

Durham E-Theses

*The Assessment Of A New Approach To Learning
Number To Achieve Arithmetical Automaticity Based
On The Use Of Dedicated Manipulatives*

ROBERT SHAW SCOTT

How to cite:

SCOTT, ROBERT SHAW (2014) The Assessment Of A New Approach To Learning Number To Achieve Arithmetical Automaticity Based On The Use Of Dedicated Manipulatives. Doctoral thesis, Durham University.

Use policy



This work is licensed under a [Creative Commons Attribution Non-commercial No Derivatives 2.0 UK: England & Wales \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/2.0/)

THE ASSESSMENT OF A NEW APPROACH TO LEARNING NUMBER TO ACHIEVE ARITHMETICAL *AUTOMATICITY* BASED ON THE USE OF DEDICATED MANIPULATIVES

by

Robert Shaw Scott

ABSTRACT

The original aim of this research was to determine whether or not a new approach to learning the basic processes of arithmetic, based on the use of dedicated manipulatives, would produce statistically significant improvements in automaticity – *the instant and accurate recall without any conscious mental effort* of previously memorised number facts. However, it was found that memorising number facts was no longer being emphasised in the participating schools of Co. Durham or Edinburgh.

The reasons identified were:

- A strongly established preference for teaching analogous procedures to calculate number facts, based on understandings of first principles.
- A general conviction that good literacy confers greater long-term benefit than good numeracy does.
- A general lack of appreciation of the potential contribution of good automaticity in improving number attainments.
- Insufficient time for memory work in overloaded curricula.

However, the new approach to learning arithmetic, using physical manipulatives, produced highly significant gains (at the 99% level) in Mental Arithmetic and General Maths, as measured using the InCAS computer adaptive programme for five to 11 years old pupils over their early years of formal number learning.

Five schools in Co. Durham and seven in Edinburgh were involved at some stage with 545 children being assessed initially, while 299 started in the Empirical Study. They attended six schools, being three each in Co. Durham and Edinburgh. Comparisons by location and also by gender were made as secondary questions.

Two Swiss schools, with a total of 23 children, were similarly assessed. Their results were not included in the Study, but they were used in terms of contextualising understandings.

The case for automaticity was made throughout the Study in the participating schools. The need for more research into the effectiveness of manipulatives in improving number attainments was identified in the literature.

**THE ASSESSMENT OF A NEW APPROACH TO
LEARNING NUMBER TO ACHIEVE ARITHMETICAL
AUTOMATICITY BASED ON THE USE OF DEDICATED
MANIPULATIVES**

BY

ROBERT SHAW SCOTT

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

AT

THE SCHOOL OF EDUCATION, DURHAM UNIVERSITY

2014

Table of Contents

	Abstract	1
	Title Page	2
	Table of Contents	3
	List of Tables	8
	List of Figures	10
	List of Illustrations	11
	List of Abbreviations and Meanings	12
	Statement of Copyright	14
	Acknowledgements	15
	Dedication	16
	Preface	17
	Part 1	19
Chapter 1	Thesis Background	19
1.1	Introduction	19
1.2	Initial Personal Reflections	19
1.3	Sumdials' Origins and Approach	20
1.3.1	Sumdials' Origins	20
1.3.2	Sumdials' Approach	21
1.3.3	Comments on the Sumdials' Approach	25
1.3.4	Summary	27
1.3	Personal Reflections (Cont.)	27
1.4	Research Questions	29
1.5	Pre-Research Studies	30
1.5.1	Introduction	30
1.5.2	Bramald's Study	32
1.5.3	Subsequent Longitudinal Study	34
1.5.4	Qualitative Study	39
1.5.4.1	Qualitative Study Observations	40
1.5.4.2	Qualitative Study Conclusions	41
1.5.5	Teachers' Scores	41
1.5.6	Pre-Study Summary	43
1.6.1	Background Conclusion	43
Chapter 2	Literature Review	44
2	Introduction	44
2.1	A Background to Arithmetic Learning	44
2.1.1	The Role of Arithmetic	45
2.1.2	A History of Teaching Arithmetic	45
2.1.3	Epistemology and Culture	47
2.1.4	Numbers and Their Representations	50
2.1.4.1	Verbal/Syntactic Systems	52
2.1.4.2	Imagistic Systems	52
2.1.4.3	Formal Notational Systems	52
2.1.4.4	Planning, Monitoring and Executive Control	52
2.1.4.5	Affective Representations	53
2.1.5	Modes of Representation	54

2.1.5.1	Uses of Representation	55
2.1.6	Manipulatives in Learning Number	57
2.1.7	Teaching Orientations	59
2.1.7.1	Connectionist	59
2.1.7.2	Transmission	60
2.1.7.3	Discovery	61
2.1.8	Attitudes to Manipulatives	61
2.1.9	The Mathematical Enterprise	63
2.1.10	Manipulatives and Learning Arithmetic	64
2.1.10.1	Building Blocks	64
2.1.10.2	Dienes Materials	65
2.1.10.3	Cuisenaire Rods	66
2.1.10.4	Number Lines	66
2.1.10.5	Comment on Manipulatives	67
2.1.11	Teacher Issues with Manipulatives	67
2.1.11.1	Concerns about Manipulatives	67
2.1.11.2	Positive Evidence	68
2.1.11.3	Benefits of Manipulatives	69
2.1.12	Conclusion	69
Chapter 3	Methodology Considerations	70
3	Introduction	70
3.1	The Requirements	70
3.1.2	Methodology	70
3.1.3	Choice of Methodology	74
3.1.4	The Sample Design	76
3.1.4.1	Study Duration	77
3.1.4.2	Sample Structure	77
3.1.4.3	Secondary Questions	77
3.1.5	Measurement Issues	77
3.1.6	InCAS – its Program and Applicability	84
3.1.7	The InCAS Process	87
3.1.8	InCAS Conclusion	89
3.2	The Actualitié	90
3.2	Introduction	90
3.2.1	Background	90
3.2.2	An Alternative Control	93
3.3	Ethics and Data Protection	94
3.4	Conclusion	95
3.5	Chapter Conclusion	95
Chapter 4	The Results	96
4	Introduction	96
4.1	Result's Abbreviations and Formats	98
4.2	Pilot Study Results	99
4.3	Empirical Study Results	102
4.3	Introduction	102
4.4.1	Main Question	102
4.4.2	Pilot Study Cohort	106
4.5	Secondary Questions	108

4.5.1	By Gender	108
4.5.2	By Location	112
4.6	School Starting Ages	116
4.7	Ability Categories	119
4.8	Conclusion	122
Chapter 5	The Discussion	123
5	Introduction	123
5.2	Discussion of Empirical Study Results	124
5.2	Introduction	124
5.2.1	Considerations	125
5.2.2	Alternative Analysis	127
5.3	Statistical Results	130
5.3.1	Contributions of the <i>Sumdials</i> ' Approach	132
5.3.2	Developed Ability	133
5.3.3	Assessing Automaticity	134
5.4	Memory Work	135
5.4.1	First Principles Argument	135
5.4.1.1	Comments	138
5.4.2	Lack of time Argument	140
5.5	Initial Conclusions	140
5.6	Secondary Questions	141
5.6.1	By Gender	141
5.6.2	By Location	142
5.6.2.1	The Settings	143
5.6.2.2	Ages at Test	144
5.6.2.3	InCAS	146
5.6.2.4	Teachers	146
5.6.2.5	Conclusion	146
5.7	Other Results	147
5.7.1	Ability Categories	147
5.7.2	Starting Ages	148
5.8	Conclusions	150
The Bridge		
Chapter 6	The Swiss Study	151
6	Introduction	151
	Part 1. The Swiss Results	151
6.1.1	Background	151
6.1.2	The Results	152
6.1.3	Discussion of the Swiss Results	154
6.1.4	Initial Reflections	156
6.1.5	Initial Conclusion	157
	Part 2. Reflections	157
6.2.1	Inquiry	157
6.2.2	Conclusion	161
Part 2		
Chapter 7	Action Research	162
7	Introduction	162
7.1.1	Considerations	162

7.1.2	Classification	164
7.1.3	Interim Conclusion	166
7.2	Research and Action Research	166
7.2	Introduction	166
7.3	Planning with Action Research	167
7.3.1	Background	168
7.3.2	Conclusion	172
7.3.3	Reflections	172
7.4	Contribution of Action Research	172
7.4.1	Background	172
7.4.2	Action Research	173
7.4.3	Sceptical colleague	174
7.4.4	Conclusion	176
7.4.5	Reflections	176
7.5	Study Limitations	176
7.5	Introduction	176
7.5.1	Background	176
7.5.2	Study Limitations	178
7.5.2.1	The Main Question	178
7.5.2.2	The Study Structures	179
7.5.2.3	Mental Arithmetic Module of InCAS	180
7.5.2.4	Teacher Capabilities	182
7.5.3	Conclusion	183
7.6	Action Research and Educational Inquiry	184
7.6	Introduction	184
7.6.1	Background	184
7.6.2	Action Research	184
7.6.3	Comment	186
7.6.4	Conclusion	187
7.7	Action Research and Personal Learning	187
7.7.1	Action Research	187
7.7.1.1	Intent	188
7.7.1.2	Process	188
7.7.1.3	Audience	188
7.7.2	Learning Curves	188
7.7.3	A Practical Example	189
7.7.4	Conclusion	190
7.8	Reflections on Action Research	190
7.8.1	Background	190
7.8.2	Classifications	191
7.8.3	Comparisons with Action Research	192
7.8.4	Conclusion	193
Chapter 8	Reflections on the Literature Review	194
8.1	Introduction	194
8.1.1	Background	194
8.2	Reflections	196
8.2.1	Comments	198
8.2.2	Interim Conclusion	199
8.3	Le Dénouement	199

8.3.1	Developing Automaticity	200
8.3.2	Comment	202
8.3.3	Confirmation	203
8.4	Conclusions	203
Chapter 9	General Review	204
9.1	Introduction	204
9.2	History	204
9.2.1	Top-Down	205
9.2.2	Starting Ages for Formal Learning	206
9.2.3	Conclusion	207
9.3	Some General Observations	208
9.3	Introduction	208
9.3.1	Government Control	208
9.3.2	Teaching Number	211
9.3.3	Interim Conclusion	215
9.4	The Contribution of Arithmetic	216
9.4.1	Understanding/First Principles	218
9.4.2	Insufficient Time	219
9.4.3	Boring	219
9.4.4	Lack of Understanding	219
9.4.5	Teacher Capabilities	219
9.4.6	Interim Summary	219
9.5	The Contribution of Automaticity	220
9.5.1	Memory Conservation	220
9.5.2	Automaticity and Mental Arithmetic	223
9.5.3	Possible Explanations	225
9.6	Memory Work	227
9.6.1	Hardwiring Number Facts	227
9.6.2	Recollections	231
9.6.3	Observation	233
9.7	A Remedy	233
9.7.1	Scaffolding	233
9.7.2	Framework	235
9.8	Contribution of Number Resources	237
9.9	Developing Number Mastery and Automaticity	238
9.9	Introduction	238
9.9.1	Explanation and Personal Reflection	238
9.9.2	Conclusions	240
Appendix 1.1	Entry-level Resource Instruction Sheet	244
Appendix 1.2	Quality Study Interview Questionnaire.	245
Appendix 5.3	School Activity Analyses	246
Appendix 6.4	Number Readiness Checklist	247
Bibliography		248

List of Tables

1.1	Comparison of Group Numbers	36
1.2	Mean Basic Arithmetic Processes Scores	37
1.3	Treatment Group Scores	38
1.4	Control Group Scores	39
1.5	Students' Descriptive Statistics	41
1.6	Teachers' Descriptive Statistics	42
1.7	Combined Descriptive Statistics	42
3.8	Subjective and Objective Schools' Assumptions	72
3.9	Internal Consistencies of InCAS	83
3.10	Sample InCAS Results Sheet	88
3.11	Pilot and Empirical Studies Timelines	89
3.12	Control School Participations	91
4.13	Pilot Study Mental Arithmetic Results	99
4.14	Pilot Study General Maths Results	100
4.15	Empirical Study Mental Arithmetic Results	102
4.16	Empirical Study General Maths Results	103
4.17	Empirical Study Developed Ability Results	104
4.18	Empirical Study Pearson's Correlations	105
4.19	Empirical Study Summary of Gains	105
4.20	Pilot Study Cohort Mental Arithmetic Gains	106
4.21	Pilot Study Cohort General Maths Gains	107
4.22	Summary of Pilot Study Gains	108
4.23	Comparison by Gender of Mental Arithmetic Results	109
4.24	Comparison by Gender of General Maths Results	110
4.25	Comparison by Gender of Developed Ability Results	111
4.26	Co. Durham and Edinburgh Pupils' Mean Ages-at-Test	112
4.27	Comparison of Mean Ages by Location	112
4.28	Comparisons by Location of Mean Mental Arithmetic Results	113
4.29	Comparison by Location of Mean General Maths Results	114
4.30	Comparison by Location of Mean Developed Ability Results	115
4.31	Comparison of Empirical and Pilot Study Mental Arithmetic Results	117
4.32	Comparison of Empirical and Pilot Study General Maths Results	118
4.33	Comparison of Empirical and Pilot Study Mean Ages-at-Test	119
5.34	Summary of Empirical Study Gains by Session	124
5.35	Empirical Study Pearson's Correlations	127
6.36	The Swiss School (1) Mental Arithmetic Results	152
6.37	The Swiss School (1) General Maths Results	152
6.38	The Swiss School (1) Developed Ability	153

	Results	
6.39	The Swiss School (2) Mental Arithmetic Results	153
6.40	The Swiss School (2) General Maths Results	153
6.41	The Swiss School (2) Developed Ability Results	154
9.42	Number Hierarchy Table	234

List of Figures

Figure 1.1	Bar chart comparing National Attainments	35
Figure 1.2	Bar chart comparing Automaticity Scores	35
Figure 1.3	Longitudinal Mean Automaticity Scores	37
Figure 2.4	Leash, Post and Behr Chart	55
Figure 4.5	Pilot Study Chart of Mean Scores	101
Figure 4.6	Empirical Study Chart of Mean Scores	103
Figure 4.7	Empirical Study Chart of M/A and D/A Scores	104
Figure 4.8	Summary Chart of Empirical Study Gains	106
Figure 4.9	Summary Chart of Pilot Study Cohort Gains	108
Figure 4.10	Empirical Study Gender Comparisons for M/A	109
Figure 4.11	Empirical Study Gender Comparisons for G/M	110
Figure 4.12	Empirical Study Gender Comparisons for D/A	111
Figure 4.13	Empirical Study Location Comparisons for M/A	114
Figure 4.14	Empirical Study Location Comparisons for G/M	115
Figure 4.15	Empirical Study Location Comparisons for D/A	116
Figure 4.16	Pilot and Empirical Studies Comparisons for M/A	117
Figure 4.17	Pilot and Empirical Studies Comparisons for G/M	119
Figure 4.18	Histogram of Initial Dataset M/A Scores	120
Figure 4.19	Corresponding Stem & Leaf Plot for M/A Scores	120
Figure 4.20	Histogram of Initial Dataset G/M Scores	121
Figure 4.21	Corresponding Stem & Leaf Plot for G/M Scores	121
Figure 5.22	Summary Chart of Empirical Study Gains	125
Figure 5.23	Boxplots of the Co. Durham and Edinburgh Pupils	145
Figure 9.24	The Arithmetic House with a Foundation	217
Figure 9.25	The Number House with no Foundation	217
Figure 9.26	Linear Approach to Memorising Number Facts	228
Figure 9.27	Swiss Roll Model for Hardwiring Number Facts	229

List of Illustrations

Illustration 1.1	Stages in Learning/Subtracting within 10	22
Illustration 1.2	The Entry-level SUMDIAL 10	23
Illustration 1.3	The SUMDIAL 20	24
Illustration 1.4	The SUMDIAL 50	24
Illustration 3.5	Set-up of laptops for InCAS Assessments	87
Illustration 5.6	Copy of the Answer	137
Illustration 9.7	Page of <i>30 Second Challenge</i> , by Lock	231
Illustration 9.8	SUMBOARD 10 for Active Number Learning	237
Illustration 9.9	A Schematic Representation of Outcomes	241

Abbreviations and Meanings

Abbreviations

The following abbreviations are used in this thesis:

InCAS Results Tables and text:

- **CEM:** The Centre for Evaluation and Measurement
- **AES:** Age equivalent score.
- **A/T:** Age-at-test.
- **D/A:** Developed Ability.
- **G/M:** General Maths.
- **M/A:** Mental Arithmetic.

Note: The full words with their capitals, as above, when used within the text denote the InCAS sessions (subjects) whereas the more general ‘mental arithmetic’, for example, implies their common usage.

The abbreviations used in *all* the results tables and figures are:

- **Effect Size** and Cohen’s suggested categorisation of effect sizes are
 - (i) A value of less than 0.2 is Trivial.
 - (ii) A value between 0.2 and 0.5 is Small.
 - (iii) A value between 0.5 and 0.8 is Medium.
 - (iv) A value of more than 0.8 is Large, (Kinnear & Gray 2011, p.183).
- **S/E:** Standard Error Differences (Equal variances not assumed).
- **p:** Significance (2-tailed) where values < 0.05 are significant at the 95% level while > 0.01 are highly significant at the 99% level.

Meanings

The following meanings will be used throughout this thesis, unless alternative ones are given:

Arithmetic: the branch of mathematics concerned with numerical calculation, such as addition, subtraction, multiplication and division.

Mathematics: a group of related sciences, including algebra, geometry, calculus, concerned with the study of number, quantity, shape and space and their interrelationships by using a specialised notation.

Number: 1. a concept of quantity that is or can be derived from a single unit, the sum of a collection of units, or zero. Every number occupies a unique position in a sequence, enabling it to be used in counting.

Maths: *Brit. Informal.* short for **mathematics**.

Numerate: able to use numbers, esp. in arithmetical operations. Compare **literate** by analogy with *literate* – **numeracy**.

Literate 3. Used to words rather than numbers as a means of expression. Compare **numerate**.

Collins English Dictionary – 3rd ed. (1994)

Automaticity: the recall of previously memorised number facts without any *conscious* mental activity. (The above dictionary only acknowledges there is such a word, but does not provide a meaning.)

(**NB.** The informal version of mathematics – maths – will be used in recognition of common practice in this country. While in North America the word ‘mathematics’ or ‘math’ generally includes arithmetic).

Statement of Copyright

The copyright of this thesis rests with the author. No quotation from it should be published without the prior written consent and information derived from it should be acknowledged.

Acknowledgements

The help of the following in enabling me to complete this thesis is gratefully acknowledged:

- Prof. Steven Higgins, my Supervisor, for single-handedly bringing one of nature's loose cannons under control.
- Mr. Peter Shannon, who became my *de facto* research assistant while providing an immense amount of insight to an outsider about the realities of schools and education.
- Dr. Eleanor Loughlin, (former) Graduate Training Officer, for all her time and support in ensuring I took advantage of the relevant training opportunities.
- Dr. Christine Merrell, Mr. Brian Henderson and Dr. Helen Cramman, all of CEM, for all their willing support that enabled me to use InCAS so effectively.
- Dr. Adetayo Kasim, of the Wolfson Research Institute, for his invaluable guidance in preparing me to carry out my statistical analysis.
- Frau Dori Fischer and Herr Jakob for their Swiss contributions with their support and encouragement.
- Mr. Gareth Cowan, formerly of CIS, for his unstinting support as he trained me to use SPSS 19.
- Mrs. Judith Watson, in the School of Education Library, for her sympathetic help with my literature searches.

Dedication

For Alison, who was very special.

Preface

The Study plan was to conduct a straightforward Randomised Control Trial (RCT), appropriate for a doctoral study, to identify the effectiveness of a new type of manipulative in helping young pupils to improve their number skills and, in particular, to achieve the automatic *recall* of the arithmetical number facts. Two RCTs that had been conducted previously became quasi-pilot studies for this one.

In the event, there were many unanticipated developments that led to the original plan being modified in both its implementation and its content. These included:

- Responding to the limited value of the control school contributions normally used in such RCTs.
- Discovering that the main question could not be answered because automaticity was no longer being emphasised.
- Taking the opportunity to assess two small Swiss primary schools during the data collection phase of the empirical study.

These developments prompted attempts to identify the possible influences behind these observations and included:

- The Swiss pupils showed clear evidence of automatic recall of arithmetical number facts.
- The general approach to the Study showed many of the characteristics of action research.

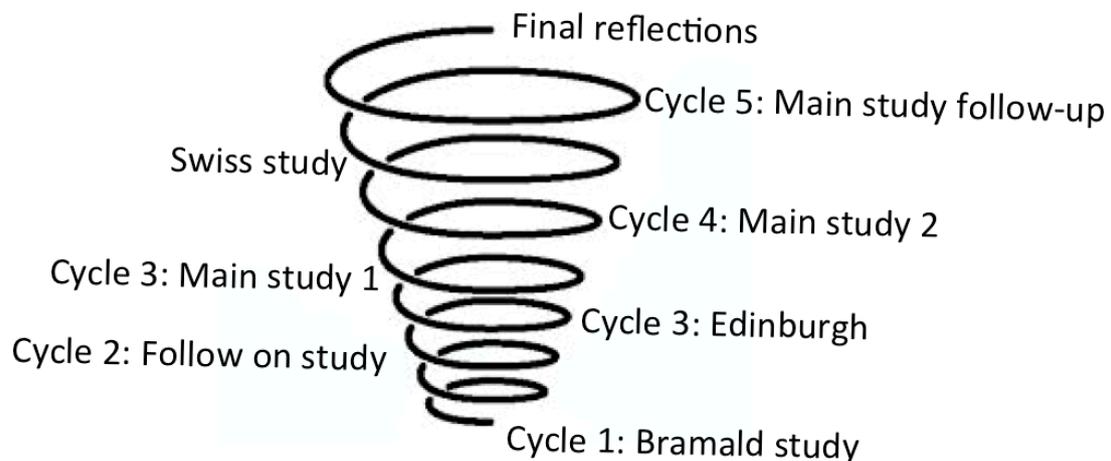
In light of these, it was concluded that the thesis should be written in two parts:

- Part 1: The usual chapters associated with RCTs of:
 - (i) Background,
 - (ii) Literature Review,
 - (iii) Methodology,

- (iv) Results,
- (v) Discussion and Conclusion.

- Part 2: A reflection on the originality and learning inherent in the action research cycles with its parallels on:
 - (i) Professional Development, and
 - (ii) Research and Development in the development of Small Businesses.

In addition, the impact of the Swiss case study will be given prominence as the bridge between these two parts to reflect its pivotal role in the development of the key ideas in this thesis. The following spiral diagram represents the stages that make-up this structure:



Part 1

1. Thesis Background

1.1 Introduction

Some personal background information will help to appreciate the way this Study developed. It includes relevant reflections on the influences that encouraged its start and continue the study into the effectiveness of the *Sumdials*' approach in achieving arithmetical automaticity of number facts. Manipulatives, in the form of the 'dials', are a key feature of the approach. The influence of action research is considered in later a Chapter.

1.2 Initial Personal Reflections

The relevant stage-setting points start with my attending junior school in South Lancashire during the late 1930s when the 3Rs of Reading, 'Riting and 'Rithmetic were the foundational subjects for all learning. My father trained as an electrical engineer before moving into general management (in continuous-process manufacturing); he was convinced that his original training provided an excellent preparation for his own career. This undoubtedly influenced the selection of science subjects at secondary school so that I could read electrical engineering at The University of Glasgow. This was followed by National Service, manufacturing, management consulting, book-printing, carton manufacture, corrugated packaging before starting my own niche packaging company in the mid 1970s.

