctly than Number. This result is based on negative ranks. The correlation results given in Table 5.4.6 indicate that these two sets of results have a positive correlation significant at 1% level.

These three sets of findings support the view that the GCSE grade does indicate mathematical ability.

5.4.3 Skills related to specified weaknesses

The data has been stored according to the following scale

Diagnosys Coding	Points awarded
Yes (skill achieved),	1
Pyes (skill not asked but assumed it would have been achieved)	2
Possible	3
Pno (skill not tested but not expected to be achieved)	4
No	5
Unasked	6

To analyse the significance an average score for each of the identified weaknesses as well as an average score for all of the skills which do not encompass the specified weakness identified in Chapter 2 Section 2.6 have been calculated. A question was awarded one mark if the log indicates yes or pyes otherwise zero marks were awarded.

Mean scores relate to 33 questions or skills. Each of the specified weaknesses relate to a specified subset of questions or skills. An overview of the skills per specified weakness is given in Table 5.4.8 below. There were 34 skills in the full data set. Skills 100-199 are level 1 and skills 200-299 are at level 2. The hierarchy of Diagnosys skills is detailed in Appendix 7. Weakness 1 Division relates to 6 questions, Weakness 2 Brackets to 8 questions whereas weaknesses 3 Indices, 4 Substitution, 5 Negatives and 6 Solving Linear Equations relate to only 1, 2, 4 and 2 questions.

Table 5.4.8 Subset of Diagnosys Skills for each specified Weaknesses

Weakness	Number	Skills
Weakness 1 Division	6	109, 112, 206, 207, 213, 216
Weakness 2 Brackets	8	110, 208, 209, 210, 211, 213, 214, 215
Weakness 3 Exponentials/indices	1	208
Weakness 4 Substituting values	2	113, 214

Weakness 5 Negative signs & values	4	101, 102, 103, 213
Weakness 6 Linear equations	2	111, 212

A mean score is calculated. A mean score with a low value suggests that an error is a misconception as detailed Chapter 2 Section 2.1. The skills which encompass each of the weaknesses specified in Chapter 2 Section 2.6 were identified from use of the log files. The format and mathematical concepts involved in each of these questions was identified. Hence for each of the six weaknesses each candidate was awarded a score between 0 and 1 where 1 would indicate complete success. The percentage total scores were also represented as a value between 0 and 1. Similarly a further six average percentage score one for each group of skills which do not encompass each of the weaknesses was calculated, that is Non-Weakness 1, Non-Weakness 2 and so forth. These scores although based on a varying number of questions were also between 0 and 1. A comprehensive analysis of responses to each subset of questions relating to the specified weaknesses is given in Appendix 26.2 .The success rates of answering questions for each of the weaknesses ranges from 97% to 13%.

Statistical Method 2.6

Significant difference between the mean scores of the data relating to each specified weakness and non-weakness will be investigated by use of Wilcoxon Rank sign test. The variance in the data sets due to the difference in the number of skills per specified weakness and non weakness indicate that matched T test would not be appropriate.

Table 5.4.9 Wilcoxon Z scores for each specified weakness and Non weakness

	Z	Sig. (2-tailed)
Division Non Weakness 1 and weakness 1	-4.755(a)	< 0.001
Brackets Non Weakness 2 and weakness 2	-5.480(a)	< 0.001
Indices Non Weakness 3 and Weakness 3	-3.973(a)	< 0.001
Substitutions Non Weakness 4 and weakness 4	-5.333(b)	< 0.001
Negative signs and values Non Weakness 5 and weakness 5	-6.569(b)	< 0.001
Solving Linear Equations Non Weakness 6 and Weakness 6	-.501(b)	< 0.001

a Based on negative ranks. b Based on positive ranks.

Each of the Z scores in Table 5.4.9 is significant at 0.01 level indicating that there is a significant difference between scores for each specified weakness and opposing Non weakness. Descriptive data is given in Table 25.2.10 in Appendix 25. The Z scores are based on negative ranks for Weakness 1 (Division), 2 (Brackets) and 3 (Indices). The Wilcoxon signed rank test (Statistical Method 2.6) Z values given in Table 5.4.9 indicate that the data subsets representing Weakness 1 (Division), Weakness 2 (Brackets) and Weakness 3 (Indices) have significantly lower success rates than the opposing Non-weakness subset. The opposite finding is indicated for Weakness 4 (Substitution) and Weakness 5 (Negative signs and Values) in that the mean values are significantly higher for these subsets than that of the opposing Non weakness. Each data set relating to a Non weakness can encompass various questions and mathematical concepts hence it cannot be deduced that Weakness 4 and 5 are indeed strengths. However we can construe that the students can undertake Substitution and handle Negative signs and values better than Division, Brackets and Indices. The Z score for Weakness 6 (Solving Linear Equations) shows that there is not a significant difference.

Hence this finding does not indicate that negative numbers and signs are a weakness. This finding does support the view that Brackets, Division and Indices were areas of specific weakness for the sample group. Findings relating to indices were limited as this subset of questions consisted of only one question.

Table 5.4.10 Wilcoxon Z scores for Weakness 1, 2 and 3

	Z	Sig. (2-tailed)
Weakness 2 (Brackets) and Weakness 1 (Division)	-2.143(a)	.032
Weakness 3 (Indices) and Weakness 2 (Brackets)	-1.226(a)	.220
Weakness 3 (Indices) and Weakness 1 (Division)	-1.846(a)	.065

a Based on positive ranks. b Based on negative ranks.

In Table 5.4.10 the only Z score significant at least at the 5% level is for Brackets and Division indicating a significant difference between the means of these scores. The Z score based on positive ranks indicates that the questions with Brackets were significantly more difficult to answer than those with Division.

5.5 Stage 3 Arrangements

5.5.1 Context

Treefrog Test Questions

Six tests were created to enable testing of the weaknesses division, brackets, indices, substitution, negative signs and values and Solving Linear Equations as identified in the QCA reports, by Matz and in the GCSE Examiners reports. The questions have been categorised to support these findings and presented to users incrementally in an intellectually progressive manner see Appendix 8 for details of the content of tests. As outlined in Chapter 4 the Stage 1 findings indicate that there is a variance in opinion of the teachers in the sample group with the research findings in that solving linear equations, division and substitution are not areas of weakness. Further investigation can enable comparison of findings with those of the teachers and those of the research and reports.

This trial will use Treefrog software (see Chapter 2). This software has been selected because of the facility to adapt the interface with users in this trial and in Stage 6. For the purpose of this trial minimal feedback and instruction has been included in the sample tests with only the first question of each test including guidance on how to solve a specific type of question. This was intended to enable users to develop an understanding of how to use the software in a specific instance. The provision of more feedback and instruction was not included within the Undergraduate trial. The objective being to enable users to display their individual approaches to questions and prevent their reiteration of suggested methods. Subsequently through the analysis of individual responses to questions presented common errors and methods of solution were identifiable a method similar to that utilized by Abidin and Hartley(1998).

Treefrog can present a range of questions to the learner. Treefrog consists of 11 types of question. Details of each type are given in Appendix 9. Specific question types are applicable for use when solving specific types of problems this includes evaluating expressions, solving equations, simplifying or factorising algebraic equations. Some of these problems can be expressed solely in words for the user to then convert into a suitable expression for solving. Some of these questions do not require the tutor to calculate the answer but to select the correct question type and input the relevant start expression. The question types related to equations require the user to present the solution in the same equation format as expected in properly constructed hand written solutions. Where the user is solving a numeric or algebraic problem the creation of equivalent expressions is required.

In considering the phrasing and presenting of questions using Treefrog software the following must be determined

- start - the numeric value or expression /algebraic equation
- value – the numeric value or expression /algebraic equation
- format of the equation or the solution of an expression
- type of equation - linear, quadratic
- type of number - real or complex
- use of brackets

- number of variables permitted

Some questions requiring similar skills were framed to reduce the risk of one off errors or an idiosyncratic response (Denscombe, 2001). It can be seen in Appendix 8 where this is the reason for inclusion of specific questions. However this does result in more questions being posed and hence the possibility that a participant may have become more despondent because of repeated lack of success.

Table 5.5.1 Distribution of questions posed

Test	1	2	3	4	5	6	Total
Number of Questions	10	15	17	7	8	5	62

5.5.2 Limitations

Stages 2, 3 and 4 were conducted in one sitting it is possible that candidates become more fatigued with the process. An overview of the number of questions set is outlined in Table 5.5.1.Other requirements of the Teacher Training Agency and other numeracy tests are the use of:

- tables of data
- money formatted data i.e. £
- % symbol included
- graphs
- comparing sets of data

Where sets of data are required to be compared for instance to deduce which set of pupil results have the largest range Treefrog is not an appropriate tool as the input mechanism within the interface does not accommodate more than one set of results and hence cannot tutor the student through the decision making process of comparing one result with another. A table outlining the purpose and limitations of Treefrog question types is included in Appendix 9.

5.5.3 Treefrog Preliminary Trial Results

Tables 25.3.1-3 presented in Appendix 25 details the results of each of the nine candidates within the Preliminary trial in terms of actual, percentage and rank values. From the results in Table 25.3.2 the rank order of the test scores as given in Table 25.3.3 can be deduced. Full details of the results attained by the nine candidates are given in Appendix 8. A comprehensive analysis of the preliminary trial and rationale for adaptation for Undergraduate trial is given in Appendix 26.1.

Errors and specific questions in the Preliminary Pilot

- Analysis of every question posed with respect of the common errors of syntax, negative signs or values, use of brackets or numerical calculations is presented in Table 25.3.4 in Appendix 25. In addition the frequency of the error and specifics of the responses made is detailed.
- Some questions did not cause an error from any of the nine candidates. All questions posed in Test 4 achieved 100% success rate and hence no analysis is included.
- These errors identified in the preliminary pilot indicated that other than syntax/formatting of answers, negative values, brackets and division are areas of misconception. The regularity of occurrence of these categories of errors was to be investigated within the Undergraduate trials.
- Furthermore these results were used to limit repetition and shorten the test in the Undergraduate trial. The subsequent adaptations to the test determined by this analysis are outlined in Table 5.5.2. The nature of errors and misconceptions is considered to be common to groups of learners and hence the analysis of the frequency of errors will inform the decision of which questions to delete from or to supplement the Undergraduate study to ensure adequate coverage of common errors relating to the weaknesses identified in Chapter 2 Section 2.6.

Adaptations following the Preliminary pilot

Table 5.5.2 Overview of changes in numbers of questions set

		Number of questions		
Test Number	Test Nature	Pre-pilot	Undergraduate	Difference
1	Number	10	7	-3
2	Number with division	15	8	-7
3	Equations	17	12	-5
4	Basic Algebra	7	6	-1
5	Brackets & expanding terms	8	8	0
6	Indices & Exponentials	5	5	0
Totals		62	46	-16

Procedures for using Treefrog

1. Candidates work at their own pace they can progress to completing the questionnaire when all tests are completed.
2. Introduce to all candidates explaining that they will be required to answer this afterwards and whilst they are using the software should consider in general terms Learning and Usability.
3. All candidates to complete the top Section of the questionnaire which contains personal details questions.
4. Demonstrate how to access Treefrog from the server.

5. Inform candidates that pen and paper should not be used but to use Treefrog itself to show working out that would be performed on paper.

5.6 Stage 3 Results

5.6.1 Data Reliability

Statistical Method 3.1

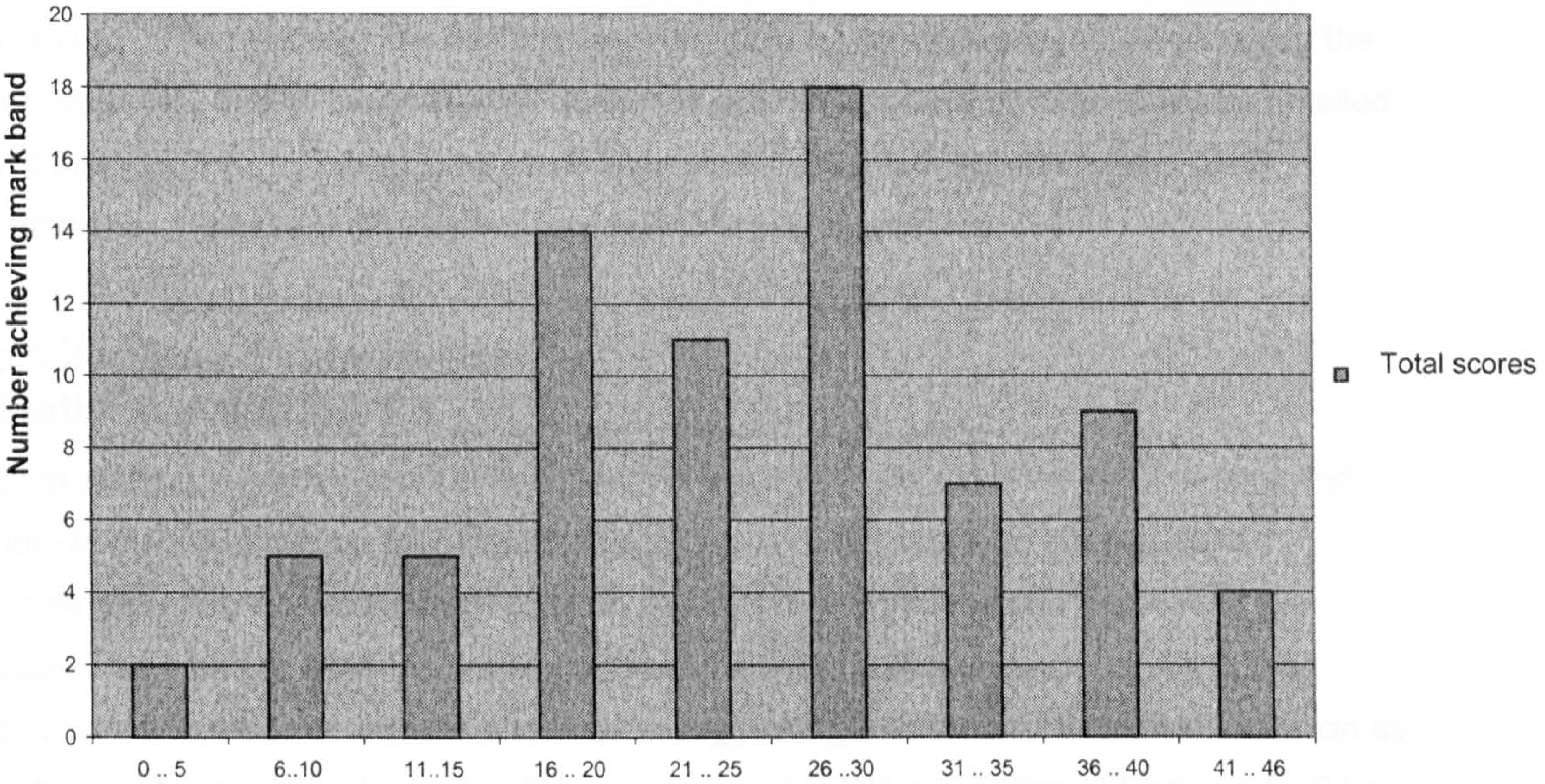
The reliability of the data derived from the Undergraduate study relating to all responses from all candidates and scores awarded to candidates relating to each of the weaknesses identified in Chapter 2 Section 2.6 will be tested by means of the Cronbach's alpha coefficient.

Table 5.6.1 Reliability coefficients of data sets

Cronbach's Alpha	N	Items
.899	12	Subsets of 6 Weaknesses and 6 Non Weaknesses
.639	2	Number of Questions tried & Number of questions answered correctly
.955	46	All questions with coded responses from all participants

The Cronbach's alpha score for all Treefrog questions $\alpha = .955$ was greater than that of all Diagnosys skills (see Table 5.4.2) α=.895. The former Cronbach score measures the reliability of the data sets of all participants in answering each question. This data stored represents whether the participant was successful or not. The data set consists of a score of 1 for success or a score of 0. The data representing Diagnosys skills is the mean score for a set of questions each scored as 1 for successful otherwise 0. The mean scores ranging between 0 and 1 for each skill data set for each participant represent rates of success. The questions posed within Diagnosys were selected by ability and performance whereas those posed within Treefrog were the same for all participants. Conversely the Cronbach's alpha score for the six weaknesses and Non weaknesses with the total percentage score within Diagnosys was $\alpha = .971$ and within Treefrog was $\alpha = .900$. These Cronbach scores measure the reliability of the success rate of all participants in responding to twelve subsets of questions.

Chart 5.6.2 Frequency Distribution of Total Scores for Treefrog questions



The Treefrog frequency distribution of success rate within the trial group displayed in Chart 5.6.2 indicates a vast range with a slight positive skew.

The results of the Tukey PP Plot indicated that the data sets did not have a normal distribution. Consequently non parametric tests for correlation and comparison have been undertaken.

Statistical Method 3.2

A further validity test to be undertaken is the correlation of Diagnosys mean scores with Treefrog mean scores.

Table 5.6.3 Correlation statistics of Diagnosys and Treefrog Percentage Scores

	Spearman's Rho Correlation	Sig. (1-tailed)	N
Diagnosys and Treefrog (%) scores	.588(**)	< 0.001	73
Algebra and Treefrog (%) scores	.590(**)	< 0.001	73
Number and Treefrog (%) scores	.485(**)	< 0.001	73

** Correlation is significant at the 0.01 level (1-tailed).

These test results indicate that there is a positive correlation significant at the 1% level between Diagnosys and Treefrog mean scores, Algebra subset and Treefrog mean scores and between Number subset and Treefrog mean scores.

The correlation of Treefrog and Diagnosys percentage scores and those for the subsets Number and Basic Algebra were undertaken (Statistical Method 3.2) to test the validity of the full range of Treefrog findings. The results of these tests as presented in Table 5.6.3 were all significant at 1% level. The strongest correlation was between Treefrog scores and Algebra scores with a correlation coefficient *.590* and Treefrog percentage score and

Diagnosys percentage score *.588* both significant at a 0.01 level. This indicates that there is a positive correlation between participant performance in Diagnosys and within Treefrog. The relatively low coefficient value could be attributed to the difference in the two systems. Within Diagnosys all questions are not posed to all candidates as detailed in Section 5.3.2. In Treefrog all candidates were presented with the full set of 46 questions. However candidates could select to skip questions.

5.6.2 Stage 3 Weaknesses

Statistical Method 3.3

Data relating to each subset of questions for each of these specified weaknesses and non-weaknesses will be ratio data calculated as a mean score for each specified weakness and non-weakness. For each participant each successful responses are scored as 1 and all other responses score 0 the mean value of each sub set is then calculated. A mean score with a low value suggesting that an error is a misconception as detailed Chapter 2 Section 2.1. The non parametric Wilcoxon Signed Rank test will be used to examine for a significant difference between each specified weakness and non weakness.

In Appendix 25 Table 25.3.11 there are statistics describing the distribution of the data relating to the subset of questions for each specified weakness and opposing Non Weakness subset are presented. Subsequently each of the subsets of questions which encompass the identified weaknesses will be summarised. Comprehensive details regarding responses to each of these subsets of questions is given in Appendix 26.2.

W1 Division

Twelve out of forty six questions included division. The questions and associated achievement rates are presented in Table 25.3.5 in Appendix 25. Levels of success range from 83% to 41%.

W2 Brackets

Eleven questions which encompassed the use of brackets are detailed with the success rates in Table 25.3.6 in Appendix 25. The range of success spans from 79% to 14%.

W3 Indices

Only two questions included the use of indices. The success rate of each of these was 11% and 6%. In the group interview and Treefrog software evaluation questionnaire it was reported that the notation ^ confused participants hence findings are unreliable. Analysis of the subsets Weakness 3 and Non Weakness 3 were not undertaken. Results of the two questions are given in Table 25.3.7 in Appendix 25.

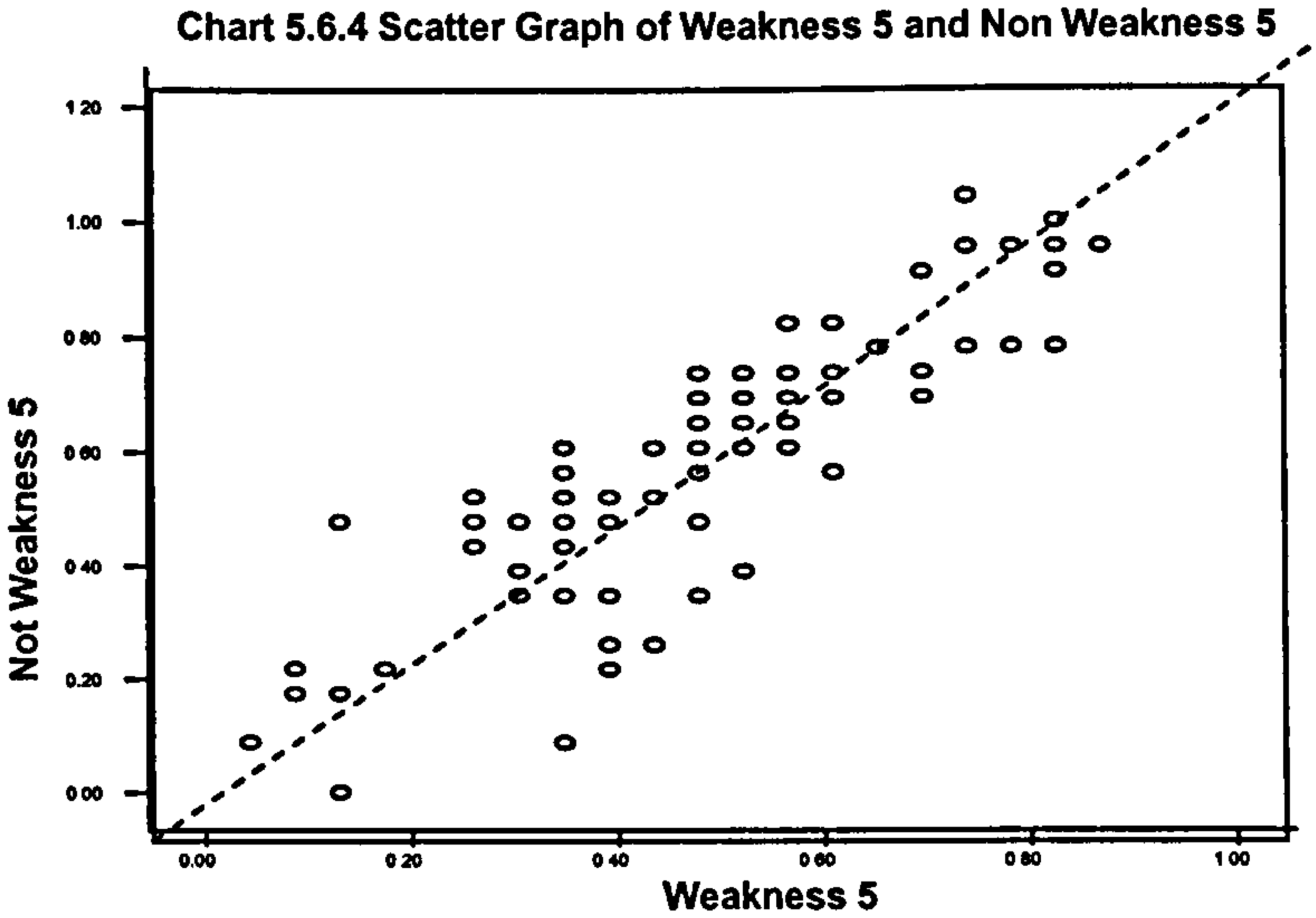
W4 Substituting values

Seven questions encompassed the concept of substitution of values. The success rate in answering these questions ranged from 78% to 97%. The results of these questions are given in Table 25.3.8 in Appendix 25.

W5 Negative signs and quantities

Half the questions in the sample included negative signs or quantities. This subset can be identified as Weakness 5. The opposing subset of questions Non Weakness 5, do not include negative signs or quantities. The success rate for subset of questions Weakness 5 ranges from 90% to 26%. The success rate for subset Non Weakness 5 ranges from 97% to 32%. These subsets do not include questions which encompassed other aspects of mathematics which pose considerable difficulties, specifically formatting, use of indices or fractional terms. The success rate of the questions in subset W5 are presented in Table 25.3.9 in Appendix 25.

Chart 5.6.4 Scatter Graph of Weakness 5 and Non Weakness 5



The scatter graph in Chart 5.6.4 indicates some linearity between Non Weakness 5 and Weakness 5. There was not an indication of linearity between any other Weakness and opposing Non Weakness. To ensure a standard approach this linearity is not considered when selecting statistical method for investigating each weakness and opposing Non Weakness. Furthermore the position of the plots being mainly above the diagonal in the chart suggests that questions without negative values and signs were less problematic to answer than those with. This linearity could indicate that understanding negative signs and values is related to general ability.

W6 Solving equations (linear)

Twenty two questions in the sample included one variable. The success rate of solving these equations ranges from 87% to 41% where the equation did not include fractional terms with brackets. The results relating to this full subset are given in Table 25.3.10 in Appendix 25.

From the box plots of each specified weakness subset with the full data set no outliers were identified. Each correlation test had N=74.

Table 5.6.5 Wilcoxon Z scores for each specified weakness and Non weakness Subset

	Z	Sig. (2-tailed)
Division Non Weakness 1 and weakness 1	-4.969(a)	< .001
Brackets Non Weakness 2 and weakness 2	-7.076(b)	< .001
Indices Non Weakness 3 and Weakness 3	-7.201(b)	< .001
Substitution Non Weakness 4 and weakness 4	-5.815(a)	< .001
Negative Signs and Values Non Weakness 5 and weakness 5	-5.204(b)	< .001
Solving Linear Equations Non Weakness 6 and Weakness 6	-.501(a)	< .001

a Based on positive ranks. b Based on negative ranks.

The results in Table 5.6.5 relate to the scores of the subset of questions with the specified weakness (Division, Brackets, Indices, Substitution, Negative Signs and Values, Solving Linear Equations) and the opposing subset of the questions not in the subset i.e without the specified weakness. Full details of the data sets and analysis of responses is given in Appendix 26.2. Each of the Z scores in Table 5.6.5 is significant at 0.01 level. The Z score for Solving Linear Equations (Weakness 6 and Non Weakness 6) indicates that there is not a significant difference between the two data sets.

As described in Chapter 5 Section 3.2 the data sets of responses to questions with the specified weakness and those of the opposing data set of questions without the specific weakness necessitated that two groups of unequal size mean scores were used. Full details of the data sets and analysis of responses is given in Appendix 26.2. The Wilcoxon z scores (Statistical Method 3.3) in Table 5.6.5 indicates that there is a significant difference in mean values between each of these pairs. However, in consideration of the rankings used the differences indicate that scores relating to the subsets Weakness 2 and Weakness 5 are significantly lower than those of Non Weakness 2 and Non Weakness 5 respectively.

The mean scores presented in Appendix 25 Table 25.3.11 show the score for subsets Weaknesses 1 (Division), 2 (Brackets) , 3 (Indices) and 5 (Negative Signs and Values) are lower than those of the opposing Non Weakness 1, 2, 3 and 5. Conversely the mean scores for subset Weaknesses 4 (Substitutions) is greater than those of the opposing subset Non Weakness 4. The full set consisted of forty six questions. The mean scores for Weakness 4 (7 questions) is significantly higher than Non Weakness 4 (39 questions). The variation in the numbers of questions in these Non weakness subsets and combinations of other areas of weakness could indicate that this finding is not significant.

These findings indicate that Weakness 2, Brackets (13 questions) and Weakness 5, Negative signs and values (23 questions) were areas of weakness for the sample group. However these findings do not support the notion that Weakness 1 Division, Weakness 3 Indices, Weakness 4 Substitutions or Weakness 6 Linear Equations were problematic for the Undergraduate group. To enable comparison of these findings with undergraduates with the performance of learners in post compulsory education questions which focus on each of the weaknesses division, brackets, substitutions, negative signs and values and substitutions will be posed in Stage 6. Due to the syntax problems incurred with questions involving indices in Treefrog questions with indices will not be included in subsequent stages.

5.7 Stage 4 - Evaluation of Stage 3 software

5.7.1 Introduction

To evaluate the impact of specific computer algebra systems (CAS) on learning we are necessarily concerned with specific educational situation(s). Through the experience of using the system users will be able to evaluate the software as a package and as a learning environment. These evaluations focused systematically on usability and learning 'heuristics' based on the work of Squires and Preece (1999) as outlined in Chapter 2 section 2.7 rather than the checklist approach which is increasing reported as an unsatisfactory instrument for evaluation. For the purposes of this study of mathematical applications the relevant criteria were selected and structured in terms of usability, learning and the synergy between them. Hence this questionnaire combines responses to statements that are coded as 'Poor', 'OK' and 'Good' with an 'open' comment. The comment providing a check on the honesty or seriousness of responses a common problem with a questionnaire approach as outlined by Robson (1999 p43). By the gathering of both a coded response to a closed question and a comment findings can encompass both quantitative data from which an overview of general opinion can be deduced as well as a collation of a range of personal perspectives.

The heuristics identified by Squires and Preece (1999) were applied to the evaluation of Treefrog presented as a questionnaire is included in Appendix 11. To enable the collection of useful and relevant responses the questionnaire was required to be written in an appropriate style and using a suitable level of language. However it was also vital that the phraseology presented queries without leading respondents.

By administering the questionnaires on a face-to-face basis during the experimental session the probability of low return rates was removed enabling a high response rate. These findings provide data for analysis regarding the performance of the software as a flexible learning tool by learners from various backgrounds and abilities. The results of this analysis can inform the enhancement of the system in terms of both learning and usability. An overview of these opinions will be used to inform the discussion within the group interview.

5.7.2 Adaptations following Preliminary Study

The questionnaire was adapted in response to the feedback obtained from all participants experience in using this form. This indicated that some statements were rather misleading. The amended and original questionnaires are in Appendix 11.

5.7.3 Analysis of findings

This questionnaire survey will result in quantitative and qualitative data. Quantitative data will be ordinal scores based on a Likeart scale. Qualitative data will comprise of the comments written by respondents. Themes within these comments will be identified and possible triangulation of results investigated. Quantitative analysis can determine strength of opinion with regard to specific heuristics. The findings of this survey will inform some of the discussion of the group interview.

5.8 Findings and Analysis of Stage 4

5.8.1 Responses on aspects of software or learning

Seventy questionnaires were completed. A complete set of comments and ratings recorded per question by all participants is detailed in Appendix 17. A summary of the quantitative is presented Appendix 17. An overview of the frequency distribution of the ratings is presented in Chart 5.8.1 below. The comments have been analysed for similarity and themes identified these have been presented in Table 5.8.2. A comprehensive analysis of the comments posted from which the summary is devolved is situated in Appendix 26.3.

Chart 5.8.1 The evaluation of the use of Treefrog overview data

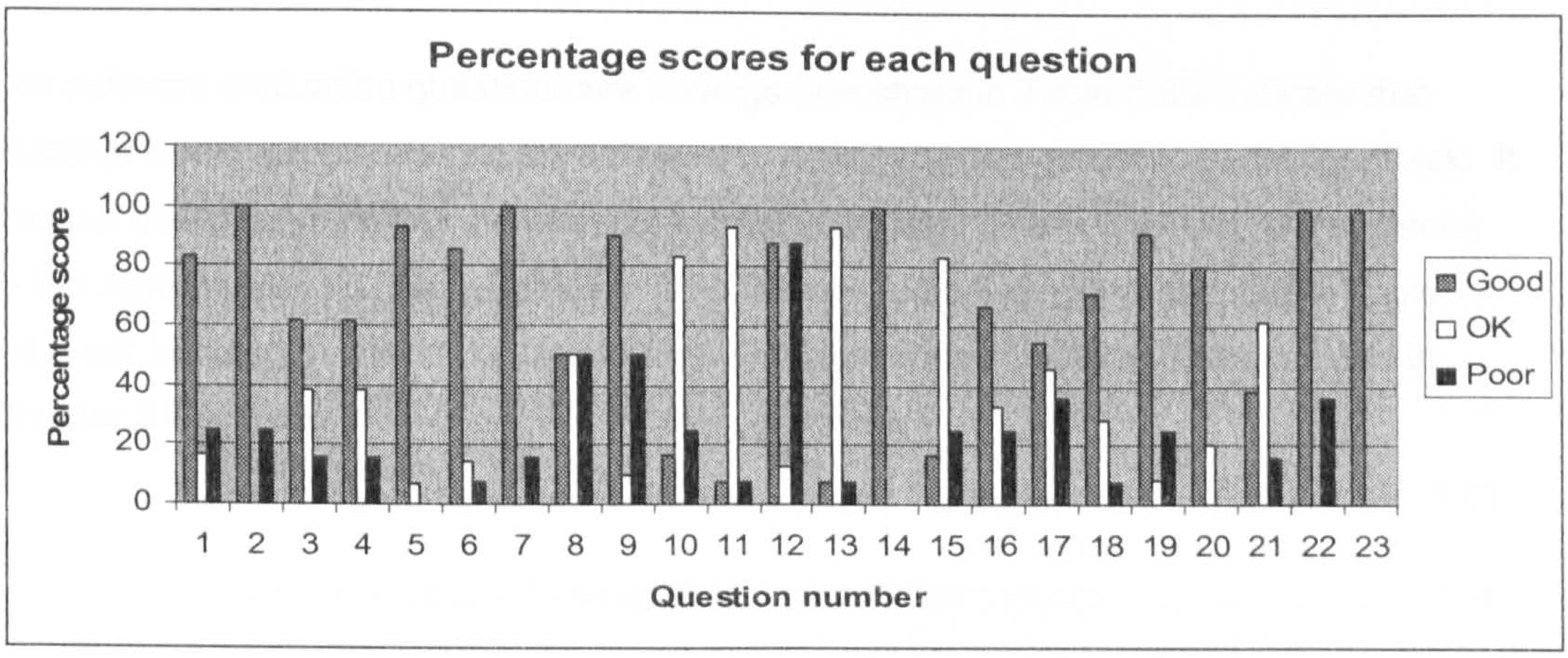


Table 5.8.2 Common Themes within comments

Theme	Number of responses
The feedback did not provide sufficient guidance for the user to be able to correct their error and learn	28
The interface was basic and simple and so could not distract from learning but was not motivating either	19
The software would not enable independent learning as there was insufficient feedback regarding errors	10
Software could support independent learning as it advises the user when they are wrong or correct	9
Exponential being represented as ^ was confusing	16

These themes will structure and guide the content of the group interview in Stage 5 of the study.

5.8.2 Analysis of results

The frequency statistics displayed in Chart 5.8.1 suggest that the users of the software generally considered the software to go 'Good' for learning. However in many instances the score awarded to specific heuristics did not correspond to the comment given. The comments indicated that many of the participants had not considered the software to be 'Good' in several aspects. The high rate of 'good' grades could be attributed to misunderstanding of the questions/heuristics or a possible influence of the Hawthorn effect as outlined in Section 5.1. Due to this inconsistency of the quantitative data this was not analysed further.

5.8.3 Software sensitising

Also, within the questionnaire feedback, there was a comment indicating that a user considered the number of tests undertaken to be tedious and another comment comparing Diagnosys to the unadapted Treefrog. These comments may suggest that the use of two computer aided learning systems concurrently impacted upon the attitudes of some users that is the use of Diagnosys had cause the users to be sensitised.

The software evaluation questionnaire findings presented in Table 5.8.2 indicate that students would have preferred the software to have provided guidance in the feedback. It was not designed to collect evidence of learners' anxiety. There is insufficient evidence to test agreement with this hypothesis. The themes identified and presented in Table 5.8.2 will be used to inform the focus of the Group interview in stage 5 as specified in Chapter 6 Section 6.3.

Chapter 6 Stage 5 - The Group interview

6.1 Introduction

Following the Undergraduate study and analysis of the results a group of participants was selected for further investigation. Bell (1992) suggests "a questionnaire can only be taken at face value, but a response in an interview can be developed and clarified". To minimise the effect of maturation this was organised during the semester in which the trial had been undertaken. All participants were able to recall their feelings and views on participating in the trial. This investigation was conducted as a group interview. This approach was selected to provide an environment in which participants felt at ease as a group of peers, individuals were less likely to show a bias because of any sense of a personal association with the interviewer. The wording of questions and statements was necessarily standardised for all participants which would be difficult to achieve in a series of individual interviews. Furthermore an interview enables a flexible approach in which the interviewer can adapt questions to responses given. Additionally it was anticipated that responses from one individual may provoke and encourage further details from other respondents. However there was a danger that individuals could overwhelm and influence the responses of others.

6.2 The group

7 students were selected who had attained the lowest scores within the Treefrog sets of questions as these participants are expected to have views on the preferred level and type of support.

7 students were selected who had either terminated questions or chosen to not undertake many questions during the trial. These participants were included to investigate why they had elected to terminate questions.

1 Mathematics Education student who had attained high scores throughout the various sets of questions posed within the trial. This participant volunteered to be involved and provided an opportunity to compare and contrast attitudes and opinions with those of an able student.

None of this group were known by the interviewer and were selected for validity of purpose as they would represent learners who require support in learning rather than those who have already attained understanding. However in that respect this group may not be truly representative of other learners within higher education. Although this group had participated in the Undergraduate trial their individual opinions within the questionnaire survey were not identifiable.

Ethics
All participants were assured of the anonymity of responses to questions. To ensure this those interviewed would be referred to using the following coding
<subject route> <number> <gender> will be used.

Subject route	DT	Design Technology Education
	Sc	Science Education
	PE	PE Education
	MA	Mathematics Education
Gender	m	male
	f	female

Bias
There is a danger that the human interaction between interviewer and interviewee will be affected by bias. The interview was conducted 'blind' in that the purpose and expectations of the trial were not shared with participants to minimise expectancy bias.

6.3 Arrangements

The interview was conducted when all the students would be expected to be on campus with a significant gap between lectures. It was located in a classroom free from the disruption of personal callers or the telephone.

Gathering together this group of students was problematic as students should only freely attend. The method used to communicate with students was email to organise a meeting in week 10. However the only students who replied to this communication were responding that they were unable to attend. No one confirmed attendance. Consequently no one attended. Subsequently the students were informed by email that there was a meeting scheduled for the following week and those invited to attend would be met at the end of a shared lecture.

Recording and verification
As the interview was conducted with a group a recording was considered to be problematic instead minutes were taken during proceedings. From the themes identified in Stage 4 Questionnaire survey as outlined in Chapter 5 section 5.8.2 the following categories of enquiry were identified:

- feedback
- guidance for correcting errors
- interactivity
- learning
- motivation
- human computer interface and learning
- syntax

This was aided by use of a prepared interview schedule giving headings under which to record responses see Appendix 12. Responses to each prompt were recorded on individual sheets headed appropriately to enable efficiency. The coding of respondents were used to further enable this.

The responses given to the prompts were recorded and then tabulated under themes. Also recorded is whether one or more respondents were involved in a reply.

6.4 Methods of Analysis

To analyse this data statements will be examined individually and as groups of commonality for support or variance of themes outlined in the hypotheses. These findings will be cross referenced with other related findings from the literature review and the questionnaire survey to establish whether there is triangulation in agreement or disagreement.

All responses and the identity of the respondents is detailed in Appendix 18. From these responses a patterns was identified and is presented in Table 6.4.1.

Table 6.4.1 Patterns of response

DT1m initiated	Explanations/feedback on the error being made would be helpful – if you put in an incorrect Answered wrongly then it is very difficult to know of a different one to give that might be right
	A wider variety of maths topics could be supported such as shape i.e. those required for school maths (these students will be required to teach KS 2 maths)
	User should have the ability to select the area of maths (student – centred)
	Not user friendly - Interface basic, unpleasant on the eye – too large a block of dense colour
	Screen could be centred – more appealing to the eye – overview of questions/results so far –could learn from errors and previous experience
	Diagrams/visual add ins within questions – more interesting to want to continue to use or to use again
	Font size/type not easy on the eye
DT1m agreed	Feedback should be supportive and encouraging – felt like a test – caused fear
	Different types of questions such as multi-choice, or to give the question and the answer but required to deduce a method of solution
	Introduction screens would have been useful – explaining the content, purpose and format requirements
	Very precise format of answers and not obvious – although not always the same, sometimes must make a whole equation with = others not (showing their lack of understanding about the types of problems and their differences)
	Different types of questions such as multi-choice, or to give the question and the answer but required to deduce a method of solution
	All students felt confident with the numeracy questions but all struggled to progress to the algebra – no connections seen – 'large jump' made in difficulty of questions

	Just like GCSE maths
	Students felt that they had not learnt any maths through using the software but that they were probably worse at maths than they thought they were although they did nt know what their errors were
Sc1M Initiated	Different types of questions such as multi-choice, or to give the question and the answer but required to deduce a method of solution
	Unexciting – boring – not encouraging to use – no desire to continue – irritating because guidance was so general
Sc1M Agreed	Explanations/feedback on the error being made would be helpful – if you put in an incorrect response then it is very difficult to know of a different one to give that might be right

6.5 Mathematical Profile of the interview group

The mathematical profile of each of the participants of the group interview consists of their performance in the Undergraduate study and their prior mathematical attainment. The summary results of the Pre Test Diagnosys, Treefrog and prior mathematical attainment (most commonly their GCSE Mathematics Grade) are presented in Table 5.5.2. Further detail relating to each participants' Diagnosys results are given in Table 25.5.1 in Appendix 25. Statistics describing each participant's performance in using Treefrog is presented in Appendix 25 Table 25.5.2. A comprehensive analysis of the interviewees behaviour in using Treefrog is given in Appendix 26.3.

Table 6.4.2 Overview Results of group interviewees

Interviewees	DT1M	PE1M	SC2F	SC1M	L1F	DTF1	PE2F	MA1M
Prior Mathematics attainment	GCSE C	GCSE C	GCSE B	GNVQ Merit	GNVQ Pass	GCSE C	GCSE C	A level B
Diagnosys (%) Score	38	29	62	56	32	32	18	88
Treefrog (%) Score	33	48	43	22	7	15	33	91

6.6 Undergraduate Trial Group interviewees errors made and progress

Candidate Sc2f, who achieved 48% despite erroneous attempts at solutions in Tests 1, 2 and 4, achieved a high level of success. However this participant did not attempt Tests 5 and 6 and had limited success in Test 3 opting to Pass on each question when an erroneous attempt had been made. This candidate had stated in the interview that learner confidence was affected negatively by doing these tests and that the experience was de-motivating for those with low level mathematics ability.

Sc2f achieved 62% in Diagnosys, which was higher than the mean result for the full sample group and had a Grade B at GCSE. Conversely this candidate described herself as being of low ability indicating a poor 'self mathematics concept'. As given in Appendix 25 table 25.5.2 this student was successful in answering twenty questions in the Treefrog trial and made forty five erroneous attempts.

L1f, the candidate with one of the lowest prior mathematics qualification obtained the lowest Treefrog score. This candidate struggled with all categories of question achieving totals of 0, 1, 0, 0, 1, 1 for each test respectively. This candidate achieved 32% in Diagnosys when the mean for the full sample group was 44%. In addition this candidate made only 5 wrong attempts but passed on forty two questions. In Appendix 18 it is outlined that this candidate could shared the view that the lack of success during the trial had a negative impact upon confidence. This candidate may have already had a low 'self mathematics concept' and hence did not continually attempt to answer the questions.

However Pef2 a candidate with GCSE Grade C achieved 5, 2, 2, 4, 2, and 0 for each test respectively but only 18% in Diagnosys. This candidate attempted all questions posed in Treefrog regardless of encountering forty seven erroneous attempts. The behaviour and ability of these candidates does indicate a mapping of the number of erroneous attempts and behaviour regarding passing on questions or exiting from tests or Non-attempting of Tests. Both L1f and Pef2 commented that feedback, which provides guidance on the methods of solution and error made, should encompass examples and demonstrations. They suggested that help should be available on request, as well as automatically, if errors were made.

From analysis of the responses given in Appendix 17 the patterns of response presented in Chapter 5 Table 5.5.1 were obtained. From these the following opinions have been identified.

The whole interview group felt that the feedback given to users should enable errors to be amended and that methods of solution should be outlined. The group felt that the feedback would be most beneficial if it gave help on specific errors made. This support could be progressively extended and complemented by introductions explaining the nature of problems posed and encompassing demonstrations and example solutions.

Furthermore, several of the group suggested that a full summary of performance in each test would be beneficial to the learner. This summary could provide details of the number and type of question successfully solved as well as details of those not attempted or those unsuccessfully attempted with an outline of the nature of the error.

The group proceeded to outline how they had felt that the experience was de-motivating and reinforced poor 'self mathematics concepts' attributable to prior mathematical experiences. They believed that all users required some success to motivate and encourage them learn.

A few interviewees suggested that the learners may assume more ownership of their learning if they were given more control by selecting areas and with the ability to select more help and support.

This group were unable to see the connection between Numeracy and Basic Algebra and believed that an understanding of the links would aid their understanding.

It was suggested that explanatory screens could be provided within the package to demonstrate a methodology in answering a question.

Many felt that the system was not sufficiently user friendly in terms of the colour and layout of the interface as a large block was of one background colour and the fonts selected were glaring on the eye and not always easily readable. The students would have preferred a more 'interesting' interface.

Following this investigation the software developed for use in Stage 6 will include:

- Feedback based on errors to outline nature of error made and enable correction (see Appendix 21)
- Explanatory Screens to demonstrate methodology
- Examples questions
- The ability to navigate backwards and forward through questions
- Format errors explained in feedback
- Supportive feedback about success to encourage users

Chapter 7 Stage 6 - Final Trial

7.1 Introduction

7.1.1 Experiment Arrangements

The study was in a sixth form college in Salford and involved 24 volunteers aged between 16 and 19. Experimental sessions were conducted during personal tutorial timetabled sessions of 45 minutes duration during a four week period. During this four week period individual students were not exposed to formal tutoring of algebra or numeracy. Within this trial participants undertook a Pre Test and a Post Test. Each test being a set of questions focusing on the solution of linear equations encompassing weaknesses 1 (division), 2 (brackets), 4 (substitution), 5 (Negative signs and values) and 6 (solving linear equations) as identified by the Undergraduate study and initially identified in Chapter 2 Section 2.6. Weakness 3 (Indices) was omitted from further study following the unreliable findings of the Undergraduate study as outlined in Chapter 5 Sections 5.4.3 and 5.6.2. These questions were presented on a paper based work sheet for both the pre and post test. Comparable questions were set for the pre and post tests. The Pre Test is presented in Appendix 19 and the Post Test in Appendix 22.

Due to the hard copy presentation of the pre test and post test undertaken all participants were necessarily provided with a standardised assessment tool hence eliminating the potential for bias.

7.1.2 The groups

The 24 participants were all enrolled on a range of programmes of study including National Qualifications Framework Level 2 and Level 3 Vocational programmes in a range of subjects such as Health and Social Care, ICT and Performing Arts as well as some who were retaking GCSE Sciences. Four of the students had achieved a Grade B, four had achieved a Grade C in GCSE Mathematics whereas the other students had achieved a grade D or E. All the students had some previous experience of using a computer however the extent of this varied from GCSE to those who had limited experience and generally did not include the use of a computer in learning mathematics or numeracy. This group was randomised in terms of both ability in mathematics and experience of ICT and hence was a representative unbiased group. This range of GCSE mathematics attainment enabled the evaluation of the usefulness of the GCSE grade as an indicator of mathematical ability.

The group was subdivided in to two matched subgroups. To ensure that there was not a biased sample in terms of ability the students were sorted in order of their total score in the Pre-test and allocated alternately to two balanced groups the control and experimental groups.

The control group, "Webfrog" used an online version of Treefrog which included only broad hints. The experimental group, "Webfrog with feedback" used an online version of

Treefrog which encompassed feedback which related to each of the identified common errors and misconceptions. Both groups would perform a post test. This enables comparison of the progress in learning of those who have used a system which provides feedback related to the error made and those who have used a system without this feedback whilst limiting other variable factors.

Use of the blind technique whereby participants were unaware of the expectations of the trial and the variation in the two versions of the software minimised the expectancy factor. Once matched on ability the students were randomised in terms of gender, course of study, age and previous GCSE result. In the context of this investigation a repeated measures approach could not be taken as students would be exposed to both computer systems. This would not be appropriate as the effective of one system could not be compared as prior experience would have an influential effect.

Each group of students were given a demonstration of the system which they were required to use. The next two consecutive sessions focused on learning using the software systems. Throughout the sessions the students' use of the software was observed.

Group	Pre Test	Investigation	Post Test
Control	X	Fixed feedback	X
Experimental	X	Adaptive Feedback	X

People within the control group did not receive the intervention under investigation that is the 'independent variable' which in this context is the use of the a system which incorporates adaptive feedback. The control group only used a system with fixed feedback.

7.1.3 Ethics

Ethics were considered throughout the duration of this study. Consent was sought from all participants involved. The consent form made participants aware of the aims of the research and the intended use of the data. All participants were aware that they could withdraw from the research process at any point with no fear of any negative effects. Verbal consent was also sought from all participants throughout the study and specifically in conducting the group interview. The purpose of the work in terms of investigating effectiveness of software in learning was discussed as well as how the information would be used. Confidentiality and anonymity of all participants was of high importance throughout the research project. All names have been removed to protect the identity of individuals. A few students were anxious about participating and elected not to, another was absent from college due to illness during the final trial and hence their involvement was terminated.

7.2 Pre Test and Post Test

The paper based Pre Test consisted of twenty questions focusing on Numeracy, substitutions in linear expressions, solving linear equations and rewriting expressions. These questions will encompass the use of brackets (Weakness 2), negative signs and values (Weakness 5) and division (Weakness 1). The map of questions posed to areas of specific weakness is given in Appendix 19. The paper based Post Test (see Appendix 22) consisted of twenty questions based on the Pre test questions. The tests were undertaken by all participants.

7.3 The Final Trial content

The Numeracy and Algebra questions presented by "Webfrog" and "Webfrog with feedback" encompassed five main areas of weakness Division (1), Brackets (2), Substitution (4), Negative signs and Values (5) and Solving Linear Equations (6) investigated in the Undergraduate study. Students in the Undergraduate study in Stage 5 had indicated that they did not recognise a connection between Numeracy and Algebra. The set of questions presented in session 1 and session 2 encompassed the following five Sections

- I Numeracy
- II Algebra
  - A Substitution
  - B Simplify
  - C Rewrite
  - D Linear equations

Each Section consisted of twelve questions which incorporated the following four question types

QT1	No weaknesses	2 questions
QT2	Negatives	3 questions
QT3	Brackets	3 questions
QT4	Brackets and Negatives	4 questions

The number of questions presented per question type reflected the conceptual difficulties incorporated. A map of errors and specified weaknesses per set of questions in session 1 and session 2 is presented in Appendix 20b. Users were not required to use a specific method for solution as equivalence of solution is used to check responses.

7.4 The feedback

Webfrog and "Webfrog with Feedback" included generalised hints for the user such as "Take Care with your calculations". Also within each system there were 'dummy' or practice questions which took the user stepwise through the interactions expected from the user in order to solve questions of that type.