Seven years as a management consultant developed my observational skills and analytical inclinations. These, with my father's dictum: 'Robin, if it is worth doing, it's worth doing well' while growing up during the Great Depression, must have influenced my development considerably. Sadly, he died as I left school leaving me to fight my own battles – with mixed results. But he convinced me, together with my mother, of the importance of integrity and they gave me a strong Christian faith.

This introduces my preparation for research. However, none took place until 20 years later when I completed a master's degree (by dissertation) at Strathclyde University on *Decision-Making and Reward* that identified the types of decision required by jobs

at different levels and relating them to appropriate remunerations. This involved some fairly limited research and analysis and this experience remained dormant for almost another 30 years. By that time, I had established my business convinced that acquiring a reputation for reliability, quality and service would lead to success. These created the word-of-mouth recommendations that helped the business to grow; advertising has not been used.

In the mid 1980s, Peter Shannon, the principal teacher of maths at a large Edinburgh comprehensive school arrived unannounced to ask if we could make an interactive resource - using a manipulative - for teaching and learning the rules of positive and negative integers (and also one for trigonometry); he had some hand-made cardboard prototypes with him. We could and did make some, albeit in different materials, and they worked well. He then returned two years later with more new resources for learning the basic arithmetic processes of addition, subtraction, multiplication and division. These convinced me (as a non-teacher) his ideas had considerable potential in helping pupils to learn number and I was pleased when he asked me to help in developing his ideas further. This coincided with my reaching normal retiring age and my son's arrival in the business, so I accepted.

Our first move was to set-up *Sumdials Limited* – named after one of his resources - because local authorities then insisted that their suppliers were VAT registered. Thus, I became a sales representative in 1996 and found myself at the bottom of a steep learning curve about primary schools. Polite interest in the new approach was the usual response, but very few orders ensued because the changing teacher pedagogies would have been unwelcome, as came to be appreciated during this Study.

1.3 *Sumdials*' Origins and Approach

The phrase '*Sumdials*' approach' has already been used and it is now described because of its central importance in this Study.

1.3.1 *Sumdials*' Origins

Traditionally, the teaching and learning of number and arithmetic uses *word*-based pedagogies. This is surprising because number is the universal language and so it should be better to learn it *directly*, rather than indirectly through the decoding of verbal explanations of arithmetical/mathematical concepts and processes.

The limitations of this approach are well illustrated by Peter's experiences in the early 1980s. He had successfully used word-based pedagogies during his previous ten years to teach the rules for positive and negative integers. However, he had to put them aside having concluded that:

- A new cohort did not have the requisite verbal skills to decode his tried-and-tested explanations
- He needed to find a different method that was independent of words.

He promised he would return to the topic three weeks later, hoping to develop a new pedagogy by then that was independent of words. He eventually alighted on the manipulative, based on the dials of early telephones, that he asked me to make and it solved the problem well. The most likely explanation was he had found an active learning method of teaching number, based on *seeing* and *doing*. In the event, it overcame the new cohort's lack of *word* skills and overcame its immediate problem.

In the process, he recognised the validity and potential of his new approach and successfully adapted it for trigonometry and, later, for learning the basic arithmetic processes of adding, subtracting, multiplying and dividing. He had named the entry-level resource, for learning to add and subtract up to 10, as the *Sumdial 10* – hence the Company's name!

1.3.2 *Sumdials*' Approach

The following schematic diagram with an original dial illustrates the structure and main features of the *Sumdials'* approach to learning arithmetic:

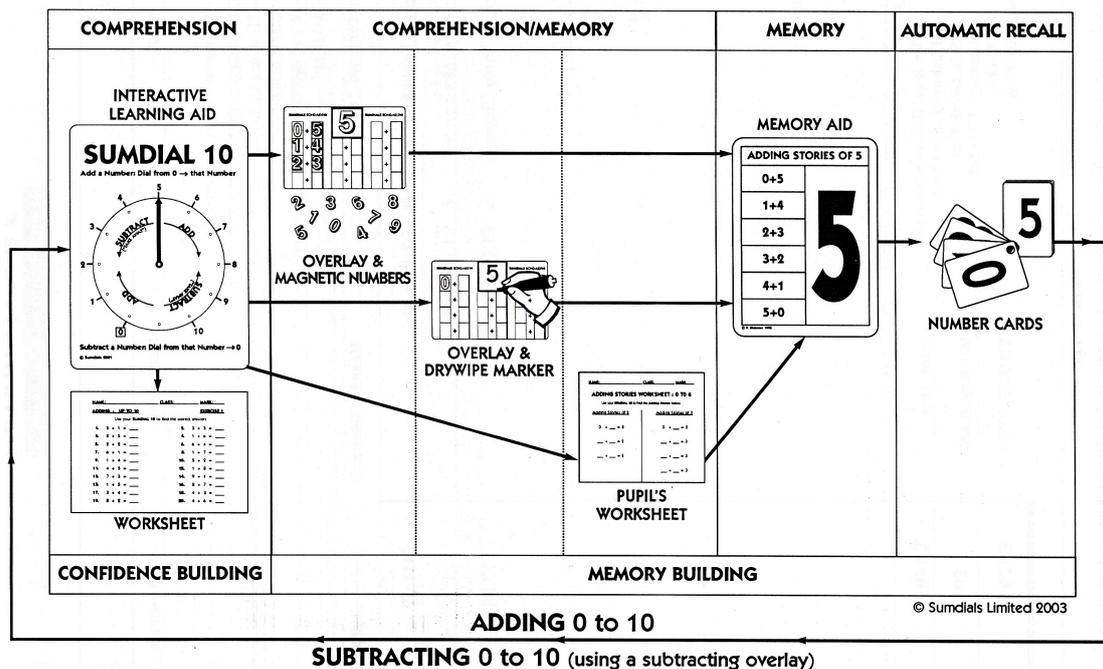


Illustration 1.1: The *Sumdials*' Approach Flowchart for leaning to add up to 10.

One comment, based on hindsight, is that such a diagram is probably more appropriate for an engineering workshop than a primary school. Having said that, the overall structure has three sections:

- 'Investigation/Discovery' aimed at 'Confidence Building' as pupils are introduced to number and become number-ready. The *Sumdial 10* is the link to the second section.
- Comprehension/Memory through Creating Number Stories as pupils use their dials to develop the basic arithmetic processes of adding, subtracting multiplying and dividing. This section uses a range of resources with the aim of "weaning" pupils off their dials.
- Memory/Automatic Recall for Memorising Number Stories and, again, further resources are used to help pupils to hardwire their number facts into their long-term memories.

The range of resources to achieve this is adaptations of long establish teaching and learning practices while only the "dials" themselves are unique. There are four such dials:

- *Sumdial 10*: Adding and subtracting within **10**.
- *Sumdial 20*: Adding and subtracting within **20**, including bridging **10**.
- *Sumdial 50*: Multiplying and dividing within **50**.
- *Sumdial 100*: Multiplying and dividing within **100**.

The redesigned entry-level dial with colour is now used to describe the steps in making a simple addition:

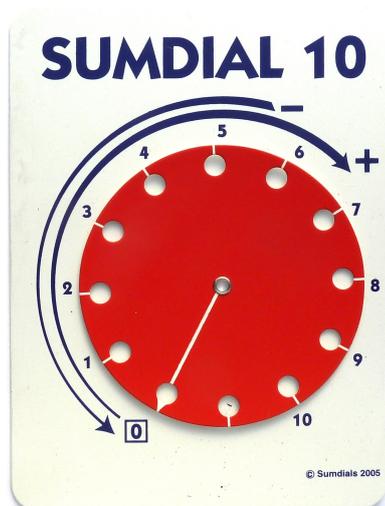


Illustration 1.2: The entry-level dial for adding/subtracting within 10

The actual ‘steps’ used to add **3** and **2** together are:

- Start by rotating the dial so that its ‘arrow’ is pointing at **0** (zero).
- Dial from **0** until the arrow points at **3** (the first addend).
- Leaving the arrow pointing at **3**, dial from **0** to **2** (the second addend) and the arrow now points at **5** - the answer.

An illustrated instruction sheet showing the actual moves of a finger is given in Appendix 1.1.

The *Sumdial 20* is now illustrated:



Illustration 1.3: The dial for adding/subtracting within 20 that includes bridging 20

It is used to establish the concept of ‘bridging 10’ when adding and subtracting within 20. It will be seen immediately that it is operated in the same way as the *Sumdial 10* to acquire the transferable skills that will operate the *Sumdial 2* correctly.

The *Sumdial 50* is also now illustrated:



Illustration 1.4: The dial for introducing multiplication/division within 50

It is used to establish the concepts of multiplication and division (up to 50) and the previously learned transferable skills can immediately be applied to demonstrate that:

- Multiplication is repeated addition.
- Division is repeated subtraction.
- Division with a remainder, when the divisor does not go exactly into the dividend, with the remainder being the number that the arrow is pointing at when the last divisor has been subtracted.

Originally, there was also a *Sumdials 100*, but it was discontinued because the accuracy of concentricity required between the dials and their backs was very difficult to achieve economically. More importantly, it was realised that the *Sumdial 50* should be sufficient to demonstrate the *principles* of multiplication and division and when that was not the case, it was unlikely that the *Sumdials 100* would achieve it and could be indicative of learning difficulties.

In summary, the operations of these dials all model well the basic arithmetic processes and the pupils can discover them for themselves once the teacher has demonstrated – not described – the steps for the *Sumdial 10*. This cuts out the need for decoding verbal explanations of the processes and it is believed intuitively that using the dials helps pupils to develop robust internal models of the arithmetic processes.

1.3.3 Comments on the *Sumdials*' Approach

The key point is that the above example treated **3** and **2** as *discrete entities* or applied them as cardinal numbers being the essence of addition. In contrast, counting-on is a different process even though it produces the same answers. However, most primary school teachers instinctively want to count-on from the larger addend when first introduced to a dial. The limitations of such an approach as counting-on are:

- Only appropriate as an initial stage in learning number, since it develops an unsound model of addition as it cannot be applied to learn multiplication.

- Error prone in that there is a tendency, having dialled in **3**, to start the counting-on for the second addend on **3** (instead of **4**) to arrive at the wrong answer and to hinder the development of pupils' confidence in manipulating numbers.

A more general point is that more than two numbers may be added together as one 'sum' provided the total is not greater than 10.

The dials are very adaptable in their uses and, for example, the *Sumdial 20* can be used to introduce:

- Bridging **10** as, for example, **8 plus 3**.
- Multiplication by dialling in **4** five times so that the arrow points at **20** to demonstrate the principle that multiplication is repeated addition.

The addition steps are reversed for subtraction and that confirms that the dials are good analogues of the basic arithmetic processes. Again, it is believed this is the feature that helps pupils to establish their own robust internal models of the basic arithmetic processes and, in turn, lay sound foundations for maths.

Arithmetic (and maths) is the universal language, so it must be better, in principle, to teach and learn it directly, rather than through word-based explanations that then have to be correctly decoded; this was the original impetus that led to the development of the *Sumdials'* approach, (p.20). To help in this, there are both pupil and teacher demonstrator versions with identical graphics: the pupil's ones are slightly smaller than A5 while the demonstrators are larger than A3. Hypothetically, this arrangement would allow the teacher of a multilingual class to teach all the basic arithmetic processes by demonstrating the steps and then signalling to the pupils to "show me" their dials step by step to check that they are using them correctly. The pupils would then be learning through visual representation and tactile affordance methods independently of language.

1.3.4 Summary

This outline of the *Sumdials'* approach to learning arithmetic should help in understanding its place in this Study and to appreciate some more general benefits when learning arithmetic.

1.4 Personal Reflections (Cont.)

The two main options to overcome the lack of orders were:

- Advertising.
- Getting independent research on the effectiveness of the new resources.

It was almost inevitable that the latter would be chosen, based on my personal aversion to advertising. More to the point, my hunch was that Peter's approach needed more development, so I commissioned the Department of Education at the University of Newcastle to carry out a six weeks study on the entry-level resource (adding up to 10). Bramald (2001), a senior lecturer, conducted it and his findings were helpful, but not conclusive enough. However, the study confirmed some aspects of the approach needed to be addressed, the main one being that the primary school teachers felt the new resources would be more suitable for secondary schools, being only black and white while they were looking for plenty of colour. The Newcastle Study is considered further below (p.30).

All the while, development of the approach continued and new resources for learning fractions, decimals and percentages were added, but it became clear (to me) that the schools were not ready for Peter's ideas. This was a time of rapid change following the introduction of the National Numeracy Curriculum (in England and Wales) and the 5 to 14 Guidelines (in Scotland). The business was never viable due to lack of sales and, almost in desperation, I decided to do a follow-up study when the pupils were completing their primary education to find out if there had been and lasting benefits from Bramald's intervention (2001). This was the case in that the treatment pupils achieved statistically significant improvements, (p.34).

My report only elicited interest in the University of Newcastle where it would be known that short-term interventions such as Bramald's interventions usually

“washout” within two or three years, as identified in the 1970s Westinghouse study and also by Sylva (1994). This indirectly led to this research study. There were several contributory factors of which the main ones are:

- My/our unshakable conviction of the huge potential of the *Sumdials*' approach.
- My/our hope that even very conservative primary school teachers would change once they knew how good the results were – a forlorn hope and a possible explanation is given (pp.61/67).
- A personal commitment, as an employer, to reverse the on-going general decline in number skills.
- Perhaps most important of all, as an outsider to the world of education from primary to tertiary, my engineering background encouraged me to approach problems as challenges to be overcome.
- Subsequently, it became apparent during my Empirical Study that this trait shared many of action research's characteristics.
- Personal experiences over thirty years earlier when a 67 years old colleague completed his PhD while another one reckoned I was more an academic than a manager at heart. These may have sown the seeds for my university research.

As further background, Peter attended his junior school in the late 1940s and, like me, received a thorough grounding in the 3Rs. We have complementary skills based on his over 40 years as a maths teacher and my own analytical traits supported by sufficient number skills to be able to contribute to our many discussions. In the process he has become my invaluable *de facto* research assistant and is a natural micro thinker while I am more of a macro thinker.

An unquantifiable influence on our work together was Peter's wife died suddenly two years before I started my Pilot Study and my wife died (cancer) just after we had started its data collection. As a result, we both shared a worthwhile project that we strongly believe will benefit pupils generally. Fortuitously, we could enlist the help of four of our grandchildren who were in the age groups of one to eight at the start of the

Study. This gave us some useful insights about primary schools and they helped us by testing our new ideas – without their realising it!

So far, some of the personal and acquired attributes that prepared me for research have been described. Now is the time to describe the responses to the external influence of the Newcastle Study and my own longitudinal Study that showed me I needed to:

- Gain a better appreciation of the teaching and learning processes through which pupils acquired their number skills.
- Become better informed about primary schools generally and develop further the *Sumdials*' approach in response to all relevant new understandings.

It was believed achieving these would make the company viable and would enhance the number skills of future generations.

1.4. Research Questions

These led to my main research question becoming:

- Does the *Sumdials*' approach to learning number, based on the use of dedicated manipulatives (dials), produce statistically significant improvements in arithmetical automaticity?
- As secondary questions, are there statistically significant differences in the number attainments:
 - (i) By gender (boys and girls)?
 - (ii) By location (between Co. Durham and Edinburgh pupils)?

This is a good point to explain my then understanding of what automaticity was, based on Peter's definition as:

- Automaticity is the instant and accurate recall of previously learned number facts *without any conscious mental effort*.

I assumed, because a part of my Master's degree involved translating academic "gobbledy-gook" into plain English, this was another example of the same tendency; I translated it to myself as: 'knowing number facts and tables' - as all my school contemporaries would have done.

However, the pre-study work by the University of Newcastle and my own is now examined because they set the stage for this research.

1.5 Pre-Research Studies

1.5.1 Introduction

Two studies of the *Sumdials*' approach to learning number preceded this research and they are considered now because they provide its context. The reason behind both studies was the conviction that the new approach to learning number supported by arithmetical automaticity was sound and would be confirmed by demonstrating that it improved results. In turn, it was believed that this would encourage primary teachers to adopt it.

The studies were:

- Firstly, Bramald's Study.
- The second Study was led by myself and was in two parts:
 - (i) A longitudinal study of Bramald's treatment cohort as it was completing its primary school education five years later. Its aim was to find out if there were any measurable differences between the surviving treatment pupils and the other pupils in their current cohorts.
 - (ii) A qualitative study to assess pupils' attitudes to numeracy in their second year of secondary schooling that involved one school in NE England and one in Edinburgh. These schools received pupils from schools that had taken part in Bramald's Study, but only some of them had taken part in it.

The structure of the *Sumdials*' teaching plan that was used in his Study was based on traditional arithmetic practice – derived from our own childhood experiences - for learning addition up to 10 and had two sections:

- Section one consisted of 12 lessons in two parts:
 - (i) Instructing pupils on how to use and apply their dials.
 - (ii) Answering adding question on worksheets.
- Section two consisted of 15 memory-work lessons.

It was expected that by the end of the six weeks' Study the majority of the pupils would be able to add up to 10 and have good automatic recall of the associated number facts. The point now is: *the treatment teachers of this Study found the approach unexceptional with its emphasis on memory work and willingly followed it* and one of them again confirmed this was the case towards the end of this Study. They also made suggestions about how it could be improved without altering its basic concept, such as adding colour and removing the printed instructions from their dials. The implications of memory work are discussed later (p.227).

In addition, as a spontaneous test, the opportunities to give the automaticity tests to 31 PGCE students and also to some (ten) teachers from two primary schools (both in North England) were taken during the second Study. The original reasons for the tests were to:

- Assess how the teachers performed, and
- Determine if there was any correlation between their scores and their ages.

The possibility that there might be a connection between their own number attainments and their attitudes to arithmetic was considered.

1.5.2 Bramald's Study

Bramald's Study involved ten classes from nine primary schools in Edinburgh and NE England. His pre-intervention testing to arrive at five treatment/control pairs was

meticulous and at the end of his Study he retested the pupils to find whether or not any differences in arithmetic achievement could be found.

‘In addition to the achievement, a measure of number bond recall at speed was also used post-treatment to see if there were any differences to be found’ (emphasis added), Bramald (2001, p.3).

His Study started with 107 pairs of treatment/control pupils from five P2 (Scotland)/ and five Y1 (NE England) classes while 80 pairs were retested at the end. He found:

- ‘The results again [as with the pre-test results] showed no significant statistical differences’ (p.7)
- ‘The [Automaticity results] experimental group clearly out-performed the control group . . . and was found to be highly significant ($t=3.77$, $p=0.00$) (p. 9)’.

In spite of this, his second Conclusion was noted:

‘A very substantial difference in pupils’ speed at correctly answering simple arithmetic number bond questions was found. Although this cannot be unquestionably attributed to the *Sumdials*’ effect, it does seem reasonable to suggest that this was the major reason for this increase. Further and more detailed investigations which include a pre-treatment measurement could corroborate this’, (p.22).

His Study also included qualitative comments based on interviews with the teachers and pupils that produced many positive comments and also constructive criticisms about the *Sumdials*’ approach, its dials and associated resources.

With hindsight, we were too naive in believing that a six weeks intervention would be long enough to achieve a measurable post-treatment gain by the treatment group. That was only long enough to complete the *Sumdials*’ teaching plan with its emphases on:

- Becoming proficient in using the dials.
- Making a start in memorising the adding number facts up to 10.

Thus, any statistical difference between the groups would be very unlikely. There would have been a better chance of measuring any differences between the two groups if a pre-treatment assessment of automaticity levels had been made; its lack

was only discovered (by me) during a progress meeting. In light of this, his qualified second Conclusion was an understandable consequence of his initial oversight, as he confirmed:

Further and more detailed investigations which include a pre-treatment measurement could corroborate this', (2001, p.22).

However, it can be argued that, since his pre- and post-study procedures were so thorough, the treatment pupils' better performance *must* have been due to the *Sumdials*' effect. This is supported by his suggestion:

'There is however enough evidence here to suggest that this is an area that could very reasonably be researched in more detail in any future study', (p.12).

Thus, the full value of this Study was not realised due to his apparent emphasis on process at the planning stage without having understood what the real aim of the Study was.

In fairness to Bramald, universities were then being encouraged to generate income as consultants from companies such as ours. However, we were only looking for results that would demonstrate the effectiveness of the *Sumdials*' approach and made no effort to understand the real issues, (p.177). I had hoped to revisit his raw data to support my argument above, but they were no longer available as his Study took place 12 years before this re-analysis.

However, our immediate practical response was to continue with the development and promotion of the *Sumdials*' approach through school visits, attending conferences and providing CPD workshops. This trait (of responding to problems as they occur) rather than walking round them became evident repeatedly in the current Study. One useful and unexpected outcome of his Study was it provided a first inkling of how messy primary school can be.

Lastly, it is mentioned that the following note was appended to Bramald's Conclusions' page before distributing his Report:

With regards to the first conclusion we mention that during the study it was realised that it should have been carried out with P1/Reception pupils, since

the pupils had already covered the ground (adding up to ten) by the time they had reached P2/Y1. (Comment inserted by *Sumdials Ltd*).

This footnote is commented upon because of its hindsight value (p.116).

1.5.3 Subsequent Longitudinal Study

Bramald's observation was, and still is, very true:

- 'They are convinced that their materials and associated programme of activities and exercises will have a beneficial effect upon pupils' learning of arithmetic.' (p.3).

Indeed, most of his report's qualitative criticisms had been implemented leading to further improvements in the *Sumdials'* approach. However, the main outcome of this conviction was to carry out a follow-up study to assess his treatment pupils' number attainments at the end of their primary schooling.

It was a spur of the moment study with no pre-planning to ensure the actual samples could provide any reliable conclusions. It took place five years after the first study and there would have been considerable turnover of pupils during that time. This allowed two groups to be created:

- Treatment pupils from the first Study.
- Control pupils, being those who had joined since the first Study.

Two measures were used:

- The pupils' national number attainments, as assessed by their schools, and
- Written automaticity tests of 10 each pre-recorded number bond questions for:

Adding up to 20,

Subtracting up to 20,

Multiplying up to 100, and

Dividing up to 100.

The mean scores of the national number attainment levels were calculated (manually) and the results were presented as bar charts in which both the Scottish (A to E) and English (1 to 5) nomenclatures are combined:

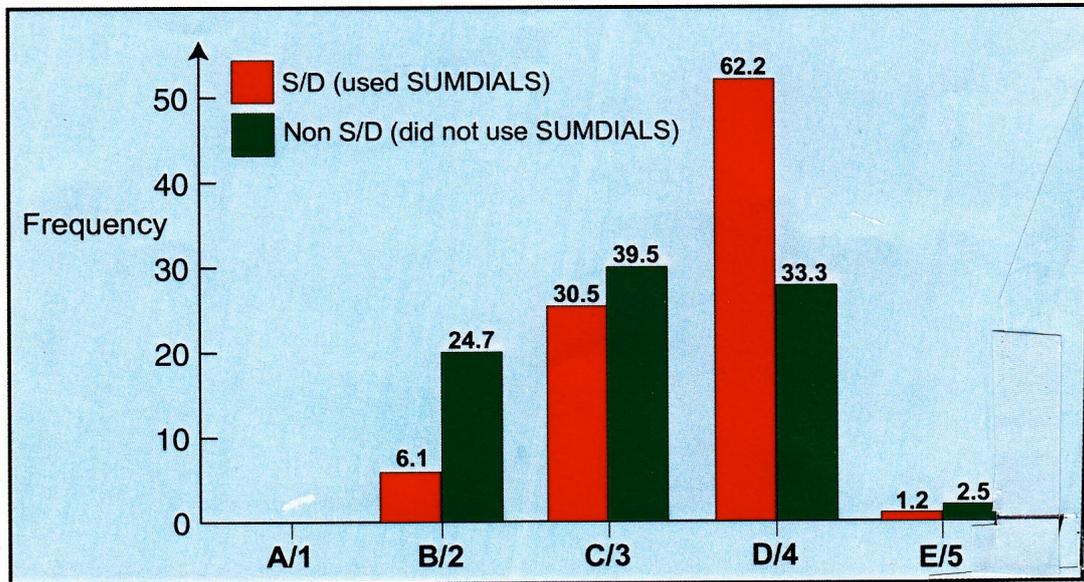


Figure 1.1: Bar chart comparing mean national assessments of treatment (red) and control (green) pupils.

The surviving treatment pupils had achieved much better results than the control pupils in their recall of their number facts, as the following bar charts show:

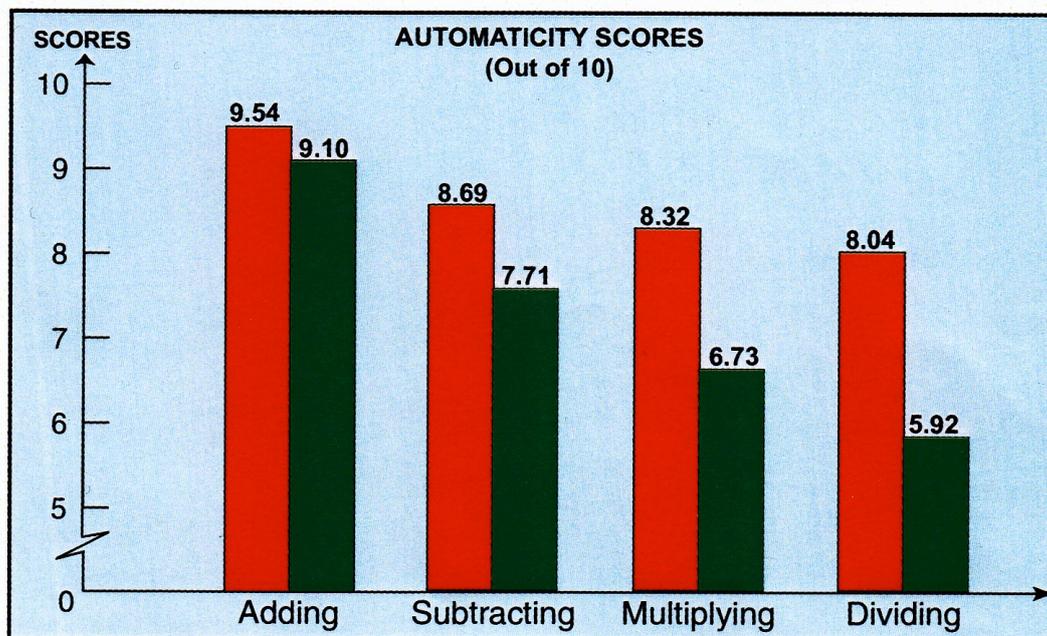


Figure 1.2. Bar chart comparing the mean automaticity scores for treatment (red) and Control (green) pupils.

This excited little interest in primary schools in contrast with The Department of Education in The University of Newcastle. This was probably because it would be known there that the effects of most interventions, such as Bramald’s Study, usually washout within two or three years. It is unlikely that there would be any mathematical anxiety amongst the treatment pupils in his Study, but it could increase as they become older to affect their attainments adversely, Dowker (2005, p.251).

However, we have always postulated that the dials contribute to the development of robust internal models of number and this could have alleviated anxiety in the treatment pupils leading to their better performances five years later. Their new internal model could explain why not only had the treatment pupils preserved their earlier gains in *adding up to 10*, but they had then used them as foundations to *obtain better results in all the other basic arithmetic processes*, as the second chart shows. This could be consistent with the findings that early performance in arithmetic is a predictor of performances in adult life. Thus Bramald’s intervention could have given the treatment pupils a ‘lift’ at a critical stage of their development to produce persisting gains, Dowker (2005, p.14).

In the event, all the test papers for this comparison were available and were compared to produce the following table:

Study Year	Treatment Group	Control Group	Totals
2007	82	81	163
2013	71	92	161
Differences	-11	+9	-2

Table 1.1: Comparisons of original Treatment and Controls Group numbers.

Ideally, the earlier (2007) and recalculated groups (2013) would have been the same. Possible explanations for the differences were errors made in either calculating the group sizes or, more likely, wrong classifications (treatment v control) had occurred. This could not be corroborated because it is now admitted that the available source classifications and calculations had a distinctly back-of-envelope appearance.

However, the original automaticity sheets were re-analysed to produce the following table and chart:

Mean Basic Arithmetic Process Scores

Basic Process	Treatment Group			Control Group			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
Add	9.62	71	0.76	9.08	92	1.71	0.39	0.200	0.007
Subtract	8.93	71	1.63	7.59	92	2.38	0.64	0.305	0.000
Multiply	8.51	71	1.84	6.65	92	2.67	0.79	0.353	0.000
Divide	8.18	71	2.26	6.51	92	7.08	0.30	0.786	0.036
TOTAL	35.24	71	6.49	29.83	92	13.84	0.48	2.12	0.000

Table 1.2: Comparison of Treatment and Control Groups' mean scores with effect sizes and p values.

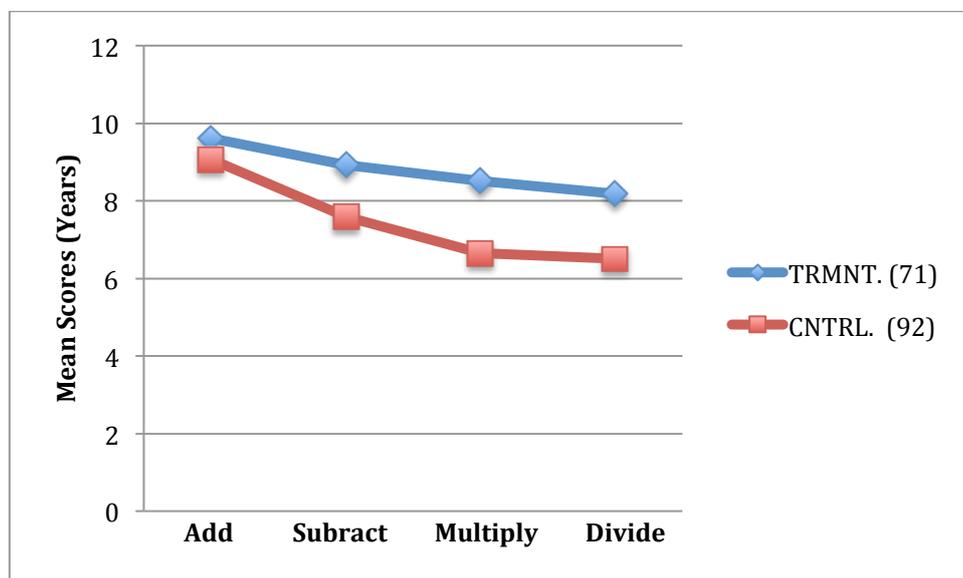


Figure 1.3: Mean Automaticity Scores by process of the treatment and Control Groups.

These results show that the treatment pupils better achievements had effect sizes in the small/medium range while the **p** values were either significant at the 95% or 99% levels. It was postulated even then that the inherent conservatism (reluctance to

change) or inertia of teachers, Brown (Thompson 1999, p.15), prevented them from responding to the potential benefits of the *Sumdials*' approach. This is considered further in the context of the Study proper (p.135).

The abbreviations used in the above table, and in *all* the results tables (in the Study proper), are:

- **Effect Size** and Cohen's suggested categorisation of effect sizes are
 - (i) A value of less than 0.2 is Trivial.
 - (ii) A value between 0.2 and 0.5 is Small.
 - (iii) A value between 0.5 and 0.8 is Medium.
 - (iv) A value of more than 0.8 is Large, (Kinnear & Gray 2011, p.183).

- **S/E:** Standard Error Differences (Equal variances not assumed).

- **p:** Significance (2-tailed) where values < 0.05 are significant at the 95% level while > 0.01 are highly significant at the 99% level.

Thus, the effect size for the mean Total Score of the treatment group is large and its p value ($p < 0.001$) is highly significant and, together with the values of the basic processes, support Bramald's suggestion that further research would be justified.

As part of the verification process, the new Mean Scores (2013) are compared with the 2007 results for both the treatment and control pupils. Their similarities are reassuring, as the following tables show:

Treatment Group

Calculations	N	Add	Subtract	Multiply	Divide	Total
Previous (2007)	92	9.54	8.69	8.32	8.04	34.59
Current (2013)	71	9.62	8.93	8.51	8.18	35.24
Differences	-	0.08	0.24	0.19	0.14	0.65

Table 1.3: Treatment Group comparative scores by subject and in total.

Control Group

Calculations	N	Add	Subtract	Multiply	Divide	Total
Previous (2007)	92	9.10	7.71	6.73	5.92	29.46
Current (2013)	71	9.08	7.59	6.65	6.51	28.67
Differences	-	0.02	0.12	0.18	0.59	0.73

Table 1.4: Control Group comparative scores by subject and in total.

These reworked data are believed to be sufficiently reliable to support the conviction that the *Sumdials*' approach had enabled the treatment pupils to achieved better overall number attainments than the control pupils did. However, to postulate on such limited statistical evidence that the treatment pupils' better national numeracy attainments would be due to a six weeks intervention five years previously could be difficult to sustain. Nevertheless, the better performances of the treatment pupils are noted.

1.5.4 The Qualitative Study

There have been on-going concerns about the “standing still” during the first two years of secondary education and this is shorthand for the lack of progress as cohorts move from primary to secondary education. In Summary, the performances in Scottish schools declined from P4 to P7 and even more so by S2 as measured in the AAP (1997). The impression was that similar outcomes were also evident in England and Wales. Our response was to do a qualitative Study involving pupils from one of Bramald's treatment primary schools in Edinburgh and one in NE England who had fed-into their appropriate secondary schools; 23 took part, only some of whom were treatment pupils. It was carried out in June (Edinburgh) and October, 2008 (NE England) at the end of their second and beginning of their third years in secondary education.

A structured interview of 15 minutes was held for each pupil with all interviews being recorded. Having welcomed the pupils, they were reminded or introduced to an

entry-level dial and tried it before answering some questions on their attitudes to mathematics. Finally, they did eight mental arithmetic questions (Appendix 1.2).

1.5.4.1 Qualitative Study Observations

The main impressions were:

- The staffs of both schools were very supportive in every respect and this was a great help.
- All the pupils were smart, well mannered and a credit to themselves, their homes and their schools. (The pupils were from limited pools, so the schools had little scope to select only the better ones).
- All pupils over-assessed their own maths abilities, generally regarded it as a boring classroom subject and had limited appreciation of its relevance to the outside world even though they knew it would be important.
- They all enjoyed using a dial during their interviews, found it helpful and one described it as “cool”! However, very few of the treatment pupils recalled how to use it and all interviewees wanted to add-on the second addend.
- Only one pupil liked maths as a subject (but was weak at it) while the popular subjects were the active ones such as PE, drama, dance, food technology and home economics.
- All pupils had been self-confident and assured, but became very tense as soon as the mini-mental arithmetic test was given; none answered all questions correctly.

In short, the pupils confirmed:

- They did not like maths,
- They did not do well in their mental arithmetic tests.
- They wanted to drop it as soon as possible.

Clearly, they saw maths as a hard subject even though they knew it was important. The differences in the mini-test scores between the former treatment and new control

pupils were slight and probably would have been nonsignificant, if they had been analysed; this suggests that the standing still syndrome is a reality.

As a final crosscheck, two of the original primary school teachers from Bramald’s Study confirmed after the Study conclusion that the mini-test questions were straightforward (no ‘catches’) and easy.

1.5.4.2 Qualitative Study Conclusion

The aim of this Quality Study was to find an explanation for the marking-time during the first two years of secondary education. One answer is the aims of the primary schools and the needs of the secondary schools are incompatible. To explain, the primary schools are no longer *preparing* their pupils for secondary schools, because they were following their own agendas. Thus, their pupils arrive at secondary school without good groundings in the basic 3 Rs of Reading, ‘Riting and ‘Rithmetic which is all the subject teachers want.

1.5.5 Teachers’ Scores

As mentioned already, the opportunities to give a group of student trainees and some teachers the automaticity tests were taken. The results are summarised in the three tables below for the:

- Students,
- Teachers.
- Students and Teachers combined.

Students’ Descriptive Statistics

TOTAL	N	Minimum	Maximum	Mean	Standard Deviation
Adding Scores	31	10	10	10.00	0.000
Subtracting Scores	31	8	10	9.81	0.477
Multiplying Scores	31	6	10	9.55	0.955
Dividing Scores	31	5	10	9.55	1.060
ALL SCORES	31	30	40	38.9	2.196

Table 1.5: Student’s Descriptive Statistics showing inferior Minimum mean scores

when compared with the older teachers.

Teachers' Descriptive Statistics

TOTAL	N	Minimum	Maximum	Mean	Standard Deviation
Adding Scores	10	10	10	10	0.000
Subtracting Scores	10	10	10	10	0.000
Multiplying Scores	10	7	10	9.3	1.252
Dividing Scores	10	6	10	9.2	1.687
ALL SCORES	10	33	40	38.5	2.915

Table 1.6: Teacher's Descriptive Statistics showing superior Minimum performances when compared with the younger students.

Both results are very similar so the two samples were combined to give a larger sample to produce this table.

Combined Descriptive Statistics

TOTAL	N	Minimum	Maximum	Mean	Standard Deviation
Adding Scores	41	10	10	10	0.000
Subtracting Scores	41	8	10	9.85	0.422
Multiplying Scores	41	6	10	9.49	1.052
Dividing Scores	41	5	10	9.46	1.227
ALL SCORES	41	30	40	38.8	2.358

Table 1.7: Combined Descriptive Statistics of PGCE students and teachers.

There is a steady decline in the Minimum Scores (from adding to division), similar to the pupils' scores in the Longitudinal Study, and an impression had built up, while testing the teachers, that there might be an inverse relationship between the teachers' ages and their automaticity scores.

In the meantime, it is noted that if further evidence from a larger sample corroborates this finding, its implications for raising number attainments generally and mental arithmetic, in particular, are disturbing. These results are discussed later (p.212).

1.5.6 Pre-Study Summary

These pre-study studies constitute a genuine longitudinal Study because some schools and some pupils took part in Bramald's study, my Longitudinal Study and the Quality Study; this possibility had not been considered at the outset. These Studies all contributed invaluable experience that proved to be good preparation for the Study proper.

However, it could reasonably be asked: What was the effect of this work on sales? The answer was: negligible! This suggests that advertising should have been tried after all, but that would have meant admitting defeat and that the evidence from these two Studies counted for nothing. Other explanations emerged during the Study proper.

1.6.1 Background Conclusion

On a personal note, I actually enjoyed doing such research and, even now, still recall our personal satisfactions when Peter had analysed our data manually to produce the bar charts. This was reinforced by the apparently willing co-operation of the schools and their interest in the results – they *really* do want to improve. Thus, I was pleased when it was suggested that I should do a proper study as part of a distance-learning PhD. This in itself was remarkable since, until recently, my experience of schools and teaching was limited to my own time as a customer of education!

What has been described so far only tells part of the story that influenced me to do research. The more important part is that it has provided me with an opportunity to show my gratitude for the many blessings over a long life in the hope that completing some worthwhile research will benefit future generations – of which my grandchildren are members.

2. Literature Review

2. Introduction

Literature Reviews, to use a building analogy, can be seen as the foundations of doctoral studies and their resulting theses as the completed buildings. To extend it, the Background Chapter can be regarded as the site preparation. As has already been explained, this thesis is in two parts and the second one, starting with its further Literature Review, can be seen as the foundation of an extension to the building – to continue the analogy.

This review has:

- A wide-ranging bibliography appropriate for answering the questions at a broad level and providing a general background to the area of mathematical learning in schools. It covers the main issues associated with learning number and the possible benefits of using manipulatives, and
- An in depth-review that is relevant to current practices associated with the teaching and learning of number that needed closer examination.
- A personal reflection that is believed to encapsulate the essential points associated with becoming members of the mathematical enterprise.

To recap, the research questions are:

- Does the *Sumdials*' approach to learning number, based on the use of dedicated manipulatives (dials), produce statistically significant improvements in arithmetical automaticity?
- As secondary questions, are there statistically significant differences in the number attainments:
 - (i) By gender (boys and girls)?
 - (ii) By location (between Co. Durham and the Edinburgh pupils)?

2.1 A Background to Arithmetical Learning

Relevant topics are now considered starting with the role of arithmetic in learning number.

2.1.1 The Role of Arithmetic

This Study needs to establish a convincing case for basic arithmetic, as an essential life skill, to correct the imbalance between it and the strong bias towards literacy that is considered later (p.47). A good starting point is that pupils are born with a natural propensity for number in anticipation that it will still be needed, according to Dehaene (1997, p.5). The need arises because numbers are all pervasive in people's daily lives to the extent that everyone is utterly dependent on number even though most people may not realise or accept how extensive this is (Butterworth 1999, p.ix).

In support of this, Krutetskii (1976, p.6) points out that the Programme of the Communist Party of the Soviet Union gave mathematics special consideration because of its contribution in enabling advances to be made in *all* sciences. It is only through its on-going development that general advances in any country can continue and its lack or inadequacy would be very restricting and inhibiting in all walks of life. His position can be viewed as a strategic one arguing generally for the development of mathematical methods and a mathematics style of thinking that will enable societies to prosper.

This is complemented at a more tactical level by Schoenfeld (in Grouws 1992, p.335) who emphasised the need to 'develop mathematical points of view' through becoming competent in applying its attributes of 'abstraction, symbolic representation and symbolic manipulation'. These he likens to tools and points out that being able to use them does not make someone a tradesman; that only happens once a structure of understanding for their appropriate trade is established. He uses this analogy to explain pupils' need to acquire number sense or 'mathematical sense-making' and in doing so become 'members of the mathematical enterprise'. When pupils join it, and only then, will arithmetic becomes their good servant rather than a hard taskmaster. Number sense is considered further below (p.63).

2.1.2 A History of Teaching Arithmetic

A short summary of the teaching and learning of arithmetic in this country describes the main events that has led to the current situation in schools today, is based on a very helpful account by Brown (in Thompson 1999, pp.3-16). She starts with the appointment of the Newcastle Commission in 1858 to enquire into primary education that found virtually no arithmetic was being taught and such as ‘was being taught, was *judged to be ineffective*’. In spite of this, the country had become very self-confident, successful and prosperous having only been taught the 2 Rs of reading and writing. In response to the Commission’s work, national expectations were introduced to the 3 Rs by including arithmetic in 1862. This was followed by the 1870 Act confirming pupils’ right to primary education.

However, the effective teaching and learning of arithmetic could only have been achieved if there had been a minimum of 25,000 properly qualified arithmetic teachers (assuming one specialist teacher for every primary school) waiting in the wings in 1862 to achieve this. As a consequence, it might reasonably be concluded that arithmetic has always been the poor relative of reading and writing even to this day. Might its teaching still be ‘judged to be ineffective’?

To return to the history of teaching and learning number in the UK, what Brown (Thompson1999, p.3-15) describes can be seen as something of a barometer of national confidence that is very relevant to this study. Simply stated: the greater the State’s intervention in education, the lower is the national confidence. The appointment of the Newcastle Commission was the earliest response to growing concerns about the perceived increases in ‘international industrial competition’. Since then, the pendulum has swung’ (her phrase) between a relatively *laissez-faire* stance on education to more control through, for example, national curricula and target setting. However, the introduction of the National Curriculum in 1999 (England and Wales) and the 5 to 14 Guidelines (Scotland) marked a step-change in government control that has continued to this day and is now almost total. Governments have supplied much funding to raise attainments, but the evidence does not unequivocally support their claims of improvement, see Tymms (2004) and Coe (2013).

Brown's final point was that a combination of 'common sense and the inertia of teachers' had always smoothed out the more extreme effects of changes of emphasis in the past and no doubt would continue to do so in the future (Thompson 1999, p.15). It can be asked if the ever-increasing government control has now made her pause for thought. Currently, arithmetic (and maths) has still to be generally accepted as an important subject at a cultural rather than an intellectual level and explanations for this are now considered within the context of the country's epistemology.

2.1.3 Epistemology and Culture

Arithmetic (and maths) is widely seen as a difficult or challenging subject according to Reynolds and Muijs (in Thompson 1999 p.17) and Dowker (2005, p.11). This may be the result of the bias in favour of language throughout the developed world. Examples of the bias, in this country, include Dowker's (2007, p.64) finding that there is a much smaller research base 'on mathematical development and difficulties than on . . . language and literacy'. Similarly, Willey *et al* (2007, p.208.) point out that local authorities' allocations for the special needs of literacy are much higher than those for numeracy. At the micro level, this is very evident during school visits where magnetic letters are invariably seen while magnetic numbers are very rare. These examples support the earlier contention that arithmetic has probably always been the poor relative of the 3 Rs in this country.

One manifestation of this is the majority of people still (wrongly) believe that good reading and writing skills are a better preparation for life than being good at arithmetic (cf. Schoenfeld in Grouws 1992, p.360 or Munn and Reason 2007, p.6). This is consistent with local authorities for education continuing to allocate more resources to literacy than numeracy according to Gross (2007, p.149) and Dowker (2007, p.64) in spite of the findings of the problems that low levels of numeracy cause (Basic Skills Agency 1997). This last point also shows that changes in beliefs do not happen quickly. Schoenfeld's observation was made in 1992 and reiterated by Malofeeva (2009, p.75). Yet, the educational and teaching establishments still do not generally accept the implications of its validity. And neither does the population at large.

A likely contributory event reinforcing such UK perceptions (that maths is a ‘hard’ subject) was the introduction of New Maths in the 1960s. This was another response to a growing concern about declining national competitiveness, following the successful launch of the Sputnik by the then USSR. So-called ‘new maths’ had many good features, but it had been devised and promoted by mathematicians (for mathematicians) and not by educationalists or employers. Research evidence suggests that the majority of the secondary school maths teacher cohort were not comfortable about delivering it (Handal and Herrington 2003), while their pupils were not adequately prepared for it; it was quietly allowed to fade away. (This could happen at that time perhaps because it was not part of a government programme, but governments learned the lesson.)

By then, lasting change to the country’s epistemology had taken place. To explain, part of epistemology is the process through which the cultural values, mores and norms are passed from one generation to the next. This included the unwritten compact between parents and schools that they would share the responsibility of teaching pupils their arithmetic skills. Part of this was helping with homework, but they could no longer do this because they were unfamiliar with the new maths approach (Lehrer and Shumow 1997). To relate a personal experience, I could get the right answers to my daughter’s maths homework, but only by using traditional methods, which did not include set theory. Her natural response was: ‘How could she be expected to do her maths homework if her father, an engineer, could not do it?’.

Thus, there were the two outcomes for parents up and down the land. The first one was the passing on of the arithmetical practices, or folklore, largely ceased because they could no longer help their pupils (with their homework) since the new approach was so different from the one they had learnt at school, analogous to ‘phenomenological primitives’ in physics, (diSessa in Javier 1987, p.83). Hughes *et al* (2007, p.142) describe this process as: ‘deskilling parents’ and it coincided with the advent of the digital age. The prevalence of computer games further weakened the bonds between the generations by displacing established board games, such as snakes and ladders, that helped pupils to develop their number senses and skills (cf. ‘mathematical enterprise’, in Grouws 1992, p.343) as well as their pupils’ social skills. The end result was to reinforce the existing convictions that arithmetic (and

maths) *is* a hard subject. One outworking of this was the pragmatic practice in some primary schools of teachers who were good at maths becoming the numeracy teacher-in-residence as they delivered their colleagues' numeracy classes who, in turn, delivered the numeracy teachers' literacy classes!