An example of a Practice Question

2z+5=11

Each of the following steps are presented to the user by selecting the 'cheat' button and feedback given based on this guidance.

Step 1	2z=11-5	Subtract 5 from both sides
Step 2	2z=6	Simplify arithmetic on right hand side
Step 3	z=6/2	Divide right hand side by 2
Step 4	z=3	

Guidance given to participants explaining the nature and structure of the content to be presented is given in Appendix 20a. "Webfrog with feedback" also included a bank of feedback statements the specific statement being automatically initiated by the user performing the related error. Eighteen error codes relate to general syntax and calculation, order of operators, algebraic manipulations, use of negative values and signs, use of brackets. A map of the feedback statements and error codes with questions posed is given in Appendix 20b. Positive feedback was also included to encourage users and inform them that responses were correct. At the end of each of the session the user was informed of their success rate by a summary report which provided an overview of the success rate within each of the Sections and an associated comment.

Number of questions correct	Comment
12 questions	Excellent work
10 or 11	Very good
8 or 9	Good well done
6 or 7	Good try
4 or 5	OK Promising result
0-3	Fair start, some more practice would help

7.5 Analysis of findings

7.5.1 Reliability of data

Statistical Method 6.1

The reliability of the Pre Test Score, Post Test Score, Pre Test weaknesses and non weaknesses scores will measured by means of the Cronbach's alpha coefficient.

Table 7.5.1 Reliability coefficients of data sets

Cronbach's Alpha	N	Items
.938	82	20 Pre Test questions, 20 Post Test questions, total score for two matched groups
.989	11	5 Weaknesses Subsets, 5 Non Weaknesses Subsets and Total mean scores
.468	4	Pre Test, Post Test, Num errors and num finished
.543	2	Num errors and num finished

The data sets which relate to all the responses to all questions and the summary data for each specified Weakness Subset and Non Weakness Subset with Cronbach's alpha scores > .9 are reliable. The first two data sets comprise of variables which represent mean values of success rates. Whereas the data sets with Cronbach's alpha scores less than .6 may not reliable. These data sets consist of more than one variable hence Cronbach's alpha coefficient describing the internal consistency of the data set based on average inter-item correlation may not be relevant. The Tukey test results of the Pre Test, five specified Weakness Subsets and Non weakness Subsets and Post test scores indicated that these data sets did not have a normal distribution. Hence Non parametric tests were selected.

Matched Groups

As outlined in Section 7.1.2 the Pre Test determined the allocation of participants to the two matched groups "Webfrog" and "Webfrog with feedback". The scores of this test, allocated group and GCSE grade of all participants is presented in Table 7.5.2. The full set of Pre Test results are detailed in Appendix 23.

Table 7.5.2 Overview Results of Pre Test and Webfrog allocation

Participant	488	2208	2289	3107	521	721	1715	827	1854	904	953	1230	624	1490	1352	1723	00004	1803	0030	312	1126	654
Pre Test Score	19	17	17	15	15	15	13	13	10	10	10	10	8	7	7	5	5	5	2	2	2	1
Rank	1	2	2	4	4	4	7	7	9	9	9	9	13	14	14	16	16	16	19	19	19	22
Webfrog or Webfrog with feedback	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
GCSE Grade	B	B	C	B	C	C	C	C	E	C	E	D	E	E	C	E	E	E	E	D	E	E
GCSE Code	25	25	15	25	15	15	15	15	0	15	0	5	0	0	25	0	0	0	0	5	0	0

- The data has been allocated to subsets by analysis of each of the Pre Test questions with regard to Weakness 1 (Division), 2 (Brackets), 4(Substitution), 5 (Negative Signs and Values) and 6 (Solving Linear Equations) as identified in section 7.2.

- For each data set relating to a specified Weakness there is an opposing Non-Weakness subset. The summary data of each of these ten subsets and the total score have been used to analyse performance.
- Mean scores have been calculated to enable comparison.
- The reliability scores of the data sets within the final trial are given in Table 7.5.1.
- The descriptive statistics including the mean values of these ten subsets and the full test is detailed in Table 25.6.1. In each of these statistical tests N=22.
- Analysis of the data relating to each of the Weaknesses in the final trial Pre Test is given in Appendix 24 tables 24.1-24.5
- Box plots describing the percentiles of each Weakness and opposing Non Weakness and each Weakness and total score indicate that there was only one outlier between Weakness 1 and Non weakness 1 hence N=21 in this instance and N=22 in all others.

7.5.2 Weaknesses Scores

Statistical Method 6.2

The correlation of the Pre Test ratio data relating to each subset of questions for each specified weakness and opposing non weakness, N=22, can be measured by the non parametric Spearman's Rho Correlation coefficient. Level of significant difference between each pair will be tested by means of the non parametric Wilcoxon signed rank test.

Table 7.5.3 Correlation of each Specified Weakness subset and opposing Non weakness

Spearman's Rho	**Division** Weakness 1 & Non weakness 1	**Brackets** Weakness 2 & Non weakness 2	**Substitution** Weakness 4 & Non weakness 4	**Negative signs and Values** Weakness 5 & Non weakness 5	**Solving Linear Equations** Weakness 6 & Non weakness 6
Correlation Coefficient	.874(**)	.884(**)	.848(**)	.817(**)	.800(**)
Sig. (1-tailed)	< .001	< .001	< .001	< .001	< .001
N	22	22	22	22	22

** Correlation is significant at the 0.01 level (1-tailed).

The results presented in Table 7.5.3 indicate that there is a positive correlation significant at the 1% level between each of the data sets representing the Weaknesses 1, 2, 4, 5 and 6 with the opposing Non Weakness data set.

Table 7.5.4 Wilcoxon Z scores for each specified weakness and Non weakness

	Z	Sig. (2-tailed)
Division Non Weakness 1 and Weakness 1	-3.629(a)	< .001
Brackets Non Weakness 2 and Weakness 2	-3.443(b)	.001
Substitution Non Weakness 4 and Weakness 4	-2.621(a)	.009

Negative Signs and Values Non Weakness 5 and Weakness 5	-3.573(b)	< .001
Solving Linear Equations Non Weakness 6 and Weakness 6	-3.548(a)	< .001

a Based on positive ranks.
b Based on negative ranks.

The Z scores in Table 7.5.4 for Weakness 1, 2, 4, 5 and 6 with opposing Non weakness are significant at 0.01 level. The Z scores indicate that there is a significant difference between scores for each specified weakness and opposing Non weakness. The descriptives relating to the distribution of the data within each of these subsets are given in Appendix 25 Table 25.6.1.

The results presented in Chapter 7 Table 7.5.3 (Statistical Method 6.2) indicate that there was a strong correlation between ability in answering all Pre Test questions with a specified weakness and the opposing subset of non weakness

- In each pair coefficient > .80 and significant at 0.01 level.

Similarly the Wilcoxon signed rank test results displayed in Chapter 7 Table 7.5.4 show that there was a significant difference in the mean scores for each specific weakness subset and opposing Non weakness subset. The mean values presented in Appendix 25 Table 25.6.1 show that those relating to each of Weakness 1 Division, 4 Substitution and 6 Solving Linear Equations > .4727 the mean value of the full data set. In addition the mean score for Weakness 1, 4 and 6 is greater than that of the opposing Non weakness subset as indicated in Table 7.5.5. These rankings indicate that Weakness 1 (Division), Weakness 4 (Substitution) and Weakness 6 (Linear equations) were not more difficult to answer. Conversely Weakness 2, Brackets and Weakness 5, Negative signs and values are weaknesses.

Table 7.5.5 Mean values

	1	2	4	5	6
Weakness	.6477	.3836	.5682	.4186	.6193
Non weakness	.4290	.5455	.4489	.6455	.3750

The detailed analysis of the Pre Test scripts presented in Appendix 26.2 and the errors made within the software trial presented in Appendix 26.3 support the findings that the concepts of using Brackets and handling negative signs and values were both areas of weakness.

These findings therefore support the view that Brackets and use of Negative signs and values are areas of difficulty for students in the sample group and that these areas of difficulty have continued from earlier stages of learning.

7.5.3 Effectiveness of systems

Statistical Method 6.3

Total scores for the pre test and post test will be calculated to measure progression. The control and experimental groups will undertake the same Post-Test. The scores of the two matched groups results will be compared by either the parametric related matched samples T-Tests or the non parametric Wilcoxon Rank sign test to investigate the impact of the adaptive feedback.

Analysis of the impact of use of the two systems "Webfrog" and "Webfrog with Feedback" is presented by means of Wilcoxon Z scores in Table 7.5.6. The results presented in this Table relate to firstly the null hypothesis that there is no significant difference between the post scores of the two groups "Webfrog" and "Webfrog with feedback" and the other three pairs to the alternative hypothesis that there is a significant difference. Data describing the distribution of the Pre Test and Post Test scores for the two matched groups is presented in Appendix 25 Table 25.6.2.

Table 7.5.6 Wilcoxon of "Webfrog" and "Webfrog with Feedback" Pre Test and Post Test

	Pre Test Scores Webfrog – Webfrog with Feedback	**Webfrog** Post Test - Pre Test Scores	**Webfrog with Feedback** Post test - Pre test scores	**Post Test Scores** Webfrog – Webfrog with Feedback
Z	-1.857(a)	-1.725(b)	-2.814(b)	-2.099(b)
Asymp. Sig. (2-tailed)	.063	.084	.005	.036

These results indicate

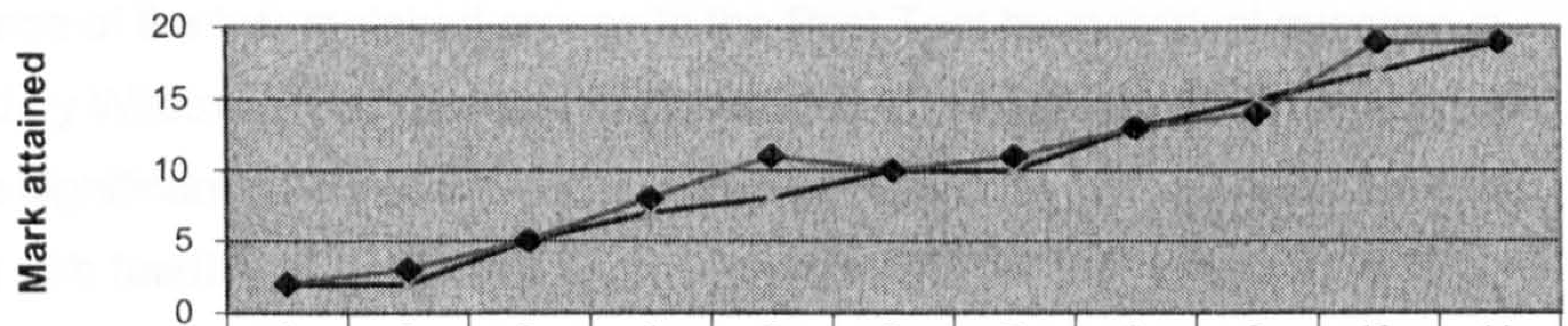
- No significant difference at the 5% level between Pre score of the matched groups "Webfrog" and "Webfrog with Feedback" which is consistent with the groups being matched.
- No significant difference between the Pre test score and post test score of the "Webfrog" group at 5% level.
- A significant difference at 1% level between the Pre Test score and Post test score of the "Webfrog with Feedback" group.
- A significant difference between at 5% level between the Post test scores of the "Webfrog with Feedback" group and the "Webfrog" group.

The Wilcoxon signed ranks tests (Statistical Method 6.4) performed on the matched groups (as indicated in Section 7.5 Table 7.5.2) of Pre Test Scores, matched groups sets of Post Test scores and a comparison of the Pre Test and Post Test scores for the "Webfrog" and "Webfrog with Feedback" groups results are shown in Section 7.5.3 Table 7.5.6. A two tailed test was necessary in comparing the Post scores and Pre scores as either could be better than the other. In comparison of Pre-Scores and Post-scores it cannot be assumed that the use of the software improved understanding and hence the Post score would be better than the Pre score, however it is unlikely to have worsened.

The two-tailed asymptotic significance estimates the probability of obtaining a Z statistic that is as extreme or more extreme in absolute value as the one displayed, if there truly is no difference between the group ranks. In this study, as presented in Section 7.5.3 Table 7.5.6, the probabilities for both the two measures of significant differences between the Pre Test and Post Test results "Webfrog" and "Webfrog with Feedback" vary. The null hypothesis stated that there would be no difference between the Post Test results of the two matched groups. Furthermore the Z scores indicate that the Post Test results after using "Webfrog with Feedback" are significantly different than those of the Pre Test. The mean values of the post scores indicates that "Webfrog with Feedback" Post score > "Webfrog" Post score and the upper and lower quartile values of the "Webfrog with Feedback" group are greater than the upper and lower quartile values of the "Webfrog" group. The spread of data is presented in Appendix 25 Table 25.6.2. The difference in value of the first quartile Q_1 was 3.00 and the difference in Q_3 was 4.0 showing a greater difference than that of the median or Q_2 difference of 1.0. Hence the results of the "Webfrog with Feedback" user group are positively skewed to higher values. In addition Chart 7.5.7 in Section 7.5.3 displays the difference in progress in learning of the group using "Webfrog" and those using "Webfrog with Feedback". These graphically demonstrate the greater improvement in the latter group. These findings indicate that there was a significant difference in the improvement between the two groups which implies that "Webfrog with Feedback" was more effective in supporting learning than Webfrog. These findings support the views discussed in Chapter 2 of Jameson (2003) in Section 5.3 that the quality of feedback is fundamental to learning and that of Kulhavey and Stock (1989) that effective feedback supports learning.

These results support the alternative hypothesis there was a significant difference in the performance of "Webfrog with feedback" before and after use of the software. The quartile, mean and median values of these two groups given in Appendix 25 Table 25.6.2 indicate that the software "Webfrog with feedback" was more effective in improving the results than the software "Webfrog". In addition these results indicate that there is a significant difference between the performance of the two groups "Webfrog" and "Webfrog with feedback" in support of the latter despite there being a significant difference between the Pre Test scores of "Webfrog" and "Webfrog with feedback" on the side of "Webfrog". Further comparison of the performance of the two groups is presented in the following two Charts in 7.5.7.

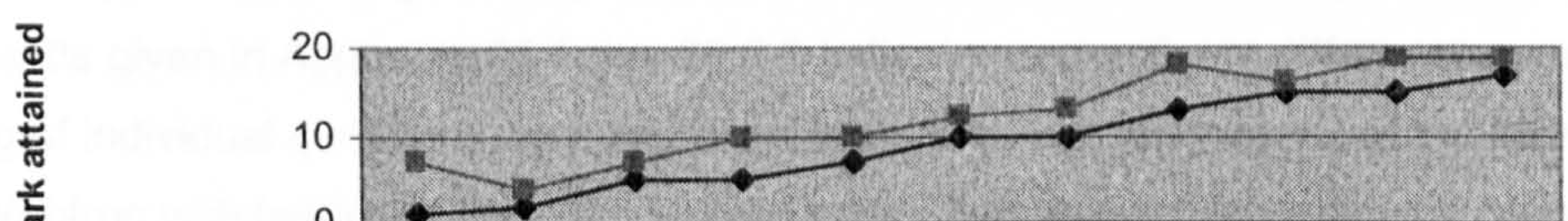
Chart 7.5.7 Progress of candidates using "Webfrog" without feedback



	1	2	3	4	5	6	7	8	9	10	11
Pre Test	2	2	5	7	8	10	10	13	15	17	19
Post Test	2	3	5	8	11	10	11	13	14	19	19

Ranked candidates

Progress of candidates using Webfrog2 with feedback



	1	2	3	4	5	6	7	8	9	10	11
Pre Test	1	2	5	5	7	10	10	13	15	15	17
Post Test	7	4	7	10	10	12	13	18	16	19	19

Ranked candidates

These results show that use of "Webfrog with feedback" significantly improved the test scores. In addition there is a significant difference between using "Webfrog with feedback" to using "Webfrog" without feedback. In that the improvement in test scores of the "Webfrog with Feedback" users is significantly higher than that of the "Webfrog" users. Hence "Webfrog with Feedback" is significantly more effective as a learner tool than Webfrog.

7.5.4 Pre and Post Tests Performance in each specific question

Mean scores of specific questions in Pre and Post Tests

The Post Test was devised so that the questions corresponded with the same specified weaknesses as those shown in the Pre Test (see Appendices 19 and 22). All participants were expected to answer all questions given.

Statistical Method 6.4

Answers to corresponding questions of the paper based pre test and post test will be compared for each of the two groups. Responses to individual questions is nominal data with a coding of 0 for wrong and 1 for correct and these can be compared by means of Wilcoxon to investigate significant difference.

Findings relating to the Wilcoxon analysis of user performance in Pre and Post Test corresponding questions for the two groups "Webfrog" and "Webfrog with Feedback" are presented in Appendix 25 in Tables 25.6.4 and 25.6.5 respectively. Comparison in performance of the two matched groups in the Post Test by individual question is presented by Wilcoxon analysis data in Table 25.6.6 in Appendix 25. These results indicate a significant difference in performance in questions 1, 4, 12 and 18 for the "Webfrog with feedback" group and Questions 4 and 16 for the "Webfrog" group. Questions 1 and 4 include Weakness 1, Division. Question 12 includes Weakness 1, Division and Weakness 6, Linear equations. Question 18 includes Weakness 2, Brackets.

These results indicate that the software with feedback had been beneficial in developing users' ability to handle Negative signs and values in three instances and Brackets in one. The results given in Appendix 25 Table 25.6.6 indicate no significant differences per ranking of individual questions were noted between the Post Test results of the "Webfrog" and "Webfrog with feedback" results. This suggests that no specific questions made a significant difference to the results of one group or the other and hence the results of the "Webfrog with feedback" group were generally better than those of "Webfrog".

Statistical Method 6.5

In Section 3.3 the appropriateness of using Spearman's Rho coefficient to measure the correlation between variables which are at least ordinal is outlined. This will be used to consider the relationship between the GCSE grade and Pre Test score. The GCSE grade will be coded using the scale detailed in Statistical Method 2.4.

The Spearman's Rho correlation coefficient of .739, significant at 1% level indicates a positive correlation between Pre Score and GCSE grade. The SPSS evidence from this test can be found in Appendix 25 Table 25.6.8. These results support the hypothesis that the GCSE grade or equivalent attained is an adequate indicator of mathematical understanding related to Numeracy and basic algebra.

7.5.5 Final Trial Pre Test weakness scores and GCSE Grade

Statistical Method 6.6

Spearman's Rho will be used to investigate the relationship between the GCSE grade and Pre test specified weaknesses subset mean scores. A mean score with a low value suggesting that an error is a misconception as detailed Chapter 2 Section 2.1.

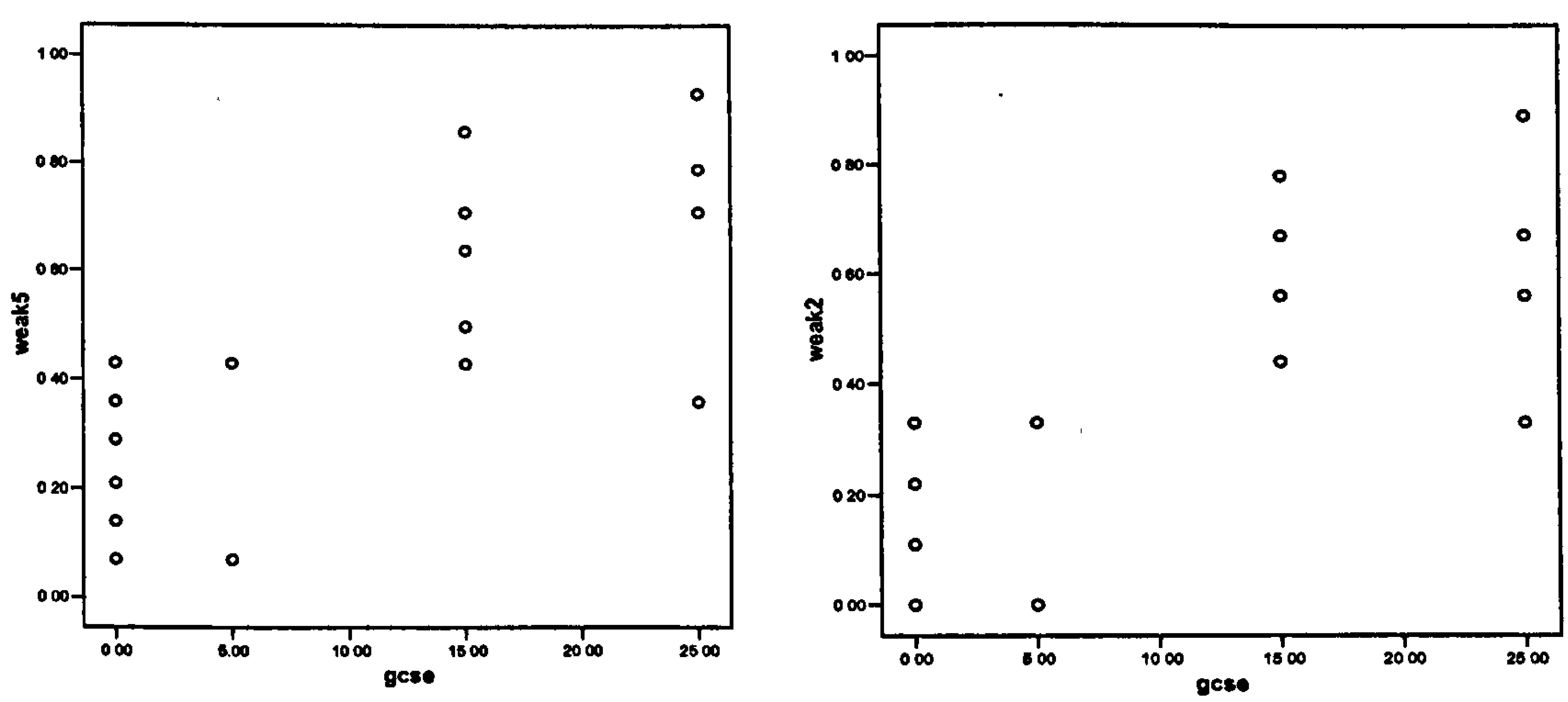
Table 7.5.8 Spearman's Rho correlation statistics between specified weaknesses and GCSE

With GCSE	**Division** Weakness 1	**Brackets** Weakness 2	**Substitution** Weakness 4	**Negative signs and values** Weakness 5	**Solving Linear Equations** Weakness 6
Correlation Coefficient	.568(**)	.759(**)	.562(**)	.767(**)	.656(**)
Sig. (2-tailed)	.006	< .001	.006	< .001	.001
N	22	22	22	22	22

The overview of the sample group given in Table 7.5.2 shows that ten candidates had achieved at least GCSE Grade C and twelve had not. The Spearman's Rho results in Table 7.5.8 indicate a positive correlation significant at 0.01 level between GCSE grade and each of the specified weaknesses. The coefficient values indicate that this is strong between Weakness 2 (Brackets) and GCSE and Weakness 5 (Negative signs and values) and GCSE as the square value of each is > .50.

Further more the scatter graphs between Weakness 2 Brackets and GCSE grade and Weakness 5 Negative Signs and Values and GCSE grade presented in Chart 7.5.9 both show a distinct difference in ability between those with at least 15 points or Grade C and those below this level. Conversely the scatter graphs of the relationship between GCSE grade and each of the weaknesses 1 (Division), 4 (Substitution) and 6 (Solving Linear Equations) did not show a pattern which would have been anticipated.

Chart 7.5.9 Weakness 2 Brackets and Weakness 5 Negative Signs and Values and GCSE



Statistical Method 6.7

Mann Whitney U Test will be used to measure differences in rank and confidence in the result between Pre Test Weakness Scores between those who have at least Grade C and those who have not achieved Grade C. This test will be used to test for a significant difference between the two groups below Grade C (Group 1) and those with at least a Grade C (Group 2).

Table 7.5.10 Comparison of Group 1 and 2 Pre Test Specified Weakness Scores

	Division Weakness 1	Brackets Weakness 2	Negative Signs & Values Weakness 5	Substitutions Weakness 4	Solving Linear equations Weakness 6	Pre Test
Mann-Whitney U	18.000	2.500	4.000	20.500	10.500	6.000
Wilcoxon W	96.000	80.500	82.000	98.500	88.500	84.000
Z	-2.880	-3.848	-3.709	-2.691	-3.316	-3.586
Asymp. Sig. (2-tailed)	.004	< .001	< .001	.007	.001	< .001
Exact Sig. [2*(1-tailed Sig.)]	.004(a)	< .001(a)	< .001(a)	.007(a)	< .001(a)	< .001(a)

The results in Table 5.6.10 indicate that the null hypothesis should be rejected and that there is a significant difference between the two groups in answering the subsets Weakness 2 (Brackets) and Weakness 5 (Negative signs and values). For each of these data subsets the Mann Whitney U value is greatly lower than that for Weakness 1 Brackets, Weakness 4 Substitutions or Weakness 6 Solving Linear Equations and full Pre Test score data set. This could suggest that the ability to handle brackets and signs could be an indicator of likely GCSE performance. That is learners who are successful in answering questions which include brackets or negative sings and values should achieve at least a Grade C in GCSE Mathematics. In addition, those who are not successful are unlikely to achieve a Grade C. Alternatively it could be anticipated that there would be a significant difference in ability to answer any questions between those who have achieved at least a Grade C GCSE and those who have not. However these results could indicate that there is not a significant difference in handling division, substitution or solving linear equations. Consequently it could be considered that Brackets and Negative signs and values are areas of weakness.

7.5.6 User performance and response

Table 7.5.11 Overview of "Webfrog" and "Webfrog with Feedback" log files

	Webfrog	Webfrog with feedback
Errors made	283	373
Questions attempted	301	329
Questions correctly answered	167	292

The detailed log files itemising errors made are presented in Appendix 24 Tables 24.7 and 24.8 and a user overview in Appendix 25 Table 25.6.3.

Behaviours observed in the student logs included attempting and reattempting to answer questions, trying only once and moving onto the next question, and stopping answering questions. Overview results relating to performance deduced from the log files are outlined in Table 7.5.11 above. This full data set of each participant given in Appendix 25 Table 25.6.7.

Rate of unsuccessful answers

From investigation of the log files three "Webfrog" users did not persevere with answering most questions. Instead answers were attempted once or twice, then, after getting the answer wrong, selected to move onto the next question. On the other hand this was not observed in the logs of any of the "Webfrog with Feedback" users. All these users made several attempts at getting an answer correct and generally got the answer correct before moving on. Less than 12% of the questions attempted in "Webfrog with Feedback" were not correctly answered whereas more than 44% of questions attempted by "Webfrog" users were not successfully answered.

Use of Hints

Hints were available to "Webfrog" and "Webfrog with Feedback" users whereby help could be requested. "Webfrog with Feedback" users requested a hint 58 times and "Webfrog" users requested a hint 41 times. Two "Webfrog with Feedback" users did not use any hints whilst three others used 16, 12 and 9 hints. One "Webfrog" user requested 13 hints whilst the others requested no more than 5.

Errors

Webfrog users in answering 301 questions made 283 errors. "Webfrog with Feedback" users in answering 329 questions made 373 errors. Despite the higher rate of errors being made by "Webfrog with Feedback" users this group answered 125 more questions correctly than the "Webfrog" group, which was almost 75% more questions answered correctly than by "Webfrog" users.

More errors were recorded when answering questions which involved a complexity of errors. For instance in answering questions which involved Brackets 45% of users made an error and in answering questions with only a negative sign or subtraction 36% of users made an error. In answering questions which included both negative values and Brackets all users made at least one error. Number of occurrences of questions encompassing these concepts would vary as all users answered a different number of questions.

Terminating

Four users terminated "Webfrog" during session 2, including one who terminated the program during session 1. No "Webfrog with Feedback" users terminated the program during either of the sessions although one user was not present at the second session.

Consequently it could be interpreted that users of the "Webfrog with Feedback" system were more motivated to attempt to answer questions despite making errors that those who used Webfrog. This could imply that the provision of feedback which gave guidance to the user resulted in users being successful which subsequently reduced their anxiety about mathematics and hence encouraged them to persevere with more questions.

Chapter 8 Discussion of findings

8.0 Introduction

A range of data sets quantitative and qualitative has been presented in Chapter 5 relating to each of the six stages:

Background investigation

- o Stage 1 coded responses from survey of Numeracy and Algebra at KS 3

Undergraduate Trial

- o Stage 2 numeric results of Pre-test (Diagnosys)
- o Stage 3 numeric results of Pilot Trial (Treefrog)
- o Stage 4 coded and open responses of Software Evaluation questionnaire
- o Stage 5 open responses of Group Interview

Final Trial

- o Stage 6 numeric results of Final Trial
  (Pre and Post Test and log of responses) of two matched groups

The data and findings derived from each of these stages will be analysed and the triangulation of findings as outlined in Chapter 4 Section 10 presented. This study has been undertaken to investigate the hypotheses 1-4 stated in Chapter 3:

- Stages 1, 2, 3 and Stage 6

Hypothesis 1: levels of difficulty in understanding specific areas of Numeracy and Algebra as identified by the QCA at KS3 and KS4 will continue into post-compulsory and higher education

- Stages 2, 3, 5 and Stage 6

Hypothesis 2: Learners who experience persistent failure experience anxiety and have a negative attitude towards mathematics.

- Stages 2, 5 and Stage 6

Hypothesis 3: the GCSE grade or equivalent attained is an adequate indicator of mathematical understanding related to Numeracy and Basic Algebra

- Stages 4, 5 and 6

Hypothesis 4: there is a difference in the effectiveness for learning between computer systems with and without feedback which focuses on common errors

Subsequently a summary of findings will outline the contribution to research made by this investigation.

8.1 Progression of errors and specific weaknesses

Hypothesis 1: difficulty in understanding specific areas of Numeracy and Algebra as identified by the QCA at KS3 and KS4 will continue into post-compulsory education

Six weaknesses identified from the KS3 SAT reports, GCSE Examiners and research findings as presented in Chapter 2 Section 2.6 have formed the basis of the investigation of the succession of areas of difficulty

- Weakness 1 Division
- Weakness 2 Brackets
- Weakness 3 Indices
- Weakness 4 Substitutions
- Weakness 5 Negative signs and values
- Weakness 6 Linear Equations

8.1.1 Stage 1 Teacher Survey

Stage 1 Electronic Questionnaire Survey study which compared the opinions of eighteen teachers of Key Stage 3 mathematics derived from experience of teaching with those of the QCA examiners derived from analysis of pupil performance in undertaking SATS. The findings as outlined in Chapter 4 Section 4.13 indicated that

- three aspects solving linear equations (weakness 6), division(weakness 1) and substitution (weakness 4) were not areas of weakness
- two aspects of indices (weakness 3) and negative signs and values (weakness 5) are areas of weakness.

The statements given in Table 4.1.1 did not provide evidence to test the notion that brackets (weakness 2) are an area of weakness.

Teachers' responses were derived from personal opinion which may not be scientifically based or informed by national reports. Individual teachers may have been informed of the QCA findings which may have informed their teaching strategies in the classroom or have disagreed with the report and ignored the findings. Conversely the teachers may not have been aware of the findings of the QCA. Depending on the validity of the QCA, GCSE examiners' and researchers' findings the inconsistency in opinion may support the recommendation of the 2004 Smith Report that to improve standards in mathematics teachers require training. Subsequently comparison of the findings of successive stages in this investigation with regard to specific weaknesses with the findings from this survey and those presented in the literature review will be undertaken.

8.1.2 Stage 2 Undergraduate Trial Pre Test (Diagnosys)

The Wilcoxon signed rank test (Statistical Method 2.6) Z values in Stage 2 given in Table 5.2.9 indicate that three aspects of mathematics Division (Weakness 1), Brackets (Weakness 2) and Indices (Weakness 3) were more problematic than general mathematics. Hence it is implied that division, brackets and indices were areas of weakness for the sample group within this set of questions. Conversely it is indicated that Substitution (Weakness 4) and Negative signs and values (Weakness 5) were not more problematic than general mathematics for the sample group. The relatively small Z score of -.501 relating to Solving linear equations (Weakness 6) was inconclusive. The Z score of -6.569 indicates a significant difference between the set of questions including Negative signs and values and those without however within this specific test there were only four skills which would have encompassed negative signs and values 12% of the full set of thirty-three skills. The Z score of -5.333 for Substitution was derived from only two skills 6% of the full set. The Diagnosys strategy for question/skill selection could result in a smaller proportion of skills than the potential four skills for Negative signs and two skills for Substitutions being presented. Should the user be successful within this limited selection but less successful within the larger opposing subset of questions which encompass a range of mathematical concepts the result would be a higher mean score for either of the subsets Negative signs and Values or Substitution than in the opposing sets of questions. However we cannot deduce that these two aspects were areas of strength as the larger proportion of skills encompassed a range of concepts but these findings do indicate that within the set of questions posed negative numbers and signs and Substitution were not areas of weakness.

In addition the Z score of -2.143 significant at 5% level presented in Table 5.4.10 indicates that the subset Brackets of eight skills were significantly more difficult to answer than the subset of Division of six skills.

8.1.3 Stage 3 The Undergraduate Pilot Trial (Treefrog)

The Spearman's Rho findings given in Chapter 5 Section 5.6.1 Table 5.6.3 indicate that there is a positive correlation between participant performance in Diagnosys and within Treefrog. The Treefrog question set consisted of fifteen number questions and twenty nine being algebraic. The correlation coefficient of Diagnosys Algebra and Treefrog of .590 is greater than the correlation coefficient of .485 of Diagnosys Number and Treefrog. Hence the relatively low coefficient value of .588 of Diagnosys and Treefrog could be attributed to the different content of the two systems.

The scatter diagram given in Chapter 5 Section 5.6.2 Chart 5.6.4 shows that mean scores for Weakness 5 (Negative Signs and Values) and Non Weakness 5 corresponded. That is, those who achieved a high mean score for Weakness 5 also achieved a high mean score for Non Weakness 5. Furthermore the position of the plots being mainly above the diagonal in the chart suggests that questions without negative values and signs were less problematic to answer.

The Wilcoxon Z scores (Statistical Method 3.3) presented in Chapter 5 Section 5.6.2 Table 5.6.5 of each of these pairs was at a significance level < .001. The Z scores and ranks indicate that Weakness 2, Brackets (13 questions) with Z=-7.076 and Weakness 5, Negative signs and values (23 questions) with Z=-5.204 were areas of weakness for the sample group. In contrast the mean scores for Weakness 1 Division (3 questions), and Weakness 4 Substitution (7 questions) are significantly higher than those of Non Weakness 1 (43 questions) and Non Weakness 4 (39 questions) respectively. Inevitably the opposing Non weakness subsets for division and substitution would encompass a range of mathematical concepts and possible weakness. The formatting of questions involving indices using the notation ^ was reported by the students in Stages 4 and 5 as being confusing therefore analysis of this data was considered unreliable. For that reason Indices (Weakness 3) was not included in the remainder of the study. The small Z score of .501 implied that there was not a significant difference between questions for Solving Linear equations (weakness 6) and those without. Consequently these findings do not support the notion that Weakness 1 Division, Weakness 3 Indices, Weakness 4 Substitutions or Weakness 6 Linear Equations were problematic for the Undergraduate group but did support the proposal that Brackets and Negative Values and signs were areas of weakness.

8.1.4 Stage 6 Final Trial (Pre Test weaknesses)

The results presented in Chapter 7 Section 7.5.2 Table 7.5.3 (Statistical Method 6.2) indicate that there was a strong correlation between ability in answering all Pre Test questions with each specified weakness and the opposing subset of non weakness

- In each pair coefficient > .80 and significant at 0.01 level.

The findings in Table 8.1.1 indicate that Weakness 1 (Division), Weakness 4 (Substitution) and Weakness 6 (Linear equations) were not more difficult than general mathematics. Indeed the mean values for Weakness 1, 4 and 6 is greater than that of the opposing Non weakness subset. The results for brackets and negative signs and values are in contrast to this.

Table 8.1.1 Mean values

	Weakness 1 Division	Weakness 2 Brackets	Weakness 4 Substitution	Weakness 5 Negative Signs and Values	Weakness 6 Solving Linear Equations
Weakness	.6477	.3836	.5682	.4186	.6193
Non weakness	.4290	.5455	.4489	.6455	.3750

Similarly the Wilcoxon results in Chapter 7 Table 7.5.4 signify that Weakness 2, Brackets and Weakness 5, Negative signs and values are significantly more problematic than general mathematics and hence are areas of weakness. Whereas Weaknesses 1, 4 and 6 results do not show that these sets of questions are more difficult to answer. The detailed analysis of the Pre Test scripts presented in Appendix 26.2 and the errors made within the software trial presented in Appendix 26.3 support the findings that the concepts of using Brackets and handling negative signs and values were both areas of weakness.

These findings therefore support the view that Brackets and the use of Negative signs and values are areas of difficulty for students in the sample group. Hence we can reasonably assume that these areas of difficulty have continued from compulsory education.

8.1.5 Progression of each specific weakness

If teachers are aware of specific difficulties faced by learners they may emphasise these as teaching points and consequently the problems and misconceptions may not continue in to subsequent stages of education. The findings of the teacher survey given in Section 8.1.1 indicate some disagreement between the scientific findings of the QCA and that of teachers.

Weakness 1 Division

The findings of Stage 2 the Undergraduate Trial Pre Test discussed in section 6.1.2 support the published findings reported in Chapter 2 that division is an area of weakness. Conversely the findings in Stage 3 Treefrog test in section 8.1.3 and the Final trial Pre test in section 8.1.4 are consistent with the findings of the teacher survey in Stage 1 that division is not an area of weakness. The finding in Stage 1 relates to less than 20% of the total number of skills presented. Although depending on user performance even fewer skills could be encountered. The Stage 1 result could suggest that division is not a general area of concern in school pupils. If division is not a weakness in school children then this concept could not proceed as an area of weakness from KS 4 in to post compulsory education. Further investigation regarding division as an area of mathematical difficulty in compulsory and post compulsory education could be informative.

Weakness 2 Brackets

Analysis of findings from three aspects of this study indicates that brackets were problematic for both sample groups of students. Specifically the results for undergraduates in Stage 2 Pre Test (Diagnosys) and Stage 3 Pilot Trial (Treefrog) presented in Sections 8.1.2 and 8.1.3 and those of learners in post-compulsory education in Stage 6 Final Trial Pre Test discussed in section 8.1.4. Hence these results indicate that Brackets were an area of weakness for students in Post Compulsory education and those embarking into higher education. These findings support the notion that Brackets identified by the QCA, GCSE examiners and academics as being an area of mathematical difficulty for school pupils continues as an area of weakness beyond compulsory education. The survey with teachers did not provide relevant information to test this judgement.

Weakness 3 Indices

Analysis of findings related to Indices in the Undergraduate Pre Test in section 6.1.2 indicated that there was a significant difference in answering questions which involved indices and those which did not. This supported the teachers' view obtained in Stage 1 outlined in Section 8.1.1 which confirmed the QCA finding that "pupils' understanding could improve their ability to perform and structure numerical and algebraic manipulations and substitutions especially when these include indices". This implies that indices are a weakness for students at university as well as Key Stage 3 pupils as identified in Chapter 2 Section 2. Unfortunately subsequent investigation was not undertaken after syntax difficulties within the Treefrog Undergraduate trial. Further investigation of understanding of indices in learners beyond compulsory education could be beneficial.

Weakness 4 Substitution

The results of the teacher survey in Stage 1 indicated that teachers did not consider general substitution to be an area of weakness. However the results of the teacher survey did indicate agreement with the QCA statement "especially when these include negative quantities in algebra". As detailed in Appendix 25 Table 25.3.8 and Appendix 19 substitution of negative and positive values were assessed in both the Undergraduate and Final Trials and so the results include this. As shown in Sections 8.1.2, 8.1.3 and 8.1.4 there is no evidence from Stages 2, 3 and 6 to support the view that Substitutions are an area of weakness for students in the Undergraduate or Final trials. This finding does suggest that Substitution does not proceed as an area of weakness from school education to beyond. Furthermore it could be construed that substitution is not a general area of difficulty in school learners but in fact it is the complexity of the question posed and specifically the inclusion of negative signs which causes the problems for learners. This view would be supported by the finding in the teacher survey (Stage 1).

Weakness 5 Negative signs and values

Findings from the teacher survey indicate that the opinion of the sample group of teachers concurred with the QCA finding that Negative Signs and Values are an area of difficulty for Key Stage 3 pupils. Findings from the Final Trial discussed in Section 8.1.4 indicate that this is an area of weakness for students in Post Compulsory education. The Undergraduate Trial Treefrog data analysed in Section 8.1.3 indicates that Negative signs and values were problematic for the undergraduate sample group. These results are consistent with the view that Negative signs and values is an area of weakness which continues from key stage 3 to beyond compulsory education and into higher education despite teacher awareness of learners' difficulties in understanding this concept.

Weakness 6 Linear Equations

In Stage 1 as outlined in Section 8.1.1 the teachers in the survey disagreed with the QCA finding that Solving Linear Equations was problematic for school learners. In addition as outlined in Sections 8.1.2, 8.1.3 and 8.1.4 there is no evidence from any of the stages of the investigation to support the view that substitutions are an area of weakness in post compulsory and university students. These findings suggest that Solving Linear Equations is not an area of difficulty for learners in post compulsory education. Furthermore the findings of the teacher survey may indicate that Solving Linear Equations may not be problematic for learners in general in key stage 3.

Summary

The published findings, the opinions of the teachers and the undergraduate and post compulsory education trials within this study support the notion that there is difficulty in understanding two specific areas of Numeracy and Algebra

1. **Negative signs and values**
2. **Brackets**

Indeed these areas of weakness were identified by the QCA at KS3 and GCSE examiners at KS4 and were also evident in the post-compulsory and undergraduate sample groups. In addition the findings of the teacher survey suggest that the lack of understanding regarding negative signs and values has progressed from Key Stage 3 education despite teacher awareness of this problem.

There was some evidence from the Undergraduate Trial Pre Test to support the notion that Indices and Division are areas of weakness in students beyond compulsory education. Yet the results from Stage 3 the Undergraduate Pilot Study and Stage 6 Final Trial did not support this view. Furthermore the findings relating to Division varied between the teachers' opinion and the published findings of the QCA, GCSE examiners and researchers. Further investigation could be beneficial in determining the extent of difficulty in dealing with Division in school pupils and those beyond compulsory education.

In addition the findings relating to Substitutions and Solving Linear equations within Stages 2, 3 and 6 of investigation suggested that these two aspects of mathematics were not more problematic than general mathematics for learners beyond compulsory education. These findings were in accordance with the teachers' view obtained in Stage 1 which contradicted the judgement of the QCA, examiners and researchers. It could be considered that Substitution and Solving linear equations are not areas of general weakness in school learners and hence these weaknesses would not proceed beyond post compulsory education. It could be that errors are made when substitutions and solving linear equations encompass negative signs and values or brackets. Further investigation could focus upon the nature and level of the difficulty for learners in compulsory education and beyond.

Consequently these findings in this study related to Negative signs and values and Brackets supports ***Hypothesis 1: "difficulty in understanding specific areas of Numeracy and Algebra as identified by the QCA at KS3 and KS4 will continue into post-compulsory education"***

8.1.6 Limitations

From these findings it is not possible to compare the extent and levels of difficulty due to the variations in the tools used.

Diagnosys is a widely tested piece of software. Not all questions are posed but marks awarded where the hierarchy would suppose success would have been gained. This may not have been wholly accurate.

Treefrog is software which is under development. Two students out of the group of nine with in the group interview (Stage 5) presented in Chapter 5 Section 6.4 commented that the software was "irritating". The entire interview group complained that "the format of answers was not obvious". It is assumed that students did not attempt questions due to lack of ability but this may not be wholly accurate. These factors could have impacted upon the success rate of students within the trial and led to a variation in the two sets of results. In addition some questions involved more than one of the identified weaknesses and hence could be erroneously answered due to misunderstanding more than one concept. To remove this factor from the analysis would be too complex. To select questions on the basis of measuring the impact of one weakness on another would require a refined focus on the content and strategy. In addition the Pre Test within the Final trial was shorter to complete and paper based. The questions posed in each of these stages although similar in content and level were not the same. All these factors could have affected the results.

8.2 Failure, anxiety and attitude towards mathematics

Hypothesis 2: Learners who experience persistent failure experience anxiety and have a negative attitude towards mathematics.

Within this investigation the opinion of the students collected in the Group interview in Stage 5 provided some information regarding their attitude towards mathematics which could be compared with their behaviour whilst using the system in Stage 3. In addition within the Final trial the log files recorded each individual user actions and responses whilst using either system.

8.2.1 Undergraduate Trial Group interviewees errors made and progress

Analysis of the behaviour of the interviewees involved in the Group Interview in Stage 5 provided specific examples. One candidate as discussed in Chapter 6 Section 6.6 failed to attempt the remaining questions after experiencing continual failure despite being previously successful. This candidate stated in the interview that their confidence was affected negatively by doing these tests and that the experience was de-motivating for those with low level mathematics ability. This could support the views of Kulhavey and Stock(1989) discussed in Chapter 2 Section 5.3 that without feedback negative views and anxiety can be promoted, of Merttens (1997) that to participate and learn users require support and of Laurilland (1999) that learning requires effective feedback. Another candidate despite having a Stage 2 mean score higher than average described herself as being of 'low ability' indicating a poor 'self mathematics concept'. This candidate had been successful in just under half of the questions but had made copious erroneous attempts. The experience of making this number of errors made could be consistent with **Hypothesis 2:** ***Learners who experience persistent failure experience anxiety and have a negative attitude towards mathematics.*** She could have already had a low 'self mathematics concept' and consequently could have expected to be required to try many times before being successful.

The candidate with a low prior mathematics qualification and Treefrog score passed on more than 90% of the questions. This candidate shared the view that the lack of success during the trial had a negative impact upon confidence. This candidate could have already had a low 'self mathematics concept' and hence did not continually attempt to answer the questions. Hence the use of the software may have reinforced their negative attitude toward mathematics.

However another candidate who had a low Pre Test Score in Stage 2 attempted all questions posed in Treefrog despite numerous erroneous attempts. There does not appear to be a relationship between the number of erroneous attempts and behaviour regarding passing on questions or exiting from tests or failing to attempt Tests.

Candidates had suggested in the Group interview in Stage 5 that help should be provided on request, as well as being automatically provide if errors were made.

The group in general had stated that the experience was de-motivating and reinforced poor 'self mathematics concepts' attributable to prior mathematical experiences.

8.2.2 Software evaluation

Within the questionnaire feedback, there was a comment indicating that a user considered the number of tests undertaken to be tedious and another comment comparing Diagnosys to the unadapted Treefrog. These comments may suggest that the use of two computer aided learning systems concurrently impacted upon the attitudes of some users as discussed in Section 5.7.3 and 5.8.3.

8.2.3 Final Trial

Behaviours observed in the student logs included attempting and reattempting to answer questions, trying only once and moving onto the next question, and to stop answering questions. Overview results relating to performance deduced from the log files are outlined in Chapter 7 Section 4.5.6 Table 7.5.11. From these results we can deduce that whilst attempting a correct answer the users of "Webfrog with feedback" made 30% more errors than those using "Webfrog", they attempted 10% more questions and answered at least 50% more correctly. The users of "Webfrog with Feedback" persevered more resulting in a greater level of success. Furthermore the difference between questions attempted and those completed correctly indicates the number of questions users selected to abandon without solution. Webfrog users did not complete the solution of 138 questions whereas Webfrog with Feedback users only failed to complete the solution of 37 questions.

These findings support the view of Merttens (1997) presented in Chapter 2 Section 5.3 that a system could provide essential support for learners who fear getting answers wrong. These learners are at risk of becoming reluctant to participate in further learning. An adaptive system can provide feedback to enable users to become successful and hence reduce the fear and risk. This type of system can prevent reinforcement of the view of poor mathematical ability and hence prevent users from giving up. The level and nature of feedback is fundamental to learning as discussed by Kulhavey and Stock (1989), Laurilland(1999), Mason and Bruning(1999) and others as outlined in Chapter 2.

In general these findings which do not measure attitude and anxiety before and after use of a system and the users' associated level s of failure or success are insufficient to be able to test the hypothesis **Hypothesis 2:** ***Learners who experience persistent***

failure experience anxiety and have a negative attitude towards mathematics.
Alternatively it could be beneficial in further study to investigate the impact that software with supportive feedback could have on improving attitudes and reducing anxiety.

8.3 Validity of GCSE Grading

Hypothesis 3: the GCSE grade or equivalent attained is an adequate indicator of mathematical understanding related to Numeracy and Basic Algebra

Within this investigation GCSE grade attained was recorded in two instances. Within the Undergraduate trial, during use of the Pre Test software Diagnosys, the GCSE grade was required by the software to determine start level. This enabled comparison of GCSE grade with performance of the 71 as presented in Chapter 5 section 4.2 and in the final trial the GCSE grade of the 22 was recorded enabling comparison with performance in the Pre Test as given in Chapter 7 5.5. Furthermore the relationship between GCSE grade attained and each weakness was investigated. In addition a comprehensive analysis of the behaviours and responses of the interview group relating to the GCSE Mathematics grade is presented in Appendix 26.3.

8.3.1 Pre Test in Undergraduate study

The Stage 2 Pre Test scores were consistent with the expectation that students with higher grades would achieve higher marks as discussed in Chapter 5 Section 5.4.2. The Spearman's Rho coefficients ρ=.479 (Statistical Method 2.4) presented in Chapter 5 Table 5.4.5 indicate a positive correlation significant at 1% level between the GCSE coding and Diagnosys scores. Hence the higher the GCSE grade the higher the Diagnosys total, Algebra and number scores. Furthermore the Spearman's Rho coefficients of .409 for Numeracy with GCSE and .472 for Algebra with GCSE suggests that Numeracy ability is less well predicted by GCSE score than algebraic ability. Correspondingly the Spearman's Rho correlation coefficient (ρ=.646) between Start Level and Total Sore given in Chapter 5 Section 4.2 (Statistical Method 2.3) shows a positive correlation significant at 0.01 level. Hence the higher the Start Level then the higher the Total Score.

8.3.2 Group interviewees

In comparing the highest mathematics qualification of this group of seven interviewees with their overall Diagnosys result it appears that the candidates who achieved as presented in Chapter 5 Table 6.4.2

- Grade C or GNVQ Pass at GCSE attained between 18% and 38%
- Grade B or equivalent at GCSE attained 56% and 62%

Furthermore in comparing the highest mathematics qualification of this group of seven with their Diagnosys Number result it appears that the candidates who achieved

- Grade C or GNVQ Pass at GCSE attained between 13% and 60%
  Grade B or equivalent at GCSE both attained 67%

and in comparing the highest mathematics qualification of this group of seven with their Basic Algebra Diagnosys result it appears that the candidates who achieved

- Grade C or GNVQ Pass at GCSE attained between 9% and 27%
- Grade B or equivalent at GCSE attained 45% and 64%

The ranking of these three sets of results are consistent with the view that Diagnosys does reflect the GCSE result. The variance in the sets of results relating to Number and Algebra and the overlap of Grade C and Grade B scores suggests that Numeracy is less well predicted than algebraic ability by the GCSE grade which corresponds to the results discussed in Section 8.3.1.

The Stage 3 Treefrog results for each test of each interviewee are presented in Chapter 6 Table 6.4.2. In comparing the highest mathematics grade previously attained with performance in the Treefrog tests it is evident that the candidates who achieved

- GCSE grade C or GNVQ Pass attained between 3 and 22 marks (7% and 48%)
- GCSE grade B or GNVQ Merit attained 20 and 10 marks (43% and 22%)
- A level grade B attained 42 marks (91%)

It is evident that there is an overlap between the scores for those who previously attained GCSE Grade B and those who attained Grade C. This could contradict the deduction outlined above that the GCSE grade may be a good predictor of algebraic ability. However a GNVQ Merit would be based on Number, Data Handling and Applying Mathematics rather than Algebra. The set of scores with the removal of the 22% relating to the student with prior attainment of GNVQ Merit would still display some overlap between Grade B and Grade C students.

These two sets of results and the scores attained of the interviewees suggest that the level represented by a GCSE Grade C may be lower than that expected by employers or higher education institutions as outlined in Section 1.1 for instance the London Mathematical Society (1995) stated that "many 'high-attaining' students are seriously lacking in fundamental notions of the subject."

. Furthermore the range of marks within this subset of results indicates that these sets of questions encompassed areas of difficulty. Hence that some specific areas of difficulty do continue into post compulsory and higher education.

8.3.3 Stage 6 Final Trial Pre Test scores and GCSE Grade

The results of Statistical Method 6.5 indicate a positive correlation between Final Trial Pre Score and GCSE grade. The overview of the sample group given in Chapter 7 Section 7.5.1 Table 7.5.2 shows that ten candidates had achieved at least GCSE Grade C and twelve had not. The results of Statistical Method 6.6 in Section 7.5.5 Table 7.5.8 show a positive correlation between each of the subsets Weakness 1, 2, 4, 5 and 6 and GCSE significant at 1% level. Only Brackets and Negative signs with GCSE grading have a correlation coefficient >.70 indicating a strong positive correlation with $R^2 > .5$. In addition in Chapter 5 the scatter graphs displayed in Chart 7.5.9 and the Mann Whitney U Test results given Table 7.5.10 indicate that Weakness 2, Brackets and Weakness 5, Negative signs and values are significantly related to the GCSE grade attained. Surprisingly the Mann Whitney results suggest that there is not a significant difference in handling Division, Substitution or Solving Linear Equations between those who have at least Grade C and those who have not. This finding could imply that GCSE grade is an indicator of the ability to handle Brackets and Negative signs and values and hence mathematical ability. The evidence in Section 8.1 shows that Brackets and Negative signs and values are areas of weakness.

8.3.4 GCSE Grade Summary

Table 6.7.1 Summary of GCSE Spearman's Rho Correlation coefficients

	GCSE – Total Diagnosys	GCSE – Diagnosys Number	GCSE – Diagnosys Algebra	GCSE – Pre Test
P	.479	.409	.472	.739
N	71	71	71	22
Sig. 2-tail	<.001	<.001	<.001	<.001

The Diagnosys spread of topics represents a broader range of the GCSE curriculum than the Pre Test content which focused on the identified weaknesses. The Diagnosys score is based on the questions posed which can vary as they are selected from a set of questions at the same level. However if a higher level is not achieved then lower level questions are presented so lower marks can be achieved. Diagnosys is a widely used diagnostic test whereas the Pre Test was derived from the findings of the Undergraduate Pilot Trial a smaller subset of concepts. Hence it could be expected that the correlation between GCSE grade and Diagnosys score would be stronger than that between GCSE grade and Pre Test. In undertaking Spearman's Rho with both sample groups a $p<0.1\%$ was recorded. The findings displayed in Table 6.7.1 show there was a considerable difference in the ρ coefficient values. This variation in value could be attributed to the difference in the size of the sample groups and the range of mathematics tested by Diagnosys and the Pre Test.

The strongest correlation is between the Pre Test and the GCSE grade in the final trial of twenty two participants whereas in the Undergraduate trial there were 71 participants. The Pre Test consisted of 30% number and 70% Algebra questions but has a much stronger correlation with GCSE grade than either of the Number and Algebra Diagnosys subsets. The GCSE grades of the Final trial group ranged between B and E with 10 Grade E participants, three Grade B, seven Grade C and two with Grade D. Whereas 79% of those of the Undergraduate Trial group as outlined in Chapter 5 Table 5.2.2 were either Grade B or C (or equivalent). Consequently the variance of the Undergraduate group GCSE grading was > 360 whilst that of the Final Trial group was < 100. The larger the variance of a data set the smaller the value of the correlation coefficient. Within the Undergraduate trial there is a stronger correlation between the total Diagnosys score and GCSE grade and the weakest between the Number score and the prior grade attained at GCSE. Hence these two sets of results relating to the correlation of the GCSE grade and performance in the trials indicate similar results.

These results indicate that in both instances there is a correlation between the GCSE grade attained and the associated mathematics test. These results imply that the GCSE grade is an adequate indicator of mathematical ability and supports the hypothesis. This is contrary to the requirements of the Teacher Training Agency, as discussed in Chapter 2 Section 3.3, that despite obtaining a Mathematics GCSE grade it is necessary for trainee teachers to undertake a Numeracy test to determine their mathematical ability. This could comply with the finding that GCSE predicted Numeracy ability less well than Algebra. The findings of Statistical Method 2.5 presented in Chapter 5 Table 5.4.6 show a two tailed positive correlation significant at the 1% level between the Total Score and each of the subcategories Number and Basic Algebra. The Wilcoxon Z scores (Statistical Method 2.5) given in Chapter 5 Table 5.4.7 indicate that there is a significant difference in the scores achieved in Number and Basic Algebra significant at 1% level. The ranking of this Z score shows that Number questions are answered more successfully than Basic Algebra questions. This corroborates the finding that only Number scores has a mean value higher than that of the full Diagnosys set. Furthermore Basic Algebra is significantly more difficult than Number. However the TTA Numeracy test is designed to indicate ability to apply mathematics to situations rather than test numerical skills directly. The GCSE mathematics grade is based on points attained from completion of coursework and examination with a weighting of 20-80. The Sunday Times (2006) reported that a candidate can achieve a pass in GCSE Mathematics with an examination score of 16% on the higher paper. Specifically a higher coursework marks combined with an overall pass mark of 28% on the higher paper or an examination score of 35%on the Intermediate paper with an overall pass mark of 43% is required. Employers, higher education and teaching organisations such as the Teacher Training Agency may expect basic numerical skills to have been mastered by those who have achieved a GCSE mathematics qualification. This notion may be supported by the introduction of functional Numeracy requirements within the soon to be introduced 14-19 Diplomas, QCA(2006). In addition the Sunday Times (2006) reports that some schools are intending to introduce the use of IGCSE Mathematics qualifications which do not encompass coursework and are similar to ‘O’ level qualifications. The findings of this study show that GCSE grade is an adequate indicator of mathematical ability correlating to the mathematical tests undertaken but that ability level may be at a lower level than expected. However the level indicated may not indicate that individuals are competent in all numerical skills.

- **These findings support and uphold the hypothesis "*the GCSE grade or equivalent attained is an adequate indicator of mathematical understanding related to Numeracy and Basic algebra*"**

8.4 Effectiveness of computer systems for learning

Hypothesis 4: there is a difference in the effectiveness for learning between computer systems with and without feedback which focuses on common errors

The analysis of the study focusing on hypothesis 4 include findings from stages 4 – 6 the software evaluation, the group interview and the final trial.

8.4.1 Software Evaluation and Group interview

From the seventy questionnaires that were completed and discussed in Chapter 5 Sections 5.7 the following factors were deduced as being identified by users for software design

- Feedback should provide guidance for syntax requirements
- Feedback should provide guidance for method
- Feedback should provide support regarding errors made
- Users should be able to easily navigate backwards and forwards within tests
- All symbols used must be consistent with users' prior learning experiences

These findings were investigated further in Stage 5 the group interview. From this investigation presented in Chapter 6 the following factors were identified for effective educational software design

- Feedback based on errors to outline nature of error made and enable correction
- Explanatory screens to demonstrate methodology
- Example questions to be included
- The ability to navigate backwards and forward through questions
- Format errors explained in feedback (e.g. '=' is required)
- Supportive feedback about success to encourage learners

These features were included in the design of the "Webfrog with Feedback" software used in Stage 6 the Final Trial.

8.4.2 Stage 6 Final Trial

The results of the Wilcoxon signed ranks tests (Statistical Method 6.3) performed on the matched groups ("Webfrog" and "Webfrog with feedback") Pre Test Scores and Post Test scores as presented in Chapter 7 Table 7.5.6 show that the impact of "Webfrog with Feedback" was significant. In addition the post scores and performance of the users of "Webfrog with Feedback" was significantly different from that of the "Webfrog" group.

The two-tailed asymptotic significance estimates the probability of obtaining a Z statistic that is as extreme or more extreme in absolute value as the one displayed, if there truly is no difference between the group ranks. In this case, as presented in Chapter 7 Table

7.5.6, the probabilities for both the two measures of significant differences between the Pre Test and Post Test results for Webfrog with feedback and two sets of Post scores vary. The null hypothesis stated that there would be no difference between the Post Test results of the two matched groups. Furthermore the Z scores indicate that the results of two the pairs Webfrog Post Test and Pre Test scores and that the two sets of Pre Scores are not significantly different. The spread of data discussed in Chapter 7 Section 7.5.3 indicate that the results of the Webfrog with feedback user group are positively skewed to higher values. In addition Chart 7.5.7 displays the difference in progress in learning of the group using Webfrog and those using Webfrog with feedback. These graphically demonstrate the greater improvement in the latter group. These findings indicate that there was a significant difference in the improvement between the two groups which implies that Webfrog with feedback was more effective in supporting learning than Webfrog. These findings support the views discussed in Chapter 2 of Jameson (2003) in Section 5.3 that the quality of feedback is fundamental to learning and that of Kulhavey and Stock (1989) that effective feedback supports learning.

Hence these findings support *Hypothesis 4: there is a difference in the effectiveness for learning between computer systems with and without feedback which focuses on common errors.*

8.4.3 Final Trial and Group interview: Errors and feedback

Eleven candidates using "Webfrog with feedback" received feedback relating to the nature of the error made relating to those identified in Chapter Section 2.6. The eleven Webfrog users were not provided with this feedback relating to the errors made. The frequency of the errors made identified by the associated error code is evidenced in Appendix 24 and a detailed analysis given in Appendix 26.3.

The effectiveness of software is best measured in terms of learner performance. The final trial test results as discussed in Chapter 7 Section 7.5.3 show that "Webfrog with feedback" was more effective than "Webfrog". Charts 7.5.7 presented in Chapter 7 Section 7.5.3 depict the improved scores for each of the participants in the two groups. It was anticipated from the findings of the group interview in consideration of the profile of the group presented in Chapter 6 Section 6.5 that systems with feedback would be effective for learners with lower ability. It is however evident from Chart 7.5.7 that all users of "Webfrog with feedback" across the ability range improved their test score. The design of this software responded to the suggestions and opinions stated in the group interview and software evaluation questionnaire comments. The corrective nature of the feedback given to users focusing on correcting specific errors made, supported the views presented in Section 8.4.1. The discussion presented in Section 8.2.3 suggests that despite making more errors users of "Webfrog with feedback" were not de-motivated.

This could imply that a system with feedback could be a more effective tool than one without. It could be deduced that a benefit of using a computer system is that students can receive immediate feedback. Hence the findings of the final trial and the software evaluation correspond to the view of Monson et al (2001) as discussed in Chapter 2 Section 5.3 that a benefit of using a computer system is that students receive immediate feedback.

8.4.4 Specific errors and feedback

The results of specific questions in Pre and Post Tests discussed in Chapter 7 Section 7.5.2 indicate that software with feedback had been beneficial in developing users' ability to handle Negative signs and values in three instances and Brackets in one. This could suggest that this software was more effective at supporting learning of Negative Signs and Brackets further investigation of this could be valuable.

To summarise, within this study it has been observed by means of the final trial Post Test scores and user behaviours that software which incorporates feedback based on the error made is more effective in supporting learning. This corroborates the views indicated in the Software evaluation and group interview. In addition it could be considered that the mapping of anticipated errors and feedback as discussed in Chapter 7 Section 7.4 has been constructive and could be beneficial to further investigations. As well the software design features identified by learners and incorporated into "Webfrog with feedback" could be considered useful to future studies.

To conclude the findings of this study support

- ***Hypothesis 4: there is a difference in the effectiveness for learning between computer systems with and without feedback which focuses on common errors***

8.5 Summary of findings

This investigation has shown that

- Brackets and negative signs continue as areas of weakness from compulsory school education to post compulsory and higher education. Comparison of GCSE grades and scores for division, substitution and solving linear equation suggest that further investigation of learners could be beneficial. Teachers' opinion regarding substitution and solving linear equations could suggest that published findings should be investigated in a study with school learners and those in education at subsequent levels.
- It would be beneficial to investigate the effect that software which supports and guides learners providing opportunities to be successful could have on personal attitudes towards mathematics. Mathematics anxiety is evidently a widespread issue which creates a barrier to success for many. Learning tools which can improve attitude and reduce anxiety would be valuable to many.
- GCSE grades do indicate a valid ranking of ability however scores attained suggest that the level indicated may be low.
- Software which anticipates errors to enable feedback based on the error made to provide guidance for users is effective in improving learning and ability to answer questions.

Chapter 9 Implications for Future research

This study was concerned with testing the following four hypotheses

Hypothesis 1: ***levels of difficulty in understanding specific areas of Numeracy and Algebra as identified by the QCA at KS3 and KS4 will continue into post-compulsory and higher education***

Hypothesis 2: Learners who experience persistent failure experience anxiety and have a negative attitude towards mathematics.

Hypothesis 3: the GCSE grade or equivalent attained is an adequate indicator of mathematical understanding related to Numeracy and Basic Algebra

Hypothesis 4: there is a difference in the effectiveness for learning between computer systems with and without feedback which focuses on common errors

The discussion presented in Chapter 8 indicates how the findings of this study have provided evidence to support all the hypotheses bar Hypothesis 2. It was recognised that the evidence and data collected were unsuitable to be able to adequately measure the effect of persistent failure on anxiety and attitude towards mathematics. However it is evident that the feeling of anxiety and negativity towards mathematical activity is a widespread concern and the findings related to the benefits of a computer system with feedback which focuses on common errors could suggest that an investigation based on a rephrasing of hypothesis 2 could be beneficial. A study which is able to measure the experiencing mathematical success on anxiety and attitude towards mathematics could be enabled by use of a system which supports and guides the user to develop their own understanding of mathematical concepts. Hence such a study which collects evidence before and after learners have had an opportunity to experience mathematical success could investigate

Learners who experience success have reduced anxiety and an improved attitude towards mathematics.

This study however has shown that the Brackets and Negative signs and values are problematic for some learners across the range of stages from Key Stage 3 and into higher education have been. Furthermore these two areas of common misconception are still evident in those who have passed GCSE mathematics. However the level of success achieved with respect to both of these concepts is significantly different in those who have achieved a pass at GCSE level (at least Grade C) and those who have not achieved Grade C.

However it is not equally evident that the levels of success related to three other areas of weakness considered in this study Substitution, Division and Solving Linear Equations are significantly different between these two groups. Hence it could be invaluable to investigate the ability of Key stage 3 students before undertaking their GCSE Mathematics. This would enable the testing of the hypothesis that

Ability in handling Brackets and Negative Signs and Values is a significantly better predictor of GCSE grade than ability in undertaking Division, Substitutions or Solving Linear Equations.

The use of division, substitution and solving linear equations are fundamental concepts in mathematics and in the application of mathematics in other subjects and in life. The results attained indicate that there was disagreement between the published findings and the views of teachers. Further investigation of Key stage 3 and Key stage 4 learners performance and errors made when undertaking division, substitutions and solving linear equations would enable exploration of these inconsistencies. This would enable the testing of the hypothesis

There are common errors and misconceptions amongst learners relating to Division, Substitution and Solving linear equations which are more problematic for learners at Key Stage 3 and 4 than general mathematics.

In addition this study showed that software with feedback providing guidance on the type of error made was significantly effective in supporting learning. Furthermore it was indicated that software with feedback designed to give feedback on anticipated errors was significantly effective when focused on Negative Signs and Values and Brackets. This suggests that the mapping of anticipated errors and feedback related to Negative signs and Values and Brackets has been more effective than that related to division, substitution and solving linear equations. The study of common errors and misconceptions relating to division, substitution and solving linear equations could enable the refinement of the feedback related to anticipated errors.

This research has compared understanding of different groups of learners in a variety of education by means of various tools. Further research could be conducted to reduce that variations and enable more control and strengthen deductions. Specifically it would be enlightening to conduct a longitudinal study of a group of learners through the national curriculum and progression into post-compulsory education or to explore the behaviours, ability and attitude of three groups of learner by means of the same software package.

A longitudinal study would enable the assessment of the depth of learning, is the understanding maintained long term? Furthermore it could be beneficial to consider, particularly in the earlier stages, does the stage at which understanding is developed have an impact upon levels of anxiety and mathematics 'self concept' of learners? However this study could have ethical implications as some students could be disadvantaged and inadequately supported. Such a study would clearly be lengthy beginning in secondary school and continuing beyond. Other variables which could influence the findings in such an instance would be difficult to control. However a study of this design would enable investigation of the impact of using software on 'deep' learning or long term understanding and the impact of the software at different ages. In consideration of the finding that some difficulties proceed from Key Stage 3 into post compulsory and higher education it would be useful to investigate if the use of software could resolve these difficulties.

In addition it may be advantageous to consider, is there a significant difference in improvement made in terms of developing understanding and reducing errors made in using Brackets and Negative signs and values when used at different stages i.e. KS3, KS4, post compulsory and HE and possibly even KS2?

The findings of this study suggest that the software features identified by the undergraduates in Stages 4 and 5 were effective in supporting the post compulsory education learners. An alternative approach could be to compare the performance of different groups of learners from Key stage 3 and 4, post compulsory education and higher education and the impact of the same software. This could further the investigation of the continuance of areas of weakness from key stage 3 and 4 into post compulsory and higher education.

Manchourini (1999) and others have indicated that the use of a computer system can be effective in learning mathematics. However this study has indicated that there can be significant differences in learning derived from the use of a learning system with adaptive feedback and a 'drill and practice' computer environment in favour of the former. Support for the hypothesis is from the widely accepted theory that learners who receive targeted feedback develop a better understanding. Furthermore software which is visually sophisticated and includes interactivity, which is supportive and relevant, could reduce anxiety and self doubt and hence motivate and encourage learners.

It could be beneficial to develop and refine this research focus to investigate:-

The effectiveness of the software with feedback for learning the correct use of Brackets and Negative signs and values, is there a significant difference in using the software for correcting specific types of errors?

Also it may be beneficial to examine the effectiveness of software with feedback for users of different abilities, questioning if there is a significant difference in learning and reducing errors made in using Brackets and Negative signs and values for users with high, mid-range, low Pre Test scores (or GCSE grades)?

This study has focused the investigation on the impact of software on a small subset of common errors and misconceptions. This research could be refined to focus on the impact of understanding of Negative signs and values and GCSE performance. In the context of the introduction of functional mathematics qualifications which at KS4 will be incorporated within the GCSE as outlined in 14-19 Education and Skills white paper DfES (2006) this could be particularly pertinent. This significant change in GCSE Mathematics follows the recommendations of Smith (2004). Within this report it is identified that "it has been widely recognised that mathematics occupies a rather special position. It is a major intellectual discipline in its own right, as well as providing the underpinning language for the rest of science and engineering and, increasingly, for other disciplines". In addition it is documented that mathematics "provides the individual citizen with empowering skills for the conduct of private and social life and with key skills required at virtually all levels of employment". The terms of reference for the Post-14 Mathematics Education Inquiry stated in Smith(2004) were to "make recommendations on changes to the curriculum, qualifications and pedagogy for those aged 14 and over ... to enable students to acquire the mathematical knowledge and skills necessary to meet the requirements of employers of further and higher education." Research which embraces the advancement of the understanding of fundamental mathematical concepts is essential for the development of individuals, society and economics.

Contrastingly the notion of using software to emulate the feedback that a teacher would provide based on the nature of an error made could be applied to a variety of learning contexts. The investigation of the effectiveness of this type of software to the learning and understanding could be extended to a wide variety of aspects of mathematics and a variety of other subjects. For instance within the teaching of English there are instances, such as in the teaching of punctuation and grammar where teachers apply this approach and hence could be programmed into a system.

Chapter 10 References

Abdin, B and Hartley, J (1998) Developing mathematical problem solving skills *Journal of Computer Assisted learning* 14, 278-291

Aczel, J & Tilley, D (1998) Algebra: rebalancing the equation. *Micromath* 14(2) 1998 11-13 ISSN: 0267-5501

Advisory Committee on Mathematics Education (2002) *Continuing Professional Development for teachers of mathematics* ACME The Royal Society

Al-Jumeily, D & Strickland, P (1998) Evaluating the Use of an Intelligent Computer Algebra System as a Learning Tool, *Evaluation of Learning Technology*, University of North London.

Armitage, S and O'Leary, R (2003) *A Guide for Learning Technologists* e-Learning Series No 4 LTSN Generic Centre

Aşkar P. and Akkoyunlu B., (1993), Kolb learning style inventory, *Science and Education, 87*, 37-47

Atkins, M (1993) Theories of learning and multimedia applications: an overview. *Research Papers in Education,* 8(2), 251-271.

BECTa (1999) *Evaluation of UK Integrated learning systems* BECTa/DfEE

Beevers, C, Paterson, J, (2003) Automatic Assessment of Problem-Solving Skills in Mathematics Source: *Active Learning in Higher Education* Volume: 4 Number: 2 Page: 127 – 144

Bell, J (1992) *Doing your research project: a guide for first-time researchers in education and social science* 2nd Edition Open University Press

Blake C, Davies C, Jones A, Morris E, Scanlon E (2003) Evaluating complex digital resources *ALT-J* Volume 11 Number 1 p 4 - 16

Bloom, B (1964) *The Taxonomy of Educational Objectives*, London: Longman

Bloom, B (1984) 'The 2 sigma problem: the search for methods of group instructions as effective as one-to-one tutoring' *Educational Researcher*, 13(6): 4-16

Boyd, W, Cullen, M Bass, D, Pittman, J & Regan, J (1998) A Response to Apparently Low Levels of Numeracy and Literacy Amongst First Year University Environmental Science Students: A Numeracy and Literacy Skills Survey *International Research in Geographical and Environmental Education* Vol 7 Number: 2 Page: 106 – 121

Branson, R (2003) cited at http://www.lifelonglearning.co.uk/mosergroup/index.htm accessed on 27/5/03

Brown, G (1997) *Assessing student learning in higher education* Routledge

Brown. S, & McIntyre, D (1981), An action-research approach to innovation in centralized educational systems, *European Journal of Science Education*, 3.

Delozanne, E., Bruillard, E. (1993). Mathematical Learning Environment: A French Viewpoint in Nwana, H. S. (eds) (1993) *Mathematical Intelligent Learning Environments* London : Intellect Mathematical Intelligent Learning Environment. S. Hyacinth. Nwana, Intellect Books: 14-32.

Cahill, R (1997) Choosing between qualitiative and quantitative approaches. *N.T Research*, 2 (3), pp 185-186

Clausen-May, T (1998) *Common Errors – Common Understanding: the development of KS3 mathematics test for fourteen-year-olds* Paper presented at BERA 1998 Conference

Coben, D, (2004) *Adult Numeracy: Review of research and related literature* National Research and Development Centre for Adult Literacy and Numeracy available at http://www.nrdc.org.uk/ accessed on 12/07/06

Cockcroft, W (1982) *Mathematics counts* Report of the Committee of Inquiry into the Teaching of Mathematics in Schools under the Chairmanship of Dr WH Cockcroft London: HMSO

Conole, G (2002) The evolving landscape of learning technology *ALT-J* Vol 10 Pt 3 4-18

Cooper, B, Harries, T (2003) Children's use of realistic considerations in problem solving: some English evidence *Journal of Mathematical Behaviour 22(2003) 451-465*

Cotton, J (1995) *The Theory of Assessment: An Introduction* KOGAN PAGE ISBN 0-7494-1709-9

Cox, M (1997) *The Effects of Information Technology on students' motivation.* Final Report NCET/King's College London 1997 ISBN 871984-28-9

Cox, M, Preston, C, Cox, K (1999) *What factors support or prevent teachers from using ICT in their classrooms?* Paper presented at the British Research Association Annual Conference, September 2-5 1999

Crook, C (1994) *Computers and the Collaborative Experience of Learning* 1st Edition Routledge

Cunningham in RAVEN, J & STEPHENSON, J (Eds), (2001) *Competence in the Learning Society* New York, Peter Lang

Dearing, R (1996) *Review of qualifications for 16-19 year olds* SCAA Publications London

Dearing, R (1997) *Higher education in the learning society.* Report of the National Committee of Inquiry into Higher Education chaired by Sir R Dearing NCIHE

Dede, C. (1998) *Learning about teaching and vice versa.* Paper presented at Conference of Society for Information Technology in Education. Washington D.C., USA

Denscombe, M (2001) *The Good Research Guide: For Small-scale Social Research Projects* Open University Press

Denzin and Lincoln, (1998) *The Landscape of Qualitative Research* Thousand Oaks Sage Publications Inc

Department for Education and Employment (1998a) *Teachers meeting the challenge of change.* DfEE Green paper

Department for Education and Employment (1998 b) *Numeracy Matters Preliminary Report* DfEE available at http://www.dfes.gov.uk/numeracy/contents.shtml accessed on 01/12/98

Department for Education and Employment (2000) *Professional Development: Support for Teaching and Learning.* DfEE 0008/2000.

DfES (2002) *National Statistics SFR* 26/2002 DfES

DfES (2003a) *Towards a unified e-Learning Strategy Consultation Document* DfES/0424/2003

DfES (2003b) *Integrating ICT into Mathematics in Key Stage 3* DfES ref 0332/2003

DfES(2003c) *The Future of Higher Education* HMSO CM 5735 accessed at http://www.dfes.gov.uk/hegateway/uploads/White%20Pape.pdf on 4/8/05

Dooley, M, Metcalf, T, Martinez, A (1999) A study of the Adoption of Computer Technology by Teachers *Educational Technology & Society* 2(4) ISSN 1436-4522

Drake, P (2002) Testing Teachers: perspectives on a pilot numeracy skills test in England *Assessment in Education: Principles, Policy and Practice* Vol 9 Number: 2 Page: 205 – 220

Donaldson, M (1963) *A study in childrens thinking.* Tavistock Publications, London.

Ecclestone K (2003). forward in Fawbert F (eds) *Teaching in Post Compulsory Education – Learning Skills and Standards* Continuum International Publishing Group

Edexcel (2000) *GCSE Mathematics (1385/1386)* Summer 2000 Publications Code UG 009272

Forsyth, I (1996) *Teaching and Learning Materials and the Internet* London: Kogan Page

Gillespie, J (2003) *Higher education and key skills* ILT Members website accessed 11/06/2003

Githua, B, Gowland Mwangi, J (2003) Students' mathematics self –concept and motivation to learn mathematics : relationship and gender differences among Kenya's secondary school students in Nairobi and Rift Valley provinces *International journal of educational development* 23 (2003) 487-499

Goos, M, Galbraith, P, Renshaw, P and Geiger, V (2003) Perspectives on technology mediated learning in secondary school mathematics classrooms *The Journal of Mathematical Behavior* Volume 22, Issue 1 , 2003, Pages 73-89

Goodman, P (2001) Eds *Technology Enhanced Learning Opportunities for change* LEA Publishers USA

Gugerty, L (1997) Non-diagnostic intelligent tutoring systems: Teaching without student models *Instructional Science* 25(6)Page: 409 – 432

Harding R(2001) What have examinations got to do with computers in education *Journal of Computer Assisted Learning* 17, 322-328

Haspekian, M (2003) *Between arithmetic and algebra: a space for the spreadsheet? contribution to an instrumental approach.* Presented at CERME 3

Hibberd, S, Litton, C, Chambers, C and Rowlett, P (2004) *Embedding CAA and Support for Mathematics in a Web-based Learning Environment* LTSN Mathstore Maths CAA series Apr 2004 accessed at http://ltsn.mathstore.ac.uk/articles/maths-caa-series/apr2004/ on 05/05/04

Ho, H, Sendturk, D, Lam, A, Zimmer, J, Hong, S, Okamato, Y, Chiu, S, Nakazawa, Y, Wang, C (2000) The affective and cognitive dimensions of mathematical anxiety *Journal for Research in Mathematics Education,* Vol 31 No 3 pp 362-379

Honey, P and Mumford, A (1992) *The manual of learning styles (3rd ed)* Peter Honey. Maidenhead.

Jameson, J (2002) Intuitive expertise in ICT graduate supervision *ALT-J* 10(3)

JMU Learning Development Team (1999) *Learning Teaching and Assessment Strategy* Internal publication supported by HEFCE

Joffe, L, Foxman, D (1988) *Attitudes and Gender Differences* Berkshire: Peter Honey

Jones, G. and Knezek, G. (1993) Non-commercial radio-satellite telecommunications: affordable options for technology educators. Cited in S. Romi (2000) Distance Learning and Non-formal Education: Existing Trends and New Possibilities of Distance Learning Experiences. *Educational Media International.* 37 (1), 39-44.

Kolb, D (1984) *Experiential Learning. Experience as the source of learning and development.* Englewood Cliffs. Prentice Hall. New Jersey.

Kramarski, B & Hirsch, C (2003) 'Using computer algebra systems in mathematical classrooms' *Journal of Computer Assisted Learning* (2003) 19, 35-40

Kulhavy, R, Stock, W (1989) Feedback in written instruction. The place of response certitude. *Educational Psychology review* 1(4), 279-308

Lakatos, I (1981) history of science and its rationale reconstructions pp.107-127 In. Hacking, I ed. *Scientific Revolution.* University Press. Oxford.

Lalonde, R, Gardner, R, (1993), 'Statistics as a Second Language? A model for Predicting Performance in Psychology Students' *Canadian Journal of Behavioural Science* 25:1, 108-125

Laurillard, D (2001) *Rethinking University Thinking*, 2nd Edition, RoutledgeFalmer

Lilley, M, Barker, T and Britton, C (2004) The development and evaluation of a software prototype for computer adaptive testing *Computers and Education* 43 109-123

London Mathematical Society (1995) *Tackling the Mathematics Problem* London Mathematical Society accessed online *at http://www.lms.ac.uk/policy/tackling/report.ps* on 28/5/03

Mackenzie, S (2002) Can we make maths count at HE? *Journal of Further and Higher Education,* 26 (2)

Manouchehri, A (1999) Computers and School Mathematics Reform: Implications for Mathematics Teacher Education, *Journal of Computers in Mathematics and Science Teaching* 18(1), 31-48

Manouchehri, A (2004) Using interactive algebra software to support a discourse community *Journal of Mathematical Behavior* 23, 37-62

Mason, B.J., Bruning, R.; (1999) *Providing feedback in computer-based instruction: What research tells us* available at http://dwb.unl.edu/Edit/MB/MasonBruning.html Centre for instructional Innovation; University of Nebraska-Lincoln;accessed on 04/11/04

Matz, M (1982). Towards a process model for high school algebra errors in D. Sleeman. *Intelligent Tutoring systems.* London, AP Academic Press

McClelland, R.J. (2002) Open/Distance Learning Materials, Research Methods for Social Sciences, Liverpool John Moores University

McClelland, R, Yolles, M (1997) The cybernetics of learning environments. CTI-AFM 8th Annual Conference. Newport Gwent. 8/9th April 1997.

McDougall, A 'Technology-supported environments for learning through cognitive conflict' *ALT-J* 10 (3)

McNamara, O, Roberts, L, Basit, T, Brown, T (2002) Rites of Passage in Initial Teacher Training: ritual, performance, ordeal and Numeracy Skills Test *British Educational Research Journal Vol. 28, No. 6*

Merttens R. (ed) (1997) *Teaching Numeracy: maths in the primary classroom.* Leamington Spa: Scholastic Ltd

Monaghan, J (1998) Moving to regular use of ICT in mathematics classes. *Micromath.* Vol 14, no. 3: Autumn 98, p10-14

Monson, R, Judd, K (2001) CalMaeth: An interactive learning system focussing on the diagnosis of Mathematical misconceptions *Journal of computers in mathematics and Science Teaching* 20(1) 19-43

Moser C (2002) *Improving literacy and numeracy A Fresh start* Basic Skills Agency Report of the working group chaired by Claus Moser
http://www.lifelonglearning.co.uk/mosergroup/

Mutch, A (2003) Exploring the practice of feedback to students *Active learning in higher education* Vol 4 Num 1 24-38

National Council for Educational Technology (1994) *IT Works* BECTa

National Council of Teachers of Mathematics. (2000) *Principles and standards for school mathematics.* Reston, VA: NCTM

Nicol, M and Anderson, A (2000) Computer assisted vs. teacher-directed teaching of numeracy in adults *Journal of Computer Assisted Learning* 16, 184-192

Nielsen, J (1992) Finding usability problems through heuristic evaluation. In Bauersfield, P., Bennet, J & Lynch, G (eds.), *Human Factors in Computing Systems* CHI'92 Conference Proceedings (pp.373-380). New York Academic Press

Nielsen, J (1994) Usability inspection methods. In Neilsen,J & Mack, R.L. (eds.) *Usability Inspection Methods.* John Wiley, p.30

Passey, D, Rogers, C, Machell, J and McHugh, G (2004) *The motivational Effect of ICT on Pupils* Research Report RR523 for Department for Education and Skills

Papert, S (1980) *Mindstorms: Children, Computers and Powerful Ideas.* New York: Basic Books First Edition

Perry, C, Ball, I (2004) Teacher Subject Specialisms and their Relationships to Learning Styles, Psychological Types and Multiple Intelligences: implications for course development *Teacher Development, Volume 8, Number 1*

Polit, D, Hungler, B (1995) *Nursing Research. Principles and Methods,* (5th ed.) JB Lippincott, Philadelphia, Pennsylvannia.

Popper, K (1981) The rationality of scientific revolutions pp80-106. In Hacking, I ed. *Scientific Revolutions.* Univeristy Press. Oxford.

QCA(1999) *Standards at key stage 3 Mathematics A report for head teachers, heads of department, mathematics teachers and assessment coordinators on the 1999 national* *curriculum assessments for 14-year-olds* accessed at http://www.qca.org.uk/downloads/6052_standards_ks3_maths_19.pdf on 04/08/05

Social and Market Strategic Research (SMSR) (2000) Report to QCA on *The evaluation of statutory assessment in 2000 Key stage 3*

QCA(1999) *Standards at Key Stage 3 Mathematics* QCA/99/517

QCA(2000a) *National curriculum programme of study: Information Communication Technology, key stage 3* accessed on line 13/6/03 at www.nconline.net.uk

QCA(2000b) National curriculum programme of study: Mathematics, key stage 3 on line 13/6/03 at www.nconline.net.uk

QCA(2000c) *Standards at key stage 3 Mathematics A report for headteachers, heads of department, mathematics teachers and assessment coordinators on the 2000 national curriculum assessments for 14-year-olds* accessed at http://www.qca.org.uk/downloads/6055_standards_ks3_maths_20.pdf on 04/08/05

QCA(2001) *Standards at key stage 3 Mathematics A report for headteachers, heads of department, mathematics teachers and assessment coordinators on the 2001 national curriculum assessments for 14-year-olds* ISBN 1 85838 461 3 accessed at http://www.qca.org.uk/downloads/6059_standards_ks3_maths_2001.pdf on 04/08/05

QCA(2002) *Standards at key stage 3 Mathematics A report for headteachers, heads of department, mathematics teachers and assessment coordinators on the 2002 national curriculum assessments for 14-year-olds* accessed at http://www.qca.org.uk/downloads/6063_standards_ks3_maths_2001.pdf on 04/08/05

QCA(2003) *Standards at key stage 3 Mathematics A report for headteachers, heads of department, mathematics teachers and assessment coordinators on the 2003 national curriculum assessments for 14-year-olds* accessed at http://www.qca.org.uk/downloads/6409_ks3_ma_standards_report_2003.pdf on 04/08/05

Ramsden, P (2000*) Learning to teach in higher education* Routledge Falmer

Reynolds, D and Muijs, D (1999) The Effective Teaching of Mathematics: a review of research. *School Leadership & Management*, Vol 19, No 3, pp 273-288

Riall, R, Burghes, D (1996) *Mathematical needs of Young Employees Centre for Innovation in Mathematics Teaching*, School of Education, University of Exeter SCAA report

Ricketts, C, Wilks, S (2002) Improving student performance Through Computer based Assessment. *Assessment & Evaluation in Higher Education*, Vol 27 ,No. 5,

Rist, R and Hewer, S (1996) What is learning technology? some definitions in Stoner, G (Eds) LTDI: *Implementing Learning Technology* LTDI Electronic Text

Robson, C (1999) *Real World Research. A Resource for Social Scientists and Practitioner-Researchers.* Blackwell. Oxford UK

Roscoe, J (2003) E-learning in higher education Optimising e-learning in HE, evolution or revolution? On ILT members website accessed on 11/06/2003

Rowntree, D (1989) *Assessing Students: How shall we know them?* Revised edition, London: Kogan Page

Sangwin, C (2002) *New opportunities for encouraging higher level mathematical learning by creative use of emerging computer aided assessment.* Available on line at http://www.mat.bham.ac.uk/C.J.Sangwin/

Saunders, G, Klemming, F (2003) Integrating technology into a traditional learning environment: reasons for risks and reasons for success *Active learning in higher education* Vol 4 (1) 74-86

Schofield, D, Ashton, H (2005) *Effective Reporting for online assessment – shedding light on student behaviour* Mathematics CAA Article accessed on line at http://mathstore.ac.uk/articles/maths-caa-series/feb2005/ [28 Feb 2005]

Seale, J (2002) *Using CAA to Support Student Learning* [online] LTSN Generic Centre Starter Guide (4) Available from http://www.ltsn.ac.uk/genericcentre/index.asp?id=17730 [30 July 2003]

Shepherd, K (2002) *Streaming audio and video for Course design* [online] LTSN Generic Centre Starter Guide (4) Available from http://www.ltsn.ac.uk/genericcentre/index.asp?id=17730 [30 July 2003]

Smith, A (2004) *Making Mathematics Count The Report of Professor Adrian Smith's Inquiry into Post-14 Mathematics Education* DfES 03/04 937764

Spooner, M (2002) *Errors and misconceptions in maths at Key Stage 2: working towards successful SATS* David Fulton Publishers

Squires, D & McDougall, A (1994) *Choosing and Using Educational Software: A Teachers' Guide*. London: Falmer Press

Squires, D, & Preece, J (1999) Predicting Quality in Educational Software: Evaluating for learning, usability and the synergy between them, *Interacting with Computers*, 11, 467-83

Stubbs, W, McClure, R, (1992), Joint address to the 1992 FEFC Conference, FEFC, Coventry, 9 – 10 December 1992

TTA (2004) *The use of ICT to support literacy and mathematics* available at http://www.standards.dfes.gov.uk/primary/publications/literacy/846217/hmi1973_nlns_primarycurric.pdf

TIMSS (1996) *Pursuing Excellence: A Study of U.S. Eighth-Grade Mathematics and Science Teaching, Learning, Curriculum, and Achievement in International Context: Initial Findings From The Third International Mathematics and Science Study* NCES Number: 97198

TIMSS (2003) *Teaching Mathematics in Seven Countries: Results from the TIMSS 1999 Video Study* NCES Number: 2003013

Trowler. P.(2001), *Policy Update (Post Compulsory) - updates material in Chapter Two of Trowler. P. (August 1998) Education Policy: a policy sociology approach* Routledge accessed at http://www.lancs.ac.uk/staff/trowler/poluppost.htm on 4/8/05

Tuckman, B (1999) *Conducting Educational Research* 5th Edition. Harcourt Brace College Publishers USA

Twining, P (2002) Conceptualising Computer Use in Education: Introducing the Computer Practice Framework (CPF), *British Educational Research Journal* 28(1): pp.95-110

Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press

Wikipedia (2005) www.wikipedia.org accessed on 16/8/05

Wilson D (1996) Self Paced Studies. *Chronicle of Higher Education*, XLII (21), Feb 2, A19-A20

Winship, J (1988) *Software Review or Evaluation: Are They Both Roses Or Is One a Lemon?* Paper presented at the Proceedings of the Australian Computer Education Conference, Perth.

Appendix 1 National Curriculum Levels

relating to **Attainment target 2: Ma2 Number and algebra**

Level 1

Pupils count, order, add and subtract numbers when solving problems involving up to 10 objects. They read and write the numbers involved.

Level 2

Pupils count sets of objects reliably, and use mental recall of addition and subtraction facts to 10. They begin to understand the place value of each digit in a number and use this to order numbers up to 100. They choose the appropriate operation when solving addition and subtraction problems. They use the knowledge that subtraction is the inverse of addition. They use mental calculation strategies to solve number problems involving money and measures. They recognise sequences of numbers, including odd and even numbers.

Level 3

Pupils show understanding of place value in numbers up to 1000 and use this to make approximations. They begin to use decimal notation and to recognise negative numbers, in contexts such as money and temperature. Pupils use mental recall of addition and subtraction facts to 20 in solving problems involving larger numbers. They add and subtract numbers with two digits mentally and numbers with three digits using written methods. They use mental recall of the 2, 3, 4, 5 and 10 multiplication tables and derive the associated division facts. They solve whole number problems involving multiplication or division, including those that give rise to remainders. They use simple fractions that are several parts of a whole and recognise when two simple fractions are equivalent.

Level 4

Pupils use their understanding of place value to multiply and divide whole numbers by 10 or 100. In solving number problems, pupils use a range of mental methods of computation with the four operations, including mental recall of multiplication facts up to 10 and quick derivation of corresponding division facts. They use efficient written methods of addition and subtraction and of short multiplication and division. They add and subtract decimals to two places and order decimals to three places. In solving problems with or without a calculator, pupils check the reasonableness of their results by reference to their knowledge of the context or to the size of the numbers. They recognise approximate proportions of a whole and use simple fractions and percentages to describe these. Pupils recognise and describe number patterns, and relationships including multiple, factor and square. They begin to use simple formulae expressed in words. Pupils use and interpret coordinates in the first quadrant.

Level 5

Pupils use their understanding of place value to multiply and divide whole numbers and decimals by 10, 100 and 1000. They order, add and subtract negative numbers in context. They use all four operations with decimals to two places. They reduce a fraction to its simplest form by cancelling common factors and solve simple problems involving ratio and direct proportion. They calculate fractional or percentage parts of quantities and measurements, using a calculator where appropriate. Pupils understand and use an appropriate non calculator method for solving problems that involve multiplying and dividing any three-digit number by any two-digit number. They check their solutions by applying inverse operations or estimating using approximations. They construct, express in symbolic form, and use simple formulae involving one or two operations. They use brackets appropriately. Pupils use and interpret coordinates in all four quadrants.

Level 6

Pupils order and approximate decimals when solving numerical problems and equations [for example, $x^3 + x = 20$], using trial and improvement methods. Pupils are aware of which number to consider as 100 per cent, or a whole, in problems involving comparisons, and use this to evaluate one number as a fraction or percentage of another. They understand and use the equivalences between fractions, decimals and percentages, and calculate using ratios in appropriate situations. They add and subtract fractions by writing them with a common denominator. When exploring number sequences, pupils find and describe in words the rule for the next term or nth term of a sequence where the rule is linear. They formulate and solve linear equations with whole number coefficients. They represent mappings expressed algebraically, and use Cartesian coordinates for graphical representation interpreting general features.

Level 7

In making estimates, pupils round to one significant figure and multiply and divide mentally. They understand the effects of multiplying and dividing by numbers between 0 and 1. Pupils solve numerical problems involving multiplication and division with numbers of any size, using a calculator efficiently and appropriately. They understand and use proportional changes, calculating the result of any proportional change using only multiplicative methods. Pupils find and describe in symbols the next term or nth term of a sequence where the rule is quadratic; they multiply two expressions of the form $(x + n)$; they simplify the corresponding quadratic expressions. Pupils use algebraic and graphical methods to solve simultaneous linear equations in two variables. They solve simple inequalities.

Level 8

Pupils solve problems involving calculating with powers, roots and numbers expressed in standard form, checking for correct order of magnitude. They choose to use fractions or percentages to solve problems involving repeated proportional changes or the calculation of the original quantity given the result of a proportional change. They evaluate algebraic formulae, substituting fractions, decimals and negative numbers. They calculate one variable, given the others, in formulae such as $V = Yr^2h$. Pupils manipulate algebraic formulae, equations and expressions, finding common factors and multiplying two linear expressions. They know that $a^2 - b^2 = (a+b)(a-b)$. They solve inequalities in two variables. Pupils sketch and interpret graphs of linear, quadratic, cubic and reciprocal functions, and graphs that model real situations.

Exceptional Performance

Pupils understand and use rational and irrational numbers. They determine the bounds of intervals. Pupils understand and use direct and inverse proportion. In simplifying algebraic expressions, they use rules of indices for negative and fractional values. In finding formulae that approximately connect data, pupils express general laws in symbolic form. They solve simultaneous equations in two variables where one equation is linear and the other is quadratic. They solve problems using intersections and gradients of graphs.

Appendix 2 Derivation of the QCA Statements questionnaire

(Questions which cite these findings formulate the questionnaire in appendix 5)

2a Analysis of Standard Assessment Tests Reports

The QCA KS 3 Mathematics SATS Report identified the following common pupil errors and misunderstandings within the specified numerical and algebraic skills. A questionnaire (included in appendix 3) based upon these errors was designed to enable investigation of the findings.

Number Skills

At level 3:

Calculator arithmetic was not attempted by 25% of pupils

62% were unable to give the missing number in 962 - ... = 476 *(Questionnaire question 1)*

92 % were unable to complete ... / 24 = 16 *(Questionnaire question 2)*

At level 5 when solving the Linear Equation typical arithmetic errors of the form

$$4y = 6 \Rightarrow y = 1/3$$

were made by the majority of pupils. *(Questionnaire question 3)*

Linear Equations

At level 5 most pupils were unable to find y from 4-2y=10-6y but 63% did know that a standard approach to collecting together like terms. Most did then undertake sensible connected steps although many made errors in manipulation or arithmetic.

(Questionnaire questions 4 & 5)

Algebra

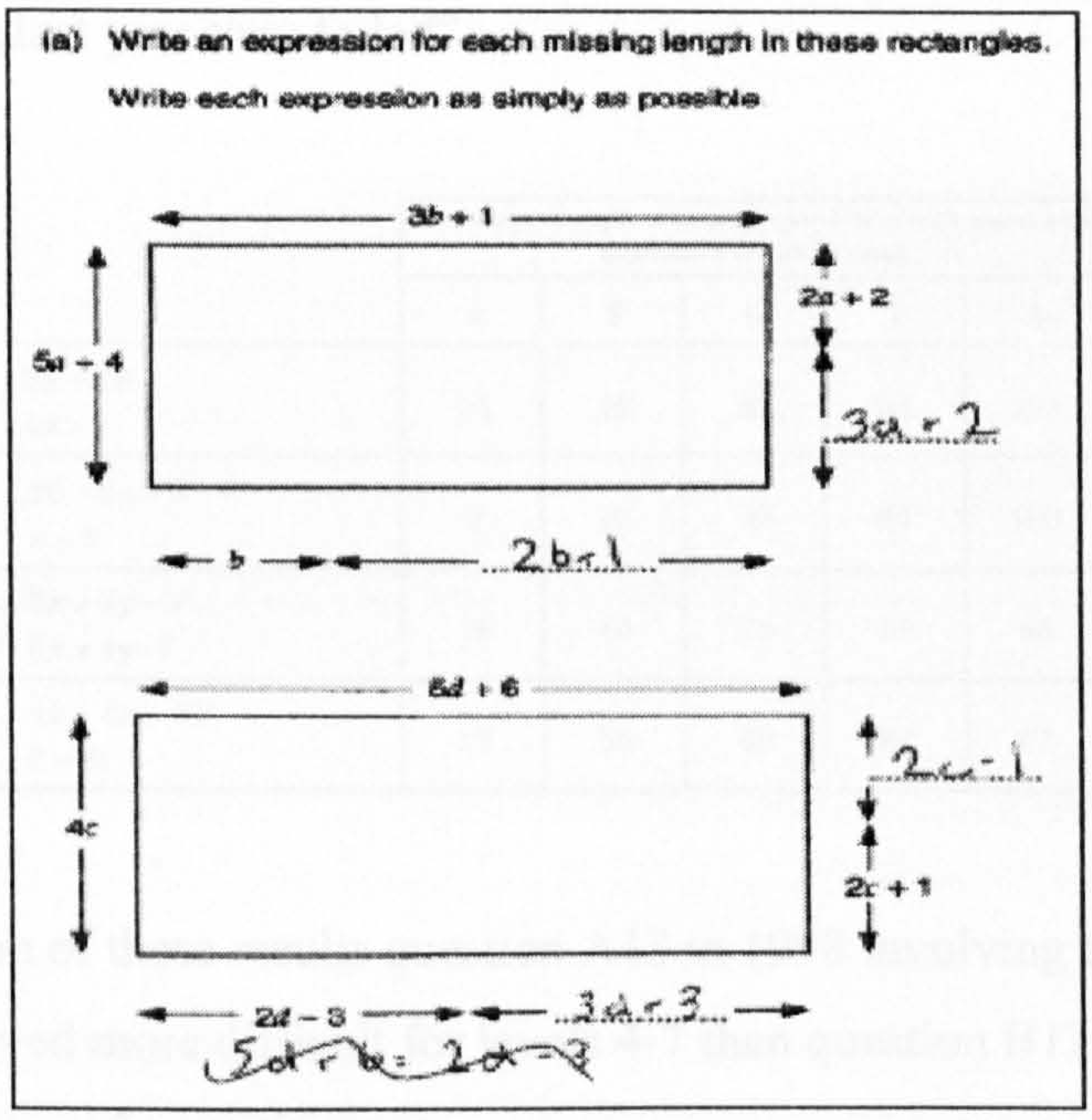
Algebraic manipulation errors when handling negative signs and negative numbers

In linear equations, level 5 $4-2y=10-6y \Rightarrow 4y = 14 \Rightarrow y=3.5$

(Questionnaire question 6)

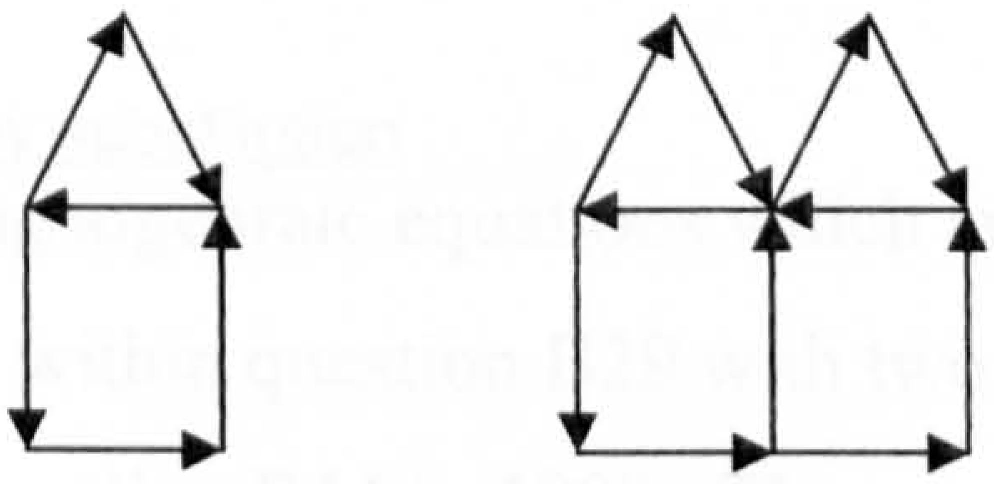
In Algebra, level 6 in 'Rectangle' shown below 62% gained both marks for the first rectangle , that is 3b+1 = b+ .... and 5a-4 = 2a+2 + ... whereas only 43% successfully solved the required equations for the second rectangle, that is 5d + 6 =

2d – 3 + …… and 4c = 2c+1 + …..both of which will require the manipulation of negative numbers. *(Questionnaire questions 7 & 8)*



In fact the report comments that generally, at higher levels, when the substitutions, manipulations and work with equations became more complex, errors appeared in pupils' answers. *(Questionnaire question 9)*

At level 6, 56% had difficulty writing an algebraic expression and 78% were not able to substitute accurately. Whilst at level 7, 66% were unable to write the required expression from a graphically given sequence of the type shown below



1 hut needs 6 matches 2 huts need 11 matches

(Questionnaire questions 10 & 11)

Also at level 7, 59% were unable to solve the equation (FINDING Y)

Solving expressions with one variable

It was noted that "the use of only one variable and the familiar nature of the expression in 1999 may have enabled good proportions of pupils at lower levels to succeed where last year they failed".