This is consistent with the point just made and also the finding that 70% of primary school teachers' preferred teaching styles were word-based leading to most pupils learning their arithmetic indirectly (Smith 1996). These confirm the general point that arithmetic is the subject that primary school teachers feel the least confident about teaching and their pupils must subconsciously pick-up this to affirm their perceptions that it must be a hard subject. Mothers would confirm this by sympathetically telling their pupils that they were 'never any good at sums' while implying that it did not do them any harm - and there is no need to try harder.

A brief consideration now of *Affordances, Constraints and Attunements* (in classrooms) concludes this stage setting. This is the title of a paper by Watson (2003, pp.103–108) that considers the influences of the interpersonal aspects (between teachers and pupils) of teaching and learning mathematics, albeit at KS3. It might seem at first glance that it has no relevance to learning arithmetic in primary schools. However, it points to the reality that teaching and learning arithmetic involves much more than simply transmitting information from teachers to pupils.

She explains that 'affordances' arose from Gibson's work in the 1950s in which he pointed out that 'learning takes place through perception of, and interaction with, an environment', such as school classes. Greeno (1998) developed this concept further when he wrote 'qualities of systems that can support interactions and therefore present possible interactions for an individual to participate in'. He then pointed out that 'within systems there are norms, effects and relations which limit the wider possibilities of the system, that is constraints ...'. His point is subconscious awareness of these encourages pupils' participation as they acquire their arithmetic skills through processes that are more akin to socialising than instructional, according to Resnick and Ford (1981, p.191).

The third word, attunement, is akin to expectation developed through familiarity with the teacher's methods and style that has prepared them for what is likely to happen at each stage, based on what has happened previously. These concepts apply at all levels of education. To cite a personal experience that occurred during a three hours long workshop on writing theses, about half way through the lecturer emphasised the importance of good grammar and specifically mentioned avoiding split infinitives. It so happened that his previous slide included one and I pointed this out. The incident defined the affordances, constraints and attunements of his workshop that I had absorbed to everyone's benefit, including the lecturer's, because up until then everyone had been fairly serious and tense in spite of his best efforts to get us to relax. My contribution achieved it and he thanked me at the end because I had reacted to the problem – without realising it.

These aspects are very relevant to learning arithmetic, especially in primary schools, with the emphasis on first becoming number-ready is most likely to be achieved through the spontaneous interactions between teachers and pupils, and between one another as they play their active learning number games (Jordan, Kaplan, Ramineni and Locuniak 2009) and as described later, (p.237).

It is hoped now that the case for the importance of arithmetic as an everyday life skill has been made and the challenge is to rehabilitate number so that it can assume its vital place in the country's epistemology.

2.1.4 Numbers and Their Representations

Most people regard the number symbols as the only representations of numbers. However, the word: represent (and its derivatives) has different meanings that are determined by their contexts. As von Glasersfeld (in Javier 1987, p.216) points out, in German there are four different words (that cannot be used interchangeably) that are determined by their contexts:

- The sketch represents (depicts) a lily.
- Jane (“mentally”) represents something to herself.
- Mr Bush represents (acts or substitutes for) the president.
- “X” represents (stands for, signifies, denotes) some unknown quantity.

The following definitions support his general point:

The term *representations* here is interpreted in a naïve and restricted sense as external (and therefore observable) embodiments of students' internal conceptualisations – although this internal/external dichotomy is artificial.

Leash, Post and Behr (in Javier 1987, p.33)

As does Goldin:

A representational system shall consist first of a collection of elements called (interchangeably) *characters* or *signs*. These elements are primitive in the sense that they do not at this point stand for or symbolise anything else.

Goldin (in Janvier 1987, p.127)

(In this Section and more generally, definitions and some citations are reproduced verbatim, trusting that what their authors have written, as respected researchers, conveys *their* intended meanings unambiguously and rewriting would be unlikely to improve them.)

These illustrate the difficulties in arriving at a universal definition of representation. This does not mean that research in cognitive and related fields has to be put on hold until a definition of representation has been agreed. Encouragingly, there is fairly general acceptance that the following components specify *representation*:

- A represented world;
- A representing world;
- Aspects of the former that are represented;
- Aspects of the latter that are represented; and
- The correspondence between them.

Kaput (in Janvier 1987, p.126)

His point is 'representing' is quite separate from what is being 'represented' and, in turn, there must be both internal and external representations, according to Goldin (in Javier 1987 p.126). Internal representations can only be inferred because they are parts of other people's internal worlds, including pupils'; they cannot be heard, seen, touched or measured and are context dependent. Goldin identifies five categories of internal representational systems (that make up Kaput's 'represented world' as above) and they are:

- Verbal/syntactic systems,

- Imagistic systems,
- Formal notational systems of mathematics,
- A system of planning, monitoring and executive control, and
- A system of affective representation.

Goldin (in Janvier (1987, p.135)

These will now be considered briefly, but first two sentences illustrate how context reduces ambiguity by diSessa (in Janvier 1987, p.131). They are:

Time flies like an arrow.

Fruit flies like a banana.

Both are grammatically correct, but the second one only makes sense when it is realised that ‘flies’ is a noun and not a verb, as in the first sentence. It greatly simplifies the task of defining ‘representation’ when the context is an implied part and avoids cumbersome definitions. Also it is pointed out that the following five categories or systems by Goldin (in Janvier 1987, p.135) relate only to Kaput’s first component (his ‘represented world’).

2.1.4.1 Verbal/Syntactic Systems

These are based, as would be expected, on natural language and their inputs come through hearing and reading, while their outputs come through speaking and writing and infer the nature of individual persons’ internal representations.

2.1.4.2 Imagistic Systems

These are the non-verbal cognitive systems of which the most important ones in mathematics education are the visual/spatial, auditory/rhythmic and kinaesthetic/tactile representations that make up the wider connotations of imagination. This suggests many different images can be processed internally to create something new and more than the sum of the individual images.

2.1.4.3 Formal Notational Systems

A distinctive feature of arithmetic (and maths) is it is a hierarchical subject that uses structured symbolic notations that include numeration systems, arithmetic algorithms and rules for symbolic manipulations, to name just three. This feature, in common with other scientific subjects, develops individual competencies in applying the rules and procedures to be measured and describe the results that emerge. Thus, it can be inferred that the internal representations of these systems are robust when the results are as they should be, while incorrect results point to unsound representations needing

to be corrected. To be able to identify precisely the right/wrong answers of arithmetic is essential, in contrast with many other subjects that include the word-based ones.

2.1.4.4 Planning, Monitoring and Executive Control

The essence of these interlinked systems is that they provide the structures for effective heuristic (trial and error) methods of solving problems. The key to this is selecting particular approaches and *self-monitoring* whether or not they are leading to fruitful solutions. When they are not, they are abandoned (executive control actions) and a different approach is tried. Such a process is likely to use imagination and, as a result, broadens the imagistic systems. Schoenfeld's model for developing self-regulating skills can be formalised when teachers, acting as 'roving consultants' in small group problem-solving sessions, are only allowed to ask the following questions:

- What (exactly) are you doing? (Can you describe it precisely?);
 - Why are you doing it? (How does it fit into the solution?), and
 - How does it help you? (What will you do with the outcome when you obtain it?)
- In Grouws (1992, p.356)

He argues that it is training, developing and mastering such self-monitoring skills that make pupils into experts able to select and reject different methods until they arrive at good solutions. However, achieving this takes time (months) because pupils initially find it very difficult to articulate their responses to even the first question, but they are well on the way to becoming part of the 'mathematical enterprise' once they can.

2.1.4.5 Affective Representations

This system is very important in that it influences attitudes about arithmetic and determines whether or not pupils come to enjoy or dislike it. This manifests itself in pupils' feelings about solving problems generally – 'I've cracked it!' - be it in arithmetic or maths. Their emotions can range, according to Goldin, through the spectrum of:

...bewilderment, frustration, anxiety, discomfort, satisfaction, pleasure, elation.

Goldin (in Janvier 1987, p.143)

Satisfaction, pleasure and elation affect pupils' attitudes about arithmetic sufficiently to become motivators and rewards in themselves. This is an example of the self-monitoring process through which they know they have successfully completed a task and is an outworking of von Glasersfeld's point (in Janvier (1987 p.15):

'Self-generated reinforcement has an enormous potential in cognitive, reflective organisms'.

Their teachers' affirmations become secondary, as he writes (in Janvier 1987, p.17):

'... if students are to taste something of the mathematician's satisfaction in doing mathematics, they cannot be expected to find it in whatever rewards they might be given for their performance but only through becoming aware of the neatness of fit they have achieved through their own conceptual construction.'

To express these in practical terms, teachers need to have the skills to give their pupils problems that they should *just* be able to solve and when they succeed give them slightly more difficult ones; solving them is a much greater motivator than awarding a 'gold star' or its equivalent. This illustrates how good affective representations can overcome the prevalent negative attitudes about arithmetic and maths.

2.1.5 Modes of Representation

So far, no attempt has been made to classify the different modes of representation (as identified by National Science Foundation funded projects) and this is now addressed using 'the five distinct modes of representation systems that occur while learning mathematics and problem solving':

Experienced-based Scripts – in which knowledge is organised round real world events that serve as general contexts for interpreting and solving other kinds of problem situations;

Manipulative models – like Cuisenaire rods, arithmetic blocks, fraction bars, number lines and, of course, 'dials', etc., in which the elements in the system have little meaning per se, but the built-in relationships and operations fit many everyday situations;

Pictures or diagrams – static figural models that, like manipulative models, can be internalised as images;

Spoken languages – including specialised sub-languages related to domains like logic;

Written symbols – that, like spoken languages, can involve specialised sentences and phrases ($x + 3 = 7$), $A \cup B = (A \cap B)$ as well as normal English sentences.

The relationship between these five modes is now shown schematically:

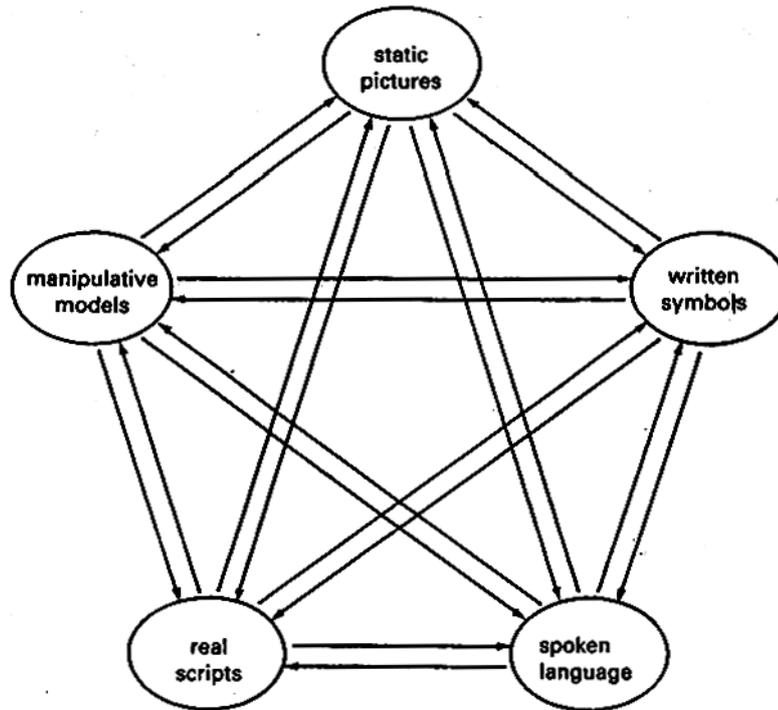


Figure 2.4: Representations and Diagram from Leash, Post and Behr (1987 p.33/4).

One application of these classifications lies in using the different modes to assess pupils' progress and understandings of arithmetic (and mathematics more broadly). This is demonstrated by their abilities to translate correctly from one representation of a problem into another one (Ainsworth, Bibby and Wood 1999). However, it is surprisingly difficult to translate from one mathematical representation to another if pupils have deficient understandings of what the aims or purposes of the procedures actually are, according to Leash, Post and Behr (in Javier 1987, p.33).

2.1.5.1 Uses of Representation

The representations pupils use can give an indication of their development progress as they restructure them in response to their surroundings, Resnick and Ford (1981, p. 113). Similarly, pupils' usages of representations of events and episodes change as their understandings develop (Bruner 1964, p.2). In this research Bruner wanted to

find out how pupils recalled earlier ones and related them to their new ones. This led to his identifying three different modes of representation:

Enactive – re-enacting past events using hands and fingers to describe them.

Iconic – using mental images to describe the main steps by, for example, only recalling the groups of four and three building block that made the total of seven without going through the enactive steps. To give an example from every day life, journey directions of how to get to a particular place will usually include only key landmarks – icons – such as traffic-lights, roundabouts, and churches while all the individual houses, shops and post-boxes are omitted.

Symbolic - the final mode in which pupils use the number words when explaining that ‘four and three equals seven’ or as number symbols by writing it as: $4 + 3 = 7$. The point here is that such symbols do not resemble the actual number of objects involved, be it building blocks or coins and so do not cue the pupil. Using the journey analogy again, the symbolic directions would be either the postcode or the co-ordinates of the destination.

In summary, the first two modes of representation use images instead of words to describe what pupils have done. The relevance of this is that progress in learning number can be assessed through the images pupils use to compensate for their limited linguistic skills. And, in the long run, pupils are more likely to develop mentally, including learning number, when they are allowed sufficient time for their physical developments to take place as its precursor and is highlighted by Blythe (in House 2011, p.131). The general subject of starting ages for formal learning is considered more fully in the second part of this review.

A similar scheme of representation was postulated that used four modes:

Idiosyncratic referring to unintelligible representations like scribbles,

Pictographic referring to drawings of building blocks and their numerosities,

Iconic using one-to-one tally marks, instead of drawings, to represent quantities,

Symbolic is the same as Bruner's symbolic mode. Hughes (1986, p.56-60) The similarities of the two classifications are reassuring while the differences are likely to be the outcomes of their different research aims. Pupils' ability to translate from one mode of representation to another is a reliable indicator of the stage they have actually reached; it is quite independent of their chronological ages. It is also a good indicator of the levels pupils have reached in their *understanding* of the arithmetic processes, according to Hughes (1986, p.111).

2.1.6 Manipulatives in Learning Number

The *Sumdials*' approach, on which the investigative part of this research study is based, is usually perceived, reasonably enough, as relating to the use of manipulatives - its dials - even though they are only one of its many resources. In light of this, manipulatives' general contribution to learning arithmetic is now reviewed.

There is a long history of using manipulatives/concrete materials to help pupils learn arithmetic. However, that does not mean teachers generally welcome them or use them properly, Szendrei (1996, p.411). Reasons include the 'maths-is-a-hard-subject' syndrome, leading to a lack of subject knowledge amongst many primary school teachers and their word-based teaching orientations (p.45). These result in a reluctance to use manipulatives - unless they make immediate and obvious transfer gains by incorporating them into their pedagogies.

The materials themselves can be divided into the two categories of:

- Everyday 'tools and artefacts or common tools'.
- Devices designed for specific 'educational purposes'.

Szendrei, (1996 p.411).

The use of common tools in the classroom has a good and continuing history in the right hands – probably those of connector teachers (p.59) – even though they can suffer from a distractor syndrome as pupils focus more on their familiar uses instead of translating from what is being demonstrated to learning new 'pencil and paper' procedures. In support of that point, a meta-analysis on the 'efficacy of teaching with concrete materials' excludes tools, defined as: rulers, scales or calculators, by Carbonneau, Marley and Selig (2012, pp.380–400).

The remarkable aspect of their meta-analysis is that eventually only 196 papers were identified, using the keywords of:

- Mathematics.
- Manipulatives.
- Concrete objects.
- Activity-based learning.
- Hands-on learning.

They were whittled down to 55 and confirm how little direct research seems to have been carried out recently on the contribution manipulatives may make in learning and teaching arithmetic.

To provide a context for such an observation, it is estimated that there are about 3,600 references from 27 chapters with a total of 640 text pages (1,000 words per page) making a total of 640,000 words in Grouws (1992). Allowance needs to be made for duplications and that could reduce the total to a range of 1,500 to 1,800 references. More to the point, the index revealed only three references to manipulatives that produced less than 3,000 words out of 640,000. In short, it can safely be concluded that research into the use of manipulatives in learning arithmetic has been very limited.

To make a personal comment now, I believe the above exercise illustrates the importance of thinking mathematically, as emphasised by Schoenfeld, (in Grouws 1992, p.335). To exemplify his point, it has provided a relevant perspective for this research on the miniscule use of manipulatives and yet took only about 30 minutes to produce (of which counting the 90 reference pages at the end of each chapter took about 20 minutes). My method was spontaneous – might it be an example of ‘seeing the world through the lens of the mathematician’ (Grouws 1992, p.341)?

Having made that observation, the meta-analysis was very thorough, and its main conclusions are:

- Manipulatives appear to improve the learning of arithmetic.
- Much more research into their effectiveness needs to be carried out.

(Carbonneau, Marley & Selig, 2012)

Relevantly, manipulatives were classified as either ‘perceptually rich’ or ‘bland’ and it is assumed that the dials are perceptually rich in that pupils require only very limited instructions to enable them to use the dials correctly.

Another meta-analysis concluded that manipulatives used in a combination of guided and direct approaches was beneficial for 3 to 6 years old children according to Malofeeva (2009, p.69) where she also confirmed that more research was needed.

She also asserts:

‘Exposing pupils to mathematics instruction early on seems to be a natural step in addressing this difficulty’ [of under achievement in the United States] (p.75) and this is discussed below.

The use of manipulatives is likely to be influenced by teachers’ orientations or styles that are now considered.

2.1.7 Teaching Orientations

Teachers’ varying attitudes to concrete resources are now considered in conjunction with their orientations, Thompson, 1999 (98-102):

- Connectionist.
- Transmission.
- Discovery.

Their characteristics – and consequences – are now considered.

2.1.7.1 Connectionist

This designation applies to those teachers whose approach to teaching and their styles is characterised by making:

- ‘Connections between different aspects of mathematics, for example, addition and subtraction or fractions, decimals and percentages’;
- ‘Connections between different representations of mathematics; moving between symbols, words, diagrams and *objects*’ (emphasis added);
- ‘Connections between pupils’ methods – valuing these and being interested in pupils’ thinking, but also sharing their methods’, by Askew in Thompson, (pp. 98-100).

These should be seen as the natural or spontaneous responses of such teachers to the situations they encounter in their classrooms; they are not learnt or acquired techniques. However, connections can only be made by those teachers who have ‘a sound subject knowledge’ that enables the appropriate connections to be made with confidence as they switch between different representations of number, e.g. equations, graphs, etc., and take their pupils with them. Thus it is natural for them to use manipulatives that model well the teaching point that is being delivered. They are likely to be seen as enthusiasts (for their subject).

These teachers will be described as: Connectors.

2.1.7.2 Transmission

This orientation is well described:

‘Teaching is believed to be most effective when it consists of clear *verbal* explanations of routines (emphasis added). Interaction between teachers and pupils tend to be question and answer exchanges in order to check whether or not pupils can reproduce the routine or method being introduced to them. What pupils already know is of less importance, unless it forms part of the new procedure’, by Askew in Thompson, (pp. 98-100).

Their emphasis is on teaching routines and procedures, rather than learning based on understanding of the number processes. Experience suggests that such teachers often view concrete resources as aids for under-achieving pupils, but otherwise they believe the clarity of their own explanations obviates the need for them. Undoubtedly, transmission teaching orientations can be effective during the early stages of

arithmetic, even though they may be seen as restrictive – hence the “talk and chalk” description.

Anecdotal evidence also suggests their attitude to concrete resources leads them to be dismissive of colleagues who use them because of a belief that arithmetic can be delivered without the need for such ‘props’ - when ‘properly taught’.

These teachers will now be referred to as: Transmitters.

2.1.7.3 Discovery

Such teachers are described as those who:

‘Tend to treat all methods of calculation as equally acceptable. As long as an answer is obtained, whether or not the method is particularly effective or efficient is not perceived as important. Pupils’ creation of their own methods is a valued process, and is based upon building up their confidence and ability in practical methods. Calculation methods are selected primarily on the basis of practically representing the operation. The mathematics curriculum is seen as being made up of mostly separate elements’, (Askew et al in Thompson 1999).

It is likely that their subject knowledge is less than the connectors probably is, otherwise they would not see the curriculum in the ways they do. Their attitudes to the use of educational materials is likely to be open because they will happily use them, but may not be clear about what their actual contributions to learning will be.

These teachers will now be referred to as: Discoverers.

Teacher orientations can be seen as part of Nisbet’s, Confucian/Socratic philosophical framework (2013, p.35) in which the Confucian tradition equates with Askew’s transmission style, while his discoverer approach is akin to the Socratic one. Dehaene’s position is babies’ minds are not blank-slates at birth while suggesting that they are not endowed with great arithmetic and maths skills already in place waiting to be developed (Dehaene 1976, p.56) and Sarama and Clements (2009, p.10) echo his earlier position. In arithmetic (and maths), both teacher orientations are less likely to contribute to helping pupils to become number-ready than through the connections and socialising that takes place as they interact with their teachers and one another.

2.1.8 Attitudes to Manipulatives

Another aspect of teacher orientations needs to be mentioned, based on Smith's study in Glasgow (p.49) where he found that their delivery preferences, loosely defined, were:

- Auditory (word based): 70%
- Visual/Practical (seeing/doing): 30%.

These findings were published shortly before Askew's classification of teacher orientations were. It is now suggested that the two classifications/orientations could be combined as:

- Auditory or Transmitter.
- Visual/Practical or Connector/Discoverer.

Personal observation at workshops suggests that teacher attitudes to the dials and, probably manipulatives in general, were 85:15% Auditory: Visual/Practical respectively. The usual responses of the former were polite interest while many of the later was an immediate: Where can I get some (dials)? The difference between the 70:30% and 85:15% ratios could be explained by the possibility that the Discoverers were undecided, being very interested, but joined the majority because they could not immediately see how their pupils would respond to them.

The general conclusion is that the majority of teachers may not be not sufficiently convinced about the potential benefits of manipulatives to justify changing their established pedagogies to include their use. However, the conclusions of the two meta-studies are that more research still needs to be carried out to confirm the likely benefits that could accrue through the use of manipulatives.

In the meantime, I now add a humbling reflection derived through assembling my thesis that has a direct read-across to changing pedagogies. As would be expected, *Word* has been used and this was straightforward when writing sections of even

chapters without any training, having “picked-it-up as I went along”. However, I had not learned how to get, for example, the correct paginations of the consecutive chapters, a cause of much frustration to be compounded when the Contents with their page numbers were compiled using Tables and all went well until the gridlines needed to be hidden! The point is I had not invested the time and effort to master *Word*, but would it have paid-off if I had? Unlikely, unless more theses or their equivalents were going to be written! This experience provided a good insight on why teachers do not lightly change their pedagogies - especially when they are so overloaded.

One lesson that can be drawn from this experience is the benefits that accrue through keeping pedagogies as simple as possible are considerable, as is believed to be the case when based on the *Sumdials*' approach, in contrast with *Word*.

2.1.9 The Mathematical Enterprise

This is the appropriate time to consider the aims behind the actual teaching and learning of arithmetic, based on the concept of Schoenfeld's ‘mathematical enterprise’.

The attribute that gains admission to the mathematical enterprise is the need to develop a particular way of seeing individuals' worlds by using good arithmetic skills ‘to make sense’ of what is happening through the two processes of:

- ‘Observing, counting, adding, subtracting, multiplying, dividing and estimating, as appropriate’.
- ‘Constructing and solving the equations that represent particular situations or events’, (though to have been by Schoenfeld, but *might* be by me).