			Percentage of pupils answering correctly at each stated level (calculated across all tiers)				
			4	5	6	7	8
1999	B11	$3k = 18$ $5k = ?$	24	49	82	98	100
1998	A13	$26 - 2n = 8$ $n = ?$	2	22	59	84	100
1999	A13	$3x + 2y = 7$ $6x + 4y = ?$	16	40	73	89	88
1998	A8	$4k + 2n = 82$ $? = 41$	48	58	69	87	67

On examination of these results question A13 in 1998 involving two separate operations proved more difficult for levels 4-7 than question B11 in 1999 which only required one type of operation.

(Questionnaire questions 13, 14, 15 & 16)

Solving proportional expressions

The relationship between the two equations in question A13 1999 required the comparison of two algebraic expressions proved significantly more problematic at lower levels than that between the two equations in question A8 1998 which required the comparison of two numerical values.

(Questionnaire question 12)

Solving by substitution

In solving algebraic equations which are not proportionally related the substitution required within question B29 with two variables is conceptually more complex than that of question B14 in 1998. The corresponding results support this conjecture of associated difficulty. In 1999 when solving by substitution given one equation from which the value of the variable can be deduced for another 78% at level 4, 54% at level 5 and 28% at level 6 were unable to achieve the correct result. Similarly in 1998 when solving by substitution where the value of the variable can be deduced in

	Percentage of pupils answering correctly at each stated level (calculated across all tiers)				
	3	4	5	6	7
1999 *Arms (c)*	5	26	64	81	95
1996 *Crosses (d)*	19	41	72	88	97

	Percentage of pupils answering correctly at each stated level (calculated across all tiers)				
	4	5	6	7	8
1999 B14 $x - 1 = 45$ $x + 2 = ?$	22	46	78	92	100
1998 B29 $m + 7 = t$ $m + 5 = ?$	0	5	16	41	69

terms of an algebraic expression to be substituted into the given equation 100% at level 3, 95% at level 5, 84% at level 6, 59% at level 7 and 31% at level 8 were unable to answer the question correctly.

(Questionnaire question 11)

Number Patterns and Sequences

In 1996 the given sequence was *(1+ 4N)* representing one central square with one square added per side for each term. In 1999 the tessellation sequence was *(1+3N)* representing a central triangle with one triangle added per side for each term. Some results are suitably comparable across several years.. When asked to explain the term *4N* within the formula the results show that at level 5 and 6 there was a marked improvement from 1996 to 1999. However, 45% at level 5 and 42% at level 6, in 1999 were unable to answer correctly. Levels 3 and 4 also reported an improvement in performance but notably at level 3, 82% of pupils and at level 4, 58% of pupils were unable to answer correctly. *(Questionnaire question 10)*

Substitute a numerical value

An expectation of level 4 attainment is that pupils will be able to calculate the nth term in a sequence. In the same questions referred to above, when asked to solve the expression for a given value of N, pupils achieved similar results with a slight improvement at levels 4 and 5 as shown in the table below.

	Percentage of pupils answering correctly at each stated level (calculated across all tiers)				
	3	4	5	6	7
1999 *Arms (b)*	9	41	75	84	94
1996 *Crosses (c)*	9	38	69	85	95

Hence in 1999, 81% of level 3 pupils, 59% of level 4 pupils, 25% at level 5 and 15% at level 6 unable to accurately substitute a numerical value into a formula without any other manipulations. *(Questionnaire question 11)*

Inverse calculations and negative signs

When pupils were asked to perform an inverse calculation to establish which term in the sequence would comprise of a given number of tiles they were required to transpose the equation which would result in a negative sign. Comparison of the two years results show that they are consistently poorer in 1999 than in 1996 at all levels.

There is a marked deterioration at levels 3 and 4, a smaller reduction at level 5 and similar results at levels 6 and 7 although these too appear weaker. These outcomes indicate that in 1999, 95% of pupils at level 3, 74% at level 4, 36% at level 5 and 19% at level 6 could not accurately manipulate equations which involve negative signs and numbers. *(Questionnaire question 6)*

Simultaneous equations

It is noticeable that the performance relating to the solution of simultaneous equations is weak at level 7 and very weak at levels 5 and 6 in all years and weak at level 8 in years 1998 and 1997. It should be noted that there are two different formats in use over the four years which may cause problems should the candidates be unfamiliar with that which is used. However candidates should be equally competent in handling either format and at level 8 comparable results were achieved in the questions set in 1999 and 1996 despite the varying format and similarily comparable results were achieved in 1997 and 1998.

	Mean mark (out of 3) awarded at each stated level (calculated across all tiers)			
	Level 5	Level 6	Level 7	Level 8
1999 $y = 2x + 1$, $3y = 4x + 6$	0.13	0.29	1.00	2.33
1998 $x + 8y = 60$, $4x - 4y = 60$	0.30	0.87	1.35	1.68
1997 $3y = x + 25$, $x = y + 15$	0.00	0.27	0.83	1.85
1996 $a + 3b = 25$, $2a + b = 15$	0.43	0.93	1.63	2.45

Interestingly whilst the pair of equations set in 1996 both involved a unitary variable, a, in the first equation and b, in the second and the highest results over the four years were achieved at all levels. In years 1998 and 1999 the sets of equations used involved only a single unitary variable in one equation. However in 1997 the pair of equations set involved one with a single unitary variable, x and the other with a pair of unitary variables, x and y. Yet at levels 5, 6 and 7 the lowest results over the four years were attained and at level 8 a much weaker result than in 1996 and 1999.

(Questionnaire questions 17 & 18)

The report concluded with the following implications for teaching and learning which relate to the areas of number and algebra:

- *Pupils working at the lower levels need increased opportunities for pre-algebra work in arithmetic to help them establish the concept of a number equation.*
- *Pupils at all levels need more opportunities to develop a clearer understanding of decimals.*
- *The successful work with pupils helping them to understand the connection between algebraic expressions and the situations they describe should continue.*
- *Pupils at all levels need help with handling indices and understanding negative quantities in both number and algebra.*

(Questionnaire questions 20 & 21)

Generally

- *At the higher levels, pupils need more opportunity to structure and organise solutions to problems. They need to construct explicit chains of reasoning in response to demands for explanations and as an introduction to proof.*

  (Questionnaire question 23)

- *At the lower levels, pupils need more opportunities to handle questions where little initial mathematical orientation is provided. They need opportunities to develop pause and reflect strategies.*

  (Questionnaire question 22)

Key findings

- *Pupils achieving level 5 showed a good understanding of algebra as generalised arithmetic.*
- *Pupils achieving level 5 and above successfully used substitutes in algebraic expressions, performed inverse substitutions, constructed algebraic expressions to represent patterns, constructed patterns to fit algebraic expressions and interpreted algebraic expressions in context.*
- *At the higher levels, when the substitutions, manipulations and work with equations became more complex, errors appeared in pupils' answers, particularly when negative numbers were involved.*
- *Pupils are generally unable to handle indices well in both number and algebra. This applies across the ability range.*

  (Questionnaire question 19)

- *Pupils could perform inverse operations and procedures well, in a variety of contexts. Pupils achieving level 3, however, did not understand the nature of number equations. In this context of* *arithmetic pre-algebra, they were unable to use inverse operations to find missing numbers.*

To summarise the following deductions and findings from this analysis of the algebraic components of the QCA KS 3 Mathematics SATS Report have been reported on:

At level 5 the majority of pupils made arithmetic errors when solving a Linear Equation

- At level 5 37% were unaware of a standard approach for collecting together like terms when solving a Linear Equation
- Algebraic manipulation errors increased to 57% at level 6 when handling negative signs and negative numbers
- At level 6, 56% had difficulty writing an algebraic expression
- At level 6, 78% were not able to substitute accurately
- At level 7, 66% were unable to write the required expression
- At level 7, 59% were unable to solve the equation
- At levels 3 & 4, comparing algebraic expressions is more problematic than comparing numerical values
- At levels 4-7, solving an equation which involves two separate operations proved more difficult than those which require only one type of operation.
- At level 3, 82%, at level 4, 58%, at level 5, 45% and at level 6, 42% were unable to explain adequately the algebraic term within an algebraic representation of a tessellation sequence
- 78% at level 4, 54% at level 5 and 28% at level 6 were unable to accurately solve by substitution when given one equation from which the value of the variable can be deduced and then used within another
- 100% at level 3, 95% at level 5, 84% at level 6, 59% at level 7 and 31% at level 8 were unable to solve by substitution of an algebraic expression
- 95% of pupils at level 3, 74% at level 4, 36% at level 5 and 19% at level 6 could not accurately manipulate equations which involve negative signs and numbers

- 81% of level 3 pupils, 59% of level 4 pupils, 25% at level 5 and 15% at level 6 unable to accurately substitute a numerical value into a formula without any other manipulations.
- Weak performance in solving simultaneous equations of whatever format

(Questionnaire questions 17 & 18)

2b Mapping of Common errors identified in QCA(2002) Standards at Key Stage 3 Mathematics

Subtraction

To help improve performance, teachers should ensure that pupils understand that subtraction is not commutative.

The most common error was to select a subtraction, but with the numbers or products in the reverse order.

(KS3a) P14

Equations, formulae and identities

Use of symbols

To help pupils improve their performance, teachers should reinforce the meanings of coefficients and symbols in algebraic terms with pupils working at levels 4 and 5.

In a mental arithmetic question in test C targeted at level 5, less than half of the pupils working at this level were able to find the value of $3x + 6$ when $x = 5$. A quarter of pupils working at the target level gave the answer 41. This suggests that these pupils had the misconception that $3x$ is 35 when $x = 5$. The need to multiply 3 by 5 had not been recognised and the unknown had been treated as a missing digit. (KS3 b)
P19

About a third of pupils working at level 5 or above were able to evaluate the expression $3m - 10$ when $m = 2$. About a quarter of pupils working at this level gave the answer 22. This suggests the misconception that $3m = 32$ when $m = 2$ hence the digit 2 is literally replaced by the variable m. P20

To help pupils improve their performance, teachers should provide opportunities to practise transformation of algebraic expressions for pupils at higher levels.

A common error, seen in a third of the responses at level 6, was a failure to multiply the 2 by *a*. The misconception that *a* multiplied by *a* is equal to $2a$ was also found in some of the responses at this level.

(KS3 c) p20

In the same test, just under half of the pupils working at level 6, the target level for the question, were able to evaluate $10k^2$ correctly, given that $k = 3$. The most common incorrect response, given by about a third of the pupils working both at this level and at level 7 , was the answer 900. This suggests that these pupils had squared $10k$ rather than squaring k and then multiplying by 10.

(KS3 c) p21

When pupils were required to use equations to write an expression responses were almost equally divided between those who multiplied $2c - d$ by 7 and those who correctly subtracted $2c - d$ from $3a + 6b$. The most common errors seen in the responses of pupils working at this level who attempted this latter method, were either a failure to use brackets or to change signs correctly.

(KS3 a, e and f)

p23

In part (a) of *Rearrange*, targeted at level 7, about a fifth of the pupils working at the target level were awarded both the marks, and a further quarter were awarded just one of the marks. This one mark was most commonly awarded for the correct expansion of the brackets. Some of the pupils working at level 7 expanded the brackets incorrectly to $p = 2e + f$

(KS3 f)

P23

Equations

To help pupils improve their performance, teachers should present linear equations with unknowns on both sides and simultaneous equations for pupils working at levels 6 and 7 to practise solving using graphical and algebraic methods.

About a quarter of the pupils working at level 6, the target level for the question, were awarded both the marks for solving the equation in part (c) of *Solving*. About two-thirds of the pupils working at level 6 showed evidence of using algebra to solve this equation, and about half of these were successful in collecting together like terms, even if they were not able to reach the correct solution.

(KS3 d)

P24

Writing an algebraic expression

In *Rectangles*, targeted at level 8, the most common error seen in the responses of the pupils working at the target level was a failure to multiply out the pairs of brackets correctly. Although the term in y_2 was included, other errors were made, as shown in the example below. (KS3 e) P24

Appendix 3 GCSE Examiners report

Foundation level

many candidates showed little facility to deal accurately with subtraction of money or time, long multiplication or division, proportional division or multiplication by integer powers of 10

Addition errors were prevalent and even more so were the subtraction errors

The majority of candidates showed a good understanding of negative numbers

In many cases the candidates' inability to subtract accurately, or their incorrect use of base 10 in time calculations, still let them down.

less than 1% of candidates were able to give the expression "$n + 3$" for pattern number n.

As always long multiplication proved to be a problem, particularly for those who attempted the traditional long multiplication method. Most candidates demonstrated little or no understanding of the process of division with 3 rem 13 and 30 only being given as solutions

Algebra still remains a weak-point in the work of candidates entered at this tier (foundation) Those few candidates who successfully expanded the brackets in (i) often spoiled their answer by suggesting $15h + 10$ simplified to $25h$.

Most candidates were able to collect 1 mark for 12×50 or £4.80 12 or £1.20. Hardly any candidate was then able to use this information to form a fraction for the % profit.

About 90% of candidates realised the need to multiply 9.60 by 200 in (a) and showed this in the space for working. However, very few were able to carry out the calculation successfully. Part (b) was usually correctly answered though there were some answers of £250 obtained by calculating 25×10.

As usual the sight of algebra with students of this level does lead to many leaving this question blank. The more able candidates made a stab at manipulating the equations but careless or inaccurate use of signs proved to the downfall for most.

confusion between nth and ninth led to a variety of numerical answers, including 50 (the ninth term), 54 (9×6) and 56 ($9 \times 6 + 2$).

Some candidates did not collect the c terms and left the answer as $7b + 2c - 6c$, while others went to the other extreme and "simplified" $7b - 4c$ to $3bc$, which received no marks.

Only a minority had the algebraic skills to write the appropriate formula

most candidates were able to perform the algebraic simplification "$12xy$" without difficulty, but some gave their answer as "$12x \times y$" or "$3x4y$" (insufficiently simplified) and others wrote the incorrect answer "$7xy$".

The first part of the method involved dividing £3 by 5. For those who actually worked out $3 \div 5 = 0.6$ (rather than the incorrect $5 \div 3 = 1.67$) some had difficulty evaluating "0.6×8", losing the accuracy mark.

most candidates had the correct method "9.60×200", but many had difficulty in evaluating this. The calculation was often set out as long multiplication, but some candidates did not take account of the place value of the "9" in 9.6 and others gave the result of multiplying a number 0 as a non-zero value. Common incorrect answers were "19200", "19.20", "3000.00", "300.00" and "18120.00". Some tried $10 \times 200 - 0.40 \times 200$, but they found this no easier. Very few used the quick method of multiplying 9.60 by 100 then multiplying the result by 2.

there were some examples of poor arithmetic here, with £6 – £4.80 worked out as "£2.80", "£5.20" or "£2.20". Very few candidates were able to work out the percentage profit correctly

many who reached "$6q - 51$" were unable to write $q = \frac{51}{6}$ or divide 51 by 6 correctly without a calculator. Answers of "8.3" (from 8 remainder 3) were often seen.

many candidates wisely rounded the given numbers to one significant figure and obtained "$\frac{3 \times 700}{0.5}$" which they simplified to "$\frac{2100}{0.5}$". A few then worked this out correctly as "4200", but the vast majority were unable to divide by 0.5 without a calculator. They realised the use of "2" was involved, but they divided by 2 instead of multiplying by 2 and gave the answer as "1050".

Very few got the right answer.. some multiplied the given numbers correctly to obtain 16×10^7, but did not convert this properly to standard form. Others used the rules for multiplying powers of numbers incorrectly and obtained

16×10^{12}. The most successful attempts were by those who converted the given numbers in standard form to the ordinary numbers 8000 and 20000. However these numbers were often multiplied incorrectly – the last three zeros of each number were crossed out and the answer was written as "8 × 20 × 1000 = 160000". Some candidates were unable to convert numbers in standard form correctly to ordinary numbers. 8×10^3 was written as 80000, 8 × 3 × 10 (=240), 8 × 8 × 8 × 10 (=5120), 80^4, 8.000 or 0.800. In part (b) some who did convert the number correctly were unable to add them accurately; answers such as 10000, 100000, 100000000, 10.000 were seen.

Most candidates found this question difficult. Often incorrect answers such as "$x = \sqrt{5(y-4)}$ were written down with no intermediate steps in the method shown, so possible marks were lost. A significant number of candidates reached the stage "$x^2 = 5y - 4$", but then gave incorrect answers such as "$x = \frac{5y-4}{x}$" or "$x = \frac{5y-4}{2}$". Some arrived at a numerical value for x.

the evaluation of 27 + 2 without a calculator caused problems, with results 13.1 (from 13 remainder 1), 13, 14.5 or 14. Working out 13.5 × 3 was also done incorrectly at time

most of the errors were made in combining the *c* terms with involved a negative number, though it was not uncommon to see an attempt to combine all four terms together.

Poor algebra was apparent in many solutions.

Many candidates revealed their lack of calculator technique, through their failure to process the figures correctly. Many failed to perform the square root before dividing by 0.306, and premature rounding was in much evidence. In this type of question candidates would benefit from setting out their working in logical way, irrespective of their use of calculator.

it was common to see a correct expansion followed by incorrect simplification. However, a significant number of candidates could not expand the brackets correctly. Only a small number of candidates understood what factorising meant, and fewer again could correctly extract even a single factor.

many candidates obtained 16×10^7 as their final answer. They did not realise that there was another step to finding the answer in standard form. A common wrong answer was 1.6×10^{12}. It was disappointing to find candidates tackling this part of the question by converting both given values to 'ordinary numbers', multiplying them and then converting back to standard form rather than using rules of indices.

many believed that $\frac{3 \times 700}{0.5}$ was equal to 1050

changing the subject of the formula, was generally well answered The common errors from weaker candidates usually came from a wrong method to eliminate the fraction correctly

Common errors were $t^6 - t^8$ as a final answer and, possibly more surprising, $t^5 - t^6 = t^{-1}$, (q5 p5)

many candidates failed to give a valid expression for the difference. The most common incorrect approach started with $n(n+1) - (n+1)$. Other weaknesses highlighted by this question involved lack of brackets in algebraic expressions and the wrong expansion of $(n+1)^2$ as n^2+1.

Many candidates failed to factorise part (a) correctly. The most common errors were $(p-q)^2$, $pq(p-q)$ and even $p(p)-q(q)$. Poor use of brackets was also apparent in answers such as $p+q(p-q)$.

the word 'product' seems not to be understood by all candidates

The main error was the application of the square root sign to the whole fraction instead of just the numerator

mishandling of the negative sign in the second bracket to give $7x - 1$.

The main difficult was how to handle the 2 terms in the variable q. Part of the reason for this difficulty may be that candidates do not have much experience of solving equations of the form.

$$\frac{2x+3}{x+4} = 5$$

The main difficulty came with the simplification of

$$12 \times \frac{1}{3}(98 - 24x)x$$

Where most candidates could not handle the combination of the multiplication sign the fraction and the bracket.

there were more successful attempts in getting a correct, but unsimplified answer. with only the best candidates realising that they could produce a lowest common denominator of $(x+1)(x+2)$ and then spot that it was possible to factorise the numerator and cancel the common factor $(x+1)$.

At Higher level

Candidates generally found some of the topics difficult. These included the .. questions involving manipulative algebra (questions 3, 13(c), 15(b), 20 and 24), correlation (question 6), estimating the mean of a frequency distribution (question 18), standard form (question 21),

There were still candidates writing divisions the wrong way, for example "10 + 25" when they intended working out "25 + 10". This sometimes led to a loss in method marks.

Many candidates make errors when rounding answers to a given number of decimal placers or significant figures. This could occur whenever requested in a question, or when the candidate chooses to round an answer. The worst cases are where candidates round off interim calculations, resulting in premature approximation.

There was a slight improvement in algebraic manipulation this year. This was evidenced in question 20 where expansions were needed, but also in manipulation of trigonometrical formulae, and derivation of algebraic expressions.

This year there was a perceptible increase in the number of candidates at this level who were unable to perform a percentage calculation correctly. Most commonly this was demonstrated when the candidate divided the amount by the percentage, rather than performing a multiplication

there were more arithmetically errors seen than in recent years. Accepting that slips are likely to occur under examination conditions there does seem to be a more serious problem in relation to work with fractions

There were a number of questions in the paper which required knowledge and use of indices. The responses of candidates to these questions seemed to show a greater weakness than in some recent years.

Appendix 4 Mapping of QCA findings with the GCSE examiners and Matz

QCA Ref	QCA finding	Matz Ref	GCSE Ref
	CONFIRMED STATEMENTS		
QCA1	Equations which involve only one variable (3k=18, 5k=?) are not problematic for the majority of pupils at level 6 and higher		F Q28a, F Q15a(i)
QCA2	At higher levels pupils need more opportunity to structure and organise solutions to problems.		I Q24
QCA3	At level 5 most pupils do know a standard approach to collecting together like terms	M27	F Q12, I Q13, I Q20, I Q3
QCA4	At level 3: most are unable to complete equations of the form … / 24 = 16	M1 M10	F Q28
QCA5	At higher levels pupils need more opportunities to handle questions where little initial mathematical orientation is given		F Q19, I Q4, I Q10
QCA6	At level 6 in re-writing an expression most are successful if positive terms only such as 3b+1 = b+ . and 5a-4 = 2a+2 + ….	M8 M9	
QCA7	At level 5 when negative numbers or signs are involved many pupils make arithmetic errors such as 4-2y=10-6y ⇒ 4y=14 ⇒ y=3.5		F Q28 b) c)
QCA8	Pupils at all levels need help with handling indices	M5, M7, M6, M11, M15, M16	F Q15a, I Q20, H Q5
QCA9	At level 3: most are unable to complete equations of the form 962 - … = 476		
QCA10	At higher levels, when the substitutions, manipulations and work with equations is more complex errors appear in pupils' answers	M2, M3	
QCA11	The solution of simultaneous equations is very weak at levels 5 & 6		

QCA Ref	QCA finding	Matz Ref	GCSE Ref
QCA12	Pupils at all levels need help understanding negative quantities in algebra	M4, M23	F Q15b
QCA13	At level 5 many pupils when solving Linear Equations make arithmetic errors of the form 4y = 6 ? y= 1/3	M26	
QCA14	At level 5 most pupils are unable to find y from 4-2y=10-6y		F 128©
QCA15	At lower levels the comparison of two algebraic expressions (3x+2y=7, 6x+4y=?) is significantly more problematic than that between two equations which involve two numerical values only (4k+2a=82, ?=41		
QCA16	A value n being deduced from an equation of the form 26-2n=8 is problematic to solve for pupils at level 6 and lower		F Q28(b)(c)
QCA17	The solution of simultaneous equations is very weak at levels 5 & 6		
QCA18	Pupils at all levels need help understanding negative quantities in number		F Q15b
QCA19	Equations which involve only one variable (3k=18, 5k=?) are problematic for the majority of pupils at level 4 & 5		F Q28 b),c)
QCA20	A value n being deduced from an equation of the form 26-2n=8 is extremely difficult to solve for pupils at level 6 and lower		F Q28 b),c)
QCA21	At level 6 most were not able to substitute accurately a numerical value into an expression		
QCA22	At levels 6 & 7, most pupils are unable to write the required expression from a sequence eg. 1 hut needs 6 matches, 2 huts need 11 matches		F Q11, I Q15, I Q6, H Q8

Appendix 5 Survey re. Common errors and weaknesses in algebra & numeracy at Key Stage 3

I am conducting research to evaluate the use of ICT to support the learning of algebraic and numerical manipulations. For the purpose of this research I am investigating the nature of common errors and weaknesses associated with this area of mathematics. The following findings and recommendations were identified in several recent Key Stage 3 Mathematics SATS Reports. It would be appreciated if you would indicate your views and opinions within your experience of teaching and learning mathematics in your institution. The levels indicated refer to those attained by pupils at the end of Key Stage 3.

	Strongly Agree	Agree	Unsure	Disagree	Strongly disagree
1. At level 3: **most** are unable to complete equations of the form 962 - ... = 476					
2. At level 3: **most** are unable to complete equations of the form ... / 24 = 16					
3. At level 5 many pupils when solving Linear Equations make arithmetic errors of the form $4y = 6 \Rightarrow y = 1/3$					
4. At level 5 **most** pupils are unable to find y from $4-2y=10-6y$					
5. At level 5 **most** pupils do know a standard approach to collecting together like terms					
6. At level 5 when negative numbers or signs are involved **many** pupils make arithmetic errors such as $4-2y=10-6y \Rightarrow 4y=14 \Rightarrow y=3.5$					
7. At level 6 in re-writing an expression **most** are **successful** if positive terms only such as 3b+1 = b+ .... and 5a-4 = 2a+2 +......					
8. At level 6 in re-writing an expression **most** are **unsuccessful** if negative numbers are required 5d + 6=2d–3 + _ and 4c=2c+1 + ...					
9. At higher levels, when the substitutions, manipulations and work with equations is more complex, **errors** appear in pupils' answers					
10. At levels 6 & 7, **most** pupils are **unable** to write the required expression from a sequence eg. 1 hut needs 6 matches, 2 huts need 11 matches					
11. At level 6 **most** were **not able** to substitute accurately a numerical value into an expression					
12. At lower levels the comparison of two algebraic expressions (3x+2y=7,6x+4y=?) is **significantly more problematic** than that between two equations which involve two numerical values only (4k+2a=82,?=41)					
13. Equations which involve only one variable (3k=18, 5k=?) are **problematic** for the **majority** of pupils at level 4 & 5					
14. Equations which involve only one variable (3k=18, 5k=?) are **not problematic** for the **majority** of pupils at level 6 and higher					

	Strongly Agree	Agree	Unsure	Disagree	Strongly disagree
15. A value n being deduced from an equation of the form 26-2n=8 is **extremely difficult** to solve for pupils at level 6 and lower					
16. A value n being deduced from an equation of the form 26-2n=8 is **problematic** to solve for pupils at level 6 and lower					
17. The solution of simultaneous equations is **weak** at levels 7 & 8					
18. The solution of simultaneous equations is **very weak** at levels 5 & 6					
19. Pupils at all levels **need help** with **handling** indices					
20. Pupils at all levels **need help understanding** negative quantities in algebra					
21. Pupils at all levels **need help understanding** negative quantities in number					
22. At higher levels pupils need **more opportunities** to handle questions where little initial mathematical orientation is given					
23. At higher levels pupils need **more opportunity** to structure and organise solutions to problems.					
Any other comments or related observations					

Thank you for your time and interest in completing the questionnaire. Please contact me by e-mail should you like any further details about using the software on which the research is focussed.

Appendix 6 . Participants Documentation

6a Guidance for participation in a research investigation

I am conducting some research concerned investigating the effectiveness of the use of an ILS to support the learning of numeracy.

Volunteers are being sought. Firstly to participate in the preliminary study investigating the nature of common errors and misconceptions associated with solving numerical and algebraic problems as well as identifying methods of solution. This would require participants to undertake a computerised numeracy pre-test and post-test and to attend several workshop sessions. In these sessions participants will utilise a mathematical tutoring system to solve algebraic and numeric problems. During each of these sessions each user will produce for analysis a log file recording all actions undertaken. Subsequently each participant will also be required to complete a questionnaire evaluating the usability of the software and the effectiveness on learning. Individual progress will be recorded and each participant will receive feedback regarding their own performance. However the data from this investigation will be grouped for analysis and used only for the following purposes:

To measure change and progression

To identify common methods, errors and misconceptions

To analyse the structure, usability and

To assess the effectiveness of the software

The findings from this preliminary study will result in changes and improvements of the ILS. It is the intention that the effects of these changes will then be assessed and evaluated by means of pre and post tests and use of the adapted system in specific sessions by another group of volunteer students.

All participants will have the right to withdraw from the investigation at any time without prejudice to access of services which are already being provided or may subsequently be provided to the participant.

6b **Participants Consent form EC3 (JMU Pilot Trial)**

LIVERPOOL JOHN MOORES UNIVERSITY

Title of project/procedure: The evaluation of an intelligent learning system as a tool for the learning of algebraically applied numerical manipulations

I, ...agree to take part in
(Subject's full name)*

this project, which is concerned with the investigation of the use of ICT to support the learning of numeracy and basic algebra. Volunteers are being sought who would be required to undertake a computerised diagnostic numeracy test and solve several sets of questions. These activities will utilise a mathematical tutoring system to assist in the solution of algebraic and numerical problems. During use of this computer system a log file recording all actions undertaken will be automatically produced for each user and each set of questions. During analysis individual progress will be recorded and mapped. Subsequently each participant will also be required to complete a questionnaire evaluating the use of the software for learning.

The data from this investigation will be grouped for analysis and used only for the following purposes: To identify common methods, errors and misconceptions

To analyse the structure and usability of the software

To assess the effectiveness for learning of the software

These details have been fully explained to me as well as described above.

Signed .... Date ....
(Subject)

I, ...certify that the details of this
(Investigator's full name)*

project have been fully explained and described in writing to the subject named above and have been understood by him/her.

Signed .... Date ....
(Investigator)

I, ...certify that the details of
(Witness' full name)
this project have been fully explained and described in writing (see above) to the subject named above and have been understood by him/her.

Signed .... **Date** ....
(Witness)

NB The witness must be an independent third party. Please print in block capitals

6c Participants Consent form EC3 (Pendleton Final Trial)

LIVERPOOL JOHN MOORES UNIVERSITY

FORM OF CONSENT TO TAKE PART AS A SUBJECT IN A MAJOR PROCEDURE OR RESEARCH PROJECT

Title of project/procedure: The evaluation of an intelligent learning system as a tool for the learning of algebraically applied numerical manipulations

I, ... agree to take part in
(Subject's full name)*

this project, which is concerned with the investigation of the use of ICT to support the learning of numeracy and basic algebra. Volunteers are being sought who would be required to undertake a diagnostic numeracy pre-test and post-test and to attend several computer based workshop sessions. The sessions will involve the use of a mathematical tutoring system to assist in the solution of algebraic and numerical problems. During each of these sessions a log may be produced for each user recording all actions undertaken. Subsequently each participant will also be required to complete a questionnaire evaluating the use of the software for learning.

The data from this investigation will be grouped for analysis and used only for the following purposes:

To identify common methods, errors and misconceptions

To analyse the structure and usability of the software

To assess the effectiveness for learning of the software

These details have been fully explained to me as well as described above.

Signed .... Date ....
(Subject)

I, certify that the details of this
(Investigator's full name)*
project have been fully explained and described in writing to the subject named above and have been understood by him/her.

Signed .... Date ....
(Investigator)

I, ... certify that the details of this
(Witness' full name)

project have been fully explained and described in writing (see above) to the subject named above and have been understood by him/her.

Signed .... **Date** ....
(Witness)

NB The witness must be an independent third party. * Please print in block capitals

Appendix 7 Diagnosis Content and Hierarchy

7a Overview of levels and themes

				DISGARDED TOPICS are shaded		
Themes	**ALL TOPICS** **Bold = selected**	**Selected**	**Level 1**	**Level 2**	**Level 3**	**Level 4**
Numbers	**101, 102, 103, 104, 105, 107, 108, 109, 112, 201, 202, 205, 206, 215, 216,** 301, 302	**101, 102, 103, 104, 105, 107, 108, 109, 112, 201, 202, 205, 206, 215, 216**			301, 302	
Powers	**106,** 203, 204, 303, 304, 305, 306, 401, 402, 403	**106**		203, 204	303, 304, 305, 306	401, 402, 403
Basic algebra	110, 111, 113, 207, 208, 209, 210, 211, **212, 213, 214**	**212, 213, 214**		207, 208, 209, 210, 211		
Algebra methods	307, 308, 311, 312, 313, 404, 407, 410				307, 308, 311, 312, 313	404, 407, 410
Equations	309, 310, 314, 315, 316, 406, 411, 412				309, 310, 314, 315, 316	406, 411, 412
Algebra & calculus	317, 341, 342, 343, 405, 408, 409, 441, 442				317, 341, 342, 343	405, 408, 409, 441, 442
Statistics	**231, 232, 233,** 234, 334, 335	**231, 232, 233**		234	334, 335	
Miscellaneous	221, 222, **223,** 321, 322, 323, 324, 421, 422	**223**		221, 222	321, 322, 323, 324	421, 422
Graphs	**151, 251,** 351, 451, 452	**151, 251**			351	451, 452
Area & volume	161, 261, 262, 263, 361, 362, 363, 461, 462		161	261, 262, 263	361, 362, 363	461, 462

Level 1 is directly relevant except 161: Area of a triangle which is not applicable as a skill every teacher would need to have. Level 2 is aligned to the level required for the QTS numeracy test however the subject content is broader than the requirements of numeracy, basic algebra and basic statistics. From Level 3 onwards the complexity of skills may be too difficult for QTS NUMERACY.

7b *Diagnosis Hierarchy*

Numbers 101,102,103,104,105,107,108,109,112,201,202,205,206,215,216,301,302

101,Multiplication of negative numbers,1}

102,Multiply negative and positive,1 → 101,Multiplication of negative numbers,1}

103,Negative Numbers,1, → 102,Multiply negative & positive,1
→ 201,Use of < and > signs, → 301,Solution of simple inequalities,3

104,Size of decimals,1 → 201,Use of < and > signs, → 301,Solution of simple inequalities,3

105,Decimal places,1 → 202,Significant figures,2 → 302,Scientific Notation,3, → 401??

*106,Definition of positive powers,1 (*EXTENSION LEVEL UNNECESSARY *) POWERS GROUP*

107,Ratios,1, → 205,Inverse ratios,2
→ 206,Cancelling numerical fractions,2

108,Factors of an integer,1, → 107,Ratios,1,
→ 206,Cancelling numerical fractions,2 → 216,Add/subtract numerical fractions,2,

OR 207 AND 213 (Unnecessary)

109,Simple fractions,1, → 216,Add/subtract numerical fractions,2, (onto 308)

110, Collect terms (simple), 1 → 211, collecting terms, 2

111,Solving a simple equation,1, → 212,Solving linear equations,2, → (onto higher levels) BASIC ALGEBRA

112,Simple calculation,1, → 214,Evaluating formula,2,
→ 215,Precedence Rules,2

113,Evaluating a simple expression,1, → 214, Evaluating formula,2, BASIC ALGEBRA

151,Coordinates,1 → *(251) Precedence Rules,2 ... (351)* *GRAPHS*
201,Use of < and > signs,2 → 301,Solution of simple inequalities,3}

202,Significant figures,2 → 302,Scientific Notation,3

205,Inverse ratios,2

206,Cancelling numerical fractions,2

212,Solving linear equations,2 *BASIC ALGEBRA*

213,Transposition of formula,2 *BASIC ALGEBRA*

214,Evaluating formula,2 *BASIC ALGEBRA*

215,Precedence Rules,2

216,Add/subtract numerical fractions,2,

223,Percentages,2, *MISCELLANEOUS*

231,Range of a set of numbers,2 *STATISTICS*

232,Mean of a set of numbers,2 *STATISTICS*

233,Simple probability (coins),2 → 334 *STATISTICS*

7c Diagnosys content for undergraduates preliminary trial

numbers,101,102,103,104,105,107,108,109,112,201,202,205,206,215,216,301,302
powers,106,203,204,304,305,306
basic algebra,110,111,113,207,208,209,210,211,212,213,214
algebra methods,307,308,312,404,410
equations,309,310,314,315,316
algebra+calculus,405,409
statistics,231,232,233,334
miscellaneous,223,324

Removed from Diagnosys Test

skill{207,Multiply algebraic fractions,2,307}
skill{208,Factors of algebraic products,2,209}
skill{222,Definition of sin and cos,2,322,422}
skill{301,Solution of simple inequalities,3}
skill{302,Scientific Notation,3,401}
skill{303,Simplification of fractions with powers,3,401}
skill{304,Rules for negative powers,3,303,305,402,403}
skill{305,Rules for fractional powers,3}
skill{306,Definition of fractional powers,3,305}
skill{307,Division of algebraic fractions,3}
skill{308,Add/subtract algebraic fractions,3,404,405}
skill{310,Relation between roots and factors,3}
skill{313,Difference of squares,3}
skill{317,Complex numbers,3,408}
skill{322,Sin and Cos formula,3}
skill{323,Definition of radians,3}
skill{324,Percentages (advanced),3}
skill{334,Conditional probability,3}
skill{335,Venn Diagrams (conditional prob.),3}
skill{341,Differentiation of powers,3,342,441,442}
skill{342,Finding Max/Min of a quadratic,3}
skill{343,Geometric Progression,3}
skill{401,Simplify with scientific notation,4}
skill{402,Logs,4}
skill{403,Arbitrary factors,4}
skill{404,L.c.d. of an algebraic fraction,4}
skill{406,Solve quad. by comp. the square,4}
skill{407,Quadratics - completing the square,4,406}
skill{408,Multiplication of complex numbers,4}
skill{409,Divide by zero (possible solution),4}
skill{412,Difficult linear equation,4}
skill{421,Deduce radius of circle,4}
skill{422,Sin and Cos as functions,4}
skill{441,Product rule,4}
skill{442,Integration of powers,4}
skill{451,Recognise formula of quad. graph,4}

7d **Mapping of questions, skills to weaknesses 1-6**

Weakness 1 Division

Diagnosys skills

109, Simple fractions
112, Simple calculation
206, Cancelling numeric fractions
207, Multiplying algebraic fractions
213, Transposition of formula
216, Add/subtract numerical fractions

	Diagnosys question and skill
109	Add: {a}1/2+3/4 {b}{1/6+1/3 {c}[1/4]+1/3 {d}[1/2]+[1/10]} {e}[1/2]+[1/6]
112	Calculate: $ {a}{17-3/4} {b}{14-2/3} {c}{13-4/2} {d}{21-5/3} {e}{20-4/3}
206	Cancel {all} possible common factors to leave the fraction in its {simplest} form. {a} 72/90 {b} 36/60 {c} 48/18 {d} 30/24 {e} 60/45
207	a}What is {$frac[6/y]$sym'$frac[y$^[2];12]} ? {b}What is {$frac[4/y$^[2]]sym'frac[3y;2]} ? {c}What is {$frac[3/y$^[3]]sym'frac[5y$^[2];6]} ? {d}What is $frac[2y/9]x [3/4y^[2]] ? {e}What is [4/y]x [y^[3]/8]} ? Give your result in simplified form.
213	{a}If {c=[ab/2]} then {b} = ? {b}If {v=[2u/3]} then {u} = ? {c}If {d=[5e/3]} then {e} = ? {d}If {s=[-3t/5]} then {t} = ? {e}If {x=[yz/4]} then {y} = ?
216	Cancel {all} possible common factors {a}[72/90]} {b}[36/60]} {c}[48/18]} {d}[30/24]} {e}[60/45]}

Weakness 2 Brackets

Diagnosys skills
110, Collect terms (simple)
208, Factors of algebraic products
209, Simple Factorisation
211, Collecting terms
213, Transposition of formula
214, Evaluating formula
215, Precedence Rules

Test	Question	TREE FROG
5	4	(z+1)/(z+4)=5/6
5	3	(y + 2)/(y - 3) = 2
3	9	17-21(s+7)
3	10	13-19(s-3)
3	7	3+21(s+4)
5	2	(ax + b) (cx + d)
3	8	3+23(s-4)
3	4	-(3x-w)
3	3	2(x+3)

More complex questions than those in Diagnosys and remainder of Treefrog

5	1	expand A(BC)
5	7	(x-5)(x-7)=3

	Diagnosys question and skill
110	Collect terms in the following expression {a}{2x-3y+1+y+4x-5} {b}{y-3x-5y+3+x-2} {c}{3u-2-u+4v-3-2v} {d}{1-3p-4-q+4p+3q} {e}{-g-3h+7-4g+5+2h}
209	Factorise the following, taking out the highest factor possible: {a}{2z-6z^[2]} {b}{10t^[2]-5t} {c}{12y-8y^[2]} {d}{6p^[2]-9p} {e}{14w-6w^[2]} Enter your answer in factorised form (with one pair of brackets)
210	Expand the bracket in: {a}{2x(x-3x$^[2])} {b}{3y(2y-1)} {c}{4z(z-2z$^[2])} {d}{2u(u$^[2]-3)} {e}{3v(v$^[3]-2v)}
215	To calculate {a} [12/6*3] {b} [15+27]/3] you press a sequence of keys on your calculator.Which one of the following would give the {WRONG} answer?

Weakness 3 Exponentials

Test	Question	TREEFROG
6	5	(A+B)^-2
6	4	$(A+B)^2$

	Diagnosys question and skill
208	Enter the other factor in the equation: {a}{12x^[3]y=3x^[2]y(?)} {b}{15pq^[2]=5pq(?)} {c}{18u^[2]v$^[2]=6uv^[2](?)} {d}{10g^[3]h=5g^[2](?)} {e}{14mn^[3]=2n^[2](?)}

Weakness 4 Substituting values

Diagnosys Skills

113,Evaluating a simple expression

214, Evaluating formula

Test & Question		Treefrog question
4	2	4x when x=6
4	1	2x when x=9
4	3	2x+3 when x=5
3	1	Find 5x when 4x=20
4	5	-3z-8 when z=-3
4	4	-3y+8 when y=2
3	5	If 3x+2y=7 find a value for 6x+4y
4	6	xy+1 when x=-4 and y=6

	Diagnosys question and skill
214	{a}If Q=[p^2+2r]/[t-1] and p=4,r=-2, t=5, {b}If {Q=[3a+b]/2c^2 and {a=5, b=-3, c=2, {c}If {Q=[2f-g^3]/4h] and f=7, g=2, h=1.5, {d}If Q=[3r^[2]-2s]/[t+4] and r=3, s=6, t=-1, {e}If Q=[x^[2]-3y^2]/[1-3z] and x=4, y=-2, z=1, then what value has Q?
113	Evaluate {a}{2+3x} if x=3. {b}{5+2x} if x=4. {c}{4+5y} if y=5. {d}{2+4z} if z=3. {e}{6+2p} if p=4.

W5 Negative signs and values

Diagnosys skills

101, Multiplication of negative numbers
102, Multiply negative and positive
103, Negative Numbers
213, Transposition of formula

Wide variety and range of Treefrog questions with half of the questions including negative values and signs.

```
Diagnosys question and skill
Multiplication of negative numbers,5}
    {a}Calculate: {(-3)*(-5)}
    {b}Calculate: {(-4*(-6)}
101 {c}Calculate: {(-6*(-3)}
    {d}Calculate: {(-2*(-7)}
    {e}Calculate: {(-4*(-5)}
Multiplying a negative and a positive number,5}
    {a}Calculate: {4*(-1.5)}
    {b}Calculate: {5*(-0.5)}
102 {c}Calculate: {6*(-0.5)}
    {d}Calculate: {7*(-3)}
    {e}Calculate: {6*(-4)}
Negative numbers
    {a}Calculate:   (-3) + (-4)
    {b}Calculate:   (-5) + (-6)
103 {c}Calculate:   (-6) + (-8)
    {d}Calculate:   (-2.5) + (-3.5)
    {e}Calculate:   (-4) + (-7)
    {a}If  {c=[ab/2]}  then  {b} = ?
    {b}If  {v=[2u/3]}  then  {u} = ?
213 {c}If  {d=[5e/3]}  then  {e} = ?
    {d}If  {s=[-3t/5]}  then  {t} = ?
    {e}If  {x=[yz/4]}  then  {y} = ?
```

W6 Linear equations

Diagnosys skills

111,
212, Solving linear equations
Could include brackets/division/negative signs and values

```
Solve the equation, and give the value of {x}:
    {a}{3x+1=13}
    {b}{4x-3=9}
111 {c}{5x+2=17}
    {d}{3-2x=9}
    {e}{2-5x=-8}
Solve the equation:
    {a}{7-3c=-5c-4}
    {b}{2d+3=-4d+18}
212 {c}{1-3t=-6t-2}
    {d}{4c+4=-2c-5}
    {e}{3-2n=6n+11}
```

Appendix 8 Treefrog Test content

Test1 Content Requirements & Questions

Num	Findings	Questions	Reason
1		90 – x = 70	Worked example
2		9-x=7	Repetition
3		92-x=72	Includes tens
4		58-x=17	Repetition
5		34-x=18	Includes carry
6	QCA 9	962 – x = 476	QCA question
7	QCA 12	9- x = 17	Double –ve
8	QCA 12	7 -x = 9	Reverse of question 1
9	QCA 12	-9 – z = -7	-ve signs
10	QCA 12	476 – z = 962 *	Reverse of QCA question

* Concept: how do we subtract a larger number from another? Subtract smaller from the larger and then negate the answer

Test 2 Content Requirements & Questions

Num	Findings	Questions	Reason
1	QCA 4	X / 8 =12	Worked example
2		X / 3 = 15	Repetition
3		X / 5 = 25	Repetition of above
4		X /15 =22	Repetition of above
5		X /24 = 16	QCA question
6	QCA 12	- X / 4 = 20	-ve sign
7		X / 6 = -18	-ve sign
8		- X / 2 = 34	Repetition of above
9		X / 3 = -15	Repetition of above
10		2X / 7 = 6	Coefficient included
11		5X / 3 = 15	Repetition of above
12		11X / 15 = 22	Factor + similar to Q4
13		8x / 7 = 56	Repetition of above
14	M1	Solve X/X	Division by itself
15		Solve a*1/a	Repetition of above
16	M5, QCA12	Evaluate 2(-3)	Finding
17	M10	Evaluate 0*a	Finding

Test 3 Rewriting expressions and word problems

Num	Findings	Questions	Reason
1	QCA 1 & 19	4x = 20 find 5x	Worked example
2		(3k=18, 5k=?)	
3	QCA 16 & 20	26-2n=8	
4	M22	2(X+3)=2X+3	
5	M23	Rewrite without brackets -(3X-W)	Negative signs
6	QCA 15	If 3x+2y=7 then complete 6x+4y=?	
7	QCA 13	If 4y= 6 find y	
8	QCA 14	Find y in 4-2y=10-6y	
9	QCA 6	Complete 3b+1 = b+	
10		Complete 5a-4 = 2a+2 + ….	
11		Complete 4c+8 = 5c + 10 + …	Include negative signs
12	QCA 3, QCA 7	Rewrite 2y=10-6y	Arithmetic errors such as 4y=14 ⇒ y=3.5
13	M8	Simplify 3+21(s+4)	
14		Simplify 3+23(s-4) to 26(s-4)	
15		Simplify 17-21(s+7)	
16		Simplify 13-19(s-3)	
17	M9	Simplify 3XY+ 4XZ	= 7XYZ error

Test 4 Substituting into algebraic expressions including word problems

Num	Findings	Questions	Reason
1		Evaluate 2x when x=9	Worked example
2	M2	Evaluate 4X when X=6	
3	M22	Evaluate 2X+3 when x=5	
4		Evaluate –3Y+8 when Y=2	
5		Evaluate –3Z-8 when Z=-3	
6	M3	Evaluate XY when X=3 and Y=5	
7		Evaluate XY when X=-4 and Y=6	
8		Evaluate XY when X=-2 and Y=-7	
9	M4	Evaluate 2(-3)	Understanding the multiplication role of brackets
10		2 QTSIN	
		Solve 2 questions	
		Evaluate questions	

Test 5 BODMAS & algebraic division and simplification of expressions

Num	Findings	Questions	Reason
1			Worked example
2	M12	Express A(BC) without brackets	A(BC) = AB*AC
3	M24	Express (AX+B)(CX+D) without brackets	$(AX+B)(CX+D)=ACX^2+BD$
4	M30	Factorise $X^2+5/6X + 1/6$	X(X+5/6) + 1/6 error
5	M13	a/(b+c)	= a/b +a/c
6	M14	(a+b)/(c+d)	= a/c +b/d
7	M17	(AX+BY)/(X+Y)	= A+B
8	M18	X/(2X+Y)	=1/(2+Y)
9	M19	(X+3Z)/(2X+Y)	=3Z/(2+Y)
10	M20	(X-3)/2X	=-3/2
11	M21	$(X^2+2XY+Y^2)/(X^2-Y^2)$	=2xy error
12	M25	Solve for x (X+1)/(X+4)=5/6	X=4,2 error
13	M26	Solve for x 2X+5=11,	X+5 = 11/2 error
14	M27	Solve for x 3X+5=Y+3	X+5=Y
15	M28	Solve for R 1/R = 1/R1+1/R2+1/R3,	R=R1+R2+R3
16	M29	Solve for x $1/X+1/X^2=3/X^2+6X^2$	
17	M31	Solve for x (X-5)(X-7)=3,	X-5=3 OR X-7=3, X=8 OR X=10 error
18	M32	Solve for x 5/(2-X)+5/(2+X)=4	5(2+X)+5(2-X) = 4 error

Test 6 Indices

Num	Findings	Questions	Reason
1		Evaluate 2^3	Worked example
2	M5	Evaluate $(-1)^3$ as –3	
3	M6	$3r^2$ as $3+r^2$	
4	M7	Rewrite $(3r)^2$	
5		Evaluate $(2)^{-2}$	Negative indices
6	M11	$(A+B)^2$	$=A^2+B^2$ error
7		$(A+B)^{-2}$	Negative indices
8	M15	Rewrite 2^{a+b}	$= 2^a + 2^b$
9		Rewrite 2^{a-b}	
10	M16	Rewrite $2^{ab} = 2^a2^b$	

Appendix 9 Treefrog details

9a Question types

Background:

TREEFROG consists of 11 types of question. Specific question types are applicable for use when solving specific types of problems this includes solving numeric expressions and solving, simplifying or factorising algebraic equations. Some of these problems can be expressed solely in words for the user to then convert into a suitable expression for solving. Some of these questions do not require the tutor to calculate the answer but to select the correct question type and input the relevant start expression. The question types which are related to equations do require the user to present the solution in the same equation format as expected in properly constructed hand written solutions. Where the user is solving a numeric of algebraic problem the creation of equivalent expressions is required.

Methodology

I will consider the following:

- start - numeric value or expression /algebraic equation
- value - numeric value or expression /algebraic equation
- format - equation or a solution of an expression
- types of equation involved: linear, quadratic
- types of number: real or complex
- use of brackets
- number of variables permitted

From analysing the TREEFROG question types and using this with sample QTS numeracy test questions I intend to deduce which equation types are relevant to the use of TREEFROG for supporting the learning relevant to the ITE student.

Other factors which are necessary to the learning relevant to preparing for the numeracy test is the use of:

- tables of data
- money formatted data i.e. £
- % symbol included
- graphs
- comparing sets of data

Where sets of data are required to be compared for instance to deduce which set of pupil results have the largest range TREEFROG is not an appropriate tool as the input mechanism within the interface does not accommodate the development of more than one set of results and hence cannot tutor the student through the decision making process of comparing one result with another.

Type	Purpose	Teacher Format	Examples	Factors to consider
Word	From word question a final value Answer does not have to be numeric Converting narrative into a question	Scenario		Can solve for fractions
		Start –word question	Find the total cost	
		Value – numeric or algebraic	£25 or 23x	
		Hint		
Wordsimp	Convert a result to it's lowest form	Scenario		Even if the answer/value expects ( ) the program will continue to solve the expression until the lowest form i.e. without ( )
		Start point – word question or expression or numeric value	Find the result or 2(2n+1) or 22	
		Value – numeric value or algebraic expression	22 or 4n+3 or 2(2n+1)	
		Hint		
Wdeval <u>Numeric</u>	Like Eval, NumEq and Simp, but the question is posed	Scenario		If used for a question with an algebraic answer then it does not recognise the expected answer eg 25x
		Start point – words / numeric expression	25% of 200 or 25*3-53	
		Value – numeric	25	
		Hint		
Transpose **equation format**	Re-arrange a general equation to give an expression for one of its variables may not be in the simplest form	Start: algebraic	d = st	More than one variable expected. If used for an equation of the form: (38+9+12+x)/4=16 then will accept x=64-59 as an answer
		Var: to be the subject	t	
		Value: not required		
		Hint		
Wdnumeq numeric ***equation format***	Solving linear equations from a "word" scenario	Start: words giving vars to use or a linear equation	let s be …. or s+2=7	Checks for the variable in the equation when setting question up, if not there will not allow the user to continue If two variables in the
		Equation: linear expression	s+2=7	
		Var: to solve for	s	

Type	Purpose	Teacher Format	Examples	Factors to consider
		Hint		equation then it is not solvable
Factor	Factorise a polynomial expression to a given form Collect like terms	Question or scenario	Factorise the following …..	Using this for a linear expression is inappropriate as this does not factorise but simplifies
		Start– algebraic expression can be linear or in higher terms of x	x2+3x+2 or x + 3+ 1	
		Value – factorised expression	(x+2)(x+1)	
		Hint		
Numeq *equation format*	Solve a linear equation to give a result of the form: variable = value where value is an integer, real or fraction in its lowest terms.	Question: scenario/guidance	Solve the following for y	Question and hint could give the same information
		Start: algebraic equation or numeric expression	3y+2=8	
		Variable: to solve for	y (can be blank)	
		Hint		
Eval Numeric	Evaluate a real expression, to an integer, real or fraction in its lowest terms.	Question: Scenario or guidance	Evaluate the following expression ….	Numeric question which requires a final numeric solution no variables involved If algebraic input i.e. a+a then does not accept the expected valid answer of 2a
		Start – numeric expression	£4.50 - £2.00	
		No value required		
		Hint		
Quadeq **expr = 0 ….. x= . or x= ..**	Solve a quadratic equation to the form variable=value1 or variable=value2, where value1 and value2 are integers, reals or fraction in their lowest terms.	Question Scenario	Solve this quadratic …. for x	Can input a linear equation for the start but this is not solvable in the question type
		Start: quadratic equation	$x^2 + 3x+2=0$	
		Variable: to solve for	x	
		No value for the answer required		
Compeval	Evaluate a complex expression to the form	Question: instructions / requirements	Solve this .. in the form x+iy	Numerical solution so can deal with real numbers a

Type	Purpose	Teacher Format	Examples	Factors to consider
Numeric	x+iy, where x and y are real	start: complex or real expression	(4 + 3i) + (3+4i) or 4+3	subset of complex numbers
		No value for the answer required		
		Hint		
Simp	Expand a polynomial expression and collect common terms together to simplify the expression	Question	Multiply out the brackets .. or gather like terms together	Question and hint could be similar or even fulfil the same purpose i.e. give direction as to what is required
		start: numeric or algebraic	2(2x+1) or (3(4+7) – 2)	
		No value required		
		Hint		

Appendix 10 Preliminary Trial Treefrog log evidence

Test1 Responses

	Question	J	M	L	S	P	F	G	Mc	Ir
1	90-z=70 Add z to both sides!	Z=20 done	90- 70=20 Syntax, omitted z 90=70+z z=20 done	20 Z=20 done	20 z=20 done	Z=20 done	z=90-70 z=90-70 z=20 done	20 20 20 20 20 20 Z=20 done	Z=20 Done	z=70- 90 z=90- 70 z=20 done
2	9-z=7 Add z to both sides!	Z=2 done	9=7+z z=2 done	Z=2 done	Z=2 done	Z=2 done	Z=2 done	Z=2 done	Z=2 Done	z=9-7 z=2 done
3	92-z=72 Add z to both sides!	Z=20 done	92=72+z z=20 done	Z=20 done	Z=20 done	Z=20 done	z=22 z=22 -z=72-92 -z=-20 z=20 z-20 z-20 z-20 z-2 z=20 done	Z=20 done	Z=20 Done	z=92- 72 z=20 done
4	58-z=17 Add Z to both sides	Z=41 done	58=17+z z=41 done	Z=41 done	Z=39 z=49 z=41 done	z=11 z=41 done	Z=41 done	Z=41 done	Z=41 done	z=58- 17 z=41 done
5	34-z=18 Add z to both sides!	Z=16 done	z=16 done	Z=16 done	Z=16 done	z=16 done	Z=16 done	Z=16 done	Z=16 done	z=34- 18 z=16 done

	Question	J	M	L	S	P	F	G	Mc	Ir
6	962-x=476 Add X to both sides	Z=486 not x X=486 done	x=486 done	X=486 done	X=486 done	x=-284 x=488 x=486 done	X=486 done	X=486 done	Z=486 X=476 X=486 done	x=962-476 x=486 done
7	9-z=17 Add Z to both sides	Z=-8 done	9=17+z z=-8 done	Z=-8 done	Z=-26 Passed	Z=-8 done	Z=-8 done	Z=-6 Z=-6 Z=-8 done	Z=28 Z=28 Z=28 Z=11 Passed	z=9-17 x=-8 z=-8 done
8	7-z=9 Add Z to both sides	Z=-2 done	7=9+z z=-2 done	Z=-2 done	Z=-16 Passed	Z=-2 done	Z=-2 done	Z=-2 done	Passed	z=7-9 z=-2 done
9	-9-z=-7 Add z to both sides!	Z=-2 done	-9=-7+z z=-2 done	Z=-2 done	Z=+2 z=-2 done	Z=-2 done	Z=-2 Done	Z=-2 Done	Z=-2 done	z=16 z=2 z=-2 done
10	476-z=962 Add z to both sides!	Z=-486 Done	z=-486 done	Z=-486 done	Z=-486 Done	Z=-486 Done	Z=-486 Done	Z=-514 Z=-514 Z=-514 Z=-514 Z=-514 Z=-514 z=-514 Z=-514 Passed	Z=-486 done	z=476-962 z=-514 Passed

Test 2: responses

	Question	J	M	L	S	P	F	G	Mc	Ir
1	x/8=12 Multiply both sides by 8	N/A	x=12*8 x=96 done	96 *Syntax* X=12*8 Passed	x=96 done	x=96 done	x=92 x=92 x=96 done	x=96 done	X=1.5 X=96 done	x=-4 x=96 done
2	x/3=15		x=15*3 x=45 done	X=45 done	X=45 done	X=45 done	X=45 Done	X=45 done	X=45 Done	X=45 done
3	x/5=25		x=25*5 x=125 done	X=125 done	X=125 done	X=125 done	X=125 Done	X=125 done	X=125 Done	X=125 done
4	x/7=56 Multiply both sides by 7		x=56*7 x=392 done	X=392 done	392 392 x=392 done	X=392 done	x=442 x=392 done	X=392 done	X=393 X=392 done	x=492 x=393 x=402 Passed
5	x/15=22		x=22*15 x=330 done	x=210 Passed	X=330 done	X=330 done	X=330 Done	X=330 done	X=330 done	X=330 done
6	x/24=16		x=16*24 x=384 done	X=384 done	X=384 done	X=384 done	X=384 Done	X=384 done	X=384 done	X=384 done
7	-x/4=22 You've really got me there ask your teacher		-x=22*4 -x=88 x=-88 done	X=-88 done	-z=-88 Passed	x=88 x=-88 done	X=-88 Done	Passed	X=-88 done	X=88 Passed
8	x/6=-18 Multiply both sides by 6 -ve sign		x=-18*6 x=-108 done	x=108 Passed	x=-108 done	-108 x=3 Passed	x=- 106 x=- 106 x=- 108 done	Passed	X=-108 done	x=18*6 x=-18*6 x=-108 done

	Question	J	M	L	S	P	F	G	Mc	Ir
9	-x/2=34 Divide both sides by -1		-x=34*2 -x=68 x=-68 done	X=-68 done	x=-68 done	Passed	X=-68 Done	Passed	X=-68 done	-x=34*2 -x=68 Passed
10	x/3=-15		x=-1.5*3 x=-15*3 x=-45 done	X=-45 done	x=-45 done	X=-45 done	X=-45 done	Passed	X=-45 done	X=45 done
11	2x/7=6 Divide both sides by 2		2x=7*6 2x=42 x=21 done	X=21 done	x=21 done	x=22 x=21 done	x-21 x=21 done	Passed	X=21 done	2x=42 x=21 done
12	5x/3=15 Divide both sides by 5		5x=15*3 5x=45 x=45/5 x=9 done	X=9 done	x=9 done	x=8 x=9 done	X=9 done	X=8 X=8 x=9 done	X=15 X=25 X=15 X=25 X=375 Passed	5x=45 x=7 x=9 done
13	11x/15=22 Divide both sides by 11		11x=15*22 11x=330 x=330/11 x=30 done	X=22 Passed	x=30 done	x=3 x=30 done	X=30 done	X=33 X=30 done	Passed	11x=330 x=30 done
14	8x/7=56 Divide both sides by 8		8x=7*56 8X=392 x=392/8 x=49 done	X=49 done	x=49 done	X=49 done	X=49 done	X=1 Passed	Passed	X=49 done
15	4x/24=16 Divide both sides by 4		4x=16*24 4x=384 x=384/4 x=96 done	X=96 done	x=96 done	X=96 done	X=96 done	Passed	X=68 X=1 Passed	x-96 x=96 done

Test 3 Responses

	Question	J	M	L	S	P	F	G	Mc	Ir
1	If 4x = 20 then find 5x No hint provided	25 done	x=20/4 5x=25 only 25 required	25 done	5x=25 25 done	20 5x=25 25 done	25 done	25 done	5x=25 5x=25 25 done	25 done
2	Find 5k given 3k=18	30 done	30 done	30 done	30 done	30 done	30 done	30 done	30 done	30 done
3	26-2n=8 Add 2 N to both sides	N=9 Done	2n=26-8 n=9 done	N=9 done	n=-9 -9 n=-9 Passed	N=9 done	N=9 done	18=2n <enter> 9=n done	N=9 done	18=-2n 18=2n n=9 done
4	2(x+3) No hint provided	2x+6 done	2*(x+3) not sure what to do next Passed	2x+6 done	2x+6 done	6xS 6x² 2x+6 done	2x+6 done	2x=3 2x=6 2x=6 2x+6 done	2x+6 done	2x+6 done
5	-(3x-w) No hint provided	w+3x w+3x w-3x done	-3x-w Passed	-3x+w done	-3x+w done	Passed	-3x+w done	-3x+w done	3x+w -3x-w -3x² +w² 3x²+w² Passed	-3x+w done
6	If 3x+2y=7 find a value for 6x+4y No hint provided	9 9 6x+4y=9 14	14 done	14 done	14 done	14 done	14 done	9 Passed	14 done	14 done
7	4y=6 Divide both sides by 4	Y=1.5 Done	Y=1.5 done	Y=3/2 done	Y=1.5 done	y=1.3 1.5 y=1.5 done	Y=3/2 Done	Passed	Y=1.5 done	Y=6/4 Y=1.5 done

	Question	J	M	L	S	P	F	G	Mc	Ir
8	4-2x=10-6x	x=1 x=1.5 done	4=10-4x x=1.5 done	X=3/2 done	x=1.5 done	X=1.5 Done	X=3/2 done	Passed	X=2 X=12 X=8 X=-8 X=1.5 done	X=-1.5 X=1.5 done
9	3+21(s+4) No hint provided	24s+96 28+s 96+24s 96+24s 3s+21s+12+84 3s+21s+12+84 3s+21s+12+84 =24s+96 3s+21s+12+84 =24s+96 Passed	3=21s+4s 3=21s+84 3+21s+84 21s+87 done	21s+87 done	s=4 24s+96 87+21s done	3+4S*21 Passed	21s+87 done	3+21s+8 4 87+21s done	24s+96 $3+21s^2+84$ 87+21s done	21s=87 21s+87 done
10	3+23(s-4) No hint provided	Passed	3+23s-92 23s-89 done	23s_89 syntax Passed	- 89+23s done	Passed	23s-89 done	3+23s- 92 23s-89 done	95+23s 3+23s+92 Passed	23s-89 done
11	17- 21(s+7) No hint provided	17s+119-21s- 147 17s+119-21s- 147=-4s-28 -42-28 =-42-28 =-42-28 Passed	17- 21s+147 17-21s- 147 21s-130 -21s-130 done	150- 21s Passed	164- 21s -132- 21s Passed	Passed	-21s-30 done	17- 21s+147 17-21s- 147 -21s- 130 done	-4s+28 Passed	21s-130 21s-130 -21s- 130 done
12	13-19(s- 3) No hint provided	Passed	13-19s-57 13-19s+57 -19s+70 done	13- 19s+57 70-19s done	80-19s 13- 19s-57 -44- 19s 19s+57 Passed	Passed	70-19s done	13- 19s+57 70-19s done	19s+70 19s-70 $19s^2+70$ Passed	-19s-52 19s+52 -19s+57 -19s-57 19s+70 19s-70 70-19s done

	Question	J	M	L	S	P	F	G	Mc	Ir
13	3xy+4xz Look for common factors.	7xyz $7x^2yz$ $7x^2yz$ 3xy+4xz $7x^2+7y+7z$ $7x^2+7y+7z$ Passed	3x(y)+4x(z) 7x(yz) 7x+(yz) 7x(y)(z) 7x(y) (7x(y)(7x(y)(Passed	X(3y+4z) done	7+2x+y z 7+2x+y +z Passed	Passed	<enter> passed	3(x+y)+ 4(x+y) Passed	12x+3y+4z 7x+3y+4z 7x+7y+7z 12x+12y+1 2z Passed	3y+4z 1y+4z 1.5y+2z 3y+4z Passed
14	What is missing in the expression 3b+1=b+.. .... No hint provided	3 3 Passed	b+2b+1 2b+1 done	1/3 Passed Question wording /syntax	2b+1 done	2*b+1 done	3 2b+1 done	3a-4 Passed	3 4 2 1 6 7 8 8 8 12 10 9 2b+1 done	2b+1 done
15	complete the expression 5a-4=2a+2+ ..	Passed	3a-6 done	3a-6 done	3a-6 done	3a-6 done	3a-6 done	3a-4 Passed	2a 3a 3a-6 done	Passed

	Question	J	M	L	S	P	F	G	Mc	Ir
16	complete the following 4c+8 = 5c + 10 + ...	Passed	-2-c done	-c-2 done	+1-3-3c No hint provided +1-3-1c -2-1c done	-2-c done	-c-2 done	-c-2 done	+2-c c-2 -4+c -4-c -2-1c done	c+2 -c-2 done
17	2y=10-6y Bring the Ys together	y=1.25 done	8y=10 y=1.25 done	y+5/4 y=5/4 done	y=1.25 done	Y=1.25 done	5/4 y=5/4 done	Passed	Passed	y=1.25 done

Test 4 Responses

	Question	J	M	L	S	P	F	G	Mc	Ir
1	2x when x=9 No hint provided	18 done	18 done	18 done	18 done	18 done	18 done	18 done	18 done	18 done
2	4x when x=6 No hint provided	24 done	24 done	24 done	24 done	24 done	24 done	24 done	24 done	16 24 done
3	2x+3 when x=5 No hint provided	13 done	13 done	13 done	13 done	13 done	13 done	13 done	13 done	13 done
4	-3y+8 when y=2 No hint provided	2 done	2 done	2 done	2 done	2 done	2 done	-6+8 2 done	14 2 done	2 done
5	3z-8 when z=-3 No hint provided	1 done	1 done	1 done	1 done	1 done	1 done	9-8 1 done	1 done	1 done
6	Xy when x=3 and y=5 No hint provided	15 done	15 done	15 done	15 done	15 done	15 done	15 done	15 done	15 done
7	Xy+1 when x=4 and y=6 No hint provided	-23 done	-4*6+1 -23 done	-23 done	25 25 -23 done	25 -23 done	-23 done	-24+1 -23 done	25 -23 done	23 -23 done

Test 5 Responses

	Question	J	M	L	S	P	F	G	Mc	Ir
1	A(BC) No hint provided	AB+AC =AB+AC Passed	A*(b*c) done	Abc done	ABC Done	ABC done	ABC Done	AXBXC Passed	ABC done	N/A
2	(ax+b)(cx+d) No hint provided	Passed	(a*x+b)*(c*x+d) a*x+b*c*x+d no hint Passed	acx²+bcx+ dax+bd done	axxcxxacdxbcxxbd <enter> Passed	Passed	acx²+axd+bcx+bd done	acx²+axd+bcx+bd done	Passed	
3	(y+2)/(y-3)= 2 Multiply both sides by Y-3	y=(y+2)/2+3 Passed	y+2=2(y-3) y=2(y-3)-2 y/2=(y-3)-2 2y=(y-3)/2 y/y-3=-2 Passed	Y=8 exit	-3 Passed	Passed	Y=8 done	Passed	Passed	
4	(z+1)/(z+4) = 5/6 Multiply both sides by Z + 4	Passed	6(z+1)/6(z+4)=5 z+1=5/6(z+4) z=5/6(z+4)-1 z²+4z=5/6-1 z/(z+4)=5/6+1 2z=5/6+4-1 Passed	Z=14 done	Passed	Passed	Z=14 done	Passed	Passed	
5	2z+5=11	Passed	2z=6 z=3 done	Z=3 done	Passed	Z=3 done	Z=3 done	Z=3 Done	Passed	
6	3z+5=z+3	Passed	2z+5=3 2z=-2 z=-1 done	Z=-1 Done	Passed	Passed	Z=-1 done	Z=1 done	3z-z+5=3 2z+5=3 2z+5-5=3-5 2z=-2 z=-1 done	

7	(x-5)(x-7)=3 Factorise	Passed	$x^2-35=3$ x=19 Passed	X=8 or x=4 done	Passed	Passed	Passed	Passed	Passed	
8	$x^2 + 5/6\,x + 1/6 = 0$	Passed	Passed	X=-1/3 or x=-1/2 done	Passed	Passed	Passed	Passed	Passed	

Test 6 Responses

	Question & Hint	**J**	**M**	**L**	**S**	**P**	**F**	**G**	**Mc**	**Ir**
1	2^3 Multiply before you add.	200 8 done	8 done	8 done	N/A	8 done	8 done	8 done	N/A	N/A
2	(-1)^3 Multiply before you add.	-1 done	-1 done	-1 done		-1 done	-1 done	Passed		
3	2^-2 Multiply before you add.	-4 Passed	Did not know ^ Passed	¼ done		¼ done	Passed	Passed		
4	(a+b)^2 No hint provided	Passed	a^2+b^2 Passed	$a^2+2ab+b^2$ done		A^2+B^2 Passed	$A^2+2AB+B^2$ done	ab^2 AB^2 AB^2 Passed		
5	(a+b)^-2 No hint provided	Passed	Passed	$1/a^2+2ab+b^2$ passed		$1/a^2+b^2$ Passed	Passed	Passed		

Appendix 11 Software evaluation questionnaires

11a Preliminary Pilot Questionnaire

Predictive Evaluation of TREEFROG using selected criteria

Details of Use & User

Date/Time Started:			Date/Time Finished:			Number of hours (est.)			
Course			Year Group			Gender			
Level of IT Experience			Grade & Highest Mathematics Qualification			Year Maths last studied			
Age Group	Pre-18		18 -30		31-40		41-50		Other

Please undertake the TREEFROG exercise of sample numeracy test questions. Then in consideration of each of the following criteria rate the usefulness of TREEFROG

LEARNING

	Good	ok	Poor	Comment
Pedagogical techniques To what extent does the learning environment should support the notion of providing help to amend errors				
Curriculum coverage The match between the subject content of the application and that required by the curriculum or learner				
Correctness of the underlying model the validity of methods of solution of the software				
Cosmetic authenticity Balance of state of the art graphics features and intrinsic benefits for learning				
Self Directed Learning the extent to which learners are empowered to progress their own learning				
Terminology Use of formal terms				
Multiple views & representations The validity of the representations / presentation relative to different learning techniques				
Reality vs artistic interpretation Graphics displayed should be conducive to learning and not distracting or misleading				

USABILITY

	Good	ok	Poor	Comment
Representation forms methods of interaction should not be cumbersome or unfamiliar, but the interface should be functional				
Interaction flow The balance between helpful informative interaction flow and feedback which intrudes and could interrupt progress.				
Navigation & Superficial complexity To what extent does the environment support movement within questions and from one question to another in support of learning				
Self direct learning To what extent does the structure of the environment support the notion of self-directed learning?				
Learners' support materials To what extent can novice users develop an understanding of the requirements of the system and hence develop the skills to maximise fully use the system.				
Symbolic representation To what extent are the symbols, icons, buttons used to represent mathematical symbols and concepts are consistent with the traditionally acknowledged conventions?				
Runtime and cognitive errors Shielding from usability obstacles such as runtime errors and malfunctions. Protection against making annoying errors such as incorrect syntax of statements				

LEARNING & USABILITY

	Good	ok	Poor	Comment
Designer/learner models correlation between method of solution and that of the system in terms of how tasks are presented and the nature of feedback				
Learner control and freedom The extent to which the environment enables you to experiment with solutions and develop a personal understanding.				
Tailoring the interface In view of the notion of you having ownership of your own progress does the interface provide flexibility for a range of learners?				
Teacher customisation of content A good feature of educational software is the tutors' facility to adapt the subject content to match either curriculum changes or changes in the needs of learners.				

Please add on the back of this sheet any other comments regarding the usefulness or barriers to use of the system.

Thank you for your time and interest to help with this research.

11b **Reworded and adapted questionnaire**

Evaluation of TREEFROG using selected criteria

Details of Use & User

Date/Time Started:			Date/Time Finished:				Number of hours (est.)				
Course			Year Group				Gender				
Level of IT Experience			Grade & Highest Mathematics Qualification				Year Maths last studied				
Age Group	Pre-18		18 – 30		31-40		41-50		Othe r		

Please undertake the TREEFROG exercise of sample numeracy test questions. Then in consideration of each of the following criteria rate the usefulness of TREEFROG

LEARNING

	Good	ok	Poor	Comment
Rate the usefulness of the software to help you to correct errors in your understanding				
How well did the software cover the subject area you need to study?				
How much did the software accept your method for solving a given question				
Rate the balance between state of the art graphics features and benefits for learning				
Do you think the software would help you to carry on doing work on your own?				
Were mathematical terms used the way you expected?				
Rate the appropriateness of the representations or presentation of different types of question				
Did you find the graphics generally helped you to learn or were they poor in that they distracted you?				

USABILITY

	Good	ok	Poor	Comment
How would you rate the methods of interaction in terms of ease of use, familiarity and functionality?				
Rate the balance between helpful informative interaction and feedback which intrudes and interrupts progress.				
To what extent does the environment support movement within questions and from one question to another to help learning?				
To what extent does the structure of the environment support the notion of self-directed learning?				
How easy was it to develop an understanding of how to use the system and hence develop the skills to maximise full use the system?				
To what extent did you recognise the mathematical symbols, icons and buttons used by the software?				
How would you rate the system's ability to shielding you from usability obstacles such as runtime errors and malfunctions?				
How would you rate the system's ability to protect you from making annoying errors such as incorrect syntax of statements?				

LEARNING & USABILITY

	Good	ok	Poor	Comment
To what extent did your methods for solving problems and those used by the system to present questions and feedback match?				
To what extent does the software enable you to experiment with solutions and develop your own personal understanding?				

Please add on the back of this sheet any other comments regarding the usefulness or barriers to use of the system.

Thank you for your time and interest to help with this research.

11c Predictive Evaluation of WEBFROG using selected criteria

Name:____________________

Details of Use & User

Course		Gender			
Level of IT Experience		Number of times used)		Grade & Highest Maths Qualification	

In consideration of each of the following criteria rate the usefulness of WEBFROG

LEARNING

	Good	ok	Poor	Comment
To what extent does the software help you to correct errors?				
To what extent is the subject content useful to your learning?				
Did the software let you use the methods you wanted to in answering questions				
How appropriate do you consider the interface to be for learning				
To what extent were you allowed to direct your own learning and pace of learning				
To what extent did you understand the terms used				
To what extent was the information presented in your way of learning?				
To what extent were the Graphics displayed should be conducive to learning and not distracting or misleading				

USABILITY

Rate the ease of interaction with the interace	Good	ok	Poor	Comment
Rate the flow and feedback in terms of usefulness and interrupting progress.				
To what extent does the environment support movement within questions and from one question to another				
To what extent does the structure of the software support self-directed learning?				
To what extent could you understand how to use the system and hence develop the skills to maximise fully use the system.				
To what extent are the symbols, icons,				

buttons used to represent mathematical symbols and concepts what you would expect?				
How would you rate the usability of the system in terms of obstacles such as runtime errors and malfunctions or annoying errors such as incorrect syntax of statements				

LEARNING & USABILITY

	Good	ok	Poor	Comment
To what extent were the tasks presented and the feedback what you would expect?				
The extent to which you could experiment with solutions and develop your understanding.				
To what extent did the interface enable you to "own" your learning?				

Please add below any other comments regarding the usefulness or barriers to use of the system.

Thank you for your time and interest to help with this research

Appendix 12 Group interview schedule

Learning

Explanations/feedback

Guidance

Feelings towards mathematics

maths topics supported

bank of questions

types of questions

Other comments relating to the mathematics content

Usability

user friendly

Motivational

Syntax

Interactivity

Other comments relating to usability of the software

Learning/Usability

Sound

Diagrams/visual add ins

Appendix 13 Mapping of Treefrog tests and Diagnosys skills

TREEFROG Test 1 Basic +, - in Solving expressions and linear equations and Test 2 Basic *, / in solving expressions and linear equations

Numbers - decimals, negative numbers, inequalities, etc. numbers,101,102,103,104,105,107,108,109,112,201,202,205,206,215,216	TREEFROG
skill{101,Multiplication of negative numbers,1}	Y
skill{102,Multiply negative and positive,1,101}	Y
skill{103,Negative Numbers,1,102,201}	Y
skill{104,Size of decimals,1,201}	Y
skill{105,Decimal places,1,202}	Y
skill{107,Ratios,1,205,206}	Y
skill{108,Factors of an integer,1,107,206,208}	Y
skill{109,Simple fractions,1,216}	Y
skill{112,Simple calculation,1,210,214,215}	Y
skill{201,Use of < and > signs,2}	Y
skill{202,Significant figures,2}	Y
skill{205,Inverse ratios,2}	Y
skill{206,Cancelling numerical fractions,2,207,216}	Y
skill{215,Precedence Rules,2}	Y
skill{216,Add/subtract numerical fractions,2,308}	Y
skill{301,Solution of simple inequalities,3}	Y
skill{302,Scientific Notation,3}	Y

Test 6 Indices

Powers,106,203,204}	
skill{106,Definition of positive powers,1,203,204}	Y
skill{203,Definition of negative powers,2}	Y
skill{204,Rules for positive powers,2}	Y
skill{303,Simplification of fractions with powers,3}	Y
skill{304,Rules for negative powers,3}	Y
skill{305,Rules for fractional powers,3}	Y
skill{306,Definition of fractional powers,3}	Y

Test 5 Distribution on multiplication (BODMAS) including algebraic division and simplification of expressions

Basic algebra - expanding brackets and collecting terms, formulae, simple factorization, linear equations, algebraic fractions	•
skill{110,Collect terms (simple),1,211,312,317}	Y
skill{111,Solving a simple equation,1,212}	Y
skill{113,Evaluating a simple expression,1,214}	Y
skill{207,Multiply algebraic fractions,2,307}	Y
skill{208,Factors of algebraic products,2,209}	Y
skill{209,Simple Factorisation,2,309,311}	Y
skill{210,Expanding one bracket,2,209,312}	Y
skill{211,Collecting terms,2}	Y
skill{212,Solving linear equations,2,309,310,314,315}	Y
skill{213,Transposition of formula,2,315}	Y
skill{214,Evaluating formula,2,316}	Y

Test 3 Rewriting expressions and word problems

Test 4 Substituting into algebraic expressions including word problems

Equations,309,310,314,315,316}	
skill{309,Simple Quadratic equations,3}	Y
skill{310,Relation between roots and factors,3}	Y
skill{312,Expanding two brackets,3,311}	Y - bodmas
skill{314,Simultaneous Equations,3}	Y
skill{315,Unusual linear equation,3}	Y
skill{316,Use of quadratic formula,3}	Y
skill{324,Percentages (advanced),3}	Y
skill{401,Simplify with scientific notation,4}	Y
skill{403,Arbitrary factors,4}	Y
skill{404,L.c.d. of an algebraic fraction,4}	Y
skill{405,Identification of common errors,4}	Y
skill{410,Substituting into a formula,4}	Y
skill{411,Solutions of a quadratic,4}	Y
skill{412,Difficult linear equation,4}	Y

Skills within DIAGNOSYS not relevant to questions posed within TREEFROG

Algebra and calculus {409}	
skill{409,Divide by zero (possible solution),4}	Y - common error

Algebra methods,307,308,311,312,313,404,407,410	
skill{307,Division of algebraic fractions,3}	Y
skill{308,Add/subtract algebraic fractions,3}	Y

234,Venn Diagrams (probability),2,335}	n- diagram
335,Venn Diagrams (conditional prob.),3}	N
skill{311,Factorising a quadratic,3,310}	N (specific use)
skill{313,Difference of squares,3}	N
skill{406,Solve quad. by comp. the square,4}	N
skill{407,Quadratics - completing the square,4}	N
Statistics - range, mean, simple and conditional probability	•
231,Range of a set of numbers,2}	Y - -ve numbers
232,Mean of a set of numbers,2}	Y
233,Simple probability (coins),2,334}	Y
334,Conditional probability,3}	Y
Miscellaneous – percentages {223}	Y

Graphs - co-ords, gradient, linear, quadratic,reciprocal graphs	N – diagrams
graphs,151,251,351,451,452}	
skill{151,Coordinates,1,251}	?
skill{251,Gradient of a straight line,2,351}	?
skill{351,Equation of a straight line,3}	?
skill{451,Recognise formula of quad. graph,4}	?
skill{452,Recognise formula of recip. graph,4}	?
Algebra+calculus,317,341,342,343,405,408,409,441,442}	
skill{441,Product rule,4}	N
skill{442,Integration of powers,4}	N
Area+volume,161,261,262,263,361,362,363,461,462}	**N**
skill{161,Area of a triangle,1,261,262}	N
skill{221,Pythagoras,2,321}	N
skill{222,Definition of sin and cos,2}	N
skill{261,Area of trapezium,2}	N
skill{262,Area and circumference of a circle,2,361}	N
skill{263,Similar triangles,2,362}	N
skill{461,Surface area of a cylinder,4}	N
skill{462,Volume Area Length relationships,4}	N
skill{421,Deduce radius of circle,4}	N
skill{422,Sin and Cos as functions,4	N
skill{323,Definition of radians,3}	N
skill{317,Complex numbers,3}	N
skill{321,Equation of a circle,3}	N
skill{322,Sin and Cos formula,3}	N
skill{341,Differentiation of powers,3,342,441,442}	N
skill{342,Finding Max/Min of a quadratic,3}	N
skill{343,Geometric Progression,3}	N
skill{361,Volume of cylinder,3}	N
skill{362,Area/Length relationship,3}	N
skill{363,Area of irregular shapes,3}	N
skill{402,Logs,4}	N
skill{408,Multiplication of complex numbers,4}	N

Appendix 14 Questionnaire Survey of Numeracy and Algebra

14a Full Responses

	Questionnaire Results																		Number of						
Questions	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	1	2	3	4	5	x	Total
1	x	1	2	3	1	4	1	4	1	2	2	2	2	2	x	5	4	2	4	7	1	3	1	2	18
2	x	1	1	1	1	2	3	1	1	1	1	1	2	2	x	5	2	2	9	5	1	0	1	2	18
3	2	3	2	1	2	4	3	2	2	2	4	4	4	2	4	5	4	4	1	7	2	7	1	0	18
4	2	4	3	2	1	2	1	3	4	1	2	4	4	4	2	5	5	1	4	5	2	5	2	0	18
5	2	2	2	1	2	1	1	2	2	1	2	2	5	1	2	1	2	1	7	10	0	0	1	0	18
6	2	2	1	1	4	1	2	1	3	1	2	2	1	1	4	5	2	1	8	6	1	2	1	0	18
7	2	1	1	3	2	1	1	1	2	1	3	1	1	1	1	1	1	x	12	3	2	0	0	1	18
8	4	2	3	3	4	4	3	2	2	4	2	3	4	2	x	5	4	x	0	5	4	6	1	2	18
9	1	2	2	3	2	2	x	2	x	1	3	2	2	3	4	2	4	x	2	8	3	2	0	3	18
10	4	3	4	2	2	5	5	3	4	4	4	4	5	3	4	5	5	3	0	2	4	7	5	0	18
11	4	5	4	4	4	5	5	3	2	5	4	5	5	4	5	5	5	3	0	1	2	6	9	0	18
12	x	2	2	4	2	1	2	x	2	3	3	3	1	3	3	5	4	x	2	5	5	2	1	3	18
13	4	5	3	4	4	4	4	2	4	4	4	4	4	5	4	5	5	2	0	2	1	11	4	0	18
14	2	1	1	1	2	2	1	2	2	1	1	1	1	1	1	1	1	2	12	6	0	0	0	0	18
15	4	4	4	3	4	4	1	3	4	4	4	3	5	4	2	5	5	4	1	1	3	10	3	0	18
16	2	4	3	2	4	2	1	2	4	2	4	2	2	3	3	5	5	3	1	7	4	4	2	0	18
17	2	4	4	2	2	3	1	3	4	1	3	4	2	4	x	5	4	4	2	4	3	7	1	1	18
18	2	5	2	1	4	1	1	1	2	4	2	1	1	3	5	3	2	2	6	6	2	2	2	0	18
19	x	2	1	4	2	3	3	2	2	2	1	2	1	1	1	2	4	2	5	8	2	2	0	1	18
20	2	2	2	4	2	2	2	2	2	2	4	4	1	1	1	5	3	4	3	9	1	4	1	0	18
21	4	2	4	4	2	4	4	3	2	4	4	4	2	x	2	5	3	4	0	5	2	9	1	1	18
22	x	2	2	2	2	2	3	2	2	1	x	2	1	3	1	1	2	2	4	10	2	0	0	2	18
23	2	2	2	2	2	2	1	2	2	2	2	1	2	2	1	1	2	2	4	14	0	0	0	0	18

Num of	Respondent																	
Responses	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	4	5	6	3	5	10	4	2	9	3	5	8	6	6	5	2	3
2	11	10	9	6	13	8	3	11	13	6	7	7	7	5	4	2	6	8
3	0	2	4	5	0	2	5	6	1	1	4	3	0	6	2	1	2	3
4	6	4	5	6	7	6	2	1	6	6	8	7	4	4	5	0	7	5
5	0	3	0	0	0	2	2	0	0	1	0	1	4	1	2	15	6	0
x	5	0	0	0	0	0	1	1	1	0	1	0	0	1	4	0	0	4
Mean	3	3	2	2	2	3	2	2	3	2	3	3	3	3	3	4	3	3
Mode	2	2	2	1	2	2	1	2	2	1	4	2	1	1	1	5	4	2

14b: Overview of levels of agreement between electronic survey and the QCA reports

Questions with majority of responses in agreement

Question 14
Equations which involve only one variable (3k=18, 5k=?) are not problematic for the majority of pupils at level 6 and higher

Question 23
At higher levels pupils need more opportunity to structure and organise solutions to problems.

Question 5
At level 5 most pupils do know a standard approach to collecting together like terms

Question 7
At level 3: most are unable to complete equations of the form ... / 24 = 16

Question 2
At level 6 in re-writing an expression most are successful if positive terms only such as 3b+1 = b+ .... and 5a-4 = 2a+2 +

Question 6
At level 5 when negative numbers or signs are involved many pupils make arithmetic errors such as 4-2y=10-6y □ 4y=14 □ y=3.5

Question 20
Pupils at all levels need help understanding negative quantities in algebra

Question 22
At higher levels pupils need more opportunities to handle questions where little initial mathematical orientation is given

Question 19
Pupils at all levels need help with handling indices

Question 1
At level 3: most are unable to complete equations of the form 962 - ... = 476

Question 18
The solution of simultaneous equations is very weak at levels 5 & 6

Question 9
At higher levels, when the substitutions, manipulations and work with equations is more complex, pupils' answers

Questions with the majority of the responses not in agreement

Question 10
At levels 6 & 7, most pupils are unable to write the required expression from a sequence eg. 1 hut needs 6 matches, 2 huts need 11 matches

Question15
At level 6 most were not able to substitute accurately a numerical value into an expression

Question 11
Equations which involve only one variable (3k=18, 5k=?) are problematic for the majority of pupils at level 4 & 5

Question 13
A value n being deduced from an equation of the form 26-2n=8 is extremely difficult to solve for pupils at level 6 and lower

Questions with majority of responses as Unsure or Variable

Question 16
A value n being deduced from an equation of the form 26-2n=8 is problematic to solve for pupils at level 6 and lower

Question 3
At level 5 many pupils when solving Linear Equations make arithmetic errors of the form 4y = 6 □ y= 1/3

Question 4
At level 5 most pupils are unable to find y from 4-2y=10-6y

Question 12
At lower levels the comparison of two algebraic expressions (3x+2y=7, 6x+4y=?) is significantly more problematic than that between two equations which involve two numerical values only (4k+2a=82, ?=41)

Question 17
The solution of simultaneous equations is weak at levels 7 & 8

Question 8
At level 6 in re-writing an expression most are unsuccessful if negative numbers are required 5d + 6=2d–3 + and 4c=2c+1 + ...

Question 21
Pupils at all levels need help understanding negative quantities in number

Appendix 15 Diagnosys Pilot Study results

15a: Individual Participant Results

START LEVEL	3	3	2	2	2	3	2	2	3	2	2	2	2	2	2	2	2	2
TOTAL	68	62	53	6	21	68	38	59	94	29	50	18	29	21	26	44	35	32
Q ASKED	18	17	25	26	17	24	31	23	18	30	28	31	18	34	26	27	29	27
Q CORRECT	11	10	9	1	4	13	10	12	16	7	12	4	5	7	6	9	8	6
NUMBERS	67	67	60	13	33	80	53	87	93	40	73	27	60	13	47	67	47	60
POWERS	67	67	67	0	0	67	0	33	100	67	33	0	0	33	0	33	67	0
BASIC ALGEBRA	91	82	55	0	18	73	18	45	91	18	27	18	9	18	18	27	18	18
STATS	33	0	0	0	0	33	33	0	100	0	67	0	0	0	0	33	33	0
MISC	0	0	50	0	0	0	100	50	100	0	0	0	0	100	0	0	0	0
GROUP	Sc	Ma	Sc	DT/IT	PE	Ma	DT/IT	Sc	Sc	Sc	DT/IT	Sc	DT/IT	Sc	DT/IT	DT/IT	Sc	DT/IT
QUAL	A-level	A-level	GCSE	GCSE	GCSE	Access	GCSE	GCSE	A-level	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	Voc	GCSE
GRADE	E	E	C	C	B	pass	C	B	C	C	C	C	C	C	C	C	pass	C

START LEVEL	3	2	2	2	2	2	2	2	2	3	2	3	3	2	2	2	2	2	2	2
TOTAL	97	47	26	26	32	29	44	82	35	76	50	68	88	62	29	59	29	18	24	44
Q ASKED	18	24	28	31	25	29	28	21	28	21	27	26	19	26	16	25	32	24	32	20
Q CORRECT	17	9	6	6	6	5	9	15	7	13	10	15	15	13	6	12	8	5	6	9
NUMBERS	93	80	40	47	40	40	73	80	67	87	73	67	100	80	47	87	33	13	33	53
POWERS	100	33	33	0	33	33	33	100	0	33	33	33	100	67	33	67	33	67	0	33
BASIC ALGEBRA	100	18	9	9	27	18	18	91	18	82	27	73	82	36	18	36	9	18	18	55
STATS	100	0	0	0	0	0	0	67	0	33	0	67	33	33	0	33	67	0	0	0
MISC	100	50	50	50	50	50	50	50	0	100	100	100	100	100	0	0	50	0	50	0
GROUP	Ma	Sc	PE	Sc	Ma	PE	Sc	Sc	PE	Ma	Sc	Ma	Ma	PE	PE	Sc	DT/IT	Sc	PE	PE
QUAL	A-level	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	A-level	GCSE	A-level	A-level	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE
GRADE	B	U	B	C	A	C	C	B	C	E	B	D	E	B	B	B	C	C	C	B

START LEVEL	2	2	3	2	2	3	2	2	2	2	2	2	2	2	2	2	3	2	3	3
TOTAL	32	82	88	62	18	82	56	79	24	26	71	18	18	53	26	18	74	56	94	74
Q ASKED	30	21	19	24	33	22	27	20	19	29	23	30	14	23	16	31	22	25	18	21
Q CORRECT	7	15	15	11	5	16	12	13	3	5	13	3	3	8	5	4	13	11	16	12
NUMBERS	53	100	100	67	20	93	73	93	53	47	93	33	33	100	33	13	73	67	93	73
POWERS	0	67	67	67	0	67	33	67	0	33	33	0	33	33	33	33	100	67	67	67
BASIC ALGEBRA	18	82	91	64	9	82	36	73	0	9	55	0	0	18	27	27	91	45	100	82
STATS	0	33	33	33	33	67	33	33	0	0	33	33	0	0	0	0	0	67	100	33
MISC	50	50	100	50	50	50	100	100	0	0	100	0	0	0	0	0	50	0	100	100
GROUP	PE	PE	Ma	Sc	PE	Ma	Sc	Sc	DT/IT	PE	DT/IT	PE	PE	PE	DT/IT	PE	Ma	Sc	Ma	Sc
QUAL	GCSE	GCSE	A-level	GCSE	GCSE	AS-level	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	A-level	Voc	A-level	A level
GRADE	C	C	B	B	C	E	C	C	C	B	B	C	C	B	C	C	B	merit	D	N

START LEVEL	2	2	2	2	2	2	2	2	2	2	2	3	2	3
TOTAL	9	68	26	50	32	41	38	53	47	26	29	82	29	71
Q ASKED	33	25	25	28	17	21	30	26	28	16	29	19	29	22
Q CORRECT	3	14	6	11	6	6	9	10	10	5	6	13	6	12
NUMBERS	7	80	40	67	47	73	60	67	73	33	33	87	47	73
POWERS	0	33	0	67	67	0	33	33	33	67	33	67	67	33
BASIC ALGEBRA	9	64	27	18	18	27	9	55	18	18	27	100	9	82
STATS	33	33	0	33	0	0	0	0	33	0	33	0	0	33
MISC	0	100	0	100	0	0	100	50	50	0	0	100	0	100
GROUP	PE	Sc	PE	DT/IT	DT/IT	DT/IT	DT/IT	PE	Sc	DT/IT	Sc	Ma	DT/IT	Ma
QUAL	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	GCSE	Overseas	GCSE	A-level
GRADE	C	C	C	C	C	C	B	C	B	C	C	B	C	E

Total Number of Participants = 71

15b: **Summary Data**

Summary data: Starting at Level 2 GCSE Grade C

	Max	Min	Mean	Mode	0-20	21-40	41-60	61-80	81+	TOTAL
TOTAL	82	21	49	59	0	5	8	3	1	17
Q ASKED	30	16	24	23	3	14				17
Q CORRECT	15	4	10	9	17					
NUMBERS	100	33	69	80	0	2	4	7	4	17
POWERS	100	0	43	33	1	11	0	4	1	17
BASIC ALGEBRA	91	9	34	18	8	3	4	1	1	17
STATS	67	0	18	0	10	5	0	2	0	17
MISC	100	0	41	0	7	0	6	0	4	17

Summary data: Starting at Level 2 with GCSE Grade B

	Max	Min	Mean	Mode	0-20	21-40	41-60	61-80	81+	TOTAL
TOTAL	82	6	34	29	8	20	8	2	1	39
Q ASKED	34	14	26	31	7	32				39
Q CORRECT	15	1	7	6						
NUMBERS	100	7	48	33	6	11	11	9	2	39
POWERS	67	0	28	0	16	13	0	10	0	39
BASIC ALGEBRA	82	0	23	18	26	8	2	2	1	39
STATS	67	0	14	0	25	12	0	2	0	39
MISC	100	0	28	0	23	0	10	0	6	39

Summary Data: Starting at Level 3

	Max	Min	Mean	Mode	0-20	21-40	41-60	61-80	81+	TOTAL
TOTAL	97	62	79	68	0	0	0	8	7	15
Q ASKED	26	17	20	18	8	7				15
Q CORRECT	17	10	14	13	15	0				
NUMBERS	100	67	83	93	0	0	0	7	8	15
POWERS	100	33	69	67	0	3	0	8	4	15
BASIC ALGEBRA	100	73	87	82	0	0	0	2	13	15
STATS	100	0	44	33	3	7	0	2	3	15
MISC	100	0	73	100	3	0	2	0	10	15

Appendix 16 Pilot Study Treefrog results

16a Individual Participants per Test Overview

TEST1.TST	engpmurp	essaande	essaben1	essabren	essacoon	essahols	essavera	essaweek	essawri1	esscbott	esscbren	essccoop
Achieved	7	7	3	6	5	6	7	7	7	7	7	7
Passed	0	0	0	1	2	1	0	0	0	0	0	0
Wrong	3	7	22	2	9	5	5	3	0	0	6	6
Exit	0	0	1	0	0	0	0	0	0	0	0	0
TEST2.TST	engpmurp	essaande	[redacted]	essabren	essacoon	essahols	essavera	essaweek	essawri1	esscbott	esscbren	essccoop
Achieved	0	8	0	6	0	8	8	8	8	7	8	7
Passed	7	0	0	2	8	0	0	0	0	1	0	1
Wrong	3	0	0	8	7	0	5	0	1	7	1	2
Exit	0	0	0	0	0	0	0	0	0	0	0	0
TEST3.TST	engpmurp	essaande	[redacted]	essabren	essacoon	essahols	essavera	essaweek	essawri1	esscbott	esscbren	essccoop
Achieved	3	11	0	3	2	5	4	0	9	9	5	10
Passed	9	1	0	8	10	7	7	0	3	3	7	2
Wrong	2	2	0	4	22	1	5	0	8	11	6	13
Exit	0	0	0	1	0	0	2	1	0	0	0	1
TEST4.TST												
Achieved	0	6	0	3	6	6	6	6	6	0	3	6
Passed	0	0	0	3	0	0	0	0	0	0	3	0
Wrong	0	0	0	11	4	0	0	0	0	0	11	4
Exit	0	0	0	0	0	0	0	0	0	0	0	0
TEST5.TST	engpmurp	essaande	[redacted]	essabren	essacoon	essahols	essavera	essaweek	essawri1	esscbott	esscbren	essccoop
Achieved	0	4	0	1	1	2	3	3	5	3	4	3
Passed	1	4	0	7	7	6	5	4	3	5	4	5
Wrong	1	5	0	6	0	1	7	18	2	3	3	4
Exit	1	0	0	0	0	0	0	0	0	0	0	0
TEST6.TST	[redacted]	essaande	[redacted]	essabren	essacoon	[redacted]	essavera	essaweek	essawri1	esscbott	esscbren	essccoop
Achieved	0	5	0	2	2	0	2	2	1	2	2	2
Passed	0	0	0	3	3	0	3	3	0	3	3	3
Wrong	0	0	0	3	3	0	3	3	0	12	9	6
Exit	0	0	0	0	0	0	0	0	1	0	0	0
TOTAL of 46	10	41	3	21	16	27	30	26	36	28	29	35
high or very low		HIGH	low						HIGH			
6 to 16	6 to 16				6 to 16							
17 to 26				16 to 26				16 to 26				
27 to 35						27 to35	27 to35			27 to35	27 to35	27 to35