These activities are also confirmed, together with the contribution made by ‘confidence and competence with numbers and measures’, Thompson (1999, p.103). In simple terms, the need is to acquire what is commonly known as good number senses such as, for example, the traditional shopkeepers had without pretending to be mathematicians. This raises the question: how is number sense acquired or taught?

An unexpected answer is provided by Resnick when she

‘states that: “several lines of cognitive theory and research point towards the hypothesis that we develop habits and skills of interpretation and meaning construction through a process more usefully conceived of as *socialization* than *instruction*” (1988, p.39)’.

Grouws (1992, p.340).

And again,

‘...becoming a good mathematical problem solver – becoming a good thinker in any domain – may be as much a matter of acquiring the habits and dispositions of interpretation and sense-making as of acquiring any particular set of skills, strategies or knowledge. If this is so, we may do well to consider mathematics education less as an instructional process (in the traditional sense of teaching specific well defined skills or items of knowledge), than as a socialization process” (1988, p. 58).

Grouws (1992,p.340).

These citations are produced verbatim because of their direct relevance to this research and are discussed in more detail later, (p.196). Suffice it to say now; number sense is unlikely to be acquired through instruction, as it is commonly understood in education.

It is immediately clear that the *connector* teacher orientation (p.59) is the one that is most likely to respond effectively to Resnick’s conclusions. One immediate observation is that a start in acquiring number sense is likely to be accompanied by becoming number-ready and, therefore, *before* formal arithmetic learning starts.

Active learning number games are likely to provide effective means of developing number senses especially when teachers are connectors or, at least, when teachers do not feel under pressure to get results and can wait until their pupils have acquired number sense and are truly number-ready and so for their introduction to arithmetic.

2.1.10 Manipulatives and Learning Arithmetic

The origins of the *Sumdials*’ approach having already been described, four other manipulatives that were designed for educational purposes are now considered.

2.1.10.1 Building Blocks

These are widely used to demonstrate one-to-one correspondence (as do fingers) when counting, counting-on and counting-back; these are natural ways of introducing pupils to counting and the ordinal concept of number. They can be readily used to introduce pupils to composition, e.g. for $4 + 3 = ?$, a group of 4 blocks can be created (through counting) and also a group of 3 blocks. They can then be combined into one group and then recounted to arrive at the answer of 7. Similarly, they can be used to demonstrate decomposition, e.g. $7 - 3 = ?$, and the principle of conservation of number.

They are widely used to very good effect, as are also counters, and are a good means of acquiring numerosities, being the links between numbers, quantities and their symbols. It will be appreciated that counting-on and counting-back are not the same processes as addition and subtraction and this is considered more fully, (p.224).

2.1.10.2 Dienes Materials

These were originally devised by Dr. Z P Dienes in the 1930s in response to his and Piaget's convictions that pupils and pupils are essentially constructivist, and not analytic, in their learning. The system has four basic pieces:

- A 1 cm "cube", usually made from wood as are all the other pieces, that represents: 1.
- "Rods" 10 cm long with a 1 cm square cross-section representing: 10.
- "Flats" 10 x 10 cm and 1 cm high representing: 100.
- 10 cm "cubes" representing: 1,000.

They are also known as *Dienes Base-Ten* materials or *Dienes Multibase Arithmetic Blocks*. *Unifix* cubes are a similar concept.

Dienes materials were a big step forward when they were introduced and are well suited to show the sizes of number and their place values. However, success in their learning is dependent on teachers having good subject knowledge so that they are emphasising particular learning-points and ensuring that their pupils make the right 'connections' (see below for a fuller explanation, between manipulating the various

blocks and particular numerical expressions (p.67). This comes about through the appropriate ‘translations’ being made, but, as has already been pointed out, this involves translating from ‘manipulative models’ to ‘written (number) symbols’, a process that can be surprisingly difficult, Behr (1987, p.10). This is the general problem with such three dimensional materials being used to represent non-dimensional numbers. A consequence of this is pupils may do their calculations correctly using Dienes’ materials, but then may not be able to apply them in real world problem solving.

Nevertheless, Dienes’ materials can be useful active learning resources, especially for the base-ten number system.

2.1.10.3 Cuisenaire Rods

Georges Cuisenaire (1891-1975) taught in a primary school in Belgium and devised his system of concrete resources in the 1920s that used differently coloured wooden rods of varying proportional lengths to represent the numbers from 1 to 10 with each rod having its own colour. This was in response to his discovery ‘of pupils’ natural inclination to play’ and possibly became the originator of active learning. However, he was the only user until Caleb Gattegno saw them in 1954. He recognised their potential and became widely used in the UK in the 1960s and early 1970s but did not live up to expectations due to a combination of poor teacher training and insufficient subject knowledge.

The effect was similar to that with Dienes materials in that many pupils became proficient in using their Cuisenaire rods without making the connections between this knowledge and the world of number. In theory, there is no reason why the rods should not be used for active learning, since that was their origin, but teacher training and good subject knowledge would be essential to get the benefits through making the right connections between them and number.

2.1.10.4 Number Lines

Number lines seem to have come into prominence with the introductions of the National Numeracy Curriculum and the 5 to 14 Guidelines in the 1990s. They can be concrete “rulers” with equally spaced counting numbers starting with zero at the left

hand end up to 10, 20 or 100; some had such scales with one on each surface of a “Toblerone” ruler. Often, traditional rulers were used as number lines while pupils also drew their own number lines on paper or dry-wipe boards that included scales appropriate to the actual sum they were answering. They were used as aids in helping pupils to count-on and count-back in single units to groups of units when multiplying – repeated addition – or dividing –repeated subtraction. There is the need to keep a tally of the number of steps being made and this can become an unwelcome complication.

The next stage in their development was empty number lines (ENL) that consisted only of lines with no scales that indicated the actual steps of a calculation. ENLs were a bi-product of Realistic Mathematics Education (RME) that developed in Holland during the 1970s under the influence of Freudenthal in response to the dissatisfaction with traditional classroom methods for learning maths that had little relevance to the realities of daily experience.

2.1.10.5 Comment on Manipulatives

The limitations of each of these approaches are summed-up by Freudenthal’s belief that

‘Cognition does not start with concepts, but rather the other way around: concepts are the result of cognitive processes’ (Thompson 1999, p.28).

It can be concluded from this that the manipulatives discussed above focus on concepts and, therefore, the understandings that were achieved were limited because the skills developed in manipulating them made limited direct contributions to learning arithmetic. This is because a translation step is needed (p.55).

2.1.11 Teacher Issues with Manipulatives

Three teacher issues have been identified about the use of manipulatives in helping pupils to learn the processes of arithmetic, according to Szendrei (1996, pp.423/4). She introduces them with the term ‘commonly shared fears’ and they are:

- The need to learn how to use them correctly.
- Having done so, will there be a worthwhile return – a better pedagogy – from making the effort?

- Most importantly, will they lead to pupils gaining a better grasp of arithmetic and, if so, how should they be used to be effective?

2.1.11.1 Concerns about Manipulatives

The word ‘fears’ is an unfortunate choice of word and, perhaps, ‘concerns’ would have been better. Having said that, the concerns could be legitimate in learning how to use, say, Dienes blocks or Cuisenaire rods effectively because it is not intuitively obvious how they should be used, as has just been explained.

In contrast, experience suggests that the *Sumdials*’ approach (with its dials) is easy to learn and to apply while the explanatory materials are appropriate and can be readily included in existing pedagogies. Moreover, the pupils generally enjoy their dials - especially the boys. Experience suggests that pedagogies are generally improved and, in turn, implies that confidence in teaching arithmetic is improved by the associated teaching plans.

However, there are two other possible explanations for teacher concerns and the first one arises from the general belief, based on experience, that a new method of teaching and learning requires three years to perfect, based on the following assessments by teachers:

- Year 1: 70% performance.
- Year 2: 90% performance.
- Year 3: 100% performance.

This applies even to very experienced teachers. Reasonably, teachers would look for convincing evidence that investment in developing new pedagogies would:

- Produce worthwhile improvements in pupils’ learning.
- Avoid adopting the latest fad such as, for example, brain gyms, the excesses of learning preferences or Gardner’s multi-intelligences at the time of Bramald’s Study.

2.1.11.2 Positive Evidence

The second one is the lack of sufficient and convincing evidence that the use of manipulatives improve pupils' arithmetic attainments. It is believed that this Study will provide such evidence, but much more will still be required: that leads to another issue, namely, the most likely users of manipulatives will be the connectors. However, they are generally perceived as being not "one of us" because of their much greater subject knowledge and enthusiasm for it. Nevertheless, an approach to overcome this deficiency is suggested later, (**Scaffolding**, p.234).

2.1.11.3 Benefits of Manipulatives

It will come as no surprise by now that relatively little has been written so far about the uses of manipulatives, but it is suggested in one paper that while they have a long history of being used in teaching arithmetic, experience of using them has not lived up to expectations. Their contribution can be summed-up as:

‘...practical number apparatus has a role in learning arithmetic through better:

- Understanding of its meaning.
- Gaining familiarity.
- Learning how to get answers efficiently in a range of ways’.

Threlfall (1996, p.11)

He stresses that manipulatives (or practical number resources) have a role in helping pupils to achieve these aims, provided they are *not* used as calculators and the warning is made against “bolting-ons” manipulators to existing pedagogies. Teachers need to be clear that their purposes comply with these aims. The general background in which these points are made suggest that reasons for using manipulatives needs greater clarity; focusing on these aims should enhance their effectiveness. This is most likely to be achieved by connector teachers with good subject knowledge, (p.59). Reassuringly, the title of his Paper also includes the word: Arithmetic.

As a reminder, the above meta-studies both made the point that more research on their usefulness is needed (p.58). It will be clear that once it becomes available it will need to be disseminated in ways that will encourage teachers so that they *want* to integrate the use of manipulatives into their existing and familiar pedagogies.

2.1.12 Conclusion

This completed the original Literature Review, but the observations made during the Pilot Study and the initial stage of the Empirical Study indicated that it was unlikely that the main question would be answered. The reason was the almost total lack of any evidence of automaticity while the pupils were being assessed: they were *calculating* their answers.

In short, the foundations have now been laid ready to create the thesis.

3. Methodology Considerations

3. Introduction

This Chapter considers the methodological issues that led to an action research oriented, objectivist approach being adopted for the Empirical Study in response to the need to obtain suitable data for statistical analysis to answer its research questions. It is in two sections being:

Section 1: The requirements for an effective research approach.

Section 2: Description of the influences that led to it being modified.

Section 1 includes a description of the InCAS computer adaptive system from Durham University's Centre for Evaluation and Monitoring (CEM) and explains why it was chosen for the data collection. Section 2 explains the reasons for dispensing with the control schools.

3.1 The Requirements

3.1.1 Methodology

This section is based on Cohen, Manion and Morrison (2007, pp.5-14). It describes how man came to understand his environment and in doing so became able to adapt to it and survive. At first, this could have been done through collective *experience* and recollection such as the migrations of animals and birds and the changing seasons with their associated weather patterns. In time, it would be *reasoned* that these were recurring and predictable events that could be codified by the equivalent of: "Red sky at night is the shepherd's delight; red sky in the morning is the shepherd's warning" to

become part of the folklore. It is likely that the older people – the survivors - would have the greatest experience of these cyclical events and, through this, would become authority figures.

It would not be surprising if they become selective by choosing only those coincidental observations that would support their hunches and convictions. The resulting tensions of such practices were eventually resolved through more rigorous approaches – *research* – by developing what is now known as the scientific method.

Again, its eight stages are:

- Hypotheses, hunches and guesses.
- Experiment designed; samples taken; variables isolated.
- Correlations observed; patterns identified.
- Hypotheses formed to explain regularities.
- Explanations and predictions tested; falsifiability.
- Laws developed or disconfirmation (hypothesis rejected).
- Generalisations made.
- New theories.

Hitchcock and Hughes (1995, p.23).

The starting stage of hunches and guesses was probably the same for both laypeople and scientists. Where they parted company was the scientists recognised the need to control variables and to avoid jumping to conclusions based on coincidence rather than demonstrable effects. Examples that have stood the test of time include Pythagoras's theorem - proving the relationship between the sides of right-angled triangles or Boyle's law – the principle that the pressure of a gas varies inversely with its volume at constant temperature – illustrate the results of this approach. Its practitioners also understood that a theory is only as good as the data that supported it. New observations or data would necessitate existing theories being modified or rejected and replaced by new and better ones.

The scientific method greatly advanced understanding within the natural (physical) sciences. However, the environment consists of much more than the natural sciences: understanding the relationships between people and how they respond or react with

one another became increasingly important as populations grew and societies evolved. The differentiating feature is that every person is unique in every respect and does not have consistent properties like the elements. In short, ‘they are not robots or puppets controlled by some external force’: they have their own feelings, hopes and anxieties, and *freewill*. These allow them to be unpredictable and to influence their own environments, often in unexpected ways. This is in contrast with the application of the scientific methods in STEM subjects - science, technology, engineering and mathematics - as the examples of Pythagoras’s theorem and Boyle’s law testify. However, the scientific method provided a good model to improve understanding of the environment generally and was adapted by researchers in the fields of human behaviour to establish the social (human) sciences that include education. Besides teaching and learning, it was also concerned with wider social concepts such as equal opportunities and the effects of deprivation on pupils’ learning in schools that are not immediately amenable to precise definitions, in contrast with the natural sciences. The above very brief historical overview provides a setting for consideration of the contrasting assumptions of the *subjective* and *objective* schools and the nature of social science. These are now summarised and then considered.

Subjective and Objective Comparisons

The Subjectivist Approach		The Objectivist Approach
Nominalism	← Ontology →	Realism
Anti-positivism	← Epistemology →	Positivism
Voluntarism	← Human Nature →	Determinism
Ideographic	← Methodology →	Nomothetic

Burrell and Morgan (1979)

Table 3.8: Assumptions made by the Subjective and Objective Schools of social science.

The concepts that lie behind these contrasting pairings are now briefly considered.

The first one concerns the differences between the assumptions of the nominalists and realists about the fundamental essence or being (ontology) of the sociological

phenomena being examined. Are their meanings only words and the *internal* creations of individual minds – nominalist – or *external* being observable without having been created in people’s minds – realist? An example from education might be the effects of parental support on their pupils’ progress at school.

The second pairing deals with the nature of knowledge, how it is acquired, how it is communicated and how it is passed on from generation to generation; this is the epistemological perspective. The contrasts here are between the view that knowledge is ‘*personal*, subjective and unique’ compared with the view that knowledge is *impersonal*, ‘hard, objective and tangible’. Researchers’ own views are likely to influence the course of interviews with participants in identifying attitudes on particular topics; in contrast, the role of the researcher for the latter is strictly that of an observer. These are also known as ‘anti-positivist’ and ‘positivist’. The research questions themselves are likely to determine the choice of the appropriate methods, as is the case in this Study.

The third pairing concerns the influences on the development of human nature itself. On the one hand, there is the view that human beings are the creators of their own environment through the exercise of freewill and creativity – voluntarism - while the contrasting view – determinism – is that human beings are products of their environments and so have no influence on it. Such views help to determine whether or not a study requires the *personal* approach of social science or the *impersonal* approach of natural science.

The final pairing determines the methodology itself because the research questions constrain the options. Thus, if the view is that the research focus is essentially *subjective* in nature, then the choice of methods is likely to be consistent with the nominalist, personal and anti-positive assumptions that require involvement by the researcher – *idiographic*. The alternative view is the research context is akin to those encountered in the natural sciences and is *objective* resulting in realist, positivist and determinist assumptions where the choices will be limited to quantitative and observational methods – *nomothetic*.

This summary provides the framework that influence researchers’ choices of methods in social science as they seek answers to their questions. Awareness of it is helpful, particularly in education where teaching and learning takes place in classrooms and

not in laboratories (this aspect is considered later). It is noted that there are other models besides the subjectivist and objectivist ones that include critical theory, feminist theory and the newly emerging complexity theory. All have their strengths and weaknesses, like the subjective and objective models, but are not considered here because they are not directly relevant to the research question.

3.1.2 Choice of Methodology

It will be apparent from this brief consideration of the subjectivist and objectivist models that there is something of a *rigid either-or labelling*. While this may be appropriate for the natural sciences with its immutable laws and proven theories, it is much less so in the social sciences where the emphasis is more on seeking to discover ‘relationships and causalities between human phenomena’, Cohen, Manion and Morrison (2007, p.11).

To revert to this Study, Bramald’s 2001 Study and the subsequent longitudinal study were unfinished business in that both gave clear indications the *Sumdials*’ approach had improved pupils’ number attainments, but they were not generally accepted as being conclusive. The aim of this Study was to determine whether or not the approach led to statistically significant improvements in automaticity; this was to be achieved by replicating the objective parts of the previous studies, but under more rigorous supervision. This dictated that the same classical treatment/control comparisons would be made again and was accepted as something of an unconsidered *fait accompli*.

However, there was a compounding factor in the choice of methodology. It was the need to carry out the pupil’ assessments during the appropriate windows in the school year. Typically, that meant doing the initially assessments immediately after the autumn half-term break and the subsequent ones immediately after the summer half-terms breaks. The effect of this is a year could be lost if a window was missed. To avoid this, urgent priority had to be given to setting up the initial assessments at beginning of the winter term even though it was not ideal. These time pressures perhaps obviated a more reflective consideration of alternative methodologies.

Nevertheless, even with the benefit of hindsight, it is very likely that the same methodology would still have been chosen, at least initially, even if more time had been available to consider other options. The two main reasons for this were:

- This Study was effectively a continuation from the two pre-studies that were both primarily quantitative studies and this strongly influenced the design of this one.
- The need to assess objectively the effectiveness of the *Sumdials*' approach developed in response to my colleague's experiences and a more general perception that teaching and learning number in schools was not 'working'.

The need to establish what pupils' attitudes to number – the qualitative/subjective aspect – was not seen at that stage as being the immediate priority.

On a personal note, I confess that I am an unrepentant objectivist, as an engineer, a management consultant and then working in manufacturing industry. I have been using the scientific method throughout my adult life and could be described as an instinctive hypothesiser, huncher and guesser! Thus, it is almost inevitable, with such a background, that I would ask *objectivist* questions.

As has already been mentioned, my Supervisor and I discussed two main approaches to quantitative studies, such as this one, and they are:

- Define the methodology at the outset and then stick rigidly to it, as would be the norm in medical research with the 'intention to treat' model.
- Define the methodology at the outset and then adapt it to any unexpected circumstances and opportunities as they occur.

Hollis and Campbell (1999).

Apparently, I used the second approach all along and he attributed this partly to my engineering background and its resulting mind-set that deals with problems and the unexpected systematically. My only comment is that it happened naturally and, to me, was the only feasible option for a longitudinal study in the messy environments of primary schools where the unexpected rules. As such it emphasised the validity and ecological validity (Bronfenbrenner, 1976) in particular of the Study, at the cost of reducing its overall reliability. However, the alternative model requires near

laboratory conditions whereas conducting research in school classrooms, as this Study did, may benefit from accommodating their typical messiness's.

A topical analogy in contemporary Edinburgh would involve planning a car journey from its west side to the centre while the tram tracks were being laid. Having made the plan, the actual journey would encounter ever-changing road closures or diversions leading to continuous modifications to the plan as the journey progressed because of the general messiness or chaos, as taxi drivers described it. But with perseverance journeys were usually completed.

Since primary schools are messy, it is now to be expected that the methodology changed as the various studies progressed. Such changes and their causes are described. Pragmatism became the determining feature of this study as workable solutions were adopted in response to the unexpected; it is not claimed they were necessarily the best ones because all options could not be considered in the available time and the emphasis was on taking action to solve challenges as they arose.

Consideration is now given to the sample design.

3.1.4 The Sample Design

The sample design had to meet the following requirements if it was to answer the study questions:

- Be a longitudinal study over at least two academic years,
- Have 16 primary schools taking part in the study (14 being the minimum for sufficient statistical power), and
- Involve schools from both England and Scotland, so as to establish a level of generalisability based on the contrasts between these two systems.

These are now considered.

3.1.4.1 Study Duration

As background, the usual experience of short interventions in teaching and learning number is that their effects wash out within two or three years. The earlier

longitudinal study strongly suggested that the *Sumdials*' approach had made a lasting impact and its confirmation would be useful. However, the requirement is that part-time distance learners would submit their theses after a minimum of six years. This means empirical studies would need to start during the third year having reviewed the relevant literature, but that did not happen with this study for several reasons. Nevertheless, the Pilot Study did start during the fourth year and that enabled three years' data to be collected, (p.96).

3.1.4.2 Sample Structure

16 schools satisfy the custom-and-practice expectations to establish sufficient 'power' for such studies (Ellis, 2010) - and provide a safety margin against up to two schools dropping out without reducing the power of studies. This allows eight pairs of comparable schools to be established, based on socio-economic indicators that avoid 'leafy suburb' and schools from areas of high deprivation being paired together. (Bramald's Study (2001) described the pairing process well.) One from each pair would become the treatment school, based on a random choice (toss of a coin), to implement the study intervention while the other school would be a control following its normal programme.

3.1.4.3 Secondary Questions

The English and Scottish education systems and curricula are different and the Study provided a good opportunity to measure differences in number attainments, if any, between them to answer one of the secondary questions of the Empirical Study. Good introductions to primary schools in Co. Durham were provided and five of them signed up along with schools from Edinburgh.

The data to answer the main question would be collected from both boys and girls and could be used to find whether or not there was any statistical differences between the genders.

Consideration is now given to the measurement pupils' number attainments.

3.1.5 Measurement Issues

This section draws heavily on the work of Bond and Fox (2001) that, in turn, was applied in the development of computer adaptive assessment systems such as the Interactive Computer Adaptive System (InCAS), as described by Merrell and Tymms (2007, p.30). As explained already, measurement is largely taken for granted in the objective studies of the natural sciences and these can range from the heights or weights of pupil (physical) to their abilities in arithmetic. Even though measurements in the natural sciences are routine, it cannot safely be assumed that such measuring is always of sufficient rigour, as is now shown.

The constructs with equal intervals used to measure heights, weights or temperatures are not *absolute* units of measurement that could be re-established *ab initio*. They are all abstractions that have been developed from very large data samples and much iteration to arrive at the ever-increasing confidence in their reliability and repeatability; this has reached such a level that their accuracy is largely taken for granted. To illustrate the reality behind what is taken for granted, imagine the proverbial Englishman and Frenchman being shipwrecked on a desert island with no common language or measuring constructs. The only data available to them would be their personal statistics expressed in either imperial or metric units. How would they create any reliable constructs to help them survive?

In contrast, temperature scales could be re-established with the aid of an un-calibrated thermometer by making a mark against the mercury level when water freezes and then making another mark when it comes to the boil. Repetition with the same thermometer would always result in the marks being in the same positions as previously. After that a scale could be constructed by dividing the distance between the two marks into whatever equal intervals were appropriate as, for example, dividing by 100 to create the Celsius scale.

However, it is only as recently as the last century that the need for suitable constructs was accepted within the social sciences so that their findings would enjoy the same standings as those of the natural sciences. This meant finding ways that could be used to measure attitudes or difficulties, for example, so that they could be expressed using linear, equal unit, additive scales. Such constructs would also have to produce

repeatable results when different investigators made the measurements. Then, and only then, would there be confidence in the results of empirical studies in the social sciences.

It is generally assumed that if the results of tests (measurements) are expressed in linear scales they must be reliable even though no consideration has been given to their suitability or validity. Teacher-set tests are likely to use spur-of-the-moment uncalibrated questions and the relative *difficulties* of the actual questions would be unknown, as would the overall difficulty of the test. Yet, awarding the *same* marks for each correct answer, totalling them and then expressing them as a percentage is widely accepted as good practice. The reality is that such totals are only raw data that need to be processed appropriately before reliable conclusions can be drawn.