TEST1	esscdunk	esscelli	esscjohn	esscmalk	esscmcqu	esscwal2	essdcoom	essdgibs	essecosg	esseearl	essewil3	essfchap	essfgart
Achieved	6	6	5	7	3	7	7	7	4	3	5	3	0
Passed	1	1	2	0	4	0	0	0	3	4	2	4	7
Wrong	9	24	14	5	16	14	13	20	9	12	2	28	1
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST2.	esscdunk	esscelli	esscjohn	esscmalk	esscmcqu	esscwal2	essdcoom	essdgibs	essecosg	esseearl	essewil3	essfchap	essfgart
Achieved	8	1	8	8	4	5	6	6	5	6	8	0	0
Passed	0	7	0	0	4	3	2	2	3	2	0	8	8
Wrong	1	13	5	0	8	11	9	27	17	10	4	3	3
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST3.	esscdunk	esscelli	esscjohn	esscmalk	esscmcqu	esscwal2	essdcoom	essdgibs	essecosg	esseearl	essewil3	essfchap	essfgart
Achieved	5	5	2	11	2	2	2	5	1	3	1	5	3
Passed	7	7	10	1	10	8	9	7	10	9	10	7	9
Wrong	1	15	17	4	12	15	5	7	15	6	4	9	13
Exit	0	0	0	1	4	3	1	3	1	0	1	2	0
TEST4.	esscdunk	esscelli	[redacted]	esscmalk	esscmcqu	esscwal2	[redacted]	essdgibs	essecosg	esseearl	essewil3	essfchap	essfgart
Achieved	6	5	0	6	5	5	0	2	6	6	6	6	6
Passed	0	1	0	0	1	1	0	4	0	0	0	0	0
Wrong	0	3	0	0	6	7	0	1	6	2	0	1	1
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST5	esscdunk	esscelli	esscjohn	esscmalk	esscmcqu	esscwal2	[redacted]	essdgibs	essecosg	esseearl	essewil3	essfchap	essfgart
Achieved	2	2	3	2	0	2	0	0	2	1	2	1	3
Passed	6	6	5	6	8	6	0	8	6	7	6	7	5
Wrong	1	7	6	8	14	8	0	0	7	2	2	12	4
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST6	esscdunk	esscelli	esscjohn	esscmalk	esscmcqu	esscwal2	[redacted]	essdgibs	essecosg	esseearl	essewil3	essfchap	essfgart
Achieved	2	2	2	2	2	0	0	2	1	1	0	2	4
Passed	3	3	3	3	3	5	0	3	4	4	5	3	1
Wrong	0	1	13	1	11	0	0	1	13	6	1	4	8
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TOTAL of 46	29	21	20	36	16	21	15	22	19	20	22	17	16
high or very low				HIGH									
6 to 16					6 to 16		6 to 16						6 to 16
17 to 26		16 to 26	16 to 26			16 to 26		16 to 26	16 to 26	16 to 26	16 to 26	16 to 26	
27 to 35	27 to35												

TEST1	essgbasf	essgjohn	essgocon	esshbier	esshcar1	esshnels	esshpeve	esshwood	essjdean	essjgre2	essjhorn	essjhoug	essjpric
Achieved	7	6	7	7	0	7	6	6	6	6	0	7	7
Passed	0	0	0	0	3	0	1	1	0	1	7	0	0
Wrong	19	20	4	11	3	1	2	38	29	7	2	2	6
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST2	essgbasf	essgjohn	essgocon	esshbier	esshcar1	esshnels	esshpeve	esshwood	essjdean	essjgre2	essjhorn	essjhoug	essjpric
Achieved	8	8	8	8	6	8	6	2	7	8	0	7	7
Passed	0	0	0	0	2	0	2	6	1	0	8	1	1
Wrong	0	2	3	0	9	2	1	29	5	0	10	9	2
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST3	essgbasf	essgjohn	essgocon	esshbier	esshcar1	esshnels	esshpeve	esshwood	essjdean	essjgre2	essjhorn	essjhoug	essjpric
Achieved	10	9	5	11	3	9	2	0	4	9	0	2	3
Passed	2	3	1	1	3	3	9	11	8	3	11	9	9
Wrong	11	25	1	17	15	8	6	8	17	5	1	5	9
Exit	2	0	0	0	3	0	1	0	0	0	0	1	0
TEST4	essgbasf	essgjohn	essgocon	esshbier	esshcar1	esshnels	esshpeve	esshwood	essjdean	essjgre2	essjhorn	essjhoug	essjpric
Achieved	6	6	6	6	5	6	6	5	6	3	6	6	5
Passed	0	0	0	0	1	0	0	1	0	0	0	0	1
Wrong	2	0	1	5	5	0	2	6	1	0	1	2	3
Exit	0	0	0	0	0	0	0	0	0	1	0	0	0
TEST5	essgbasf	essgjohn	essgocon	esshbier	esshcar1	esshnels	esshpeve	esshwood	essjdean	essjgre2	essjhorn	essjhoug	Essjpric
Achieved	4	8	4	5	0	6	2	2	1	7	1	2	0
Passed	4	0	3	3	2	1	6	6	7	1	7	6	0
Wrong	19	33	15	26	2	6	1	10	3	2	3	0	0
Exit	0	0	0	0	1	0	0	0	0	0	0	0	0
TEST6	essgbasf	essgjohn	essgocon	esshbier	esshcar1	esshnels	esshpeve	esshwood	essjdean	essjgre2	essjhorn	essjhoug	essjpric
Achieved	2	5	3	2	0	5	2	1	3	3	2	2	0
Passed	3	0	2	3	0	0	3	4	2	2	3	3	0
Wrong	13	6	4	6	0	0	1	8	13	1	3	2	0
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TOTAL of 46	37	42	33	39	14	41	24	16	27	36	9	26	22
high or very low	HIGH	HIGH		HIGH		HIGH				HIGH			
6 to 16					6 to 16			6 to 16			6 to 16		
17 to 26							16 to 26					16 to 26	16 to 26
27 to 35			27 to35						27 to35				

TEST1	essjtayl	esswil2	esskbeva	esskchas	esskcoon	esskcosg	esskcrai	esskgali	esskheav	esskqui1	esslhugh	esslhull	esslhyde
Achieved	5	7	6	3	6	0	6	7	3	7	7	7	6
Passed	2	0	1	0	1	7	0	0	4	0	0	0	1
Wrong	2	9	5	0	9	4	20	8	9	0	5	4	6
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST2	essjtayl	esswil2	esskbeva	esskchas	esskcoon	esskcosg	esskcrai	esskgali	esskheav	esskqui1	esslhugh	esslhull	esslhyde
Achieved	8	8	7	4	0	1	8	8	0	8	2	8	8
Passed	0	0	1	4	8	7	0	0	8	0	6	0	0
Wrong	3	1	5	10	0	0	7	8	16	1	13	3	0
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST3	essjtayl	esswil2	esskbeva	esskchas	esskcoon	esskcosg	esskcrai	esskgali	esskheav	esskqui1	esslhugh	esslhull	esslhyde
Achieved	11	5	9	3	3	0	7	3	3	7	3	7	5
Passed	1	7	3	9	8	11	5	1	9	5	8	5	7
Wrong	7	27	5	18	1	1	16	5	4	3	15	7	3
Exit	0	0	0	0	0	0	1	1	0	0	1	0	0
TEST4	essjtayl	esswil2	esskbeva	esskchas	esskcoon	esskcosg	esskcrai	esskgali	esskheav	esskqui1	esslhugh	esslhull	esslhyde
Achieved	6	2	6	5	4	0	6	6	6	6	5	6	6
Passed	0	4	0	1	2	6	0	0	0	0	1	0	0
Wrong	0	8	0	3	3	0	0	2	0	1	7	2	0
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST5	essjtayl	esswil2	esskbeva	esskchas	esskcoon	esskcosg	esskcrai	esskgali	esskheav	esskqui1	esslhugh	esslhull	esslhyde
Achieved	7	2	3	2	1	1	0	4	3	2	2	2	2
Passed	1	6	5	6	7	7	8	4	5	6	6	6	6
Wrong	5	6	9	19	2	0	5	7	4	2	0	3	0
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TEST6	essjtayl	essjwil2	esskbeva	esskchas	esskcoon	esskcosg	esskcrai	esskgali	esskheav	esskqui1	esslhugh	esslhull	esslhyde
Achieved	4	0	2	0	2	1	0	2	0	1	0	2	0
Passed	1	0	3	5	3	4	0	3	5	4	5	3	0
Wrong	3	0	2	7	1	0	0	5	6	5	7	2	0
Exit	0	0	0	0	0	0	0	0	0	0	0	0	0
TOTAL of 46	41	24	33	17	16	3	27	30	15	31	19	32	27
high or very low	HIGH					low							
6 to 16					6 to 16				6 to 16				
17 to 26		16 to 26		16 to 26							16 to 26		
27 to 35			27 to35				27 to35	27 to35		27 to35		27 to35	27 to35

TEST1	esslmull	essloliv	essmdenn	essmpar1	essmpimb	essmwil1	essnevan	esspbode	essphilt	esspmitc	essrcull
Achieved	7	1	7	6	7	1	6	0	7	7	7
Passed	0	0	0	1	0	0	1	0	0	0	0
Wrong	0	22	4	1	6	1	20	0	4	2	3
Exit	0	1	0	0	0	1	0	7	0	0	0
TEST2	esslmull	essloliv	essmdenn	essmpar1	essmpimb	essmwil1	essnevan	esspbode	essphilt	esspmitc	essrcull
Achieved	7	1	7	7	8	4	7	1	8	8	8
Passed	1	7	1	1	0	4	1	3	0	0	0
Wrong	1	8	7	0	5	12	27	16	14	3	5
Exit	0	0	0	0	0	0	0	1	0	0	0
TEST3	esslmull	essloliv	essmdenn	essmpar1	essmpimb	essmwil1	essnevan	esspbode	essphilt	esspmitc	essrcull
Achieved	8	0	3	4	4	1	5	0	4	6	9
Passed	4	7	9	8	8	2	7	0	8	6	3
Wrong	12	10	9	3	9	8	5	0	6	9	3
Exit	0	2	0	0	0	2	1	0	0	0	0
TEST4	esslmull	essloliv	essmdenn	essmpar1	essmpimb		essnevan	esspbode	essphilt	esspmitc	essrcull
Achieved	5	4	0	6	6	0	6	6	6	6	6
Passed	0	2	0	0	0	0	0	0	0	0	0
Wrong	0	6	0	1	2	0	2	2	0	0	0
Exit	1	0	0	0	0	0	0	0	0	0	0
TEST5	esslmull	essloliv	essmdenn	essmpar1	essmpimb	essmwil1	essnevan	esspbode	essphilt	esspmitc	essrcull
Achieved	5	1	1	3	2	0	1	0	2	3	7
Passed	3	7	7	5	6	0	7	0	6	5	1
Wrong	6	2	0	0	0	0	8	0	2	4	7
Exit	0	0	0	0	0	0	0	0	0	0	0
TEST6	esslmull	essloliv	essmdenn	essmpar1	essmpimb	essmwil1	essnevan	esspbode	essphilt	esspmitc	essrcull
Achieved	0	0	0	0	0	0	0	0	0	0	0
Passed	0	0	0	0	0	0	0	0	0	0	0
Wrong	0	0	0	0	0	0	0	0	0	0	0
Exit	0	0	0	0	0	0	0	0	0	0	0
TOTAL of 46	32	7	18	26	27	6	25	7	27	30	37
high or very low											
6 to 16		6 to 16									HIGH
17 to 26			16 to 26					6 to 16			
27 to 35	27 to35			16 to 26			16 to 26				
					27 to35				27 to35	27 to35	

TEST1	essrmclo	essrmiln	essrmoff	essrthom	essshick	esssmars	esssmci1	esssmill	essswils	essvmcna	esszsyke
Achieved	7	5	5	7	7	5	6	2	7	7	7
Passed	0	2	2	0	0	2	1	5	0	0	0
Wrong	9	18	23	2	18	6	6	1	10	14	9
Exit	0	0	0	0	0	0	0	0	0	0	0
TEST2	essrmclo	essrmiln	essrmoff	essrthom	essshick	esssmars	esssmci1	esssmill	essswils	essvmcna	esszsyke
Achieved	8	8	2	8	8	0	8	6	8	7	8
Passed	0	0	6	0	0	8	0	2	0	1	0
Wrong	1	0	18	2	0	5	0	5	5	5	3
Exit	0	0	0	0	0	0	0	0	0	0	0
TEST3	essrmclo	essrmiln	essrmoff	essrthom	essshick	esssmars	esssmci1	esssmill	essswils	essvmcna	esszsyke
Achieved	11	5	2	9	11	2	5	1	11	1	4
Passed	1	10	9	3	1	10	7	11	1	10	8
Wrong	5	4	2	5	3	26	3	21	9	3	9
Exit	0	1	1	0	0	2	0	0	0	0	0
TEST4	essrmclo		essrmoff	essrthom	essshick	esssmars	esssmci1	esssmill	essswils	essvmcna	esszsyke
Achieved	6	0	4	6	6	6	6	6	6	3	6
Passed	0	0	2	0	0	0	0	0	0	3	0
Wrong	0	0	3	1	0	3	0	2	0	6	1
Exit	0	0	0	0	0	0	0	0	0	0	0
TEST5	essrmclo	essrmiln	essrmoff	essrthom	essshick	esssmars	esssmci1	esssmill	essswils	essvmcna	esszsyke
Achieved	8	5	2	3	8	1	4	3	8	0	1
Passed	0	3	6	5	0	7	4	5	0	8	7
Wrong	9	3	1	1	7	5	7	13	3	0	3
Exit	0	0	0	0	0	0	0	0	0	0	0
TEST6	essrmclo	essrmiln	essrmoff	essrthom	essshick	esssmars	esssmci1	esssmill	essswils	essvmcna	esszsyke
Achieved	0	0	0	0	0	0	0	0	0	0	0
Passed	0	0	0	0	0	0	0	0	0	0	0
Wrong	0	0	0	0	0	0	0	0	0	0	0
Exit	0	0	0	0	0	0	0	0	0	0	0
TOTAL of 46	40	23	15	33	40	14	29	18	40	18	26
high or very low											
6 to 16	HIGH				HIGH				HIGH		
17 to 26			6 to 16			6 to 16					
27 to 35		16 to 26						16 to 26		16 to 26	16 to 26
				27 to35			27 to35				

16b Pilot Study Treefrog Results per Test September 2001

TEST 1

	total	%		Reason	Source
Q1			90 – x = 70	Worked example	
Achieved	69	87			
Passed	6	8			
Exit	4	5			
Num of students	79				
Q2			58-x=17	Repetition	
Achieved	62	78			
Passed	10	13			
Exit	7	9			
Num of students	79				
Q3			962-x=476	QCA question	QCA 9
Achieved	58	73	At level 3 **most** are unable to complete equations of the form 962-...= 476 (level 3 is very low would expect to be attainable)		
Passed	14	18			
Exit	7	9			
Num of students	79				
Q4			9-z=17	double -ve	QCA 12
Achieved	55	70	Pupils at all levels need help understanding negative quantities in algebra		
Passed	14	18			
Exit	10	13			
Num of students	79				
Q5			7-z=9	reverse of q1	QCA 12
Achieved	62	78	Pupils at all levels need help understanding negative quantities in algebra		
Passed	7	9			
Exit	10	13			
Num of students	79				
Q6			-9-z=-7	negative signs	QCA 12
Achieved	60	76	Pupils at all levels need help understanding negative quantities in algebra		
Passed	9	11			
Exit	10	13			
Num of students	79				
Q7			476-z=962	negative signs	QCA 12
Achieved	52	66	Pupils at all levels need help understanding negative quantities in algebra		
Passed	16	20			
Exit	11	14			
Num of students	79				
Total for TEST1					
Achieved	76				
Passed	14				
Exit	11				
Ach, Pass or Exit	100				
Total number	79				

TEST 2

	TOTAL	%			
Q1			x/8=12	Worked example	QCA 4
Achieved	56	75	**At level 3: most** are unable to complete equations of the form ... / 24 = 16		
Passed	18	24			
Exit	1	1			
Num of students	75	100			
Q2			x/5=25	Repeat eg	QCA 4
Achieved	62	83	**At level 3: most** are unable to complete equations of the form ... / 24 = 16		
Passed	12	16			
Exit	1	1			
Num of students	75	100			
Q3			x/7=56	Repeat eg	QCA 4
Achieved	58	77	**At level 3: most** are unable to complete equations of the form ... / 24 = 16		
Passed	16	21			
Exit	1	1			
Num of students	75	100			
Q4			x/15=22	Repeat eg	QCA 4
Achieved	58	77			
Passed	16	21			
Exit	1	1	**At level 3: most** are unable to complete equations of		
Num of students	75	100	the form ... / 24 = 16		
Q5			-x/4=22	negative sign	QCA 12
Achieved	49	65			
Passed	24	32			
Exit	2	3	Pupils at all levels **need help understanding**		
Num of students	75	100	negative quantities in algebra		
Q6			x/6=-18	negative sign	QCA 12
Achieved	48	64			
Passed	25	33			
Exit	2	3	Pupils at all levels **need help understanding**		
Num of students	75	100	negative quantities in algebra		
Q7			x/3=-15	negative sign	QCA 12
Achieved	56	75			
Passed	17	23			
Exit	2	3	Pupils at all levels **need help understanding**		
Num of students	75	100	negative quantities in algebra		
Q8			5x/3=15	negative sign	QCA 12
Achieved	50	67			
Passed	22	29			
Exit	3	4	Pupils at all levels **need help understanding**		
Num of students	75	100	negative quantities in algebra		

Total for Test 2		
Achieved	73	
Passed	25	
Exit	2	
Ach, Pass or Exit	100	
Num of students	75	

TEST3

Q1	Num	%	Find 5x when 4x=20	Worked example	QCA 1 (confirmed) & 19 (refuted)
Achieved	61	84	Equations which involve only one variable (3k=18, 5k=?) are **not problematic** for the **majority** of pupils at level 6 and higher Equations which involve only one variable (3k=18, 5k=?) are **problematic** for the **majority** of pupils at level 4 & 5		
Passed	11	15			
Wrong					
Exit	1	1			
Total	73	100			
Q2			26-2n=8	Apply worked example	QCA 16 (unsure) & 20 (refuted)
Achieved	55	75	A value n being deduced from an equation of the form 26-2n=8 is problematic to solve for pupils at level 6 and lower A value n being deduced from an equation of the form 26-2n=8 is **extremely difficult** to solve for pupils at level 6 and lower		
Passed	16	22			
Exit	2	3			
Total	73	100			
Q3			2(x+3)	expanding brackets	Matz 22
Achieved	55	75			
Passed	17	23			
Exit	1	1			
Total	73	100	2x+3 expected		
Q4			-(3x-w)	double -ve & expanding brackets	Matz 23
Achieved	50	68			
Passed	21	29			
Exit	2	3			
Total	73	100	3x-w expected		
Q5			3x+2y=7 find a value for 6x+4y	comparison +ve only	QCA 15 (unsure)
Achieved	52	71	At lower levels the comparison of two algebraic expressions (3x+2y=7, 6x+4y=?) is significantly more problematic than that between two equations which involve two numerical values only (4k+2a=82, ?=41		
Passed	18	25			
Exit	3	4			
Total	73	100			
Q6			4-2x=10-6x	transpose with -ve	QCA 14 (unsure)
Achieved	31	42			
Passed	39	53			
Exit	3	4			
Total	73	100	At level 5 most pupils are unable to find y from		
Q7			3+21(s+4)	brackets	Matz 8
Achieved	23	32			
Passed	46	63			
Exit	4	5			
Total	73	100	24(s+4) expected		
Q8			3+23(s-4)	"-ve operator with brackets	Matz 8

Achieved	24	33			
Passed	44	60			
Exit	5	7			
Total	73	100	26(s-4) expected		
Q9			17-21(s+7)	"-ve sign	Matz 8
Achieved	19	26			
Passed	48	66			
Exit	6	8			
Total	73	100	"-4(s+7) expected		
Q10			13-19(s-3)	double -ve	Matz 8
Achieved	20	27	"-6(s-3) expected		
Passed	47	64			
Exit	6	8			
Total	73	100			
Q11			3b + 1 = b + ...	rewriting an equivalent expression	
Achieved	15	21			
Passed	52	71			
Exit	6	8			
Total	73	100			
Q 12			4c+8=5c+ 10+ ...	rewriting an equivalent expression including negative quantities	QCA 6 (confirmed)
Achieved	13	18			
Passed	54	74	At level 6 in re-writing an expression **most** are **successful** if positive terms only such as 3b+1 = b+ . and 5a-4 = 2a+2 +		
Exit	6	8	Finding is opposite to assumption		
Total	73	100	But formatting may be misleading		

Total for TEST3

	%
Achieved	48
Passed	47
Exit	5
Ach, Pass or Exit	100
Total number	73

TEST4

	Total	%		Reason	Source
question 1 out of 6			2x when x=9	worked example	Matz 2
Achieved	64	96	2+9 expected		
Passed	1	1			
Wrong					
Exit	2	2			
Number of students	67	100			
Question 2			4x when x=6	Repetition of example	Matz 2
Achieved	65	97	4x=46x expected		
Passed	1	1			
Wrong					
Exit	1	1			
Number of students	67	100			
Question 3			2x+3 whenx=5	Repetition of example	Matz 22
Achieved	63	94	2x+3 expected		
Passed	3	4			
Exit	1	1			
Number of students	67	100			
Question 4			-3y+8 when y=2		Matz 3 (but easier)
Achieved	60	90	"-ve coefficient get harder 6+8 expected		
Passed	6	9			
Exit	1	1			
Number of students	67	100			
Question 5			-3z-8 when z=-3		Matz 3
Achieved	54	81	more complex two -ve signs "-9-8 expected		
Passed	12	18			
Exit	1	1			
Number of students	67	100			
Question 6			xy+1 when x=-4 and y=6		Matz 3 including -ve
Achieved	52	78	more complex two variables 24+1 expected		
Passed	14	21			
Exit	1	1			
Number of students	67	100			

Total means for test4	
Achieved	89
Passed	9
Exit	2
Ach, Pass or Exit	100
Number of students	67

TEST5

	Total	%		Reason	Source
Q1			expand A(BC)	worked example	Matz 12
Achieved	55	79	Ab*BC expected		
Passed	15	21			
Exit	0	0			
Total	70	100			
Q2			(ax + b) (cx + d)	Brackets	Matz 24
Achieved	23	33	Expanding brackets		
Passed	46	66	ACX2+BD expected		
Exit	1	1			
Total	70	100			
Q3			(y + 2)/(y - 3) = 2		Matz 14 ??
Achieved	17	24	Fractions and transposition		
Passed	51	73			
Exit	2	3			
Total	70	100			
Q4			(z+1)/(z+4)=5/6		Matz 14 ???
Achieved	10	14	Fractions and transposition		
Passed	58	83			
Exit	2	3			
Total	70	100			
Q5			2z+5=11		Matz 26
Achieved	46	66	Transposition with coefficient		
Passed	22	31	z+5=11/2 expected		
Exit	2	3			
Total	70	100			
Q6			3z+5=z+3		Matz 27
Achieved	29	41	transposition with coefficient and constants		
Passed	39	56	z+5=z expected		
Exit	2	3			
Total	70	100			
Q7			(x-5)(x-7)=3		Matz 31
Achieved	9	13	expanding and factorising		
Passed	59	84	X-5=3 OR X-7=3, X=8 OR X=10		
Exit	2	3			
Total	70	100			
Q8			$x^2 + 5/6\,x + 1/6 = 0$		Matz 30
Achieved	4	6	Fractions as coefficients		
Passed	64	91	X(X+5/6) + 1/6		
Exit	2	3			
Total	70	100			

TOTAL %

Achieved	34
Passed	63
Exit	2
Ach, Passed or Exit	100
Number of students	70

TEST 6

	Total	%			
Q1			2^3	Worked example	
Achieved	57	88	indices 6 expected for error		
Passed	7	11			
Exit	1	2			
	65	100			
Q2			-1^3	Negative value	Matz 5
Achieved	50	77	negative reciprocal with index "-3 expected		
Passed	14	22			
Exit	1	2			
	65	100			
Q3			2^-2	Negative indice	Matz 7
Achieved	7	11	Negative indices "-4 expected Formatting problems with ^		
Passed	56	86			
Exit	2	3			
	65	100			
Q4			$(A+B)^2$	Brackets	Matz 11
Achieved	13	20	Indices on brackets A^2+B^2 expected		
Passed	51	78			
Exit	1	2			
	65	100			
Q5			(A+B)^-2		Matz11
Achieved	4	6	Negative indices and brackets probably A^-2+B^-2 expected		
Passed	59	91			
Exit	2	3			
	65	100			

Totals for test 6	
Achieved	40
Passed	58
Exit	2
Ach, Passed or Exit	100
Num of students	65
Non participants	10

16c Ranking of Questions by Achievement

	Number	Question	Achieved	% Passed	Exit	No.	Reason	( )	"-ve	Transp	Weaknesses	Reason for inclusion
4	2	4x when x=6	97	1	1	67	Matz 2					4x=46x expected
4	1	2x when x=9	96	1	3	67	Matz 2					6+8 expected
4	3	2x+3 whenx=5	94	4	1	67	Matz 22				reverse of q1	Equations which involve only one variable (3k=18, 5k=?) are **not problematic** for the **majority** of pupils at level 6 and higher
4	4	"-3y+8 when y=2	90	9	1	67	Matz 3 (but easier)		y		double -ve	At lower levels the comparison of two algebraic expressions (3x+2y=7, 6x+4y=?) is significantly more
6	1	2^3	88	11	2	65	Worked example				negative signs	X(X+5/6) + 1/6
1	1	90 – x = 70	87	8	5	79	Worked example		y	y	negative signs	probably A^-2+B^-2 expected
3	1	Find 5x when 4x=20	84	15	1	73	QCA 1 (confirmed) & 19 (refuted)				"-ve sign	A^2+B^2 expected
2	2	x/5=25	83	16	1	75	QCA 4	y			Repeat eg	"-4(s+7) expected
4	5	"-3z-8 when z=-3	81	18	1	67	Matz 3		y	y	Repeat eg	Pupils at all levels need help understanding negative quantities in algebra
5	1	expand A(BC)	79	21	0	70	Matz 12				negative sign	At level 6 in re-writing an expression **most** are **successful** if positive terms only such as 3b+1 = b+ . and 5a-4 = 2a+2 +
1	5	7-z=9	78	9	13	79	QCA 12		y	y	Repeat eg	
1	2	58-x=17	78	13	9	79	Repetition		y		negative sign	"-4 expected
4	6	xy+1 when x=-4 and y=6	78	21	1	67	Matz 3 including -ve		y		negative sign	
2	3	x/7=56	77	21	1	75	QCA 4	y			Repetition of example	26(s-4) expected
2	4	x/15=22	77	21	1	75	QCA 4				negative sign	26(s+4) expected
6	2	-1^3	77	22	2	65	Matz 5		y	y	worked example	z+5=z expected
1	6	-9-z=-7	76	11	13	79	QCA 12		y	y	"-ve coefficient	ACX2+BD expected
3	2	26-2n=8	75	22	3	73	QCA 16 (unsure) & 20 (refuted)		y	y	Apply worked example	**At level 3: most** are unable to complete equations of the form ... / 24 = 16
3	3	2(x+3)	75	23	1	73	Matz 22			y	Worked example	**At level 3: most** are unable to complete equations of the form ... / 24 = 16
2	7	x/3=-15	75	23	3	75	QCA 12			y	two -ve signs	Pupils at all levels **need help understanding** negative

	Number	Question	Achieved	% Passed	Exit	No.	Reason	( )	"-ve	Transp	Weaknesses	Reason for inclusion
												quantities in algebra
2	1	x/8=12	75	24	1	75	QCA 4				two variables	Pupils at all levels need help understanding negative quantities in algebra
1	3	962-x=476	73	18	9	79	QCA 9		y	y	expanding brackets	At level 3 **most** are unable to complete equations of the form 962-...= 476
3	5	If 3x+2y=7 find a value for 6x+4y	71	25	4	73	QCA 15 (unsure)			y	comparison +ve only	
1	4	9-z=17	70	18	13	79	QCA 12		y	y	transpose with -ve	
3	4	-(3x-w)	68	29	3	73	Matz 23		y	y	"-ve operator with brackets	X-5=3 OR X-7=3, X=8 OR X=10
2	8	5x/3=15	67	29	4	75	QCA 12			y	brackets	6 expected for error
1	7	476-z=962	66	20	14	79	QCA 12		y	y	"-ve sign	
5	5	2z+5=11	66	31	3	70	Matz 26	y			Expanding brackets	A value n being deduced from an equation of the form 26-2n=8 is problematic to solve for pupils at level 6 and lower
2	5	"-x/4=22	65	32	3	75	QCA 12		y	y	double -ve	z+5=11/2 expected
2	6	x/6=-18	64	33	3	75	QCA 12		y	y	rewriting an equivalent expression	Ab*BC expected
3	6	4-2x=10-6x	42	53	4	73	QCA 14 (unsure)	y	y		worked example	Pupils at all levels **need help understanding** negative quantities in algebra
5	6	3z+5=z+3	41	56	3	70	Matz 27	y	y		Transposition with coefficient	2+9 expected
3	8	3+23(s-4)	33	60	7	73	Matz 8	y	y	y	transposition with coefficient and constants	**At level 3: most** are unable to complete equations of the form ... / 24 = 16
5	2	(ax + b) (cx + d)	33	66	1	70	Matz 24	y			Fractions and transposition	"-9-8 expected
3	7	3+21(s+4)	32	63	5	73	Matz 8				expanding and factorising	
3	10	13-19(s-3)	27	64	8	73	Matz 8		y		indices	"-6(s-3) expected
3	9	17-21(s+7)	26	66	8	73	Matz 8		yy		Worked example	Pupils at all levels **need help understanding** negative quantities in algebra
5	3	(y + 2)/(y - 3) = 2	24	73	3	70	Matz 14 ??	y	yy		double -ve & expanding brackets	Pupils at all levels **need help understanding** negative quantities in algebra
3	11	3b + 1 = b + ...	21	71	8	73	QCA 6 (confirmed)	y			Fractions as coefficients	2x+3 expected

	Number	Question	Achieved	% Passed	Exit	No.	Reason	( )	"-ve	Transp	**Weaknesses**	**Reason for inclusion**
6	4	$(A+B)^2$	20	78	2	65	Matz 11	y		y	negative reciprocal with index	Equations which involve only one variable (3k=18, 5k=?) are **not problematic** for the **majority** of pupils at level 6 and higher
3	12	4c+8 = 5c + 10 + ...	18	74	8	73	QCA 6 (confirmed)	y	yy		Fractions and transposition	At level 5 most pupils are unable to find y from
5	4	(z+1)/(z+4)=5/6	14	83	3	70	Matz 14 ???	y		y	Negative indices	
5	7	(x-5)(x-7)=3	13	84	3	70	Matz 31				Indices on brackets	24+1 expected
6	3	2^-2	11	86	3	65	Matz 7	y			Negative indices and brackets	Pupils at all levels need help understanding negative quantities in algebra
6	5	(A+B)^-2	6	91	3	65	Matz11					3x-w expected
5	8	$x^2 + 5/6\,x + 1/6 = 0$	6	91	3	70	Matz 30				QCA question	Pupils at all levels need help understanding negative quantities in algebra

Appendix 17: Software Evaluation Questionnaire Results

17a Treefrog Individual participant codes awarded

Participant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Route															dt							
Gender		f		f	f	m	f	f	m	f	f		f	m	f	f	f	f	m		f	f
Highest Maths Qualification	g				g	g	g	g	g		AS	g	g	A	g	g	A	g	g		g	g
Maths grade	c		c	c	c	c	b	c	c	c	u		b	e	c	c	n	c	c		b	c
Year maths studied	94		99	99	99		0	99		99	99	95	99	1	0	99	1	98	98		99	92
Age group		2	2	2	2	2		2	2	2	2	2	2	2	2	2	2	2	2		2	3
IT experience		poor		ibt2	av	av	av		nil	av	A		yr 9	g	av	g	yr 9	g	poor		g	
Rate the usefulness of the software to help you to correct errors in your understanding	2	1	1	2	2	1	2	2	2	1	2	1	3	1	2	3	2	2	2	2	3	2
How well did the software cover the subject area you need to study?	1	2	1	2	3	1	3	2	2	1	1	2	3		1	1	2	2	1	2	2	3
How much did the software accept your method for solving a given question	1	2	2	1	2	2	2	2	2	2	2	1	2	2	2	2	1	2	1	1	2	3
Rate the balance between state of the art graphics features and benefits for learning	1	2	3	2	2	1	2	2	2	1	3	2	2	2	1	2	2	2	3	2	2	3
Do you think the software would help you to carry on doing work on your own?	2	2	3	2	1	1	2	3	3	2	2	1	2	3	1	3	1	2	2	2	3	2
Were mathematical terms used the way you expected?	2	2	2	2	1	1	3	2	2	2	3	1	2	1	2	2	2	2	2	2	1	3
Rate the appropriateness of the representations or presentation of different types of question	1	2	2	1	2	1	2	1	2	2	2	2	2	2	1	2	1	2	3	2	1	2
Did you find the graphics generally helped you to learn or were they poor in that they distracted you?	2	3	2	2	2	2	2	2	2	1	3	2	1	2	1	2	2	2	2	3	1	1
How would you rate the methods of interaction in terms of ease of use, familiarity and functionality?	2	2	1	1	2	1	2	1	2	1	2	2	2	1	2	2	1	2	2	3	2	2

Participant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Rate the balance between helpful informative interaction and feedback which intrudes and interrupts progress.	2	2	2	1	2	2	2	3	3	2	1	2	3		1	3	1	3	2	3	2	2
To what extent does the environment support movement within questions and from one question to another to help learning?	3	2	2	1		2		2	2	2	3	2	2		1	2	2	1	1	2	2	2
To what extent does the structure of the environment support the notion of self-directed learning?	1	2	1	1	2	2	2	2	2	2	2	1	2	1	1	1	1	2	2	2	2	3
How easy was it to develop an understanding of how to use the system and hence develop the skills to maximise full use the system?	1	1	1	1	2	1	2	1	2	2	3	1	2	1	1	1	2	1	3	2	2	2
To what extent did you recognise the mathematical symbols, icons and buttons used by the software?	2	2	1	2	2	2	2	2	2	2	2	2	1	1	2	1	1	3	1	3	1	2
How would you rate the system's ability to shielding you from usability obstacles such as runtime errors and malfunctions?	1	1	1	2	2	2	2	2	2	2	2	3	2		1	2	2	2	2	1	2	
How would you rate the system's ability to protect you from making annoying errors such as incorrect syntax of statements?	2	1	2	2	2	1	3	2	2	1	2	2	2		2	1	2	2	3	2	1	
To what extent did your methods for solving problems and those used by the system to present questions and feedback match?	2	2	2	2	2	1	2	3	3	1	2	2	2	1	2	2	2	2	3	2	2	2
To what extent does the software enable you to experiment with solutions and develop your own personal understanding?	3	1	2	1	2	1	2	2	2	1	2	2	1	3	2	2	2	2	3	2	2	3

Participant	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
Route					dt		sc	sc	dt		dt										pe		
Gender	f	m	f	m	m	m	f	f	f	f	f	m	m	m	m			f	m	f	f	f	f
Highest maths qual	onc	g	g	g	g	g	g	g	g	g	g	a	g	g	gnvq		g	g	g		g	g	g
Maths grade	pass	c	c	c	c	b	c	c	c	c	c	c	c	c			b	c	c		c	c	c
Year maths studied	89		92	99	87	99	0	97	96	99	99	92	97	95	1		99	88	95		99	99	
Age group	3	2	2	2	2	2	2	2		2	2	3	2	2	2		2	3	2	3	2	2	2
IT experience	hnd		poor	A	av	av			p		p	no qual	p	good	av		clait		p		yr 9	av	av
Rate the usefulness of the software to help you to correct errors in your understanding	3	3	3	3	3	2	2	3	2	2	2	3	2	2	2	3	3	3	2	2	3	2	3
How well did the software cover the subject area you need to study?	3	2	2	1	2	1	1	2	2	1	2	1	2	2	3	1	1	2	3	2	3	3	2
How much did the software accept your method for solving a given question	1	3	3	1	2	2	2	2	1	1	3	2	2	2	3	2	2	2	3	3	2	3	2
Rate the balance between state of the art graphics features and benefits for learning	2	3	2		2	2	2	2	2	3	2	3	2	2	2	1	1	1	3	2	2	2	2
Do you think the software would help you to carry on doing work on your own?			3	1	3	2	2	2		3	2	3	2	3	3	1	1	2	3	1	2	2	3
Were mathematical terms used the way you expected?			2	2	2	1	1	1	3	1	2	2	3	2	1	2	2	1	3	2	1	3	1
Rate the appropriateness of the representations or presentation of different types of question	3	2	2	2		2	1	3	2	2	2	2	2	1	2	1	3	2	2	2	1	3	2
Did you find the graphics generally helped you to learn or were they poor in that they distracted you?	1	3	2	1	2	2	2	2	1	3	2	2	1	1	3	1	2	2	3	2	1	2	1
How would you rate the methods of interaction in terms of ease of use, familiarity and functionality?	2	3	2		2	2	2		2	1	2	2	2	1	3	1	2	2	2	2	2	3	3
Rate the balance between helpful informative interaction and feedback which intrudes and interrupts progress.	1	3	3		2	2	1	2	2	1	2	3	2	3	2	1	2	2	1	2	3	2	2

Participant	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
To what extent does the environment support movement within questions and from one question to another to help learning?	3		3		2	2	2	2	2	2	2	3	2	2	3	2	2	2	3	2	2	2	3
To what extent does the structure of the environment support the notion of self-directed learning?	1		3		2	2	2	2	2	3	2	2	2	3	2	1	1	1	3	2	2	2	3
How easy was it to develop an understanding of how to use the system and hence develop the skills to maximise full use the system?	2		1		3	3	2	1	2	1	2	2	2	2	2	1	2	2	1	2	1	2	3
To what extent did you recognise the mathematical symbols, icons and buttons used by the software?	2		1		2	2	1	1	2	1	2	2	2	1	1	2	2	2	1	2	2	1	1
How would you rate the system's ability to shielding you from usability obstacles such as runtime errors and malfunctions?	2		2		2	2	2		2	2	2	1	1	1	1	1	1	2	1	2	2	2	3
How would you rate the system's ability to protect you from making annoying errors such as incorrect syntax of statements?	3				2	1	2		2	2	2	3	1	1	2	2	3	2	3	3	2	3	3
To what extent did your methods for solving problems and those used by the system to present questions and feedback match?	2		3		3	2	2	2	2	2	2	2	1	3	3	1	2	3	2	2	2	2	3
To what extent does the software enable you to experiment with solutions and develop your own personal understanding?	2		3		3	1	1	2	2	2	2	3	2	3	1	3	3	2	3	2	3	3	3

Participant	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
Route			pe													pe		
Gender	f		f	f	f	f	f		f	m	f	m	m		f	f	f	m
Highest maths qual	a		as		a	g	found	g	a	g	g	g	g	g	a	g	cse	a
Maths grade	e		c	d	b	c	pass	c	e	b	c	c	b	b	d	c	1	b
Year maths studied	1		1	0	1	99	95	98	1	0	98	99	99	98	1	99	99	97
Age group	2		2	2	2	2	3	2	2	2	2	2	2	2	2	2	3	2
IT experience	p		g		A		av	g			g	some	clait	none			access	high
Rate the usefulness of the software to help you to correct errors in your understanding	2	3	2	2	3	2	3	3	2	2	2	2	3	3	2	2	2	1
How well did the software cover the subject area you need to study?	3		2	2	2	2	2		2	2	2	1	1	2	2	2	3	2
How much did the software accept your method for solving a given question	1	1	2	3	3	1	2	3	3	2	2	2	2	2	2	1	1	3
Rate the balance between state of the art graphics features and benefits for learning	3		3	2	2	1	2	2	2	2	2	1	2	2	2	2	3	2
Do you think the software would help you to carry on doing work on your own?	1	3	2	2	3	1	2			2	2	2	2	2	3	1	2	2
Were mathematical terms used the way you expected?	2	1	2	2	2	2	1	3	2	2	2	2	1	2	1	2		3
Rate the appropriateness of the representations or presentation of different types of question	2	1	2	2	2	2	2		2	2	2	2	2	2	2	1	3	2
Did you find the graphics generally helped you to learn or were they poor in that they distracted you?	2		2	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2
How would you rate the methods of interaction in terms of ease of use, familiarity and functionality?	2	2	1	2	1	2	2	2	2	2	3	2	1	1	2		1	2
Rate the balance between helpful informative interaction and feedback which intrudes and interrupts progress.	2	1		2	3	2	2	3	3	2	1	3	3	1	2		3	2

Participant	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
To what extent does the environment support movement within questions and from one question to another to help learning?	2	2		2	3	3		2	2	2	2	2	1	1	2		2	2
To what extent does the structure of the environment support the notion of self-directed learning?	2	1	2	2	1	2	3	1	2	2	2	2	2	1	2		2	2
How easy was it to develop an understanding of how to use the system and hence develop the skills to maximise full use the system?	1	1	2	2	2	1	1	2	2	2		1	2	1	1		1	2
To what extent did you recognise the mathematical symbols, icons and buttons used by the software?	1	1	2	2	1	2	1	2	1	2		1	2	2	1			2
How would you rate the system's ability to shielding you from usability obstacles such as runtime errors and malfunctions?	2		1	2	2	2	2	2	1	2	1	2	1	1	1		3	2
How would you rate the system's ability to protect you from making annoying errors such as incorrect syntax of statements?	3	3	3	2	2	2	2	3	1	2	2	3	3	2	2		3	2
To what extent did your methods for solving problems and those used by the system to present questions and feedback match?	2	2	2	2	3	2	2	3	3	2	2	2	2	1	2		1	1
To what extent does the software enable you to experiment with solutions and develop your own personal understanding?	1	3	2	2	3	2	3	3	2	2	2	3	2	1	2		2	2

Participant	64	65	66	67	68	69	70
Route		pe		pe	MA	pe	pe
Gender	m	f		f	f	f	f
Highest maths qual	Scot highr	KS3	G	g	a	g	g
Maths grade	b	pass	C	b	e	b	b
Year maths studied	99	99	99	99	99	99	99
Age group	2	2	2	2	2	2	2
IT experience	av	KS3	clait			Nil	
Rate the usefulness of the software to help you to correct errors in your understanding	3	3	2		2		
How well did the software cover the subject area you need to study?	3	2	2	2	2	2	2
How much did the software accept your method for solving a given question	3	2		3	2	3	3
Rate the balance between state of the art graphics features and benefits for learning	3	2	2	3	2	3	3
Do you think the software would help you to carry on doing work on your own?	3	2	2	1	2	2	2
Were mathematical terms used the way you expected?	3	1	2	2	2	2	2
Rate the appropriateness of the representations or presentation of different types of question	3	2	2	2	2	2	2
Did you find the graphics generally helped you to learn or were they poor in that they distracted you?	3	2	2	2	2	2	2
How would you rate the methods of interaction in terms of ease of use, familiarity and functionality?	2		2	2	2	2	2
Rate the balance between helpful informative interaction and feedback which intrudes and interrupts progress.	3		2	2	2	2	2

Participant	64	65	66	67	68	69	70
To what extent does the environment support movement within questions and from one question to another to help learning?	3		2	2	2	2	2
To what extent does the structure of the environment support the notion of self-directed learning?	3		1	2	2	2	2
How easy was it to develop an understanding of how to use the system and hence develop the skills to maximise full use the system?	1		3	2	2	2	2
To what extent did you recognise the mathematical symbols, icons and buttons used by the software?	3		2	1	2	2	2
How would you rate the system's ability to shielding you from usability obstacles such as runtime errors and malfunctions?			2	1	2	2	2
How would you rate the system's ability to protect you from making annoying errors such as incorrect syntax of statements?	3		2	2	2	3	3
To what extent did your methods for solving problems and those used by the system to present questions and feedback match?	3		3	2	2	2	2
To what extent does the software enable you to experiment with solutions and develop your own personal understanding?	3		3	2	2	3	3

17b TREEFROG Evaluation COMMENTS

LEARNING

q1	Rate the usefulness of the software to help you to correct errors in your understanding	
13	did nt explain where you were going wrong	p
16	does not show how to work out answers	p
23	did nt help if you did nt know where to start	p
24	it is easier and less confusing to work out on paper	p
25	not enough explanation given when you consistnely get something wrong	p
26	I did not understand the tips	p
27	if wrong should give an answer	p
30	needed more explanation	p
34	the actual error should be highlighted	p
38	more information needed on where errors were made	p
39	more confusing	p
43	not enough information was given	p
45	did not show correct method of achieving answer	p
47	if you knew the answer good if you didn't the working answers were hard to understand	p
50	did not give correct answer to me or move on to the next question	p
53	help button didn't help me	p
58	limited help more would be helpful	p
59	answers not given with working out (P)	p
64	didn't explain what it wanted done	p
65	did not understand what I'd done wrong	p
1	it works more as a tool for testing rather than teaching	o
4	could explain where you go wrong	o
9	does nt give explanation	o
20	should show how the answer is found	o
28	helpfule hints for correcting answers	o
32	if often suggested hints	o
35	I am extremely tired so my mind was having great difficulty	o
46	it was annoying	o
57	good but did not show what was wrong properly	o
61	some help available	o
10	you can go back at any time	g

q2	How well did the software cover the subject area you need to study?	
3	very good covered all areas	g
4	wider range of maths eg %	o
6	it helped me understand it better	g
10	fractions are not one of my strong points	g
13	only seemed to cover algebra	p
14	have nt started course	
16	I realise how much I need to know for the course	g
27	it showed me that I need to brush up on powers and algebra	o
28	large range of questions	g
35	I'm sorry	o
38	questions covered all aspects	g
39	overall spread	g
42	find it easier to write working out than looking at a screen	p
43	I do PE	p
46	they were similar questions	p

q2	How well did the software cover the subject area you need to study?	
47	? Don't know what the subject area is, just started the course	
51	too much algebra	o
57	covered a wide range	g

q3	How much did the software accept your method for solving a given question	
8	did n't always give helpful tips for solving certain problems	o
10	sometimes you had to put it one way that seemed unuseful	g
23	did nt accept knowing the answer straight away for some questions	g
25	nowhere to put working out on screen	p
34	"order"	o
38	often answers were interpreted wrong	o
42	needed to learn how to use graphic in order to answer questions	p
46	all the time	g
50	if I gave correct answer sometimes it would tell me to do a method that made no sense	o
61	standard ways explained or help given if error occurred	g
62	was good when computer highlights in red to indicate an incorrect answer	g
63	not helpful enough when question incorrect	p
64	no difference between straight to answer or all stages	p
66	more info was needed	
67	no room to do calculations	p
69	no room to work out	p
70	no room to work out	p

q4	Rate the balance between state of the art graphics features and benefits for learning	
3	improvement on graphics could be made	p
10	large clear lettering easy to read	g
11	should be more fun - more colourful maybe have sounds and graphics	p
13	basic graphics only covers a narrow area of mathematics	o
14	clear and precise, but not as good as if written work	o
16	it is kept simple so that there is not too much information given	o
20	are the graphics state of the art?	o
26	?	
27	it was nt garish or stressful on the eyes it needs more user friendly facilities OK	o
38	very well presented	g
39	simple	g
42	needed to learn how to use graphic in order to answer questions	o
46	it was harder to use the computer	p
47	?	
50	maybe but it was quite boring	o
64	bad graphics leanred only a little	p
67	did n't help if wrong or explain what to do	p
69	didn't help correct mistakes	p
70	didn't explain well enough and no room to work out	p

q5	Do you think the software would help you to carry on doing work on your own?	
1	On my own I would look for exercises with more explanation	o
6	I think tis kind of coursework would be understood by most	g
10	yes but only as much as a text book	o
18	yes	

q5	Do you think the software would help you to carry on doing work on your own?	
19	yes as easily obtainable	o
20	only to a limited capacity	o
21	no explanations to answers	p
24	no I found it very un user friendly	
25	I would feel discouraged using this because it does n't give feedback	p
26	yes the right maths	g
27	no the other test package was much better	p
28	sometimes I like to miss out the middle step	o
31	yes	
32	due to lack of graphics found it dull	p
34	no needs written explanations of how to do problems plus examples	p
38	very quick and well presented testing	o
41	not at all	p
42	in own time	g
45	didn't really understand how to use functions	p
46	yes at least it tells you when you've gone wrong	g
47	no, not really	p
50	yes	
51	hints confused me	g
53	eventually	
60	I would prefer the work from in text book	p
64	no its boring and fiddly	p
65	as long as it explains mistakes which are made	o

q6	Were mathematical terms used the way you expected?	
7	not always	p
11	no I cam across symbols I have never seen before eg ^	p
31	no did nt understand some	p
41	not really	p
64	no.no explaining	p
4	don't know what ^ meant	o
8	some were difficult to understand	o
10	yes they were how I thought they would be	o
17	the only query was ^	o
19	yes	o
26	most of the time, some were confusing	o
27	yes	o
38	they were as expected	o
39	not always	o
42	if writing solution would have given all forms instead of one	o
46	never used ^ well I can't remember using it	o
59	^ did not know what it meant	o
61	ones I was familiaer with seemed set as usual	o
66	yes	o
6	yes used them before	g
47	yes	g
18	no did not understand ^	
23	^ used for power might confuse	
24	yes	
50	yes	
56	no some symbols did not recognise eg ^ to the power of	

q7	Rate the appropriateness of the representations or presentation of different types of question	
23	confusing	p
39	at all times very confusing	p
62	not enough variety of mathematics	p
64	badly set out ways or showing answers	p
13	narrow	o
14	sometimes not too clear	o
26	all questions were presented the same	o
46	changing z to x was annoying	o
48	possibly more variation in the questions	o
6	the questions varied as did the level of hardness	g
38	all aspects of subject covered	g
27	each question should have help (advice), answer and process and another question option	

q8	Did you find the graphics generally helped you to learn or were they poor in that they distracted you?	
11	did not distract me but a bit boring	p
20	did not help to learn	p
32	the background was plain and uninteresting	p
64	very dull	p
6	I was sometimes distracted	o
14	prefer to work by hand	o
16	too much info on the screen can distract attention	o
27	graphics are not important in this package, only in the case of eye stress from bright colours	o
28	they did n't really catch my attention so did n't help or hinder me	o
34	no distraction but no aid either	o
39	fine	o
42	will get easier with practice	o
58	did n't make any difference	o
65	they didn't distract me but the help button didn't work	o
10	I found the large clear screen and lettering helpful	g
13	basic you concentrated purely on the task	g
22	no distraction from graphics	g
26	they were simply no distractions (good)	g
38	not distracting	g
61	didn't distract quite basic graphics	g
2	What graphics?	
18	distracting slightly	
47	? Neither they were just normal	

Plus comments on usability for q 9 – 16 and Learning and Usability q17 and q18

Appendix 18 Group Interview Responses

Coding of participants <subject route> <number> <gender>

Subject route	DT	Design Technology Education	Sc	Science Education
	PE	PE Education	MA	Mathematics Education
Gender	m	male	f	female

Learning

DT1m, Sc1m	Explanations/feedback on the error being made would be helpful – if you put in an incorrect response then it is very difficult to know of a different one to give that might be right
PE1m, L1f	Guidance on methods of solution and style of answer needed – otherwise guessing – should be able to use any method (group felt sure that only one method was ever acceptable – misconception which actually related to their misunderstanding of the differences between equations and statements)
L1f, PE2f	this guidance could be available at all times when answering a set of questions – on screen partially – all steps given when needed – should have examples, demos, feedback relating to the actual error made not just general comments
L1f, Sc2f, Sc1m	Confidence destroyed – unenjoyable experience (for those with low level maths – not a factor for those with a higher level of understanding) – reminded users of 'bad' experiences in school maths – demotivated –boring – not drawing in
L1f, All	Feedback should be supportive and encouraging – felt like a test – caused fear
DT1m, Sc1m	A wider variety of maths topics could be supported such as shape i.e. those required for school maths (these students will be required to teach KS 2 maths)
DT1m, Sc1m, L1f	User should have the ability to select the area of maths (student – centred)
MA1m, Sc1m, Sc2f, L1f	Larger bank of questions – greater variety of difficulty – could be levelled so all will get some success – choice could be either manual by the user or automatically calculated on user's performance (like Diagnosys)
L1f, PE1m, PE2f, Sc2f	No results given at the end of the test – summary of questions which are correct (number / mathematical content) those partially correct with type of errors and those completely wrong or unattempted – feedback relating to errors
Sc1m, DT1m	Different types of questions such as multi-choice, or to give the question and the answer but required to deduce a method of solution
ALL	All students felt confident with the numeracy questions but all struggled to progress to the algebra – no connections seen – 'large jump' made in difficulty of questions
ALL	Just like GCSE maths
ALL except MA1m	Students felt that they had not learnt any maths through using the software but that they were probably worse at maths than they thought they were although they did nt know what their errors were

Usability

DT1m, ALL	Not user friendly - Interface basic, unpleasant on the eye – too large a block of dense colour
Sc1m, PE1m	Unexciting – boring – not encouraging to use – no desire to continue – irritating because guidance was so general
DT1m, ALL	Font size/type not easy on the eye
PE1m, ALL	Very precise format of answers and not obvious – although not always the same varied – sometimes must make a whole equation with = others not (showing their lack of understanding about the types of problems and their differences)
L1f, PE2f, DT1m	Introduction screens would have been useful – explaining the content, purpose and format requirements
L1f, PE1m, PE2f,	Demonstration/example questions taking the user step by step through answering a question would be useful – maybe should be able to select whether you require this support
DT1m	Screen could be centred – more appealing to the eye – overview of questions/results so far –could learn from errors and previous experience
Learning/Usability	
ALL	Sound not important could give the option
DT1m + ALL	Diagrams/visual add ins within questions – more interesting to want to continue to use or to use again

Appendix 19 Pre Test for Final Trial

Calculate the following

1. $1 + 3 - 4*2$
2. $3 - (4 + 2)$
3. $4 + 2 * (8 - 1) - 3$
4. $-5 + 4 * 5$
5. $3 - 2 * (-4 - 2)$
6. $2 * (4 + 2) / 3$

Solve the following Linear expressions

7. Find the value of $-3y + 8$ when $y = 2$
8. Find $5x$ when $4x = 20$
9. Find the value of $-3z - 8$ when $z = -3$
10. If $3x + 2y = 7$ find a value for $6x + 4y$

Solve the following equations

11. Find the value of x when $90 - x = 70$
12. Find the value of x when $x/5 = 25$
13. Find the value of z when $7 - z = 9$
14. Find the value of z when $2z + 5 = 11$
15. Find the value of z when $3z + 5 = z + 3$
16. Find the value of z when $9 + z = 17$

Simplify or rewrite the following expressions

17. $3 + 3(s - 4)$
18. $2(x + 3)$
19. $-(3x - w)$
20. $1 - 2(s - 3)$

Appendix 20 Webfrog with Feedback

20a Rationale

<u>Section 1 Numerical Calculations</u>

In this set of questions you should step by step perform addition, subtraction and multiplication in the correct order.
Some questions posed will include brackets, some will include negative values.
In the following 'dummy question clicking on the Cheat button in this practice question you can see how to use the software.