The diagnostic value of being able to translate from one to another mode has already been described (p.55) and its principles can be adapted to arrive at measures of degrees of difficulty that holds the key to reliable number assessments. Golf handicaps illustrate some of the principles. The point in common is that golf handicaps are measures of players' *abilities*.

Now each golf course is unique (like each school), while the majority of other ball games such as football, rugby or tennis are played on standard pitches or courts of prescribed dimensions. Usually, a golf course has 18 holes of varying lengths between about 125 and 575 yards to give a total length within a range of 6,000 to 7,500 yards. Each course is given a Standard Scratch Score (SSS), based on its length. This is the number of shots that a "scratch" golfer - one who would be *expected* to make no mistakes - would play in a round of golf. However, golf courses, besides being of different lengths, are also of varying *difficulties* and SSSs are adjusted up or down by a small number of strokes, typically no more than three, in recognition of this. Such adjustments are based on analyses of a large number of returns (total marks) over time and are continually reviewed – or recalibrated - in light of new ones. Again, SSSs are the equivalent of correctly answering all the questions in a test.

As in an arithmetic test, not every golfer is *expected* to achieve a scratch score – no mistakes - and to allow for this an initial handicap is calculated as the difference between the SSS and the mean of three returns (scores/tests). This can range from scratch (none) up to a maximum of 24 shots. A golfer's handicap is reviewed after every competition (test) and it may be left unchanged or adjusted up or down, based on his actual score. Handicaps are expressed on a *linear* scale (equal intervals equivalent to one shot) and are an inverse measure of a golfer's golfing *ability*, i.e. the lower the handicap, the better the golfer. Again, a handicap indicates a golfer's *ability* and how he is *expected* to score in competitions (tests).

The main observations that can be drawn from this system that are relevant in measuring pupils' abilities in arithmetic are:

- The three emphasised words 'ability', 'difficulty' and 'expected' are used deliberately, being the accepted terminology associated with developing arithmetic tests.
- SSSs are based on a large number of returns from members and are kept under continuous review to maintain confidence in them.
- In contrast, individual players make relatively few returns and so there is less confidence in the reliability of *individual* handicaps – even though a player could answer between 70 and 110 questions (shots) in a round of golf.
- However, the linear handicap construct is based on raw data and gives the ordinal values of players' handicaps, but as is now explained, it does not accurately indicate differences in abilities. For example, the abilities of 23 and 22 handicap golfers are very similar and there would be an almost 50:50 chance that either of them would win when playing against one another. However, the differences in ability between 1 and scratch (0) are greater and the scratch golfer would be expected to win most times when playing against a golfer with a handicap of 1. This is a generally acknowledged and accepted limitation that causes little harm – handicaps only apply to amateur golfers. The point is the equal intervals do not equate to equal differences in abilities

(Note: 24 is the maximum handicap that is awarded (to men), but the actual mean of 24 handicap members in a club (school) would be higher because many of them would not be expected to play to 24.

- A golfer's handicap is nationally recognised so a member takes his handicap with him when he joins another club (school) and it is a measure of one *attribute* only, i.e. his ability to play golf and tells his new club nothing about his other attributes. If the course at his new club is more difficult than his previous one, his handicap would likely be adjusted upwards (moved to a lower class).
- *Unexpectedly* good scores are most likely to occur with new golfers who are improving rapidly. This leads to their re-handicapping (moving to a higher class), rather than the more usual single shot adjustments, to bring them more into line with how they would be *expected* to score in future. Also a player usually has lucky shots (equivalent perhaps to good guesses in a test) in a round that *could* also lead to a reduced handicap (moved to a higher class).

A similar approach is available to measure pupils' arithmetic *abilities*, but it has one important difference. It is that the *difficulties* of number items (questions) used to measure abilities have been established and then expressed on equal interval constructs. This has been achieved by experienced teachers and researchers compiling banks of items of varying difficulties and pre-testing them using pupil of known abilities who would *all* be expected to answer *some* of the items correctly while *none* would be expected to answer *all* of them correctly. This confirms the range of difficulties is right, whereas if some pupil cannot answer any items or some can answer all of them, it cannot be known by how much they are either too difficult or too easy.

Easier items will have a higher proportion of correct answers and, hence, show an inverse indication of their difficulties. It is unlikely that the initial results will be exactly as expected: some items will get 'erratic' scores that are unexpectedly poor or good. This signals the possibility that their wording, for example, may be ambiguous, so changes and further testing would be needed before they are ready for use. The

mention of ‘rewording’ illustrates an important point and it is the original wording was possibly beyond the pupils’ *verbal* abilities with the result that two attributes would have inadvertently become conflated.

When pupils sit the tests, their answers are marked to get their individual scores and the *ordinal* raw data. These are then processed using an approach such as Rasch measurement (Long, Wendt & Dunne, 2011) to arrive at equal interval constructs of the difficulties of items achieved through log transforms and appropriate Rasch software that checks the acceptability of the items. Besides preserving the original ordinal order, they give improved indications of the relative difficulties of the individual items. The actual difficulties can then be expressed as, say, the age equivalent scores (AESs) of pupil with known abilities who would have a 50:50 chance of getting the correct answers. This is in contrast with traditional marking where adding two integers is treated as having the same difficulty as a long-division item and both correct answers are awarded the same marks. While Rasch measurement would provide golf handicaps that would reflect golfers’ actual abilities more accurately, little would be gained since the present arrangements are widely accepted.

This account has brought out some important principles that enable repeatable measurements to be made with confidence even when using different investigators. They include:

- The larger the banks of tested items, the greater can be the confidence that the difficulties of individual items are reliably calibrated and, in turn, used with increased confidence to assess the abilities of individual pupil.
- Much iteration has been necessary to get this far and as more items and data are added the greater will become the confidence in the constructs. This is the process that was started in the natural sciences by Gauss - his “bell” curve – during the early 1700s.
- Each question must be such that it only measures *one* attribute when assessing the arithmetic abilities of pupil.

It was concluded, based on this description, that the InCAS computer adaptive program, developed by The Centre for Evaluation and Monitoring (CEM) at The University of Durham, would measure pupils' arithmetic abilities well. In passing, such assessment programs are becoming more widely used and this trend is likely to continue and the publication of "Key Stage 2 testing, assessment and accountability review: Final Report", by Lord Bew in July, 2011 supports this assertion. For example, it recommends that: 'Maths should continue to be externally tested'. The InCAS program meets this well because the assessments are derived from data collected *directly by computers*, as described later (p.97). Also the data are available in subsequent years to measure pupils' *progress* and this satisfies another Bew recommendation to give: '... at least as much weighting to progress as attainment'.

The essential requirement is that the data are reliable, be it for schools or research studies. The following table shows that InCAS performs well in this respect. It is also important that there is an internal consistency to the items. Cronbach's alpha, a psychometric instrument, confirms this in that only scores in the range of 0 to 1 are considered and scores greater than 0.70 signify acceptable internal consistency.

Internal Consistencies of InCAS

Session	Item	Person
Picture Knowledge	1	0.89
Non-verbal Ability	0.96	0.86
General Mathematic	1	0.97
Mental Arithmetic	0.99	0.96

Table 3.9: The internal consistencies of the InCAS scores are high.

These confirm that InCAS data are likely to be reliable with such scores at both item and person (pupil) levels. (Cronbach's alpha: *experiment-resources.com*).

In short, the InCAS computer adaptive assessment system is a very appropriate way to measure pupils' initial abilities and the subsequent changes in their arithmetic abilities, as a longitudinal study would require.

3.1.6 InCAS – its Program and Applicability

The InCAS program can assess the following Sessions (subjects) for pupils' chronological ages between five and eleven years old:

- Reading.
- General Maths.
- Spelling.
- Mental Arithmetic.
- Attitudes.
- Developed Ability.

The relevant ones for this Study are:

- General Maths.
- Mental Arithmetic.
- Developed Ability.

Mental Arithmetic and General Maths selected themselves because the data collected from them after statistical analysis should answer the questions of the Empirical Study. Developed Ability was included in the Empirical Study at the suggestion of the University's Centre for Evaluation and Monitoring (CEM) and its value is described below, (p.86). Comments on its contribution are made later, (p.104).

The General Maths questions are curriculum-based and cover the four areas of:

- *Number 1.* This includes counting, informal arithmetic (i.e. a number problem presented as: 'Here are 6 ice creams, if 3 are taken away how many will be left?'), partitioning and place value, fractions and decimals.
- *Number 2.* This deals with sorting, patterns, formal arithmetic, problem solving and algebra.
- *Measures, Shape and Space.*
- *Handling Data.*

Mental Arithmetic, in like manner, assesses the four basic arithmetic processes of:

- Addition
- Subtraction
- Multiplication
- Division

These complement the General Maths scores. The relevant feature is the time taken to answer each question is recorded. It can be inferred that if it is more than, say, four or five seconds (to read the question and then manipulate the mouse/touchpad to the selected answer – one of four choices) - it has been *calculated* and not recalled automatically. This would allow assessment to be made of the effectiveness of the intervention to improve automaticity. Subsequently, it was found that accessing the individual times was not straightforward, (p.131).

InCAS can be run on school computer networks or using stand-alone computers, usually laptops. Eight laptops would allow about 40 pupil doing the three sessions to be assessed in a day and would be very attractive for this Study. The typical elapsed time to assess a pupil is between 30 and 40 minutes.

Other InCAS attractions include:

- It has a growing track record with its substantial and ever-expanding database as about 120,000 pupils are assessed each year. At such a size, it becomes a *de facto* national ‘control’ based on pupils’ *ages-at-test*, rather than their school years. This feature reduces the need for control schools where the majority of the pupils being assessed would also be following their normal programmes; this aspect is considered more fully, (p. 93).
- It would have high acceptability to pupils following their early exposure to computers, especially at home. Thus they would be comfortable about using keyboards and touch-pads. Their assessments would be seen as “quizzes” and not tests. This contrasts with written answers that can be more stressful (Terzis & Economides, 2011).
- Another important benefit is teachers do not have to set and mark tests; this is generally very welcome and understandable due to shortages of time.

Teachers find the diagnostic value of the assessments in identifying ‘gaps’ in pupils’ knowledge very useful and the greater objectivity behind the scores helpful. Independent and ‘objective’ scores also provide a welcome support for teachers when meeting ambitious parents who may hold unrealistic beliefs about their own pupils’ abilities.

The Developed Ability session should provide a more general domestic/socio-economic measure of the pupil that would be helpful to teachers as they assess pupils’ overall strengths and weaknesses, (Merrell & Tymms, 2006). It has two parts:

- Picture knowledge (vocabulary) assessed by presenting a word (mainly of everyday objects) together with five pictures, one of which corresponds to the given word to be selected as the correct answer.
- Non-verbal ability uses the Problems of Position (POP) test developed by David Moseley (1976). A split screen is used on which patterns of up to six dots are shown on the left half and are “filled-in” using the keypad. Then, the same pattern is included within many more apparently random dots on the other half of the screen; the matching dots from the two screens have to be selected.

Pupils have six minutes to answer as many questions as they can from each part.

The claimed value of this session is the knowledge for either part is not usually learned through formal subject teaching and is typically *acquired* through everyday social activities, both inside and outside class rooms. Thus they provide an indication of pupils’ abilities as they develop and very low scores can be an indication of social deprivation or unsatisfactory domestic situations leading to limited development of their intra-personal skills (managing themselves) and inter-personal skills (dealings with other people). Importantly, they can also indicate whether or not pupils have reached the appropriate thresholds that would allow them to benefit from formal subject learning, (Tymms, 2010, EDUC: 00225)).

The results from all sessions are expressed as Age Equivalent Scores (AESs) using the YY: MM format. This then allows pupils' Ages-at-Tests (A/Ts) (also in the YY: MM format) and AESs to be compared directly to get indications of how individual pupil, classes or schools are progressing,) and are easy to interpret, according to Merrell & Tymms, (2007, p.31). Thus, a pupil has made progress when the differences between AESs and A/Ts are positively greater than previously and *vice versa*. This meets one of the Bew's recommendations. The general point is that much of the complicated statistical work is done "behind the scenes" to arrive at formats that are easy to understand – and are sound and reliable: this is important for teachers, schools and parents, and also researchers.

3.1.7 The InCAS Process

The three main InCAS stages in assessing pupil are:

- Preparing the computers for assessments.
- Assessing the pupil
- Processing the data.

The initial preparation for stand-alone laptops, before going to a school, includes entering into the laptops the school's name, a class identifier and then the pupils' bibliographic data (first name, surname, DOB and gender). Three letter passwords, one for each session, are then generated for each pupil and uploaded into *each* laptop. Computer prompts make these processes straightforward. Normally, the actual assessments are conducted in a 'spare' schoolroom and the best desk/table arrangement is a straight line to minimise distracting eye contacts between pupils.



Illustration 3.5: Set-up of the laptop computers ready for InCAS assessments.

It takes about 30 minutes for two people to set-up eight laptops. Stand-alone computers offer great flexibility in that pupil can be assessed in whatever order suits the teachers and this can include changing from one class to another to accommodate PE lessons – another example of the messiness of primary schools.

The overall assessment procedures are simple. Pupils are collected from their classroom, settled at their laptops and their unique passwords are entered. Their personal data are displayed for their confirmation before starting their assessments. This is repeated for each session and then they return to their classroom once they have completed their assessments. The pupils do their assessments at their own speeds and the resulting data are generated entirely by the individual pupil interacting with their laptops without *any* other interventions and, especially, by their teachers who are not present. The pupils mostly enjoy their “quizzes”.

On completion, the assessment data are uploaded onto storage devices, such as memory sticks, and then sent to CEM for processing, usually within two days. The results are available in a variety of formats of which the standard one meets most needs. The following screenshot illustrates a typical layout:

Sample InCAS Results Sheet



Name	Age (Yrs:Mths)	Age Equivalent Scores (Yrs:Mths)			
		Number1	Number2	MSS	Data
John McCall	8:0	7:4	7:2	8:0	6:5
Kevin Murray	7:11	8:6	10:0	7:5	7:0
Chris Murray	7:8	7:3	6:5	7:1	7:8
Caroline Murray	8:5	8:0	8:6	7:7	9:4
Journe Murray	7:11	6:11	7:3	5:3	6:6
Shamh McFarlane	8:0	6:4	7:0	7:6	6:1
Sophie McCreave	8:3	7:10	7:11	7:11	8:10
Colin McCreave	8:2	8:1	8:4	8:4	8:3
James McCreave	7:7	6:6	6:11	7:8	7:5
Ben Wilson	7:10	5:4	7:0	6:5	4:2
Madison Pringle	7:11	9:1	8:3	8:5	8:11
Kyle Sanderson	7:8	7:1	7:1	7:11	8:3
Carlin Smith	7:8	8:2	8:0	6:4	7:2
James Smith	8:2	7:6	9:9	7:10	7:2
Oliver Smith	7:8	6:7	6:9	7:3	7:8
Mark Wake	7:7	6:4	5:0	6:7	6:4
James Wake	8:4	7:1	7:7	7:7	9:0

Table 3.10: The layout of the InCAS G/M modules and assessment scores.

The following table sets out the main stages with their timings for the Empirical Study:

Pilot and Empirical Studies Timelines

Date	Activity	Remarks
Autumn 2009	Find and 'sign-up' 16 primary schools for Study.	Schools to be from Co Durham and Edinburgh.
Winter 2010	Pilot Study with first year pupil in Edinburgh Schools.	Easier to monitor, being local to researcher.
Jan. to June, 2010	Implementation fidelity visits to Pilot Study Schools.	Liaise with staffs of schools and monitoring.
June, 2010	Re-assess and review results for each session by each school.	Check how the controls/ experimentals compare.
Autumn, 2010	Assess first year pupil of all schools in main Study and second year pupil of P/S	Include staff training of experimental schools on the intervention.
June, 2011	Re-assess, analyse and review all	Review assessments, staff

	assessments of each control and experimental School.	feedbacks and progress of Study.
Autumn, 2011 to June, 2012	Continue implementation fidelity visits to all experimental Schools.	Include general liaison Brief staffs and respond to feedbacks.
June, 2012	Re-assess, analyse and review all assessments of each control and experimental School.	General review of results, progress and findings.
Autumn, 2012 to June, 2013	Continue implementation fidelity visits to all treatment Schools.	A study of three years duration makes for a stronger longitudinal study.
June, 2013	Re-assess, analyse and review all assessments of each control and treatment School.	Final analysis of data and write-up. Summaries of findings and thanks to all participating schools.

Table 3.11: Time Line for the Pilot and Empirical Studies,

3.1.8 InCAS Conclusion

The merits of the InCAS computer adaptive assessment system are many with the main ones being:

- Assessments made by computer with no teacher interventions.
- High reliability of scores (at class levels, especially).
- Ease and flexibility of use.
- High acceptability to pupils.
- Speed of assessments.
- Independent reports.
- Good diagnostic value for teachers.

For these reasons, it was concluded that InCAS would be an essential resource for this Study.

So far, the methodological issues that influenced the design of the Empirical Study have been described and are consistent with an objectivist approach of collecting data for statistical analysis from pupil in their early years of formal education at 16 schools.

The *actualitié* is now considered.

3.2 The *Actualité*

3.2 Introduction

This section describes the influences that led to the modification of the plan for the Empirical Study. There were no major and unexpected developments, but a series of mostly trivial events contributed to an overall impact on the course of the Study itself. They centred on the question of the value of the control schools to the research design and the difficulties of determining school's choices about their involvement in the research.

Most of these events were part and parcel of the messiness of primary schools. The relevant ones are described to explain the changes in the Study Plan.

3.2.1 Background

The original plan was to have both treatment and control schools in the Pilot and Empirical Studies, believing both were necessary, but this proved not to be so in the case of control schools. Some of the reasons for changing the design are now given.

As background, the word messy has already been introduced to describe primary schools and this is due to their general unpredictability from many influences, both internal and external, that result in the unexpected becoming the norm. The staffs accept this, as do the pupil because it is all they have ever known. This makes for an exciting environment for research, but demands flexibility and perseverance on the part of the researcher.

To explain the change, the usual incentive to become a control school, at the time of Bramald's Study, was to offer resources and staff training that the treatment schools had received, once their results were known. This was no longer sufficient incentive due to the many pressures schools operate under now – especially, when new pedagogies would be required. Thus agreement to be a control school usually became dependent on the personal goodwill of the head teacher or deputy-head teacher believing it would help a research study. However, they are susceptible to pressures from their staff responding to the disturbances that undertaking the assessments caused; teachers are more concerned about immediate needs.

The effects of these is borne out by the following table:

Control School Participations

Assessment Times (Months)	6	9	21	33
Control Pupil	154	Not tested	76	15

Table 3.12: Number of Control Pupil assessed each Time (Month)

Only the main explanations for the declining involvement of the control schools are now given and include:

- During the elapsed time of the Study, two out of four deputy-head teachers left and it was decided not to re-assess a third school to avoid being a distraction as it prepared for a forthcoming HM Inspection.
- The head teacher of the remaining school thought her school was still a treatment school until she discovered after the month 9 re-assessments had been arranged that her teachers had unilaterally decided at the outset not to take part in the research without informing her. However, the school (15 pupils) is in a high deprivation area and was re-assessed as a goodwill gesture in the hope that it would return to the fold. It did not. (A nice example of the messiness of schools today?).
- The time and effort in preparing beforehand, conducting the assessments and then getting the resulting data processed is considerable and can only be justified by their contribution to the research.
- The clinching episode (from the Empirical Study) occurred when one school agreed (through its numeracy co-ordinator) to take part in the Study as a treatment school, but withdrew at the last minute. Its reason for taking part in the first place was its numeracy results had to be improved and it was believed the *Sumdials'* approach would achieve this. However, shortly after the research team had made the initial assessments, the head teacher had decided to use another intervention (Directed Mentoring) and it was accepted that the class could only continue in the Study as a control. In the event, the class was re-assessed but eventually dropped out through no replies to e-mails – another example of the varying pressures on schools
- To include a positive point, the resources (from that school) were returned and then supplied to a new experimental school at very short notice that, in the

event, became very supportive of the Study. The way in which this school coped with the unexpected illustrated well that there could be a positive side to the challenges of working with schools.

It will be realised that the above episodes are just the tip of an iceberg and there were continuing lesser ones such as double-booking assessment days, planned assessment classrooms being used by another class, assessment rooms with no power supplies for the laptops and many more.

However, the over-riding reality is that control schools, based on such experiences, only make a limited contribution towards studies such as these. The reason is control schools can no longer fulfil the key requirement for being a control, namely continuing with their *normal* teaching and learning number programmes to provide a homogeneous group against which the treatment classes can be compared. This follows the many introductions of their own new initiatives to help them achieve their targets in the current highly pressurised school environments that contrasts with the apparently calmer era of Bramald's Study.

Put simply, control schools only participate as a goodwill gesture and what they do, or do not do, is beyond the control of the research team. In contrast, researchers using the treatment/control (placebo) classical double-blind model that is widely used in medical research studies control *both* the experimental and control groups. This is the crucial difference. The alternative to the use of control schools is now considered.

3.2.2 An Alternative Control

Many problems arose in this study with implementing the initial plan to use a randomised control design based on 16 schools with eight pairs of treatment/control pairs, four being in Co. Durham and four in Edinburgh. The three main difficulties were:

- The logistical load associated with organising and running such a large study properly was too great for the research team to manage.

- Only seven schools could be recruited in Edinburgh for the Pilot Study of which three became treatment and four control schools. Thus, it became an opportunity sample from the start and not a randomised sample with a consequent degrading of its results
- Assessing the control schools was imposing an unwelcome burden on the research team and they did not always appreciate their own results – the ‘halo’ effect.

This meant a replacement control facility has to be found quickly while continuing to use the InCAS assessment system that had been acquitting itself well.

As a general rule in quantitative studies, the greater the size of the sample the more confidence there can be in its results and indeed the plan to use 16 schools met the criterion of having more than 14 schools to ensure sufficient statistical power. Nevertheless, the many and varied problems that had been encountered, especially with the control schools, ruled out consideration of replacing them by more committed schools.

The alternative that presented itself was the InCAS database made-up of all the approximately 120,000 primary school pupils who have been assessed each year for many years; it is used to assess the first six years of primary education. At such a size, it can be seen as a truly representative and stable sample that has become a *de facto* national standard. One effect of this is that a class, school or group of schools can be assessed against it with complete confidence in its consistency. To explain, there will be multiples of approximately of 60 pupils on its database for each year of its life with *exactly the same age* as each pupil within the assessment group to compare their scores with its mean scores for each subject. As a rule of thumb, a sample size of 60 or more cases is usually regarded as being sufficient. Such groupings are always available even though there will be a degree of randomness about the dates arranged for each assessment.

CEM is confident in the reliability of the mean scores for classes or greater. There would be nearly 300 pupils taking part in the Empirical Study and, importantly, this would give complete confidence in the mean scores of the individual participating classes. All the individual scores are provided in the results sheets and can be used readily for diagnostic purposes. A cautionary point is an individual pupil is *a sample*

of one who could have a ‘good’ day with better than expected score or *vice versa* with a lower score. Having said that, the *trend* of her scores when compared with the mean scores of her age group over successive assessments would be informative.

In summary, adopting pupils’ ages-at-test as the study control was to be an expedient, but it was believed that the InCAS database constituted a more homogeneous and reliable control than could be expected from only several not very committed control schools that would be naturally pursuing their own agendas. Many benefits would accrue from this and they include:

- A greatly simplified management and assessment process.
- Studies with greater confidence in the results through eliminating the uncertain extraneous effects of control schools based on smaller samples.