Practice Question

4+2*(8-1) – 3

Step 1: 4+2*7-3	calculate brackets first
Step 2: 4+14-3	perform multiplication before addition of subtraction
Step 3:18-3	add together all positive terms
Step 4: 15	perform final subtraction

<u>Section 2 Substituting values</u>

In this set of questions you should step by step use the given value for the algebraic variable and then calculate the numerical expression as done in Section 1.
Some questions posed will include brackets, some will include negative values.
In the following 'dummy question clicking on the Cheat button in this practice question you can see how to use the software.

Practice Question

Find the value of -3z-8 when z=-3

Step 1: -3*-3-8	substitute z with -3
Step 2: +9 – 8	minus times minus is positive
Step 3: 1	perform addition and subtraction to calculate the value

<u>Section 3 Simplifying expressions</u>

In this set of questions you should step by step simplify these expressions by GATHERING algebraic terms together and REMOVING brackets.
Some questions posed will include brackets, some will include negative values.
In the following 'dummy question clicking on the Cheat button in this practice question you can see how to use the software.

Practice Question

Simplify – 1 + 2(x+3)

Step 1 -1 + 2x + 6	Multiply the brackets by 2
Step 2 2x + 6 -1	Collect the numerical terms together
Step 3 2x+5	Caclulate the numerical terms

Section 4 Solving Equations

In this set of questions users should step by step COLLECT the algebraic terms TOGETHER on one side of the equation and the numeric terms together on the other.
These are equations so they must include an EQUAL TO SIGN "="
Some questions posed included brackets, some included negative values.
In the following 'dummy question clicking on the Cheat button in this practice question the user can see how to use the software.

Practice Question
2z+5=11
Step 1 2z=11-5 Collect all numerical values on one side of the equals and algebraic terms all on the other
Step 2 2z=6 Calculate the numerical terms
Step 2 z=6/2 Find value of only one z hence divide both sides by 2
Step 3 z=3 Calculate the numeric terms to find value of one z

20b Questions and feedback

I Numeracy

Questions posed	Relevant errors
QT1: No weaknesses	
1 1 + 2 * 3	E3 Add before multiply E2 Calculation errors
2 1 + 2*4 + 1	
QT2: Negatives	
1 -1 + 2 * 3	E3 Add before multiply E4 subtracting larger number from smaller E5 swop –ve and +ve E6 –ve *-ve E2 Calculation errors
2 1 – 2*4 + 1	
3 -2 * -2 + 1	
QT3: Brackets	
1 2 (3 + 4)	E3 Add before multiply E7 Calculating brackets first E8 adding the bracket multiplier E9 multiplying first term only E10 multiply numeric terms only E2 Calculation errors
2 (2+3)3 +1	
3 (2+3)(2 + 3)	
QT4: Negatives and brackets	
1 1 - (2 + 3)	E3 Add before multiply E4 subtracting larger number from smaller E5 swop –ve and +ve E6 –ve *-ve E7 Calculating brackets first E8 adding the bracket multiplier E9 multiplying first term only E10 multiply numeric terms only E2 Calculation errors
2. -2 + (2 - 3)	
3. -2 +3(1 - 3)	
4. +1 – 2 (-1 – 2)	

II Algebra a) Substitutions

<table>
<tr><td></td><td>Generic errors</td></tr>
<tr><td>QT1: No weaknesses</td><td></td></tr>
<tr><td>1 4x when x=6</td><td rowspan="2">E2 Calculation errors
E17 Substituion meaning addition
E18 substitution meaning replacing in position</td></tr>
<tr><td>2 2x+3 when x= 5</td></tr>
<tr><td>QT2: Negatives</td><td></td></tr>
<tr><td>1 -3y+8 when y=2</td><td rowspan="3">E4 Positive value minus a larger number becomes a negative larger number
E5 Negative and positive signs are interchanged
E6 Negative multiplied by a negative is negative
E17 Substitution meaning addition
E18 Substitution meaning replacing
E2 Calculation errors</td></tr>
<tr><td>2 -3z-8 when z=-3</td></tr>
<tr><td>3 -2x +6 when x=2</td></tr>
<tr><td>QT3: Brackets</td><td></td></tr>
<tr><td>1 if 3x+2y=7 find a value for 6x+4y</td><td rowspan="3">E7 Not calculating brackets first
E8 Adding the value of the multiplier to the values in the brackets
E9 Multiply first term in brackets only
E10 Multiply numeric terms only
E17 Substitution meaning addition
E18 Substitution meaning replacing
E2 Calculation errors</td></tr>
<tr><td>2 if 2x+4y = 10 find a value for x+2y</td></tr>
<tr><td>3 If x+3y=7 find a value for 7x+21y</td></tr>
<tr><td>QT4: Negatives and brackets</td><td></td></tr>
<tr><td>1 If 2x+5y=3 find a value for –(4x+10y)</td><td rowspan="4">E4 Positive value minus a larger number becomes a negative larger number
E5 Negative and positive signs are interchanged
E6 Negative multiplied by a negative is negative
E7 Not calculating brackets first
E8 Adding the value of the multiplier to the values in the brackets
E9 Multiply first term in brackets only
E10 Multiply numeric terms only
E17 Substitution meaning addition
E18 Substitution meaning replacing
E3 Calculation errors</td></tr>
<tr><td>2 If x-y=2 find a value for (–x+y)</td></tr>
<tr><td>3 If 2y+3x=6 find a value for -4y-6x</td></tr>
<tr><td>4 If 9x-6y=15 find a value for -3x+2y</td></tr>
</table>

ii Algebra b) Simplify expressions

	Errors
QT1: No weaknesses LIMITED	
1 1 + X + 3	E3 Add before multiply E2 Calculation errors
2 1 + X + 2X	
QT2: Negatives	
1 1 + X − 3	E3 Add before multiply E4 subtracting larger number from smaller E5 swop –ve and +ve E2 Calculation errors
2 1 − X + 2X	
3 -2X + 2 − 3X − 1	
QT3: Brackets	
1 2(x+3)	E3 Add before multiply E7 Calculating brackets first E8 adding the bracket multiplier E9 multiplying first term only E10 multiply numeric terms only E2 Calculation errors
2 3+21(s+4)	
3 3(2a)	
QT4: Negatives and brackets	
1 –(3x-w)	E3 Add before multiply E4 subtracting larger number from smaller E5 swop –ve and +ve E6 –ve *-ve E2 Calculation errors
2 -9(s-3)	
3 17 -2 (s+7)	
4 7 -3 (2-t)	

ii algebra c) Rewrite

E1 syntax errors could be relevant to all these questions as in solving Linear equations

	Errors
QT1: No weaknesses	
1 3b+1= b+	E2 Calculation errors E11 Omit algebraic terms E12 Adding unlike terms eg algebraic and numeric values E13 Coefficient swopping between algebraic and numeric terms E14 Confusion between equivalent and not equal to
2 4c +2 = 2c + 2 + ...	

QT2: Negatives	
1 4c+8=5c+8 + …..	E2 Calculation errors E3 Add before multiply E4 subtracting larger number from smaller E5 swop –ve and +ve E11 Omit algebraic terms E12 Adding unlike terms eg algebraic and numeric values E13 Coefficient swopping between algebraic and numeric terms E14 Confusion between equivalent and not equal to
2 b-1 = 2b + …..	
3 4c + 8 = 5c +10 + ….	
QT3: Brackets	
1 4c+2=2(c+1) + ….	E2 Calculation errors E7 Calculating brackets first E8 adding the bracket multiplier E9 multiplying first term only E10 multiply numeric terms onl E11 Omit algebraic terms E12 Adding unlike terms eg algebraic and numeric values E13 Coefficient swopping between algebraic and numeric terms E14 Confusion between equivalent and not equal to
2 2(a+2) = a+1 + ….	
3 3(2a+1) =	
QT4: Negatives and brackets	
1 3(a-2) = 4(a+1) + …	E2 Calculation errors E7 Calculating brackets first E8 adding the bracket multiplier E9 multiplying first term only E10 multiply numeric terms onl E11 Omit algebraic terms E12 Adding unlike terms eg algebraic and numeric values E13 Coefficient swopping between algebraic and numeric terms E14 Confusion between equivalent and not equal to
2 -2 (2a+3) = 5a+10 + ..	
3 5s-7 = 5(2s+2) + ….	
4 4(c+2)= 5 (c-2) + ….	

II Algebra d) Solve linear equations

As these are EQUATIONS **E1 Syntax errors** are possible in all questions

	Errors
QT1: No weaknesses	
1 x+2 =5	E3 Add before multiply E2 Calculation errors E15 +ve –ve inverse transposition errors
2 3X+16=82	

QT2: Negatives)	
1 90-x = 70 2 7-z = 9 3 -9-z = -7	E3 Add before multiply E2 Calculation errors E4 Positive value minus a larger number becomes a negative larger number E5 Negative and positive signs are interchanged E6 Negative multiplied by a negative is negative E15 +ve –ve inverse transposition errors
QT3: Brackets	
1 3+(s+2) = 6 2 3(2y+1) = 3y 3 5(y+2) = 6(y+1)	E3 Add before multiply E2 Calculation errors E7 Calculating brackets first E8 adding the bracket multiplier E9 multiplying first term only E10 multiply numeric terms only E15 +ve –ve inverse transposition errors E16 * / inverse transposition errors
QT4: Negatives and brackets	
1 3 – (s + 2) = 1 2 3 + (s - 2) = 2 3 3 – (s – 2) = 1 4 4 + 3(x – 4) = 2(1-x)	E3 Add before multiply E2 Calculation errors E4 Positive value minus a larger number becomes a negative larger number E5 Negative and positive signs are interchanged E6 Negative multiplied by a negative is negative E7 Calculating brackets first E8 adding the bracket multiplier E9 multiplying first term only E10 multiply numeric terms only E15 +ve –ve inverse transposition errors E16 * / inverse transposition errors

Appendix 21 Webfrog with Feedback Errors and Feedback Mapping

Error Codes	Feedback Responses for the errors
General	
E1 Syntax or format errors (i.e 20 not x=20)* ***this type of error should be tested for in every question type**	If this question is an **equation** so an answer of the form **"x =** ***number*"** is required **OR** if it is an expression **no "=" is required**
E2 Calculation errors** ** **this is the catch all error**	**Take care** with **all** your **calculations**
Order of operators	
E3 Add before multiplying	Remember always **multiply before adding or subtracting**
Misunderstanding negative sign	
E4 Positive value minus a larger number becomes a negative larger number	When **subtracting a larger value from a positive number** the answer is **negative and smaller** eg 1-3=-2
E5 Negative and positive signs are interchanged	We **cannot swop negative and positive signs** eg 1-2 ≠ 2-1 because -1 is **not the same** as 1. You could use a **Number line.**
E6 Negative multiplied by a negative is negative	Remember a **negative value multiplied by a negative value** gives a **positive** eg -2 * -2 = + 4
Brackets	
E7 Not calculating brackets first	Remember always **calculate** the **brackets first**
E8 Adding the value of the multiplier to the values in the brackets	Remember **multiply every term** in the **brackets** by the value outside **do not add** the bracket to this value
E9 Multiply first term in brackets only	Remember when multiplying brackets **every term in the brackets** is multiplied not just the first eg 2(x+1) = 2*x + 2*1
E10 Multiply numeric terms only	Remember when multiplying brackets every term in the brackets is **multiplied** numbers and algebraic terms

Algebraic	
E11 Omit algebraic terms	When rewriting equations we **cannot leave out** the **algebraic terms**
E12 Adding unlike terms eg algebraic and numeric values	Remember we **cannot add** add numbers and algebraic terms together they are **unlike terms** i.e. **different** like apples and oranges
E13 Coefficient swapping between algebraic and numeric terms	We **cannot swop** the coefficients eg. $2b+3 \neq 2 + 3b$
E14 Confusion between equivalent and not equal to	**Equals** = is **not the same** as equivalent hence 3c+1 is equivalent to c+1/3 but **not equal to**
Linear equation solving	
E15 +ve –ve inverse transposition errors	When taking terms from LHS to RHS of an equation the sign changes eg if $2x+1 = 4$ then $2x = 4 - 1$ or $2x+1=4x$ then $1=4x-2x$
E16 * / inverse transposition errors	When we multiply or divide one side of the equation we must do the same to the other eg $2x =1$ then divide both sides by 2 so $x=1/2$
Substitutions	
E17 Substituion meaning addition	**2x** means 2 multiplied by x not (2+x)
E18 substitution meaning replacing in position	**2x when x=6 means 2 multiplied by x do not put a 6 in the place of x i.e. not 26**
Feedback for correct responses:	**POSITIVE FEEDBACK Well done that is correct**

Appendix 22 Final Trial Post Test

Calculate the following

1. $1 + 4 - 3*2$
2. $2 - (3 + 4)$
3. $3 + 2 * (8 - 2) - 2$
4. $-3 + 4 * 8$
5. $1 - 4 * (-5 - 2)$
6. $2 * (6 + 3) / 3$

Solve the following Linear expressions

7. Find the value of $-3y + 7$ when $y = 3$
8. Find $4x$ when $5x = 20$
9. Find the value of $-2z - 6$ when $z = -2$
10. If $2x + 3y = 7$ find a value for $4x + 6y$

Solve the following equations

11. Find the value of x when $80 - x = 10$
12. Find the value of x when $x/4 = 22$
13. Find the value of z when $8 - z = 10$
14. Find the value of z when $3z + 4 = 13$
15. Find the value of z when $3z + 5 = z + 3$
16. Find the value of z when $9 + z = 17$

Simplify or rewrite the following expressions

17. $2 + 3 (s - 4)$
18. $2 (x + 2)$
19. $- (-3x + w)$
20. $1 - 3(s - 2)$

Appendix 23 Pre Test and Post Test Results

Question	40500000624	40500001230	40500000030	20300001126	40500001854	40500001723	40500001490	40500000953	40500001803	40500000654	40500000312	40500000004	40500003107	40500001352	40500002208	40500002289	30400000488	40500001715	40500000827	40500000721	40500000904	40500000521	Question total	Question Rank
Pre Test (1=correct and 0=wrong)																								
1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	1	1	1	0	0	1	0	0	14	4
2	0	1	0	1	1	0	1	1	1	0	0	0	1	1	1	1	1	1	0	1	1	1	13	6
3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	0	0	1	5	16
4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0	1	0	1	5	16
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	2	18
6	1	1	1	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	16	2
7	1	1	0	0	1	0	0	1	0	0	0	0	1	0	1	1	1	1	1	0	1	1	10	9
8	1	0	0	0	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	0	1	14	4
9	0	1	0	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1	0	1	0	1	8	12
10	1	1	0	0	1	0	0	1	0	0	1	1	1	0	1	0	1	1	1	1	1	1	12	8
11	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	17	1
12	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	1	1	1	1	1	0	0	9	10
13	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	9	10
14	1	1	0	0	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	13	6
15	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0	1	7	13
16	1	1	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	16	2
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	19
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	6	14
19	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	1	1	0	6	14
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
Total	8	10	2	2	10	5	7	10	5	1	2	5	15	7	17	17	19	13	13	15	10	15		
Rank	13	9	19	19	9	16	14	9	16	22	19	16	4	14	2	2	1	7	7	4	9	4		
wfrog1 or 2	1	2	1	1	1	2	2	1	2	2	2	1	2	1	2	1	1	1	2	2	2	1		
Post Test (1=correct and 0=wrong)																								
1	1	1	0	0	1	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	16	3
2	1	1	1	0	1	0	1	1	0	1	1	0	1	0	1	1	1	1	0	1	1	1	15	7
3	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	0	0	0	6	17
4	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	18	2
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	0	0	4	19
6	0	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	16	3
7	1	1	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1	0	1	1	1	1	11	12
8	1	1	0	0	1	1	1	1	0	1	0	0	1	1	1	1	1	0	1	1	1	1	15	7
9	1	1	0	0	1	1	0	1	0	1	0	0	0	0	1	1	1	1	1	1	1	1	13	10
10	0	0	0	0	0	1	1	1	0	0	0	0	1	0	1	1	1	0	1	1	1	1	10	13
11	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	19	1
12	1	1	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	16	3
13	1	1	0	0	1	1	0	1	1	0	0	0	1	0	1	1	1	1	1	1	0	1	13	10
14	1	1	0	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	16	3
15	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0	1	7	16
16	0	1	0	0	0	0	0	0	1	0	0	0	1	0	1	1	1	1	1	1	0	1	9	14
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	6	17
18	1	1	0	0	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	15	7
19	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	0	8	15
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	3	20
Total	11	12	2	3	11	10	10	10	7	7	4	5	16	8	19	19	19	13	18	19	13	14		
Rank	11	10	22	21	11	13	13	13	17	17	20	19	6	16	1	1	1	8	5	1	8	7		

Appendix 24 Final Trial Pre Test Weaknesses

Table 24.1 Results of Weakness 1 questions: 6, 8, 12, 14

Question	40500000624	40500001230	40500000030	20300001126	40500001854	40500001723	40500001490	40500000953	40500001803	40500000654	40500000312	40500000004	40500003107	40500001352	40500002208	40500002289	30400000488	40500001715	40500000827	40500000721	40500000904	40500000521	Question total	Question Rank
6	1	1	1	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	16	2
8	1	0	0	0	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	0	1	14	4
12	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	1	1	1	1	1	0	0		
14	1	1	0	0	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	13	6
Σ	3	3	1	0	3	2	2	4	2	0	0	2	4	2	4	4	4	4	4	4	2	3		

Table 24.2 Results of Weakness 2 questions: 2, 3, 5, 6, 17, 18, 19,20

Question	40500000624	40500001230	40500000030	20300001126	40500001854	40500001723	40500001490	40500000953	40500001803	40500000654	40500000312	40500000004	40500003107	40500001352	40500002208	40500002289	30400000488	40500001715	40500000827	40500000721	40500000904	40500000521	Question total	Question Rank
2	0	1	0	1	1	0	1	1	1	0	0	0	1	1	1	1	1	1	0	1	1	1	13	6
3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	0	0	1	5	16
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	2	18
6	1	1	1	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	16	2
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	19
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	6	14
19	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	1	1	0	6	14
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
Σ	1	2	1	1	2	1	2	2	2	0	0	1	4	2	5	6	7	3	4	4	4	5		

Table 24.3Results of Weakness 4 questions: 7,8,9,10

Question	40500000624	40500001230	40500000030	20300001126	40500001854	40500001723	40500001490	40500000953	40500001803	40500000654	40500000312	40500000004	40500003107	40500001352	40500002208	40500002289	30400000488	40500001715	40500000827	40500000721	40500000904	40500000521	Question total	Question Rank
7	1	1	0	0	1	0	0	1	0	0	0	0	1	0	1	1	1	1	1	0	1	1	10	9
8	1	0	0	0	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	0	1	14	4

9	0	1	0	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1	0	1	0	1	8	12
10	1	1	0	0	1	0	0	1	0	0	1	1	1	0	1	0	1	1	1	1	1	1	12	8
Σ	3	3	0	0	3	1	2	3	1	0	1	2	3	1	4	3	4	4	3	3	2	4		

Table 24.4 Results of Weakness 5 questions: 1, 2, 3, 4, 5, 6, 7, 9, 11, 13, 15, 16, 17, 19, 20

Question	40500000624	40500001230	40500000030	20300001126	40500001854	40500001723	40500001490	40500000953	40500001803	40500000654	40500000312	40500000004	40500003107	40500001352	40500002208	40500002289	30400000488	40500001715	40500000827	40500000721	40500000904	40500000521	Question total	Question Rank
1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	1	1	1	0	0	1	0	0	14	4
2	0	1	0	1	1	0	1	1	1	0	0	0	1	1	1	1	1	1	0	1	1	1	13	6
3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	0	0	1	5	16
4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0	1	0	1	5	16
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	2	18
7	1	1	0	0	1	0	0	1	0	0	0	0	1	0	1	1	1	1	1	0	1	1	10	9
9	0	1	0	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1	0	1	0	1	8	12
11	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	17	1
13	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	9	10
15	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0	1	7	13
16	1	1	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	16	2
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	19
19	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	1	1	0	6	14
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
Σ	4	6	1	2	6	3	5	5	3	1	1	2	10	5	11	12	13	7	7	9	6	10		

Table 24.5 Results of Weakness 6 questions: 8, 10, 11, 12, 13, 14, 15, 16

Question	40500000624	40500001230	40500000030	20300001126	40500001854	40500001723	40500001490	40500000953	40500001803	40500000654	40500000312	40500000004	40500003107	40500001352	40500002208	40500002289	30400000488	40500001715	40500000827	40500000721	40500000904	40500000521	Question total	Question Rank
8	1	0	0	0	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	0	1	14	4
10	1	1	0	0	1	0	0	1	0	0	1	1	1	0	1	0	1	1	1	1	1	1	12	8
11	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	17	1
12	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	1	1	1	1	1	0	0		

13	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	9	10
14	1	1	0	0	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	13	6
15	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0	1	7	13
16	1	1	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	16	2
Σ	5	5	0	0	6	4	3	6	3	0	2	5	8	3	8	7	8	8	8	8	5	7		

Table 24.6 means for each weakness and total score

weakness	40500000624	40500001230	40500000030	20300001126	40500001854	40500001723	40500001490	40500000953	40500001803	40500000654	40500000312	40500000004	40500003107	40500001352	40500002208	40500002289	30400000488	40500001715	40500000827	40500000721	40500000904	40500000521
1	0.75	0.75	0.25	0	0.75	0.5	0.5	1	0.5	0	0	0.5	1	0.5	1	1	1	1	1	1	0.5	0.75
2	0.22	0.33	0.11	0.11	0.33	0.22	0.33	0.33	0.33	0.00	0.00	0.11	0.56	0.33	0.67	0.78	0.89	0.44	0.56	0.56	0.56	0.67
4	0.75	0.75	0	0	0.75	0.25	0.5	0.75	0.25	0	0.25	0.5	0.75	0.25	1	0.75	1	1	0.75	0.75	0.5	1
5	0.29	0.43	0.07	0.14	0.43	0.21	0.36	0.36	0.21	0.07	0.07	0.14	0.71	0.36	0.79	0.86	0.93	0.50	0.50	0.64	0.43	0.71
6	0.625	0.625	0	0	0.75	0.5	0.375	0.75	0.375	0	0.25	0.625	1	0.375	1	0.875	1	1	1	1	0.625	0.875
all	0.4	0.5	0.1	0.1	0.5	0.25	0.35	0.5	0.25	0.05	0.1	0.25	0.75	0.35	0.85	0.85	0.95	0.65	0.65	0.75	0.5	0.75

Table 24.7

Pre test score	1	15	2	5	13	15	10	10	17	5	7	
Candidates	654	721	312	1803	827	3107	904	1230	2208	1723	1490	totals
errors												
[error,multiply,2,by,all,expressions,in,the,bracket]	0	0	0	0	0	0	0	0	0	2	0	2
[error,multiply,before,adding,or,subtracting]	13	7	4	5	0	4	0	0	3	0	4	40
[error,ok]	0	1	0	0	0	0	0	0	0	0	0	1
[error,exactly,one,variable,please]	0	0	2	0	0	0	0	0	0	0	0	2
[error,wrong,variable]	0	0	0	0	0	1	0	0	0	0	0	1
[error,please,enter,an,equation,involving,s,e.g.,s + 1 = 9]	0	0	0	2	0	0	1	0	0	0	0	3
[error,please,enter,an,equation,involving,x,e.g.,x + 1 = 9]	0	0	0	0	0	0	0	0	0	1	0	1
[error,please,enter,an,equation,involving,y,e.g.,y + 1 = 9]	0	0	0	0	0	2	0	0	0	0	0	2
[error,please,enter,an,equation,involving,z,e.g.,z + 1 = 9]	2	0	1	0	0	0	0	4	0	0	0	7
[error,sign,error,on,the,left - hand,side]	0	0	0	2	2	1	3	2	0	0	0	10
[error,[sign,error,on,the,right - hand,side]]	0	0	0	0	0	0	2	0	0	0	0	2
[error,sign,error]	7	2	16	24	1	5	0	1	3	19	8	86
[error,subtracting,a,larger,number,from,a,positive,one,gives,a,i	0	0	1	0	0	0	0	0	0	0	0	1
[error,you,subtracted,7,on,the,left,hand,side,but,you,added,7,on,the,right,hand,side]	0	0	0	1	0	0	0	0	0	0	0	1
[error,you,added,9,on,the,left,hand,side,but,you,subtracted,9,on,the,right,hand,side]	0	0	0	0	0	0	1	0	0	0	0	1
[error,Syntax error on right hand side]	0	0	0	0	0	0	0	0	0	1	0	1
[error,Syntax error on left hand side]	0	0	0	0	0	0	0	0	0	0	7	7
[error,please,type,your,answer]	0	0	1	0	1	0	1	0	2	1	1	7
[error,take,care,with,all,your,calculations]	8	11	16	20	4	24	13	16	19	56	11	198
Total errors made												
Errors	30	21	41	54	8	37	21	23	27	80	31	

Table 24.8 Webfrog with feedback users feedback

Candidates	1715	1126	1854	953	4	1352	488	2289	30	521	624	
errors												
[error,multiply,2,by,all,expressions,in,the,bracket]	0	0	0	0	0	0	0	0	0	1	0	1
[error,multiply,before,adding,or,subtracting]	0	0	3	4	1	3	1	6	3	7	3	31
[error,ok]	0	0	0	0	0	0	0	0	0	0	0	0
[error,exactly,one,variable,please]	0	0	0	0	0	0	0	0	0	0	0	0
[error,wrong,variable]	0	0	0	0	0	0	0	1	0	0	0	1
[error,please,enter,an,equation,involving,s,e.g.,s + 1 = 9]	0	0	0	0	0	0	0	0	0	5	0	5
[error,please,enter,an,equation,involving,x,e.g.,x + 1 = 9]	0	0	0	0	0	0	0	0	1	0	0	1
[error,please,enter,an,equation,involving,y,e.g.,y + 1 = 9]	0	0	0	0	1	0	0	0	0	0	0	1
[error,please,enter,an,equation,involving,z,e.g.,z + 1 = 9]	0	0	0	0	0	0	0	0	0	2	0	2
[error,sign,error,on,the,left - hand,side]	0	0	0	2	1	0	0	0	0	2	0	5
[error,[sign,error,on,the,right - hand,side]]	0	0	0	0	0	0	0	0	0	0	0	0
[error,sign,error]	0	0	6	6	16	0	0	0	6	6	14	54
[error,subtracting,a,larger,number,from,a,positive,one,gives,a,i	0	0	0	0	0	0	0	0	0	0	0	0
[error,you,subtracted,7,on,the,left,hand,side,but,you,added,7,on,the,right,hand,side]	0	0	0	0	0	0	0	0	0	0	0	0
[error,you,added,9,on,the,left,hand,side,but,you,subtracted,9,on,the,right,hand,side]	0	0	0	0	0	0	0	0	0	1	0	1
[error,Syntax error on right hand side]	0	0	0	0	0	0	0	1	0	0	0	1
[error,Syntax error on left hand side]	0	0	0	0	0	0	0	0	0	0	0	0
[error,please,type,your,answer]	0	0	0	2	2	0	1	1	1	3	0	10
[error,take,care,with,all,your,calculations]	7	0	13	25	26	9	7	16	21	18	11	153
Total errors made												
Errors	7	0	22	39	47	12	9	25	32	45	28	
total Questions answered correctly												
[finished,well,done]	0	0	13	25	30	10	37	69	12	66	10	

Appendix 25 SPSS evidence

Table 25.1.1: Summary of survey results per question

	Number of								Numbers				
Statement	1	2	3	4	5	X	Total		Agree	Disagree	Unsure	No	Sample Size
14	12	6	0	0	0	0	18		18	0	0	0	18
23	4	14	0	0	0	0	18		18	0	0	0	18
5	7	10	0	0	1	0	18		17	1	0	0	18
7	12	3	2	0	0	1	18		15	0	2	1	17
6	8	6	1	2	1	0	18		14	3	1	0	18
22	4	10	2	0	0	2	18		14	0	2	2	16
2	9	5	1	0	1	2	18		14	1	1	2	16
19	5	8	2	2	0	1	18		13	2	2	1	17
18	6	6	2	2	2	0	18		12	4	2	0	18
20	3	9	1	4	1	0	18		12	5	1	0	18
1	4	7	1	3	1	2	18		11	4	1	2	16
9	2	8	3	2	0	3	18		10	2	3	3	15
4	4	5	2	5	2	0	18		9	7	2	0	18
16	1	7	4	4	2	0	18		8	6	4	0	18
3	1	7	2	7	1	0	18		8	8	2	0	18
10	0	2	4	7	5	0	18		2	12	4	0	18
15	1	1	3	10	3	0	18		2	13	3	0	18
13	0	2	1	11	4	0	18		2	15	1	0	18
11	0	1	2	6	9	0	18		1	15	2	0	18
17	2	4	3	7	1	1	18		6	8	3	1	17
21	0	5	2	9	1	1	18		5	10	2	1	17
8	0	5	4	6	1	2	18		5	7	4	2	16
12	2	5	5	2	1	3	18		7	3	5	3	15

Table 25.1.2: Type of response per respondent

	Respondent																	
Num of Responses	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	4	5	6	3	5	10	4	2	9	3	5	8	6	6	5	2	3
2	11	10	9	6	13	8	3	11	13	6	7	7	7	5	4	2	6	8
3	0	2	4	5	0	2	5	6	1	1	4	3	0	6	2	1	2	3
4	6	4	5	6	7	6	2	1	6	6	8	7	4	4	5	0	7	5
5	0	3	0	0	0	2	2	0	0	1	0	1	4	1	2	15	6	0
x	5	0	0	0	0	0	1	1	1	0	1	0	0	1	4	0	0	4
Mean	2.6	2.7	2.4	2.5	2.5	2.7	2.2	2.2	2.5	2.3	2.8	2.7	2.5	2.5	2.6	3.8	3.4	2.5
Mode	2	2	2	1	2	2	1	2	2	1	4	2	1	1	1	5	4	2

Table 25.1.3: Summary of proportional results per question

Statements	Sample Size	Agree	Disagree	Unsure/No	Statements	Sample Size	Agree	Disagree	Unsure/No
14	18	100	0	0	14	17	100	0	0
23	18	100	0	0	23	17	100	0	0
5	18	94	6	0	7	16	100	0	19
7	17	88	0	12	2	15	100	0	20
22	16	88	0	13	22	15	100	0	27
2	16	88	6	6	5	17	94	6	0
6	18	78	17	6	6	17	88	13	6
19	17	76	12	12	19	16	86	14	19
1	16	69	25	6	9	14	82	18	43
9	15	67	13	20	1	15	79	21	20
18	18	67	22	11	12	14	78	22	57
20	18	67	28	6	18	17	75	25	6
4	18	50	39	11	20	17	75	25	6
12	15	47	20	33	16	17	62	38	24
16	18	44	33	22	4	17	60	40	12
3	18	44	44	11	3	17	53	47	12
17	17	35	47	18	17	16	46	54	25
8	16	31	44	25	8	15	45	55	40
21	17	29	59	12	21	16	36	64	19
10	18	11	67	22	10	17	15	85	24
15	18	11	72	17	15	17	14	86	18
13	18	11	83	6	13	17	13	88	6
11	18	6	83	11	11	17	7	93	12

Table 25.2.1 Participants scores per mathematical area

ID	Maths Qualifi-cation	Maths Qual Grade	Numbers	Powers	Basic algebra	Algebra Methods	Equations	Algebra & Calculus	Statistics	Miscell	Lives used	Success in using lives	Time used	Total	Q Asked	Correct answers	Age group
S	GCSE C	1	65	33	18	0	0	0	0	50	5	1	40	31	27	9	>27
Mc	GCSE C	1	59	33	27	0	0	0	0	50	1	1	36	31	37	10	> 27
G	GCSE C	1	71	0	36	0	0	0	25	0	3	1	27	33	36	11	24-27
Ir	GCSE B	2	53	50	45	0	0	0	0	50	4	1	41	35	30	10	20-21
M	GCSE B	2	53	17	36	0	0	0	50	100	0	0	36	35	35	18	> 27
J	BTEC	1	71	33	27	0	0	0	0	100	5	1	37	37	36	12	> 27
P	GCSE C	1	88	33	36	0	0	0	50	100	1	1	38	50	34	15	>27
F	A level C	4	76	100	100	40	80	50	50	100	0	0	34	79	33	22	>27
L	A level B	5	94	67	100	60	80	50	50	50	2	0	28	79	32	21	20-21
	Mean		70	41	47	11	18	11	25	67	2	1	35	46	33	14	
	Max		94	100	100	60	80	50	50	100	5	1	41	79	37	22	
	Min		53	0	18	0	0	0	0	0	0	0	27	31	27	9	

Table 25.2.2 Frequency and distribution of Diagnosys test results

	Max	Min	Mean	Mode	0-20	21-40	41-60	61-80	81+
Diagnosys Total	82	6	38	26	8	27	16	5	2
Numbers	100	7	54	33	6	14	16	16	6
Powers	100	0	32	33	18	25	0	14	1
Basic Algebra	91	0	26	18	35	12	6	3	2
Statistics	67	0	14	0	37	17	0	4	0
Miscellaneous	100	0	32	0	31	0	17	0	10
Q ASKED	34	14	26	25	10	48			
Q CORRECT	15	1	8	6					

Table 25.2.3 Parametric Correlation of Total Diagnosys Total score and Starting level

Pair	Spearman's rho Correlation Coefficient	Sig. (1-tailed)	N
Start level and Total	.646	.000	73

Table 25.2.4 Non parametric Wilcoxon

	Gcse – totaldia	gcse – number	gcse – algebra
Z	-6.764	-7.078	-3.933
Asymp. Sig. (2-tailed)	.000	.000	.000

Table 25.2.5 Paired Samples Descriptive Statistics Diagnosys

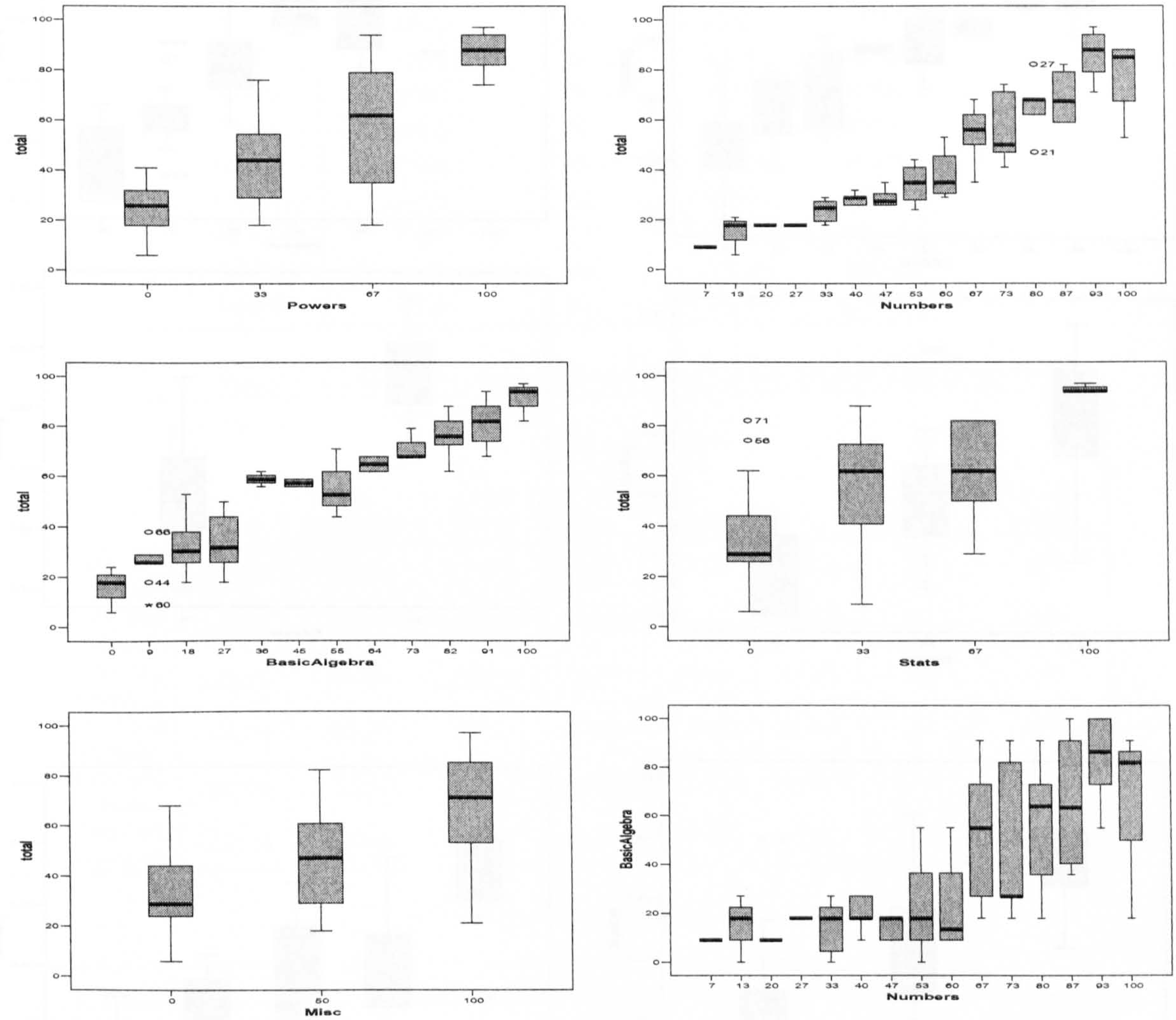
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Total	46.01	71	23.657	2.808
	Numbers	59.21	71	24.693	2.931
Pair 2	Total	46.52	73	23.703	2.774
	Powers	39.70	73	29.862	3.495
Pair 3	Total	47.59	70	23.492	2.808
	Basic Algebra	39.81	70	30.905	3.694
Pair 4	Total	46.38	71	23.721	2.815
	Stats	20.58	71	27.234	3.232
Pair 5	Total	46.52	73	23.703	2.774
	Misc	40.41	73	42.211	4.940

Table 25.2.6 Paired Samples Test Diagnosys

	Paired Differences							
		Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
Pairs	Mean			Lower	Upper	t	df	Sig. (2-tailed)
Total - Numbers	-13.197	10.762	1.277	-15.744	-10.650	-10.333	70	.000
Total – Powers	6.822	21.659	2.535	1.768	11.875	2.691	72	.009
Total – Basic Algebra	7.771	13.128	1.569	4.641	10.902	4.953	69	.000
Total - Stats	25.803	21.648	2.569	20.679	30.927	10.043	70	.000
Total - Misc	6.110	33.356	3.904	-1.673	13.892	1.565	72	.122
Numbers – Basic Algebra	21.233	22.572	2.642	15.966	26.499	8.037	72	.000

Statistical Method 2.5

Chart 25.2.7 Box plots of Percentile values of Total percentage score with each subgroup percentage score



Statistical Method 2.8

Chart 25.2.8 Box plots of percentile values of Weakness subsets with total mean score and Basic Algebra with Numbers

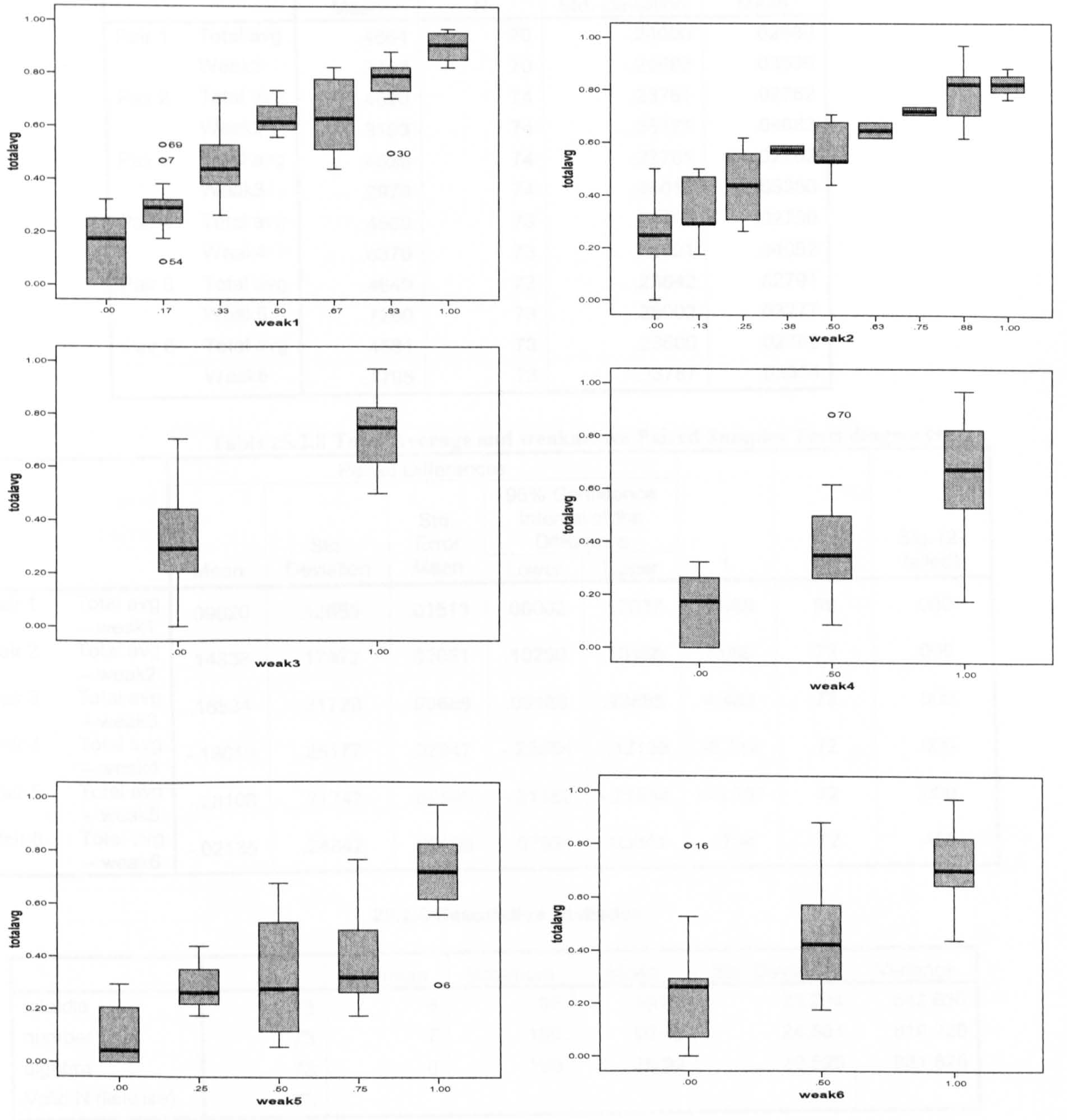


Table 25.2.7 Weakness Descriptive Statistics Treefrog

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Total avg	.4664	70	.24000	.02869
	Weak1	.3762	70	.29582	.03536
Pair 2	Total avg	.4626	74	.23761	.02762
	Weak2	.3193	74	.35121	.04083
Pair 3	Total avg	.4626	74	.23761	.02762
	Weak3	.2973	74	.46019	.05350
Pair 4	Total avg	.4569	73	.23402	.02739
	Weak4	.6370	73	.34621	.04052
Pair 5	Total avg	.4649	73	.23842	.02791
	Weak5	.7260	73	.28002	.03277
Pair 6	Total avg	.4581	73	.23600	.02762
	Weak6	.4795	73	.33787	.03954

Table 25.2.8 Total Average and weaknesses Paired Samples Tests diagnosys

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Total avg – weak1	.09020	.12655	.01513	.06002	.12037	5.963	69	.000
Pair 2	Total avg – weak2	.14338	.17472	.02031	.10290	.18386	7.060	73	.000
Pair 3	Total avg – weak3	.16534	.31729	.03688	.09183	.23885	4.483	73	.000
Pair 4	Total avg – weak4	-.18010	.25177	.02947	-.23884	-.12135	-6.112	72	.000
Pair 5	Total avg – weak5	-.26108	.21747	.02545	-.31182	-.21034	-10.257	72	.000
Pair 6	Total avg – weak6	-.02135	.24842	.02908	-.07931	.03661	-.734	72	.465

25.2.9 Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
totaldia	73	6	97	46.85	23.294	542.630
number	73	7	100	60.42	24.884	619.220
algebra	73	0	100	38.92	30.526	931.826
Valid N (listwise)	73					

Table 25.2.10 Treefrog Weakness and nonweakness Distribution statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Total avg	.4664	70	.24000	.02869
	Weak1	.3762	70	.29582	.03536
Pair 2	Total avg	.4626	74	.23761	.02762
	Weak2	.3193	74	.35121	.04083
Pair 3	Total avg	.4626	74	.23761	.02762
	Weak3	.2973	74	.46019	.05350
Pair 4	Total avg	.4569	73	.23402	.02739
	Weak4	.6370	73	.34621	.04052
Pair 5	Total avg	.4649	73	.23842	.02791
	Weak5	.7260	73	.28002	.03277
Pair 6	Total avg	.4581	73	.23600	.02762
	Weak6	.4795	73	.33787	.03954

Table 25.3.1: Preliminary trial candidates **Absolute results of candidates in each TREEFROG Test**

		Candidates								
Test	**Total possible**	**J**	**M**	**L**	**S**	**P**	**F**	**G**	**Mc**	**I**
1	10	10	10	10	8	10	10	9	8	9
2	15	N/a	15	11	14	13	15	8	11	12
3	17	9	13	14	13	11	16	11	11	15
4	7	7	7	7	7	7	7	7	7	7
5	8	0	3	8	1	2	6	3	2	N/a
6	5	2	2	4	N/a	3	3	1	N/a	N/a
TOTAL	62	28	50	54	43	46	57	39	39	43

Table 25.3.2: Percentage results of candidates in each TREEFROG Test

Test	**Number of questions**	**J**	**M**	**L**	**S**	**P**	**F**	**G**	**Mc**	**I**	**Total % success**
1	10	100%	100%	100%	80%	100%	100%	90%	80%	90%	93
2	15	N/a	100%	73%	93%	87%	100%	53%	73%	80%	83
3	17	53%	76%	82%	76%	65%	94%	65%	65%	88%	74
4	7	100%	100%	100%	100%	100%	100%	100%	100%	100%	100
5	8	0%	38%	100%	13%	25%	75%	38%	25%	N/a	39
6	5	40%	40%	80%	N/a	60%	60%	20%	N/a	N/a	50
All	62	60%	81%	87%	75%	74%	92%	63%	68%	88%	