3.3 Ethics and Data Protection

In all research studies in education ethical issues are important to ensure that the pupils and teachers involved are not harmed or disadvantaged in any way (Cohen and Morrison 2013, p.57). An incidental advantage of the changed methodology was that all of the schools and teachers chose to be part of the Study and willingly provided information and access. Missing out on the intervention did not therefore disadvantage the pupils of the control classes.

The School of Education Ethics Committee granted Ethical permission for the research while the Head Teachers in all of the schools agreed to be part of the project. Although the *Sumdials*’ approach is an intervention, it is designed to enable pupils to meet the arithmetical objectives of the mathematical curriculum in England and Scotland.

Data collected on the pupils was stored anonymously and handled in accordance with the Data Protection regulations outlined by Durham University.

3.4 Conclusion

It can reasonably be concluded that adapting the methodology to use ages-at-test as the control would facilitate the completion of the Empirical Study. A rigid adherence

to the planned methodology would probably have required binding contracts to keep the control schools on board; that would have been a very undesirable development – and probably counter-productive.

3.5 Chapter Conclusion

It will now be apparent that the use of the word “messy” to describe primary schools is apt, as this has become their inherent characteristic. It is largely the result of unrealistic *external* expectations. Thus, the successful completions of the Pilot and Empirical Studies were rewards in themselves. They were also a tribute to the school staffs that made their mainly constructive supporting contributions.

The next Chapter provides the Results of the Pilot and Empirical Studies together with other unplanned Results.

4. The Results

4. Introduction

The original plan was to carry out two randomised treatment/control studies, consisting of:

- A Pilot Study involving primary schools in Edinburgh only (for logistical reasons).
- The Empirical Study with schools from both Co. Durham and Edinburgh.

The aim of the Pilot Study was to verify that the InCAS computer adaptive system was:

- Appropriate for collecting quantitative attainment data from pupils in their early primary school years,
- Acceptable to the staff and pupils of the participating schools, and
- Easy to administer.

It scored well on all three counts for the first two assessments of the Pilot Study and this confirmed that the InCAS system would be appropriate for the Empirical Study.

Again, the main question is:

- Does the *Sumdials*' approach to learning number, based on the use of dedicated manipulatives (dials), produce statistically significant improvements in arithmetical automaticity?
- As secondary questions, are there statistically significant differences in the number attainments:

(iii) By gender (boys and girls)?

(iv) By location (between Co. Durham and Edinburgh pupils)?

However, the usefulness of the control schools came into question during the Pilot Study, and as a consequence, they were not used in the Empirical Study (p.93). Both studies should therefore be classified as within-subject longitudinal studies in which the pupils' progress would be measured against their Ages-at-Test (A/T). In the event, the implication of a comparative quantitative analysis was that there could be more confidence in the overall validity of the results from within-subjects studies (Kroesbergen and Van Luit 2003) such as this one had become without the control schools.

The change did not affect treatment schools since the participating teachers:

- Only wanted to know how *their* pupils were responding to the interventions.
- Were always aware of the relative ages of their own pupils, especially in the earlier years’, and could have argued that their inferior performances were due to their being in more disadvantaged catchment area.

InCAS provides both subject scores *and* ages-at-test (A/T) simultaneously. Thus, the subject and A/T **gains** are easier to compare than differences between the treatment and control schools would have been. A/Ts are derived from dates-of-birth (DOBs) that are *always* up-dated automatically at each assessment; a pupil can only be assessed if her DOB is already in the system. This makes the data quality more robust, particularly when the pupils are tracked over time.

Originally, it was planned that the Empirical Study would be a longitudinal study over two years, but data from three academic years would become available by assessing the original Pilot Study cohort in parallel; this was carried out to increase confidence in the overall findings (p.76).

Lastly, the opportunity to assess two small rural schools in Switzerland was taken and their results are provided for comparison only (p.151) and are not included in the Study results.

4.1 Results’ Abbreviations and Formats

The results of the data analysis are now summarised using both tabular and chart representations and, again, using the following abbreviations and criteria:

- The two assessment sessions (subjects) in both Studies were:
 - (i) Mental Arithmetic (M/A),
 - (ii) General Maths (G/M), with
 - (iii) Developed Ability (D/A) – Empirical Study only.
- All assessments are expressed as Age Equivalent Scores (AESs), but ‘scores’ is also used in the texts.

- *Mean* Mental Arithmetic and General Maths AESs, A/Ts, and Standard Deviations (S/Ds) are shown together with Effect Sizes and T Test values (**p**).
- Cohen's suggested categorisation of Effect Sizes are
 - (i) A value of less than 0.2 is *trivial*.
 - (ii) A value between 0.2 and 0.5 is *small*.
 - (iii) A value between 0.5 and 0.8 is *medium*.
 - (iv) A value of more than 0.8 is *large*, (Kinnear & Gray, 2011 p.183).
- S/E: Standard Error Differences (Equal variances not assumed).
- **p**: Significance (2-tailed) where values < 0.05 are significant at the 95% level while < 0.01 are highly significant being at the 99% level.
- The same tabular formats are used throughout and the results will be presented in the same sequence within studies of:
 - (i) Pilot Study:
 - Mental arithmetic (M/A).
 - General Maths (G/M).
 - (ii) Empirical Study:
 - Mental arithmetic (M/A).
 - General Maths (G/M).
 - Developed Ability (D/A).
- 1st September (the start of school year) was the equivalent to zero for both Studies and thereafter the months were counted cumulatively to arrive at the actual assessment months. The values that were used in the Studies were: 3, 6, 21 and 33. They should be seen as indicative in that the actual months could be for:
 - (i) 3 - late October in Scotland or November in England (due to different starting times).
 - (ii) 6 - late January/February (for Pilot Study only).
 - (iii) 9, 21 and 33 – late May/early June (Scotland) or late June/early July (England) being the times when formal classroom

activities were being replaced by traditional end of session ones such as sports days.

4.2 Pilot Study Results

The Pilot Study results are now presented.

Pilot Study Mental Arithmetic Results

Month	Pilot Study M/A			Pilot Study A/T			Effect Size	T Test	
	Mean	N	S/D	Mean	P	S/D		S/E	p
6	3.342	109	2.484	5.639	111	0.315	-1.30	0.237	0.000
9	3.861	87	2.399	5.976	90	0.312	-1.24	0.252	0.000
GAINS	0.519			0.337			0.06		

Table 4.13: Mental Arithmetic (M/A) mean AESs compared with mean Ages-at-Test.

These results provided the first evidence of the disparity between the mean Mental Arithmetic and Ages-at-Test with *large* negative effect sizes and **p** values that are highly significant at the 99% level. This was to be a recurring feature of the Results, as is noted now and considered later (p.101).

Pilot Study General Maths Results

Month	Pilot Study G/M			Pilot Study A/T			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	P
6	5.467	110	1.262	5.639	111	0.315	-0.19	0.261	0.154
9	6.129	88	1.157	5.976	90	0.312	0.18	0.129	0.221
GAINS	0.662			0.337			0.37		

Table 4.14: General Maths (G/M) mean AESs compared with mean Ages-at-Test (A/T).

This table confirmed the previous manual calculations that the General Maths gain was, indeed, approaching one month for each elapsed month of the Study. This supported the decision to proceed with the Empirical Study even though the gain in effect size was small and the **p** value was non-significant.

One explanation is the mean General Maths score was two months less than the Age-at-test at the beginning while it was nearly two months more at the end. Since the elapsed time was only nominally three months (it was actually five months) the likelihood of statistically significant changes would be small. However, the participating head teachers were very encouraged by the General Maths gains and would not have readily accepted the effect size argument as a reason for not proceeding with the Empirical Study.

The Mental Arithmetic, General Maths and Ages-at-Test scores are now presented in a composite chart:

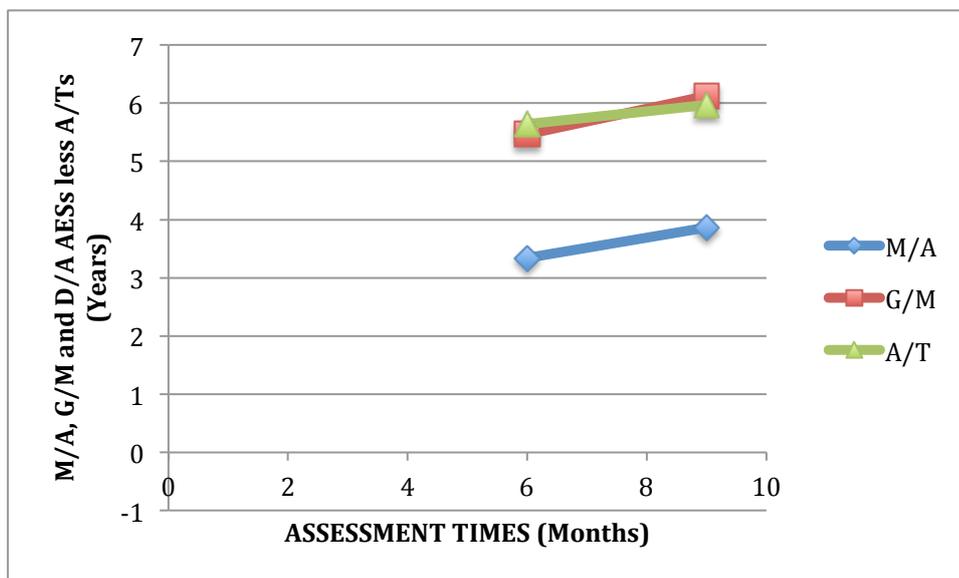


Figure 4.5. Pilot Study mean M/A, G/M and A/T Scores at start and finish of the Pilot Study.

These tables, with the accompanying chart, show that:

- Both Mental Arithmetic and General Maths scores made greater gains than the increases in the pupils' Ages-at-Test during the Study.
- The Mental Arithmetic scores remained over two years behind the mean Ages-at-Test and the mean General Maths scores.
- The mean General Maths score gained 6.2 months compared with 4.0 months for Mental Arithmetic and that means the gap between them had increased.
- These were the first indications that Mental Arithmetic had become the neglected number subject and the implications of this are discussed later, (p. 126).

Generally, this Study achieved high acceptability for everyone and, very importantly, the pupils enjoyed the “quizzes” - to the extent some even tried to have a second attempt. However, that would not have been possible because their personal passwords can only be used once during each subject assessment.

4.3 Empirical Study Results

4.3 Introduction

The Empirical Study was conducted using the same Pilot Study methodology with seven schools participating, three being from Edinburgh and four from Co. Durham. In addition, the assessment of Developed Ability (D/A) was included and is a type of IQ test that assesses acquired (not taught) cognitive and social skills.

4.3.1 Main Question

The results for the main question are now summarised using the same tabular representations for each subject and a composite chart representation for Mental Arithmetic and General Maths subjects. The first table shows the mean **Gains** in Mental Arithmetic and Ages-at-Test (in years); again, all tables include the effect size and T Test **p** values.

Empirical Study Mental Arithmetic Results

	Empirical Study M/A	Empirical Study A/T	Effect	T Test
--	---------------------	---------------------	--------	--------

Month	Mean	N	S/D	Mean	N	S/D	Size	S/E	p
3	3.884	190	1.941	5.542	190	0.388	-1.21	0.135	0.000
9	5.104	165	1.674	6.153	166	0.373	-0.86	0.127	0.000
21	6.964	145	1.480	7.137	147	0.375	-0.16	0.122	0.159
GAINS	3.120			1.601			1.05		

Table 4.15: Mental Arithmetic (M/A) mean AESs compared with mean Ages-at-Test.

An initial observation is the pupils improved their mean Mental Arithmetic scores at a rate of almost twice (95%) that the mean increases in their mean Ages-at-Test, as occurred in the Pilot Study. The effect size is large and the **p** value is non-significant having almost closed the highly significant gap (at the 99% level) at the start of the Study. Put simply, the pupils had gone from being 21 months behind their mean Ages-at-test at the beginning of the Study to only two months behind by the end (18 months later) or had *gained* almost one month for each elapsed month. These are considered with the General Maths gains.

The following table summarises the **Gains** in mean General Maths and Ages-at-Test (in years).

Empirical Study General Maths Results

Month	Empirical Study G/M			Empirical Study A/T			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
3	5.257	190	1.085	5.536	192	0.388	-0.34	0.077	0.000
9	5.876	165	0.860	6.153	166	0.373	-0.42	0.068	0.000
21	7.573	147	0.882	7.137	147	0.375	0.60	0.073	0.000
GAINS	2.307			1.601			0.94		

Table 4.16: General Maths (G/M) mean AESs compared with mean Ages-at-Test (A/Ts).

The General Maths gain in **p** value is highly significant at the 99% level while the effect size is large.

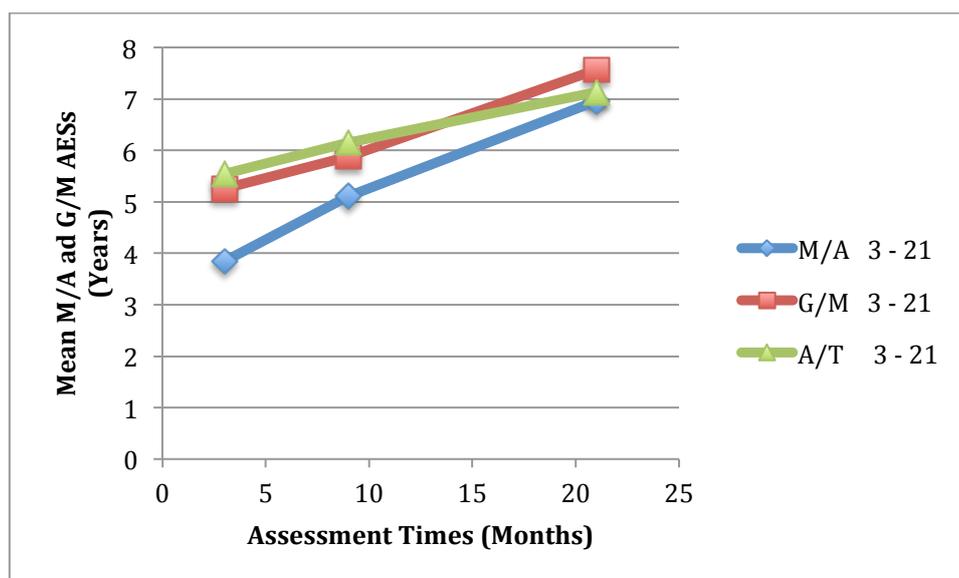


Figure 4.6. Empirical Study Mean M/A, G/M and A/T AESs over Study.

This chart shows that after the initial “catch-up” in mean Mental Arithmetic, its subsequent gains were slightly greater than the General Maths gains, while both were greater than the increases in mean Ages-at-Test.

And, lastly, the following table summarises the mean **Gains** in Developed Ability and Ages-at-Test (in years).

Empirical Study Developed Ability Results

Month	Empirical Study D/A			Empirical Study A/T			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
3	3.979	184	1.958	5.536	192	0.388	-1.11	0.142	0.000
9	5.292	166	1.938	6.153	166	0.373	-0.62	0.147	0.000
21	7.086	146	1.674	7.137	147	0.375	-0.04	0.136	0.717
GAINS	3.107			1.601			1.07		

Table 4.17: Developed Ability Results mean AESs compared with mean Ages-at-Test.

Developed Ability has no direct bearing on the main question. It would normally be expected to increase in line with pupils’ chronological ages and certainly not at nearly twice their rate, as happened. It will be seen in the following chart that Developed Ability and Mental Arithmetic have unexpectedly increased in step:

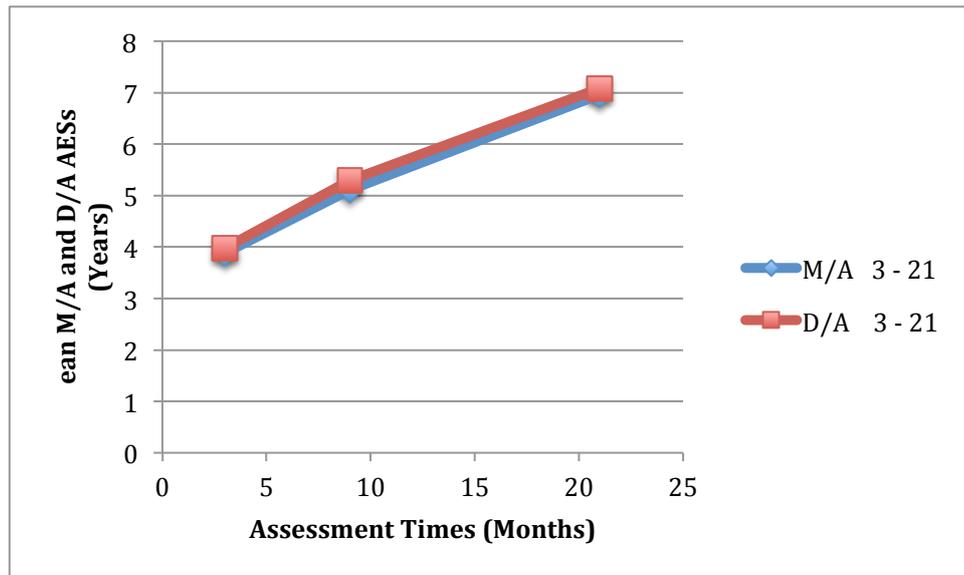


Figure 4.7: Empirical Study comparing Mean M/A and D/A AESs during the Study.

Comment is now limited to suggesting that Mental Arithmetic was the “driver” and, in response, the Pearson’s correlations between the three sessions of Mental Arithmetic, General Maths, Developed Ability and General Maths (I &2) modules 1 and 2 only (its reason being explained later, p.128) were determined. Their values are presented in descending order in the table below:

Empirical Study Pearson’s Correlations

Pairings	N	Correlations
G/M v G/M (1&2)	148	0.953**
G/M v D/A	148	0.671**
G/M (1&2) v M/A	148	0.627**
G/M (1&2) v D/A	146	0.627**
G/M v M/A	146	0.610**
M/A v D/A	144	0.429**

** Correlation is significant at the 0.01 level (2-tailed).

Table 4.18: Correlations between M/A, G/M, G/M sub-group and D/A AESs.

These are very strong correlations and indicate that sessions (subjects) are inter-related. The implications of this table will be discussed in the wider context of the related studies than have taken place so far (p.127).

Also, the overall gains that were achieved during the Empirical Study are summarised in the following table and composite chart:

Empirical Study Summary of Gains by Session

Session (Subject)	Gain in AESs	Age Increase	Gain over Age (Years)	Rate of Gain (Months/year)
Mental Arithmetic	3.10	1.60	1.52	11.4
General Maths	2.31	1.60	0.71	5.3
Developed Ability	3.11	1.60	1.51	11.3

Table 4.19: Summary of Gains prepared for Discussion of the Empirical Study Results.

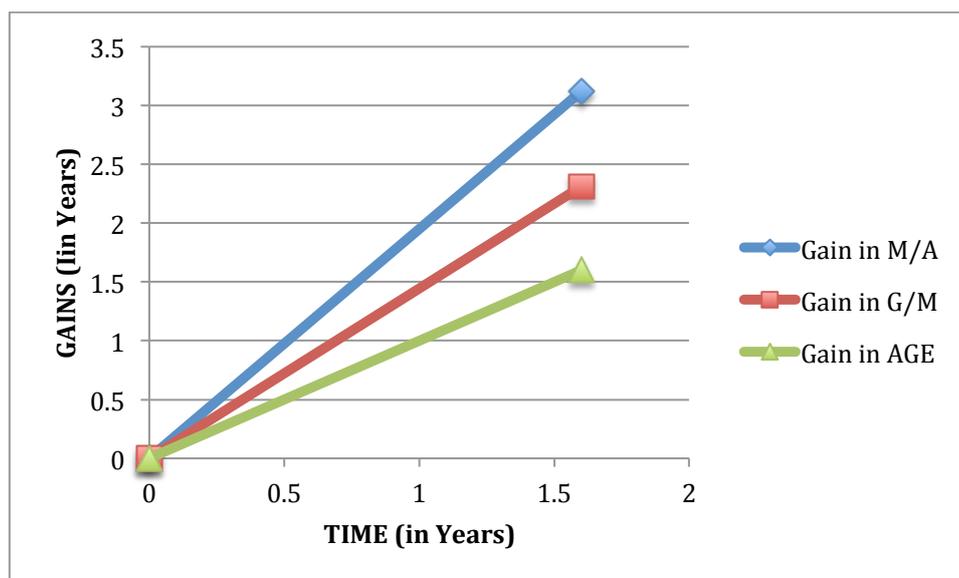


Figure 4.8: The Study gains in chart format (the D/A gain is not shown separately since it is virtually the same as the M/A gain).

Suffice it to say now that all the subject gains are very remarkable and especially mental arithmetic within the context of this Study.

4.4.2 Pilot Study Cohort

The assessments of the Pilot Study cohort continued in parallel with the Empirical Study cohort and that meant it ran from months 6 to 33 or for an additional year. The

results are now represented in both tabular and chart formats for M/A and G/M using the same procedures.

Pilot Study Cohort Mental Arithmetic Results

Month	P/S Group			A/T			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		E/S	p
6	3.342	109	2.48	5.643	109	0.315	-1.30	0.237	0.000
9	3.861	87	2.40	5.975	87	0.313	-1.23	0.261	0.000
21	6.114	85	1.95	6.898	85	0.423	-0.55	0.220	0.001
33	6.924	80	2.22	7.953	80	0.325	-0.65	0.260	0.000
GAINS	3.582			2.311			0.65		

Table 4.20: Summary of the M/A gains made by the Pilot Study Cohort.

These results are similar to those of the Empirical Study, while the rate of gain slowed in the third year after the catching-up that could have taken place during the second year, having been nearly 28 months behind at the start. Overall, mental arithmetic seems to be something of a neglected subject.

Pilot Study Cohort General Maths Results

Month	P/S Group			A/T			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
6	5.467	110	1.26	5.643	109	0.315	-0.19	0.122	0.154
9	6.129	88	1.16	5.975	87	0.313	-0.18	0.261	0.000
21	6.504	86	1.11	6.893	85	0.423	-0.46	0.134	0.006
33	8.143	79	1.15	7.953	80	0.325	0.22	0.136	0.170
GAINS	2.676			2.311			0.41		

Table 4.21: Summary of the G/M gains made by the Pilot Study Cohort.

The feature of particular interest is that there is evidence of loss of momentum in the second year (perhaps because of because of concentration on Mental Arithmetic) while the pick-up again could be attributed to a focus on memory work. However, the

overall gain was still in the right direction and worthwhile even if the **p** value is non-significant. The results in these tables are now re-presented in the following chart:

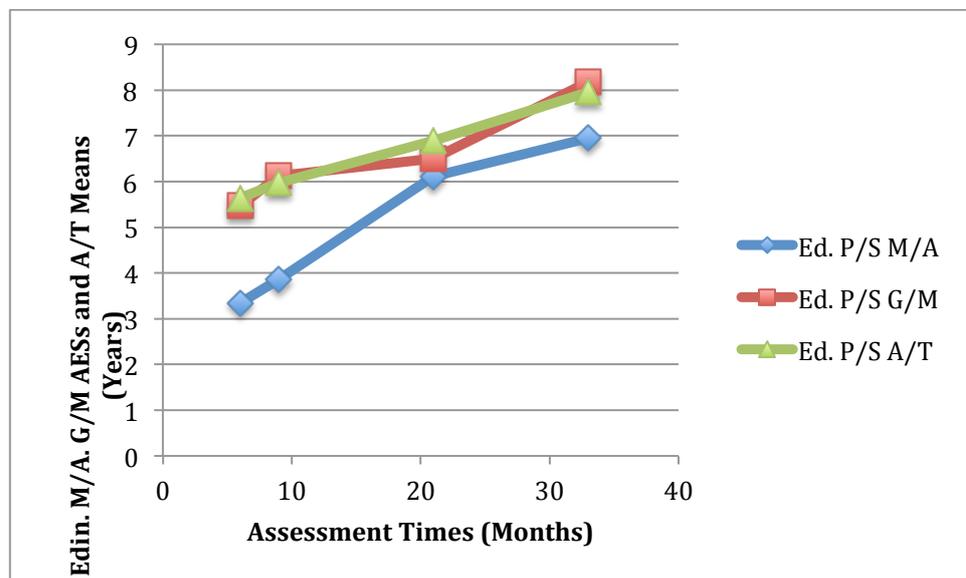


Figure 4.9: Pilot Study Cohort Mean M/A, G/M AESs and A/Ts at each assessment.