Table 25.3.3: Rank order of each candidate in each TREEFROG Test

Test	J	M	L	S	P	F	G	Mc	I	Rank of tests	Number of questions
1	1	1	1	8	1	1	6	8	6	2	10
2	N/a	1	6	3	4	1	8	6	5	3	15
3	9	4	3	4	6	1	6	6	2	4	17
4	1	1	1	1	1	1	1	1	1	1	7
5	8	3	1	7	5	2	3	5	N/a	6	8
6	4	4	1	N/a	2	2	6	N/a	N/a	5	5
All tests	9	4	3	5	6	1	8	7	2		

Table 25.3.4 Analysis of Preliminary Pilot Participants errors

Test 1						
Question	**Response**	**Expected**	**Syntax**	**-ve**	**Numeracy**	**Number of learners**
1	90-70=20 or 20	Z=90-70	✓			4
	Z=70-90			✓		1
3	Z=22 repeatedly	Z=20			✓	1
4	z=39, z=49	Z=41			✓	1
6	z=486	-x=476-962 or x=962-476	✓			2
	X=-284			✓		1
7	x=-8 (used x not z)	z=-8 from 9-z=17	✓			1
	Z=-26*			✓		1
	z=28, z=11			✓		1
	z=-26			✓		1
	Z=-6				✓	1
8	z=-16	z=-2 from 7-z=9		✓		1
9	z=+2	z=-2		✓		1
	z=16, z=2			✓		1
10	z=-514	z=486 from 476-z=962		✓		2

Test 2							
Question	**Response**	**Expected**	Syntax	-ve	/	Number	Number of learners
1	96	x=96	✓				1
	x=92					✓	1
	X=1.5				✓		1
	X=-4				✓	✓	1
4	392	X=392	✓				1
	X=442					✓	1
	X=393					✓	1
	X=492,x=393,x=402					✓	1
5	X=210	X=330				✓	1
7	-z=-88	X=-88	✓		✓		1

Test 2							
	X=88				✓		1
8	X=108	X=-108		✓			1
	-108		✓				1
	X=3 (divided not multiplied)				✓	✓	1
	X=-106					✓	1
9	-x=68	X=-68	✓			✓	1
11	X=22	X=21				✓	1
	x-21		✓				1
12	X=8	X=9				✓	2
	X=15					✓	1
13	X=22	X=30				✓	1
	X=3					✓	1
	X=33					✓	1
14	X=1	X=49				✓	1
15	X=68	X=96				✓	1

Test 3							
Question	**Response**	**Expected**	Syntax	-ve	( )	Numeracy	Number of learners
1	5x=25	25 only required	✓				3
3	N=-9	N=9		✓			1
	18=-2n			✓			1
4	$6x^2$	2x+6				✓	1
	2x=6		✓				1
5	W+3x	+w-3x		✓			3
6	9	14				✓	2
7	Y=1.3	Y=1.5				✓	1
8	X=2	X=1.5				✓	1
9	24s+96	21s+87			✓	✓	3
	3+4S*21				✓	✓	1
10	23s_89	23s-89	✓				1
11	17s+119-21s-147	-21s-130			✓	✓	1
	17-21s+147			✓		✓	2
	150-21s					✓	1
	164-21s			✓		✓	1
	-4s+28				✓	✓	1
	21s-130			✓			1

Test 3							
12	13-19s-57	-19s+70		✓			2
	80-19s			✓			1
	19s+70			✓			2
13	7xyz	X(3y+4z)				✓	2
	7+2x+yz					✓	1
	12x+3y+4z					✓	1
	3y+4z					✓	1
14	3	2b+1					3
	3a-4						1
15	3a-4	3a-6					2
16	+1-3-3c	-2-c					1
	+2-c				✓		1
17	Y+5/4	Y=5/4	✓				1

Test 5							
Question	**Response**	**Expected**	Syntax	()	/	Numeracy	Number of learners
1	AXBXC	**Abc**	✓				2
2	a*x+b*c*x+d	acx²+axd+bcx+bd					2
3	Y=(y+2)/2+3	Y=8		✓	✓		2
4	6(z+1)/6(z+4)=5	Z=14		✓	✓		1
7	x^2-35=3 or x=19	X=8 or x=4	✓	✓		✓	1

Test 6						
Question	**Response**	**Expected**	Syntax	[ ]	Numeracy	Number of learners
4	a^2+b^2	$a^2+2ab+b^2$		✓		3
5	$1/a^2+2ab+b^2$	$1/(a^2+2ab+b^2)$		✓		2

Table 25.3.5 Success rates of questions with W1 Division

Test	Number	Question	% Achieved	% Passed	Number Exited	Sample Size
2	2	x/5=25	83	16	1	75
2	3	x/7=56	77	21	1	75
2	4	x/15=22	77	21	1	75
3	2	26-2n=8	75	22	3	73
2	7	x/3=-15	75	23	3	75
2	1	x/8=12	75	24	1	75
2	8	5x/3=15	67	29	4	75
5	5	2z+5=11	66	31	3	70
2	5	"-x/4=22	65	32	3	75
2	6	x/6=-18	64	33	3	75
3	6	4-2x=10-6x	42	53	4	73
5	6	3z+5=z+3	41	56	3	70

Table 25.3.6 Success rate of questions with brackets

Test	Question No	Question	% Achieved	% Passed	No. Exit
5	1	expand A(BC)	79	21	0
3	5	-(3x-w)	71	25	4
3	4	2(x+3)	68	29	3
3	8	3+21(s+4)	33	60	7
5	2	(ax + b) (cx + d)	33	66	1
3	10	17-21(s+7)	27	64	8
3	9	3+23(s-4)	26	66	8
5	3	(y + 2)/(y - 3) = 2	24	73	3
3	11	13-19(s-3)	21	71	8
6	4	$(A+B)^2$	20	78	2
5	4	(z+1)/(z+4)=5/6	14	83	3

Table 25.3.7 Success rate of questions with indices

Test	No.	Question	% Achieved	% Passed	No. Exit	Exit No.
6	3	2^-2	11	86	3	65
6	5	(A+B)^-2	6	91	3	65

Table 25.3.8 Success rate of substituting values questions

Question Number	Question	% Achieved	% Passed	No. Exit	Sample Size
2	Find 4x when x=6	97	1	1	67
1	Find 2x when x=9	96	1	3	67
3	Find 2x+3 when x=5	94	4	1	67
4	Find -3y+8 when y=2	90	9	1	67
5	"-3z-8 when z=-3	81	18	1	67
6	xy+1 when x=-4 and y=6	78	21	1	67
T3q1	Find 5x when 4x=20	84	15	1	73

Table 25.3.9 Success rate of questions with negative signs and values

	Number	Question	% Achieved	% Passed	Exit	No.
4	4	"-3y+8 when y=2	90	9	1	67
1	1	90 – x = 70	87	8	5	79
4	5	"-3z-8 when z=-3	81	18	1	67
1	5	7-z=9	78	9	13	79
1	2	58-x=17	78	13	9	79
4	6	xy+1 when x=-4 and y=6	78	21	1	67
6	2	-1^3	77	22	2	65
1	6	-9-z=-7	76	11	13	79
3	2	26-2n=8	75	22	3	73
1	3	962-x=476	73	18	9	79
1	4	9-z=17	70	18	13	79
3	4	-(3x-w)	68	29	3	73
1	7	476-z=962	66	20	14	79
2	5	"-x/4=22	65	32	3	75
2	6	x/6=-18	64	33	3	75
3	6	4-2x=10-6x	42	53	4	73
5	6	3z+5=z+3	41	56	3	70
3	8	3+23(s-4)	33	60	7	73
3	10	13-19(s-3)	27	64	8	73
3	9	17-21(s+7)	26	66	8	73
5	3	(y + 2)/(y - 3) = 2	24	73	3	70
3	12	4c+8 = 5c + 10 + ...	18	74	8	73
05	7	(x-5)(x-7)=3	13	84	3	70

Table 25.3.10 Success rate of questions for solving linear equations W6

	Number	Question	% Achieved	% Passed	Exit	No.
1	1	90 – x = 70	87	8	5	79
2	2	x/5=25	83	16	1	75
1	5	7-z=9	78	9	13	79
1	2	58-x=17	78	13	9	79
2	3	x/7=56	77	21	1	75
2	4	x/15=22	77	21	1	75
1	6	-9-z=-7	76	11	13	79
3	2	26-2n=8	75	22	3	73
2	7	x/3=-15	75	23	3	75
2	1	x/8=12	75	24	1	75
1	3	962-x=476	73	18	9	79
1	4	9-z=17	70	18	13	79

	Number	Question	% Achieved	% Passed	Exit	No.
2	8	5x/3=15	67	29	4	75
1	7	476-z=962	66	20	14	79
5	5	2z+5=11	66	31	3	70
2	5	"-x/4=22	65	32	3	75
2	6	x/6=-18	64	33	3	75
3	6	4-2x=10-6x	42	53	4	73
5	6	3z+5=z+3	41	56	3	70
3	8	3+23(s-4)	33	60	7	73
5	3	(y + 2)/(y - 3) = 2	24	73	3	70
5	4	(z+1)/(z+4)=5/6	14	83	3	70

Table 25.3.11 Treefrog Weakness and nonweakness Distribution statistics

	Mean	Std. Deviation	Mean difference	Std. Error Mean
Weakness 1	.6565	.31933		.03712
Not Weakness 1	.8649	.33067	0.2084	.03844
Weakness 2	.2956	.28845		.03331
Not weakness 2	.6220	.21508	0.3264	.02484
Weakness 3	.0733	.24236		.02799
Not weakness 3	.5579	.21899	0.4846	.02529
Weakness 4	.7981	.31517		.03639
Not weakness 4	.4899	.22730	-0.3082	.02625
Weakness 5	.4910	.20100		.02321
Not weakness 5	.5815	.23549	0.0905	.02719
Weakness 6	.4487	.34744		.03934
Not weakness 6	.4383	.25482	-0.0104	.02885

Table 25.4.1 Overview of questionnaire quantitative data

			Total figures				**Percentage figures**		
	Num	Question	Good	OK	Poor	Nil	Good	OK	Poor
Learning	1	Rate the usefulness of the software to help you to correct errors in your understanding	7	37	23	3	10%	55%	34%
	2	How well did the software cover the subject area you need to study?	17	38	12	3	25%	57%	18%
	3	How much did the software accept your method for solving a given question	15	37	17	1	22%	54%	25%
	4	Rate the balance between state of the art graphics features and benefits for learning	9	44	15	2	13%	65%	22%
	5	Do you think the software would help you to carry on doing work on your	13	34	18	5	20%	52%	28%

			Total figures				Percentage figures		
	Num	Question	Good	OK	Poor	Nil	Good	OK	Poor
		own?							
	6	Were mathematical terms used the way you expected?	18	39	10	3	27%	58%	15%
	7	Rate the appropriateness of the representations or presentation of different types of question	13	48	7	2	19%	71%	10%
	8	Did you find the graphics generally helped you to learn or were they poor in that they distracted you?	14	47	8	1	20%	68%	12%
	9	How would you rate the methods of interaction in terms of ease of use, familiarity and functionality?	15	45	6	4	23%	68%	9%
	10	Rate the balance between helpful informative interaction and feedback which intrudes and interrupts progress.	12	35	18	5	18%	54%	28%
	11	To what extent does the environment support movement within questions and from one question to another to help learning?	6	44	11	9	10%	72%	18%
	12	To what extent does the structure of the environment support the notion of self-directed learning?	17	41	8	4	26%	62%	12%
	13	How easy was it to develop an understanding of how to use the system and hence develop the skills to maximise full use the system?	26	33	6	5	40%	51%	9%
	14	To what extent did you recognise the mathematical symbols, icons and buttons used by the software?	24	37	3	6	38%	58%	5%
	15	How would you rate the system's ability to shielding you from usability obstacles such as runtime errors and malfunctions?	19	39	3	9	31%	64%	5%
Usability	16	How would you rate the system's ability to protect you from making annoying errors such as incorrect syntax of statements?	9	34	19	8	15%	55%	31%
Learning	17	To what extent did your methods for solving problems and those used by the system to present questions and feedback match?	8	44	14	4	12%	67%	21%

	Num	Question	Total figures				Percentage figures		
			Good	OK	Poor	Nil	Good	OK	Poor
	18	To what extent does the software enable you to experiment with solutions and develop your own personal understanding?	10	33	23	4	15%	50%	35%

Table 25.5.1 Diagnosys Results of group interviewees

	DT1M	PE1M	SC2F	SC1M	L1F	DTF1	PE2F	MA1M
Total percentage scores per category	GCSE C	GCSE C	GCSE B	GNVQ Merit	GNVQ Pass	GCSE C	GCSE C	A level B
All Questions	38	29	62	56	32	32	18	88
Numbers	53	60	67	67	47	47	13	100
Powers	0	0	67	67	67	67	33	67
Basic Algebra	18	9	64	45	18	18	27	91
Statistics	33	0	33	67	33	0	0	33
Miscellaneous	100	0	50	0	0	0	0	100
Number of questions	31	18	24	25	29	17	31	19
Number Correct	10	5	11	11	8	6	4	15

Table 25.5.2 Treefrog Results of group interviewees

		DT1M	PE1M	SC2F	SC1M	L1F	DT1F1	PE2F	MA1M
Treefrog Test 1	Achieved	7	5	7	7	0	0	5	6
	Passed	0	2	0	0	7	0	2	0
	Wrong	13	14	6	3	4	0	23	20
	Exit	0	0	0	0	0	7	0	1
Treefrog Test 2	Achieved	6	8	7	0	1	1	2	8
	Passed	2	0	1	7	7	3	6	0
	Wrong	9	5	2	3	0	16	18	2
	Exit	0	0	0	1	0	4	0	0
Treefrog Test 3	Achieved	2	2	3	3	0	0	2	9
	Passed	9	10	9	9	11	0	9	3
	Wrong	5	17	9	2	1	0	2	25
	Exit	1	0	0	0	0	0	1	0
Treefrog Test 4	Achieved	0	0	5	0	0	6	4	6
	Passed	0	0	1	0	6	0	2	0
	Wrong	0	0	3	0	0	2	3	0
	Exit	0	0	0	0	0	0	0	0
Treefrog Test 5	Achieved	0	3	0	0	1	0	2	8
	Passed	0	5	0	1	7	0	6	0
	Wrong	0	6	0	1	0	0	1	33
	Exit	0	0	0	7	0	0	0	0

Treefrog	Test 6 Achieved	0	2	0	0	1	0	0	5
	Passed	0	3	0	0	4	0	0	0
	Wrong	0	13	0	0	0	0	0	6
	Exit	0	0	0	0	0	0	0	0
TOTAL Achieved		15	22	20	10	3	7	15	42
		33%	48%	43%	22%	7%	15%	33%	91%
Total Passed or Exited		12	20	11	17	42	7	26	4
Total Wrong Attempts		27	20	45	9	5	18	47	86

Table 25.6.1 Descriptives of final trial Pre test full data set each weakness and non weakness subset

	Full test	weak 1	Non weak 1	weak 2	Non weak 2	weak 4	Non weak 4	weak 5	Non weak 5	weak 6	Non weak 6
Mean	.4727	.6477	.4290	.3836	.5455	.5682	.4489	.4186	.6455	.6193	.3750
Std. Deviation	.27504	.35067	.26747	.24832	.31492	.33790	.27047	.26372	.34327	.34606	.26085
Std. Error Mean	.05864	.07476	.05702	.05294	.06714	.07204	.05766	.05622	.07319	.07378	.05561
Mean difference		-0.175	0.0437	0.0891	-0.0728	-0.0955	0.0238	0.0541	-0.1728	-0.1466	0.0977
Minimum		00	.06	.00	.09	.00	.06	.07	.00	.00	.00
Maximum		1.00	.94	.89	1.00	1.00	.94	.93	1.00	1.00	.92
Variance		.123	.072	.062	.099	.114	.073	.070	.118	.120	.068

Table 25.6.2 Distribution Statistics of Webfrog and Webfrog with feedback Pre test and Post Test

	Mean	Std. Deviation	Minimum	Maximum	Percentiles			Variance
Pair					25th	50th (Median)	75th	
Pre Test Webfrog	9.82	5.741	2	19	5.00	10.00	15.00	32.964
Pre Test Webfrog with feedback	9.09	5.504	1	17	5.00	10.00	15.00	30.291
Post Test Webfrog	10.45	5.733	2	19	5.00	11.00	14.00	32.873
Post test Webfrog with feedback	12.27	5.217	4	19	7.00	12.00	18.00	27.218

Table 25.6.3 Overview of Webfrog Users Log files

Webfrog Users ID	30	1126	4	1352	624	953	1854	1715	521	2289	488
Pre test Score	2	2	5	7	8	10	10	13	15	17	19
Number of errors made	33	0	48	12	28	25	42	11	48	27	9
Number finished	11	0	30	10	9	13	25	0	66	69	36
Post Test Score	2	3	5	8	11	10	11	13	14	19	19
Webfrog with feedback Users ID	654	312	1803	1723	1490	1230	904	827	3107	721	2208
Pre test Score	1	2	5	5	7	10	10	13	15	15	17
Number of errors made	56	41	54	30	31	23	21	8	21	37	27
Number finished	80	18	45	8	17	15	6	62	16	46	72
Post Test Score	7	4	7	10	10	12	13	18	16	19	19

Table 25.6.4 WEBFROG Pre and Post Test corresponding questions

	Post - Pre Q1	Post - Pre_Q2	Post - Pre_Q3	Post - Pre_Q4	Post - Pre_Q5	Post - Pre_Q6	Post - Pre_Q7	Post - Pre_Q8	Post - Pre_Q9	Post- Pre_Q10	Post - Pre_Q11	Post - Pre_Q12	Post - Pre_Q13	Post - Pre_Q14	Post - Pre_Q15	Post - Pre_Q16	Post - Pre_Q17	Post - Pre_Q18	Post- Pre_Q19	Post - Pre_Q20
Z	-.447(a)	.000(b)	.000(b)	-2.121(c)	-1.000(a)	-.447(a)	-1.000(a)	-.577(a)	-1.000(c)	-1.342(a)	-1.000(c)	-1.633(c)	-1.414(c)	-1.000(a)	.000(b)	-2.236(a)	-.577(c)	-1.732(c)	-1.000(c)	-1.414(c)
Asymp Sig	.655	1.000	1.000	.034	.317	.655	.317	.564	.317	.180	.317	.102	.157	.317	1.000	.025	.564	.083	.317	.157

a Based on positive ranks.
b The sum of negative ranks equals the sum of positive ranks.
c Based on negative ranks.
d Wilcoxon Signed Ranks Test

Table 25.6.5 WEBFROG WITH FEEDBACK Pre and Post Test corresponding questions

	Post - Pre Q1	Post - Pre_Q2	Post - Pre_Q3	Post - Pre_Q4	Post - Pre_Q5	Post - Pre_Q6	Post - Pre_Q7	Post - Pre_Q8	Post - Pre_Q9	Post- Pre_Q10	Post - Pre_Q11	Post - Pre_Q12	Post - Pre_Q13	Post - Pre_Q14	Post - Pre_Q15	Post - Pre_Q16	Post - Pre_Q17	Post - Pre_Q18	Post- Pre_Q19	Post - Pre_Q20
Z	-2.000(a)	-.577(a)	.000(b)	-2.449(a)	-1.732(a)	.000(b)	-1.414(a)	-1.000(a)	-1.342(a)	.000(b)	.000(b)	-2.000(a)	-.577(a)	-1.732(a)	.000(b)	-1.732(c)	-1.732(a)	-2.236(a)	.000(b)	-1.000(a)
Asymp Sig	.046	.564	1.000	.014	.083	1.000	.157	.317	.180	1.000	1.000	.046	.564	.083	1.000	.083	.083	.025	1.000	.317

a Based on negative ranks.
b The sum of negative ranks equals the sum of positive ranks.
c Based on positive ranks.
d Wilcoxon Signed Ranks Test

Table 25.6.6 Webfrog and Webfrog with feedback POST TEST Questions

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
Z	-1.732(a)	.000(b)	.000(b)	.000(b)	-1.414(a)	-.577(a)	-.816(a)	-1.000(a)	.000(b)	-1.732(a)	.000(b)	-1.000(a)	.000(b)	-1.000(a)	.000(b)	-1.414(a)	.000(b)	-1.414(a)	-1.414(a)	-1.000(c)
Asymp. Sig. (2-tailed)	.083	1.000	1.000	1.000	.157	.564	.414	.317	1.000	.083	1.000	.317	1.000	.317	1.000	.157	1.000	.157	.157	.317

a Based on negative ranks.
b The sum of negative ranks equals the sum of positive ranks.
c Based on positive ranks.
d Wilcoxon Signed Ranks Test

Table 25.6.7 Final trial number of errors

User	Pre Test	Number of errors	Number correct
654	1	30	8
312	2	41	18
1723	5	80	28
1803	5	54	45
1490	7	31	17
1230	10	23	15
904	10	21	6
827	13	8	62
3107	15	37	46
721	15	21	16
2208	17	27	72
30	2	32	12
1126	2		
1352	5	12	10
4	7	47	30
624	8	28	10
953	10	39	25
1854	10	22	13
I1715	13	7	
2289	15	25	69
521	17	45	66
488	19	9	37

Table 25.6.8 Final trial number of errors

	Pre Test Score – GCSE Grade Coded
Z	-3.923(a)
Asymp. Sig. (2-tailed)	.000

a Based on negative ranks.
b Wilcoxon Signed Ranks Test

Table 25.6.9 Final trial Pre and Post test correlation results

			post1_sco	pr1_scor	pre_scor	post_sco
Spearman's rho	post1_sco	Correlation Coefficient	1.000	.975(**)	.970(**)	.917(**)
		Sig. (2-tailed)	.	.000	.000	.000
		N	11	11	11	11
	pr1_scor	Correlation Coefficient	.975(**)	1.000	.993(**)	.959(**)
		Sig. (2-tailed)	.000	.	.000	.000
		N	11	11	11	11
	pre_scor	Correlation Coefficient	.970(**)	.993(**)	1.000	.952(**)
		Sig. (2-tailed)	.000	.000	.	.000
		N	11	11	11	11
	post_sco	Correlation Coefficient	.917(**)	.959(**)	.952(**)	1.000
		Sig. (2-tailed)	.000	.000	.000	.
		N	11	11	11	11

** Correlation is significant at the 0.01 level (2-tailed).

Appendix 26.1 Preliminary trial adaptations

26.1a Stage 2 Pilot Study – Pre Test, Diagnosys Results

Preliminary study

The participants in the Preliminary trial expressed the view that the Diagnosys test was too lengthy and that it detracted effort from the use of Treefrog. The results shown in appendix 25 Table 25.2.1 show that the time used in the Preliminary study ranged between 27 and 41 minutes. To shorten the length of time required to undertake the Pre Test, mathematical skill areas were removed on the basis of lack of relevance to the six areas of weakness and the level of difficulty. The higher level skills at Level 3 and 4 skills areas except percentages were removed. Other skills removed were selected on the lack of relevance to the focus of the study. The selection of skill areas is shown in the Hierarchy of Diagnosys given in Appendix 7. The maximum number of mathematical skills (as detailed in Appendix 7)across the areas in which questions could be posed by Diagnosys in the Pilot trial was reduced from 48 to 34 as indicated in Chapter 5 Table 5.2.2 in Chapter 5 section 2.1.

26.1b Stage 3 Treefrog

Adaptations following preliminary trial

The data presented in Appendix 25 tables 25.3.1, 25.3.2 and 25.3.3 indicates that the nine participants in the preliminary trial ranged from those who were successful in answering 60% of questions attempted to those who were able to answer 90% of questions correctly. Despite their success feedback from they indicated that they felt that some questions were repetitive and that there were too many questions to answer.

In consideration of learner performance the number of questions within the tests was reduced. The number of questions was reduced to 75% by analysis shown in Chapter 5 table 5.3.1. These results indicated that the seven questions in Test 4 did not pose any difficulties for any of the candidates whilst Test 1 caused few difficulties with only 2 candidates recording an erroneous response to two questions and a further two erroneously answering one question. Hence the number of questions posed could be reduced. However test 4 only consisted of seven questions and test 1 of ten hence limiting the extent of a possible reduction. Tests 2 and 3 proved to be generally achievable with all candidates achieving more than half marks with only one candidate failing to achieve above 70% success in Test 2 and only one failing to achieve at least 65% in Test 3. Tests 2 and 3 consisted of 15 and 17 questions respectively hence an absolute reduction of 7 and 5 questions was feasible. The questions posed in tests 5 and 6 proved problematic for the sample group. One candidate achieved maximum success and another 75% success in test 5 and hence consequently raised the overall success recorded whereas the other six candidates achieved less than 40% success.

The nature of errors and misconceptions is common to similar groups of learners and it is intended to be able to use these results to anticipate errors and hence use these as teaching points. Each of the sets of questions posed were analysed for frequency of common errors relating to syntax, use of negative values and signs, and numerical calculations the findings of which are presented in appendix 25 Table 25.3.4. Duplication of specific types of question was also identified. A number of questions posed within each test were refined depending on these findings.

All the questions posed in Test 1 were maintained within the pilot test.

Four questions were omitted from test 2 for the Pilot. These questions were problematic for only one learner in the Pilot and the associated type of error was duplicated in another question which had recorded a higher commonality of error.

Three questions were removed due to duplication in test 4 and one removed as participants reported that the solution methodology was beyond their recall having not studied mathematics for many years. Hence it could not test their understanding effectively but did cause some anxiety. Within the Pilot Test 4 all participants achieved 100% success rate hence there were no errors to analyse.

All questions posed in Test 5 and 6 were maintained for the pilot study. Hence the number of questions posed within the Pilot was reduced from 62 to 46.

- syntax and formatting of answers,
- negative signs and values,
- brackets
- division

the latter three being common areas of misconception. The regularity of occurrence of these categories of errors was therefore investigated within the pilot.

Appendix 26.2 Analysis of Weaknesses

26.2a Diagnosys

Weakness W1 Division

Within the Diagnosys Pre Test posed there were six skill tests or questions which involved division 109, 112, 206, 207, 213 and 216. Only 7 of the 73 participants were considered successful in Skill 207 Multiplying algebraic fractions. This skill was presented through algebraic questions of the form $\frac{6}{y} \mathrm{x} \frac{y^2}{2}$. A third of the participants were judged successful in tackling Skill 213 which tested transposition of algebraic equations to find an expression for one of the terms such as

```
If   c=[ab/2]   then   b = ?
```

These results suggest that algebraic division was a problematic aspect of mathematics for many of the participants. Due to the hierarchical nature of the software not all participants would have been presented questions representing each of these skills. The fractional representation of many of these questions may also have been a factor in the success rate of these questions.

Skills which relate to weakness W2 Brackets

Within the Diagnosys Pre Test posed there were eight skill tests or questions which involved division (110, Collect terms (simple), 208, Factors of algebraic products, 209, Simple Factorisation, 210, 211, collecting terms, 213, Transposition of formula, 214, Evaluating formula and 215, Precedence Rules). Out of a group of 73 participants the number of those who were successful in response to these skills varied between 33 and 17. Skill 110 achieved the highest success rate whilst skills 209 and 215 recorded the lowest success rate. Skill 110 presented questions of the form

```
Collect terms in the following expression
{a} (2x-3y)+1+y+(4x+5)
```

skill 209 was tested by questions of the form

```
Factorise the following, taking out the highest factor
possible:
{2z-6z²}
```

and skill 215 questions of the form

```
To calculate  [15 + 27] / 3
you press a sequence of keys on your calculator. Which
one of the following would give the {WRONG} answer?
```

The success rate for these questions suggests that manipulating brackets with mathematical operators and rewriting expressions in the simplest form using brackets was problematic for many students. The factorising in skill 209 does require understanding of the terminology and the use of factors as well as the manipulation of brackets. Removing brackets from an expression in skill 110, where the use of brackets did not effect the meaning of the expression, was less problematic but this question also involved the handling of two different algebraic terms.

Skills which relate to weakness W3 Exponentials

There is only one Diagnosys skill, 208 which tests the ability to answer questions which involve exponentials. This question was presented to all participants. Twenty two participants out of a group of 73, 30% were successful in tackling questions of the form:

Find the missing term $12x^3y=3x^2y(?)$

Overall success rate for all other skills is 48% for all participants. This result infers that that handling exponentials was problematic for a large number of the participants in the sample group. The comparison of mean scores for weakness 3 and Not Weakness 3 compares the result for this one skill with 33 other skills. Hence the data set Not Weakness 3 encompass a broad range of mathematical concepts and potential errors.

Skills which relate to weakness W4 Substituting values

There were two Diagnosys skills (Skill 113, Evaluating a simple expression and Skill 214, Evaluating formula) which tested the ability to answer questions which involve substituting values. Skill 113 was tested by questions of the form

$6+2p$ if $p=4$

and skill 214 by questions of the form

$Q=[p^2+2r]/[t-1]$ and $p=4, r=-2, t=5$

Of the seventy three participants sixty four were successful in skill 113 and thirty in skill 214. These results imply that substituting values in simple expressions was not problematic for the majority of the sample group. However substituting values into more complex expressions including division, brackets and negative signs and values was more difficult for the sample group.

Skills which relate to Weakness 5 Negative signs and values

There were four Diagnosys skills (101, Multiplication of negative numbers, 102, Multiply negative and positive, 103, Negative Numbers and 213, Transposition of formula) which directly involved the handling of negative signs and values. The total number of successes of each of these skills out of a sample group of 73 was 57, 65, 68 and 27 respectively. These results indicate that transposing a formula is problematic for many respondents and this affected the mean score for this weakness for a large proportion of the sample group.

Skill which relates to Weakness 6 Linear equations

Only two skills (111and 212) relate to Solving linear equations. At level 1 typical questions posed were of the form 3x + 1=13. Questions at level 2 could also include brackets, division or negative signs and values being of the form 7 - 3c = -5c - 4. Success rates were 75% for skill 111 and 21% for skill 212 indicating that the complexity of the question with the inclusion of negative values impacted upon the participants' ability to answer the question correctly. Simple linear equations were not problematic for the majority of the sample group.

26.2b TREEFROG

Weakness 1 Division

The questions and success rates in responses to each within this data set are detailed in Appendix 25 table 25.3.5. Three questions required the respondent to transpose the division into multiplication to calculate a value for x these achieved the highest success rate. Questions which required the transposition of multiplication into division were more problematic with success rates dropping from 83% to 67%. The complexity of also including a negative value resulted in 65% pass rate. Question 2 of Test 2 (x/5=25) required the respondent to transpose the division into multiplication to calculate a value for x rather than transposing multiplication into division. As these operations are the inverse of each other similar levels of success would be expected, in fact 84% of 73 achieved the former and 83% of 75 the latter.

The greatest level of difficulty with pass rates of only 42% and 41% of the sample group related to equations which required transposition terms and division of terms.

W2 Brackets

A significant reduction in achievement for the previous set of questions, from (68-79% to 14-33%) was recorded for the set of questions as presented in Appendix 25 Table 25.3.6. This set of results indicates that the majority of the sample who answered the questions posed were unable to achieve success. Question 2 in Test 5 and Question 4 in Test 6 required an expression within a bracket to be multiplied by another expression within another bracket. However questions 10 and 9 from Test 3 both included the product of a numeric with an algebraic linear expression in brackets combined with the subtraction or addition of a number yet only achieved 27% and 25% success rates respectively from a sample of size 73. In comparing question 11 of Test 3 (which encompasses two negative signs in the product of a numeric with an algebraic linear expression in brackets combined with the subtraction of this product from a numeric) with question 8 of Test 3 (which involved only positive terms) the success rate differed from 21% to 33% of 73. This finding suggests that students do require more support in understanding negative terms.

Questions 3 and 4 of Test 5 were similar both encompassing a fractional algebraic term and a numeric. The numeric term within the second question was expressed as a fraction rather than integer and the success rate being 10% lower in this instance. Other questions posed which are expressed fractionally were not algebraic but numeric, see table 25.3.5. The success rate for answering these questions ranges from 83% to 75% if the equation only includes one division operator to be transposed. However on including an additional multiplication operator to transpose, as in question 8 the success rate drops to 67%. Other variations in achievement rates could be attributed to the associated difficulties of multiplying different integers. For instance, 56 x 7 and 12 x 22 may be perceived to be more difficult.

However three questions (Test 5 question 1, Test 3 question 5 and Test 3 question 4) which were less complex tested the respondents understanding of the conceptual significance of brackets. The second, involving the subtraction of a positive and negative term within the brackets. The third, encompassing multiplication of an algebraic and a numerical term within the brackets. Similar rates of success of around 70% were recorded however as expected the worked example question (expand A(BC)) recorded the highest level of achievement.

W4 Substituting values

As outlined in Appendix 25 Table 25.3.8 six questions posed required the respondent to substitute a numerical value for x into an expression. A success rate of at least 90% was achieved within the sample group who attempted this set of questions for four of these questions. These findings suggest that the majority were able to perform this operation accurately. Questions which only required the use of one operator, multiplication were slightly more successful than when the use of two operators, multiplication and addition was required. The introduction of the concept of a negative quantity showed a slight reduction in success from 94-97% to 90% when used and 81% and 78% when negative values and subtraction were combined. Responses showed at reduction in success rate of at least 9% which translates to six candidates.

Another question which included substitution was presented in Test 3 (Find 5x when 4x=20) which was attempted by a larger sample group of 73 but only 84% of which achieved success with 15% Passing. This question had the added requirement of division given the value of a multiple of x rather than directly giving the value for x.

Negative signs and quantities

Twenty three questions of the forty six presented involved negative signs or quantities. An overview of the results of this subset of questions are given in Appendix 25 Table 5.3.9. In an expression (xy+1 when x=-4 and y=6) which encompasses two operators, multiplication and addition as well as two substitutions the success rate obtained was 78% of 67 respondents. However when the question also included brackets only 33% of respondents were successful. However in another question with brackets but without a negative quantity there was the same success rate. These findings suggest that learners may not require more help to understand negative quantities as suggested by the QCA. The results also show that complexity of questions can impact upon achievement

26.2c Final Trial Pre Test

Final trial analysis

The results presented in answering questions which encompass division or weakness 1 are presented in Appendix 24 Table 5. These results related to weaknesses 1 indicate that question 15 which included division was the most problematic question with less than a third of the full group answering the question correctly. Other questions which included division were answered by at least two thirds of the group. The solution of Question 15 "Find the value of z when 3z + 5 = z + 3" involved transposition and negative signs as well as division.

Brackets were problematic for the majority of the group with the majority erroneously answering six of the nine questions. These questions were ranked in the lowest seven.

Questions which encompassed weakness 4 the substitution of values varied in terms of difficulty. All questions in this subset were accurately answered by the majority of the group. The greatest difference in success rate being between 15 out of 22 participants answering "Find 4x when 5x = 20" rather than only 10 correctly responding to "If 2x + 3y = 7 find a value for 4x + 6y.

Weakness 5, Negative signs and values, was encompassed into twelve of the twenty questions with varying levels of complexity. Question 20 was the poorest answered question with only 3 out of the group of 22 correctly responding whereas eighteen answered question 4 correctly. Seven of the group answered at least eight of the twelve questions correctly whilst eight were unable to answer more than three. Question 20 combined the use of a negative sign with brackets. Question 4 encompassed the use of a negative quantity with multiplication of two other terms.

Questions which involved Solving linear equations, weakness 6, were generally answered well by the sample group. Detail regarding the success of responses per participant to this subset of questions is given in Table 24.5. Fifteen of the twenty two answering at least 4 of the 8 questions correctly. Eight participants answered at least three quarters of these questions correctly.

Appendix 26.3 Software comments and errors made

26.3a Stage 4 Software evaluation questionnaire analysis

Evaluation of Software for Learning

The quantitative data presented in appendix 25 Table 25.4.1 indicated that in all cases less than 30% of the sample group considered the software to be 'Good' for learning. However only in one regard (Question 1) did more than a third of those who responded consider the software to be 'Poor'; furthermore at least a fifth of the responses to questions 3, 4, and 5 rated the software as poor.

Question 1 caused a significant level of qualitative response amongst participants many expressing the view that the lack of guidance regarding the nature of the error, the method of solution and the correct answer was poor. These responses are consistent with the responses obtained quantitatively.

Similarly twelve participants stated in response to question 5 that they found the software unhelpful to use due to the limited feedback provided. Conversely ten others stated that they believed that the software would enable them to continue working on their own and a further three suggested that they also would do so if the software provided more feedback.

Few comments were returned in response to question 2 and most were not fully relevant. The quantitative data given in appendix 25 Table 25.4.1 was inconclusive in indicating opinion. These two findings suggest that the learners were not best placed for answering this query. Four participants when asked about the software's acceptance of their method of solution (question 3) stated their disappointment that in their view there was no provision for working out. However within the interface the input window could function as a working area as specific responses were not required only equivalence. This suggests that some users were not adequately prepared in the use of the software. In a few incidences this attitude was repeated in responses to other questions. One respondent stated that the software always accepted their method of solution and two others stated that they were required to use different methods whilst a further two respondents stated that their methods of solution were accepted. Due to the design of the structure of the objects within Treefrog it is not possible for the software to reject any response which is valid and hence require a specific method of solution. The contradictions within these responses and the normal distribution of the quantitative data suggest that these findings were also inconclusive.

Responses to question 4 appear mixed with several expressing the view that the simplicity and plainness of the interface was beneficial whilst conversely the responses of several others was that the interface should be improved to be more 'interesting' and easier to use. Once again a few respondents have stated their opinion that the feedback was inadequate as it did not provide support to amend errors. However many comments made no reference to the effect upon learning. Furthermore a variety of opinions regarding the level of support and distraction presented by the graphics within the package were recorded in response to question 8. Several participants stated that the graphics were not distracting but two respondents felt that they were. However a few stated that the interface was unlikely to support or motivate learning. These findings appear inconclusive.

There was a limited number of commentary responses to question 7 regarding the representations used suggesting that participants were without a strong opinion. In response to question 6 eight participants specifically stated that they were not familiar with ^. Not all participants attempted the set of questions involving indices. However there was again a response regarding the requirement for supportive and relevant feedback.

Evaluation of software Usability

Full results of the survey quantitative and qualitative comments are presented in Appendix 17. The quantitative data presented in appendix 25 Table 25.4.1 relating to a range of opinions it is indicated that in two respects (Questions 10 and 16) at least a quarter of those who responded considered the software to be 'poor', in four respects (Questions 9, 13, 14 and 15) less than a tenth considered the software to be poor and furthermore at least a tenth of the responses to questions 11 and 12 rated the software as poor.

Responses relating to questions 9, 13, 14 and 15 suggest that 90% of the users rated the software as at least satisfactory and with regard to questions 11 and 12 at least 80% satisfactory. Answers to questions 10 and 16 show a high level of dissatisfaction with the helpfulness of the interaction including that relating to the required syntax of the statements made. The comments given relating to these questions indicate that these trainees found the feedback provided by the software unhelpful not because it was disruptive or distracting but because it was limited and did not clearly indicate the required format of responses.

With respect to question 9 a range of comments were received from those who felt the software was easy to use to those who clearly had felt some frustration in using the software. These comments suggest that these trainees had not used the available worked answer for the first question in each test to direct and instruct them on either the syntax of questions or use of the software. The comments relating to dissatisfaction in using the system as queried in question 13 were similar. One individual comment did indicate that there was a lack of understanding of the difference between an equation and an expression. Only 9% of trainees grading of question 9 and 13 supported the view that communication with the software was poor. Responses to question 14 regarding to the use of symbols indicated two potential difficulties, the use of ^ to express indices and the division symbolic representation. The latter could be expressed in both formats and the former changed to utilise a conventional representation. The comments recorded in response to question 15 suggest that the trainees did not fully understanding the focus of the question but were considering the general usability of the software. In fact one participant did comment in the 'Other' section that they had experienced a runtime error but this comment was not repeated by any other candidates.

With regard to the usability of the environment a range of comments were received including those who felt that the network of questions posed was good as the nature of these questions was progressive and that it was beneficial to see a worked example. However several commented on the weakness of not being able to navigate backwards within questions posed, the method of moving on was clumsy and once again better explanations of errors would be beneficial. However at least 80% indicated that they were satisfied with the usability of the system.

Evaluation of Software for Learning and Usability

Two questions posed related to the evaluation of the usability and learning of Treefrog. The results of these two questions, as given in Appendix 25 Table 5.4.1, are skewed towards the users lack of satisfaction. Recurring comments regarding the lack of guidance and instruction provided from the feedback are dominant. This suggests that the learners made a link between obtaining helpful guidance and enabling personal experimentation to further their understanding and development.

Some responses also indicated the misconception that the software required the users to utilise a specific method of solution and these comments may relate to the syntax requirements of the system. For instance the awareness of the difference in requirements for an equation starting with a format of "x = .." and an expression which does not require this start is necessary.

A couple of responses indicated the preference for a working out area despite the provision that the solution window could be used to undertake all steps. However some students did indicate that they appreciated the ability to work at their own pace, questions could be retried and some feedback was received instantly.

Other Evaluation Comments

All respondents were provided with the opportunity to give general comments relating to the experience of participating in the trial. Responses ranged from "I found the program enjoyable and useful and being able to retry the questions" to "disheartening". Several comments focused on the inadequacies of the feedback reiterating those stated earlier.

26.3b Stage 5 Group Interview

Profile of Research interviewees

The mean Diagnosys result for the whole sample group was 44% whereas the mean for the seven low attaining students as defined in Chapter 4 section 4.1.3 was 38%. The group also included one interviewee with a Grade B in A level Mathematics and a Diagnosys score of 88%. Appendix 25 table 25.5.2 shows that the group were most successful in Number achieving a mean score of 51%. It should be noted that one candidate achieved a significantly lower mark of 13 which reduced this mean down from 57%. This candidate achieved the lowest overall mark having been asked 31 questions only 4 were answered correctly. Basic algebra was attempted by all seven participants but was considerably less successful than number in that the overall mean achieved was 28% almost half that of number. Four candidates achieved less than 20% in this section.

Behaviours and view of the system.

Candidate Sc1m who achieved a GNVQ Merit did not undertake all the tests or questions, not attempting Tests 4 and 6, exiting from Test 5 and passed on many questions in Test 2 and 3 hence obtained a low score of 10 marks. Similarly candidate Sc2f who achieved a Grade B at GCSE did not undertake Test 4 and passed on many questions in Test 3 and 5. This candidate also made numerous wrong attempts at solutions – 14 in test 1 and 17 in test 3. This candidate achieved a score of 22. However the former candidate had made a few erroneous attempts. Comments from Sc1m during the interview indicate that he did not enjoy using the software and a wider choice of question types would have been beneficial. He also agreed with others that explanations based on the error being made would have been helpful if guidance was given regarding the nature of a correct response. Candidate DT1m made limited attempts to tackle questions posed in Treefrog. Every question in Test1 was exited from and Tests 3, 5 and 6 were not attempted. Test 4 achieved 100% success however in Test 2 many erroneous attempts caused most questions to be uncompleted. During the group interview this candidate clearly stated their view of the weaknesses of the system. DT1m believed that feedback and explanations focusing on the error made should have been provided, the system should be user centred hence providing more choices for the user in terms of mathematics area studied and if more help was wanted. Also stated was a view that the basic appearance of the interface was a weakness.

Group Interviewees errors made and mathematical ability

Candidate Sc2f who achieved 48% despite erroneous attempts at solutions in Tests 1, 2 and 4 achieved high levels of success. However this participant did not attempt Test 5 and 6 and had limited success in Test 3 selecting to Pass on each question if an erroneous attempt was made. This candidate had stated in the interview that learner confidence was effected negatively by doing these tests and that the experience was de-motivating for those with low level mathematics ability. As Sc2f achieved 62% in Diagnosys which was higher than the mean result for the full sample group and had a Grade B at GCSE to describe herself as being of low ability indicating a poor 'self mathematics concept'.

L1f, the candidate with one of the lowest prior mathematics qualification obtained the lowest Treefrog score. This candidate struggled with all categories of question achieving totals of 0, 1, 0, 0, 1, 1 for each test respectively. This candidate achieved 32% in Diagnosys when the mean for the full sample group was 44%. However Pef2 a candidate with GCSE Grade C achieved 5, 2, 2, 4, 2, and 0 for each test respectively but only 18% in Diagnosys. This candidate attempted all questions posed in Treefrog regardless of encountering erroneous attempts. There does not appear to be a correlation between number of erroneous attempts and behaviour regarding passing on questions or exiting from tests or non-attempting of Tests. Both of these candidates commented that feedback which provides guidance on the methods of solution and error made should encompass examples and demonstrations. They suggested that help should be available on request as well as automatically if errors were made.

From analysis of the responses given in Appendix 25 Table 25.5.1 the patterns of response presented in Chapter 5 Chapter 5 Table 5.5.1 were obtained. From these the following opinions have been identified. All the group felt that the feedback given to users should enable errors to be amended and that methods of solution should be outlined. The group felt that the feedback would be most beneficial if it gave help on specific errors made. This support could be progressively extended and complemented by introductions explaining the nature of problems posed and encompassing demonstrations and example solutions. Furthermore several of the group suggested that a full summary of performance in each test would be beneficial to the learner. This summary could provide details of the number and type of question successfully solved as well as details of those not attempted or those unsuccessfully attempted with an outline of the nature of the error. The group proceeded to outline how they had felt that the experience was de-motivating and reinforced poor 'self mathematics concepts' attributable to prior mathematical experiences. They believed that all users required some success to motivate and encourage them learn. A few interviewees suggested that the learners may assume more ownership of their learning if they were given more control by selecting areas and with the ability to select more help and support.

This group were unable to see the connection between Numeracy and Basic Algebra and believed that an understanding of the links would aid their understanding. It was suggested that through explanatory screens this could be provided within the package. Many felt that the system was not sufficiently user friendly in terms of the colour and layout of the interface as a large block was of one background colour and the fonts selected were glaring on the eye and not always easily readable. The students would have preferred a more 'interesting' interface.

26.3c Stage 6 Final trial Errors made

Order of operators (essential to W1, W2, W5 and W6)

Responses to 33 questions by seven candidates in the sample group of 11 users of Webfrog with feedback and responses to 21 questions by ten users of Webfrog indicated the common misconception that operators have equal order of priority and hence expressions can be simplified by the order in which they are written. Subsequently all Webfrog with feedback users responded appropriately to the feedback and proceeded to amend the erroneous response with 80% leading to correct answers and 20% where the user did not proceed to answer the question accurately. Webfrog users amended accurately responses to 30% of questions in which they had responded with this error.

Brackets *(w2)*

Six candidates out of the 11 users of Webfrog with feedback displayed some confusion over the symbolism of brackets. Four of these (312, 1803, 721 and 3107) displayed the error only in initial questions and amended responses appropriately, two others (1723 and 654) were not consistent in their response to the error message. Errors displayed show a lack of understanding of the representation and significance of brackets. Three candidates showed a misunderstanding of the need to multiply all parts of the brackets by the coefficient outside. A further error, the priority to calculate the brackets first was shown by all six candidates. Of the thirty instances of this error the feedback given to users led to 70% of subsequent responses being correct answers to questions. Four users of Webfrog displayed a difficulty with the use of brackets and then without suitable feedback proceeded to pass on the question. Subsequent questions with brackets were passed.

Negative values and signs (w5)

Seven users of Webfrog gave erroneous responses to questions which involved an error relating to negative values and subtraction. Two Webfrog users then proceeded with input which corrected this error on every occasion, three users proceeded with corrected input on some occasions and the other users did not proceed to respond accurately in any question where the response included an error related to negative values. All of the eleven Webfrog with feedback participants displayed an error regarding the interchanging of negative and positive signs for values. Some of these errors could be mathematical 'slips' whereas some could indicate misconceptions. The following 6 candidates only had success following the error message 3107 in 6 questions, 1723 and 2208 in 3, 827 in 2 and 1490 and 1230 in 1. Participants 654, 1803, 721 showed a mixed response to the feedback. Candidate 654 was unsuccessful in 2 questions but successfully amended the responses in 6, similarly 1803 and 721 were unsuccessful in one question but successful in five. One candidate 312 recorded 2 problem questions and in answering these they did not amend the error in response to the given feedback. There were 32 responses to questions out of 38 where the error messages related to sign errors, which is an 84% success rate.

A typical unsuccessful answer, related to a candidate's misunderstanding that subtraction is not commutative, was recorded in answering Q9 which required the candidate to respond -4. Initially the candidate transposed the digits 5 and 1 to respond 5-1 but in this recorded answer the error message 'sign error' failed to adequately guide the candidate with regard to their erroneous input.

One other error relating to the misconception that subtraction is not commutative was shown in inaccurate transposition of an equation. For instance this was noted when candidate 721 failed to reduce transpose the equation to the answer s=1 with an initial manipulation error of 3+s+6=2. The error messages of sign error on the left-hand side and sign error on the right hand side did not adequately support the learner in amending their responses. Other candidates did not present this error.

Equation format errors (w6)

Six Webfrog with feedback candidates made an error relating to the format required for answers to equations. A typical error made was

3+2=5

to calculate the value of the variable s when the expected response was

3+2=s

5=s

This indicates a weakness with regard to understanding of the terminology equation and expression. There were ten instances of this error seven of which were correctly amending following the feedback message.

Calculation errors

All Webfrog with feedback participants received feedback with regard to calculation errors within at least one question. This feedback was the most frequently given by the system with 198 instances recorded by Webfrog with feedback users and 154 by Webfrog users.

This feedback corresponds to three types of responses. Firstly the expected error where a participant has miscalculated numerically for instance

$3x=66$

$x=33$ rather than $x=22$

Most of these instances where invoked were the user had erroneously attempted to complete the answer to a question in one step. A variety of errors could be involved including a combination of more than one error hence no other feedback was possible. Where this occurred the error message was successful in prompting the user to reattempt the question. In the worst cases candidates responded in a 'trial and error' manner rather than approaching the question stepwise as recorded in other responses. Other instances of this feedback correspond to where the user has made errors which relate to misunderstanding order of operators, brackets or the use of negative values and subtraction or a combination of these. The mechanism for determining this type of error does not encompass where there is more than one error or misconception.

There are some instances where participants confused operators such as addition and multiplication. This type of error could have been anticipated and hence more guidance could have been fed back to the user.

There were 46 responses within Webfrog with feedback to questions (66%) which resulted in this feedback which guided the user appropriately to achieve a correct answer and 22 where the feedback was not sufficiently effective. This error was recorded in responses to 56 questions by Webfrog users, 20% or thirteen of these were then answered correctly. The frequency of the occurrences of this error varied from once to seven times in a response to a specific question.

Syntax errors

Seven times Webfrog users erroneously failed to input an answer but selected to check the answer. This error was amended with the support of other students using software or the researcher. One further syntax error made was of an erroneous expression whereby two operators were adjacent in 1-+2+4 in response to +1-2(-1-2). Without guiding feedback the user did not correct the response but repeated the same response and then passed on the question.

Six Webfrog with feedback participants made errors relating to software syntax in that expressions were incomplete such as 10- or no response was inputted but the check button invoked. Each instance of this was correctly adapted following the feedback guidance that a syntax error had been made. Participants may have mistakenly or prematurely pressed the check button.