The gains shown in the above tables and chart are now re-presented in the following summary table:

Summary of the Pilot Study Cohort Gains

Subject	Gain	Age Increase	Gain over Age
M/A	3.582	2.311	1.271
G/M	2.676	2.311	0.365

Table 4.22: Summary of the Pilot Study Cohort gains in M/A and G/M scores for consideration in the Discussion Chapter.

Clearly, these gains are not as great as those of the Empirical Study. It will be recalled that the Edinburgh cohort has been analysed alone because of the concerns of the experienced teacher that her pupils were starting their formal number learning before they were ready for it. These results are discussed later, (p.148).

4.5 Secondary Questions

The results are now given for the two secondary questions of:

- Are there statistically significant differences in the number attainments:

- (v) By gender (boys and girls)?
- (vi) By location (between Co. Durham and Edinburgh pupils)?

4.5.1 By Gender

This question was included in response to the general perception that boys are better than girls with number. However, it was not known at the time that this was widely studied territory (e.g. Else-Quest, Hyde, & Linn, 2010), but confirmation of the null hypothesis would also indicate that the dataset itself is a typical sample and, in turn, there can be confidence in any conclusions drawn from it.

The results are summarised using the same tabular and chart representations, as previously, but with only the starting and finishing data being shown since the aim was now to confirm that the dataset was representative.

Comparison by Gender of Mental Arithmetic Results

Month	Empirical Study Boys			Empirical Study Girls			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
3	3.840	104	2.060	3.849	86	1.799	0.00	0.284	0.976
21	7.012	78	1.648	6.910	68	1.257	0.07	0.241	0.675
GAINS	3.172			3.061			0.07		

Table 4.23: The comparisons of mean M/A scores by gender revealing only slight differences.

While the increase in the boys' mean Mental Arithmetic score was greater than that of the girls, the effect size is trivial and the **p** values are non-significant. This is apparent when the scores are presented in chart format:

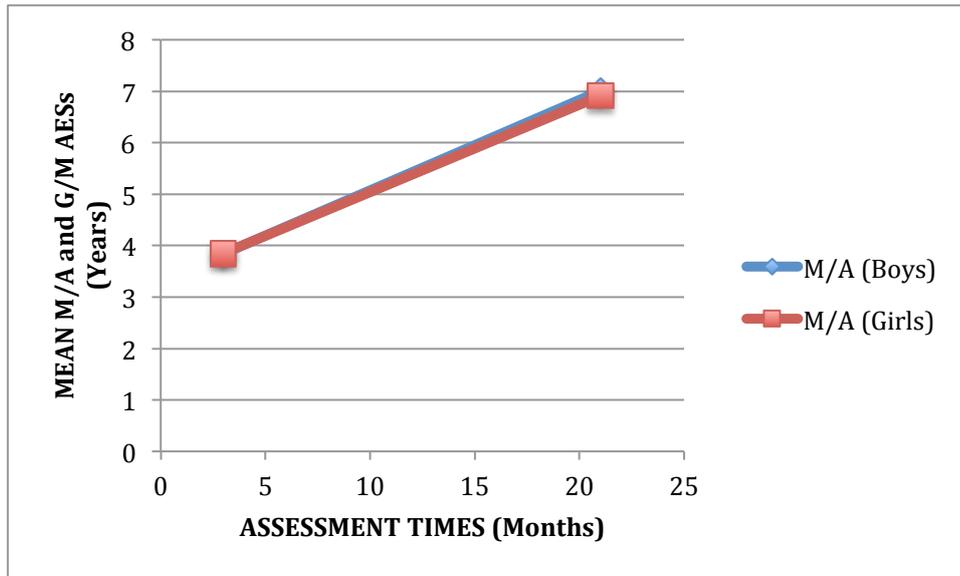


Figure 4.10. Empirical Study Mean Gender M/A Scores at start and finish.

The following table summarises the General Maths results

Comparison by Gender of General Maths Results

Month	Empirical Study Boys			Empirical Study Girls			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
3	5.244	104	1.114	5.273	86	1.054	0.04	0.158	0.583
21	7.583	80	0.976	7.541	68	0.860	0.05	0.151	0.785
GAINS	2.339			2.268			0.01		

Table 4.24: The comparisons of mean G/M scores by gender reveals only slight differences.

Again, the Boys' mean gain was higher than the Girl's, but the differences on both the effect size – trivial – and the **p** values were non-significant. Again, this is confirmed when they are represented in chart format:

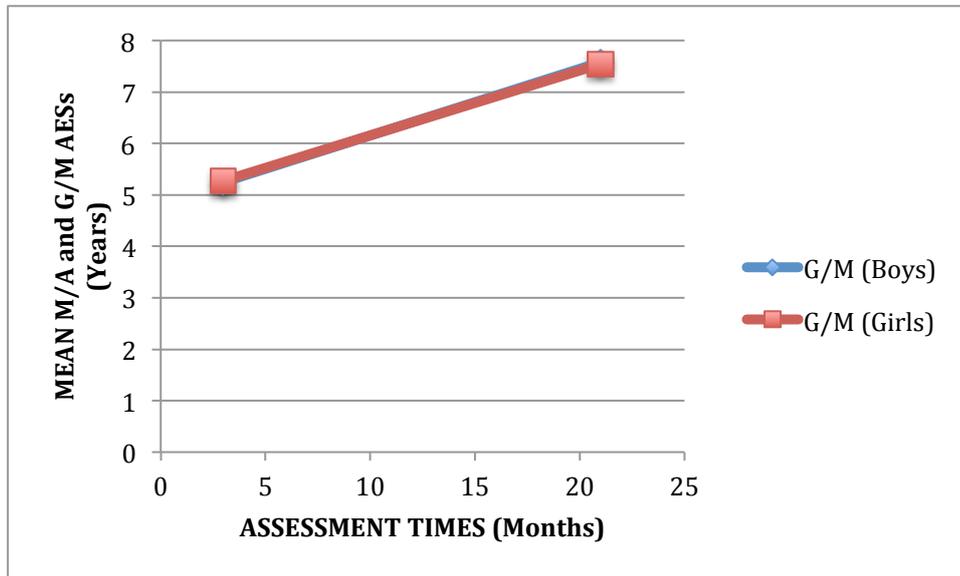


Figure 4.11 Empirical Study Mean G/M Gender Scores at start and finish.

Thus the null hypothesis is confirmed.

Comparison by Gender of Developed Ability Results

Month	Empirical Study Boys			Empirical Study Girls			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
3	4.027	102	2.002	3.919	82	1.911	0.05	0.291	0.712
21	7.199	79	1.757	6.954	67	1.574	0.15	0.278	0.380
GAINS	3.172			3.035			0.10		

Table 4.25: The comparisons of mean D/A scores by gender reveal only slight differences.

Again, the Boys' mean scores were higher than the Girl's, but the differences on either the effect size – trivial – and the p values were non-significant. For the sake of completeness, these are also now represented in chart format.

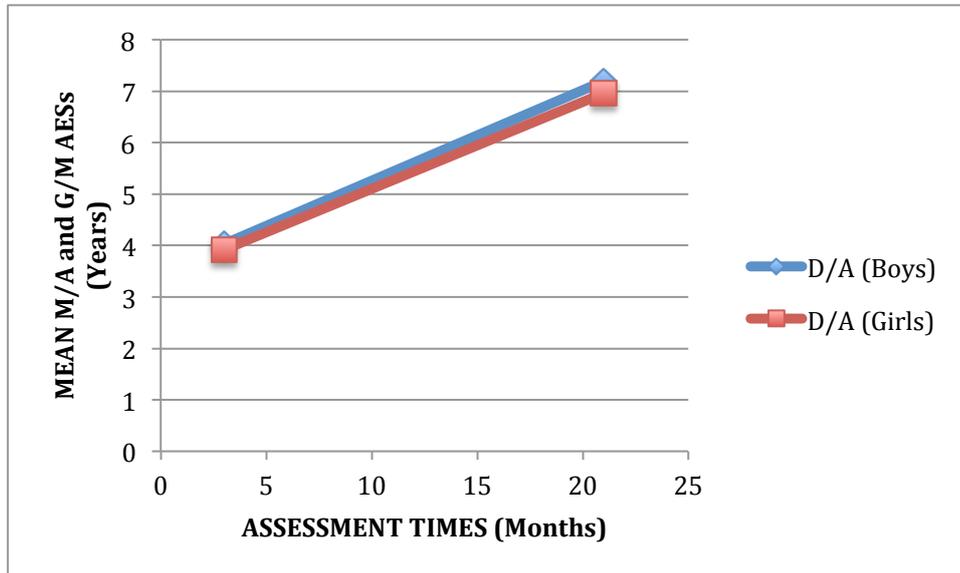


Figure 4.12. Empirical Study Mean Gender D/A Scores at start and finish.

A hint for the persisting perception of boys being better than girls may be given by all the boys' mean scores of this Study being higher than the girls' were - even though they were not statistically significant. This may represent the effects of cultural expectations (Gunderson, Ramirez, Levine and Beilock 2012). Having said this, the two extreme scores were achieved by boys and this is consistent with the observation 'Males are more likely to be *extremely* good at mathematics' and the general tendency for more boys than girls to participate in higher mathematics, made by Dowker, (2005, p.7). Perhaps the positive performances of the exceptional boys are noticed, while being balanced by slightly more of them at the bottom of the distribution.

4.5.2 By Location

The Co. Durham and Edinburgh schools have different starting dates, so their mean Ages-at-Test were extracted and are presented in the following tables:

Co. Durham and Edinburgh mean Ages-at-Test

Month	Empirical Study Co. D.			Empirical Study Edin.			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
3	5.745	111	0.303	5.250	81	0.303	1.64	0.443	0.000
9	6.322	97	0.312	5.915	69	0.312	1.31	0.500	0.000
21	7.273	85	0.345	6.952	62	0.335	0.92	0.057	0.000
GAINS	1.528			1.702			-0.72		

Table 4.26: The effects of the different starting ages on the mean A/Ts of Co. Durham Edinburgh schools

The effect sizes each year are large throughout and the **p** values are highly significant throughout and these are considered shortly. However, a closer analysis shows that the mean differences in the relative Ages-at-Test were changing over the course of the Study, as the following table reveals.

Comparison of Mean Ages by Location

Month	Mean Ages		Differences	
	Co. Durham	Edinburgh	Years	Months (Approx.)
3	5.745	5.250	0.495	6
9	6.322	5.915	0.407	5
21	7.273	6.952	0.321	4

Table 4.27: The reducing trend of differences in mean A/Ts during the Study.

These changes can perhaps be attributed to:

- Timetabling issues in that the assessments were made when it was practically convenient to the participating schools (within reason) rather than being driven by strict Study timings.
- A greater number of older Co. Durham pupils and younger Edinburgh pupils could have left during the Study.

The possible influences of these age factors are considered later, (p.149). In the meantime, the actual results are now given using the same tabular formats as previously, starting with the M/A results.

Comparisons by Location of Mean Mental Arithmetic Results

Month	Empirical Study Co. D.			Empirical Study Edin.			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
3	4.564	111	1.907	2.834	79	1.494	0.99	0.247	0.000
9	5.533	97	1.468	4.492	68	1.767	0.56	0.261	0.000
21	7.314	84	1.537	6.490	62	1.248	0.58	0.231	0.000
GAINS	2.750			3.656			-0.41		

Table 4.28: The comparisons between the Co. Durham and Edinburgh mean M/A AESs.

Both the Co. Durham and the Edinburgh pupils achieved remarkable gains during the elapsed time (approximately 18 months) with the Edinburgh pupils gaining one month for each elapsed month of the Study. This could be attributed to a “catch-up” effect from a very low starting point. This might therefore be further evidence of the lack of emphasis on mental arithmetic in the Edinburgh schools.

The effect size was large at the outset and had been reduced to medium by the end of the Study, while the *p* values remained highly significant at the 99% level.

The results are now displayed in chart representation:

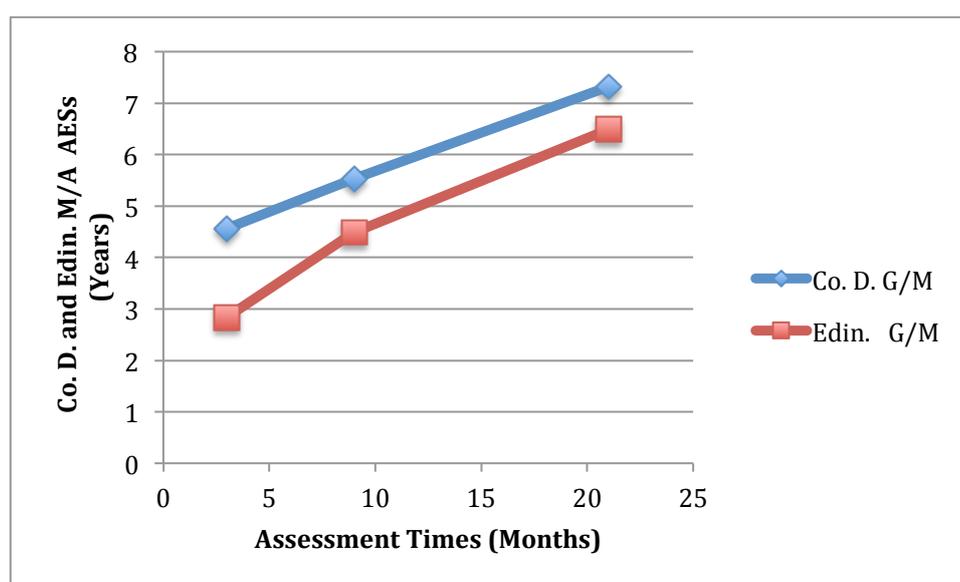


Figure 4.13: Empirical Study comparing the Co. Durham and Edinburgh M/A

AESs.

This chart shows that after the initial catch-up gain made by the Edinburgh pupils' mental arithmetic the subsequent gain on the Co. Durham pupils became much less.

Comparison by Location of Mean General Maths Results

Month	Empirical Study Co. D.			Empirical Study Edin.			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
3	5.521	111	1.156	4.885	79	0.852	0.61	0.334	0.000
9	6.082	97	0.794	5.581	68	0.872	0.60	0.333	0.000
21	7.862	85	0.939	7.160	63	0.730	0.81	0.137	0.039
GAINS	2.341			2.275			0.20		

Table 4.29: The comparisons between the Co. Durham and Edinburgh mean G/M AESs.

Again, the effect size became large while the p values remained significant at the 95% level by the end of the Study.

The results are now displayed using a chart representation:

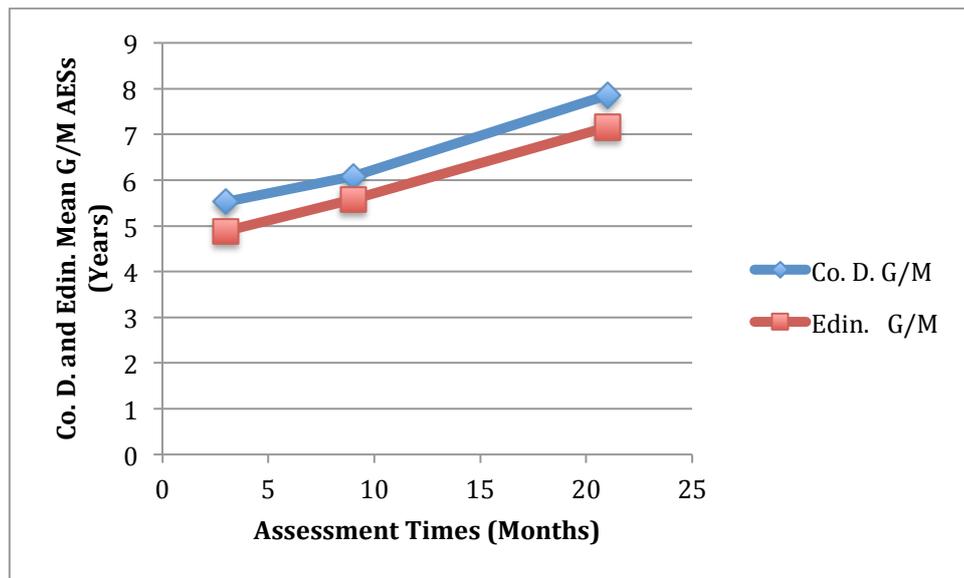


Figure 4.14: Empirical Study comparing the Co. Durham and Edinburgh G/M AESs.

The trajectories confirm that the gains of Edinburgh Schools were slightly greater than those of the Co. Durham Schools.

Overall, in light of both the Mental Arithmetic and General Maths scores the null hypothesis can be rejected because the Co. Durham Schools' effect size was medium/large and their **p** value was significant at the 95% level compared with the Edinburgh Schools. This was unexpected and is discussed later, (p.142).

Comparison by Location of Mean Developed Ability Results

Month	Empirical Study Co. D.			Empirical Study Edin.			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
3	4.259	111	1.736	3.553	73	2.199	0.37	0.306	0.022
9	5.793	97	1.685	4.588	69	2.061	0.65	0.301	0.000
21	7.490	84	1.575	6.539	62	1.661	0.59	0.272	0.001
GAINS	3.231			2.986			0.22		

Table 4.30: The comparisons between the Co. Durham and Edinburgh mean D/A AESs.

The results are now displayed in chart representation:

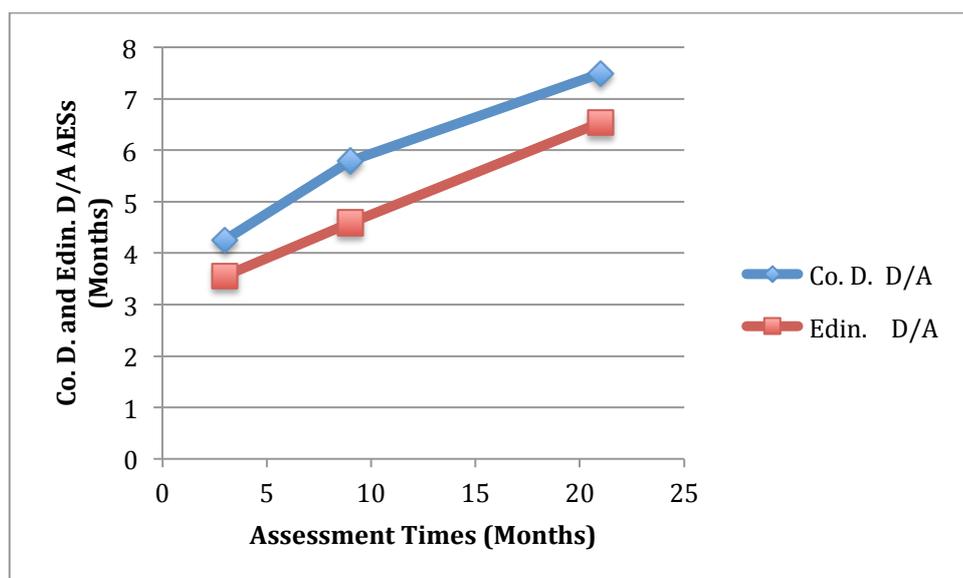


Figure 4.15: Empirical Study showing the D/A AESs of Co. Durham and Edinburgh.

It is noted that the Co. Durham gains are slightly greater than the Edinburgh ones. Such overall gains would not normally be expected, as has already been mentioned, with a medium effect size and with a highly significant **p** value at the 99% level by the end of the Study. This is discussed because, again, Developed Ability usually develops with age rather than through teaching and the shortages of time make it unlikely that teachers would be encouraging better general knowledge to improve these scores.

All these findings are discussed more fully (p.133).

4.6. School Starting Ages

Since starting ages for formal subject learning had emerged as something of a current issue in Edinburgh, as will be discussed (p. 144), the mean scores in the Pilot and Empirical Study cohorts are now compared, although it was not part of the original plan. Only Months 9 and 21 are analysed because the Pilot cohort was initially assessed at its Month 6 while the Empirical cohort was initially assessed at its Month 3 (one year later); no sound basis could be found that would accommodate these different assessment times.

The following table summarises the Mental Arithmetic results together with a comparative chart.

Comparison of Empirical and Pilot Studies Mental Arithmetic Results

Month	Empirical Study M/A			Pilot Study M/A			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
9	4.538	69	1.796	4.544	86	1.796	0.00	0.690	0.993
21	6.490	62	1.248	6.080	85	1.948	0.24	0.264	0.122
GAINS	1.952			1.536			0.24		

Table 4.31: The comparative gains of the Empirical and Pilot Studies M/A AESs.

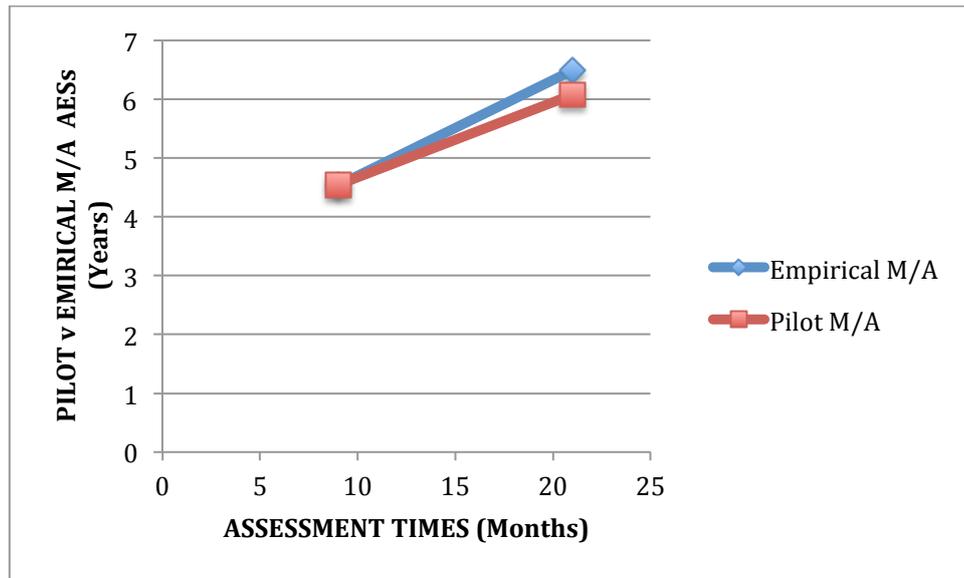


Figure 4.16: Edinburgh Pilot v Empirical Studies comparisons in M/A AESs.

The month 9 mean ages were almost identical, but the Empirical cohort subsequently made better progress. However, the effect size was trivial and the p value was non-significant.

The comparisons for General Maths age equivalent scores are now represented in both tabular and chart formats.

Comparison of Empirical and Pilot Studies General Maths Results

Month	Empirical Study G/M			Pilot Study G/M			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
9	5.599	69	0.878	6.130	88	1.150	-0.10	0.162	0.001
21	7.160	63	0.730	6.538	85	1.084	0.65	0.149	0.000
GAINS	1.561			0.408			0.75		

Table 4.32: The comparative gains of the Empirical and Pilot Studies G/M AESs.

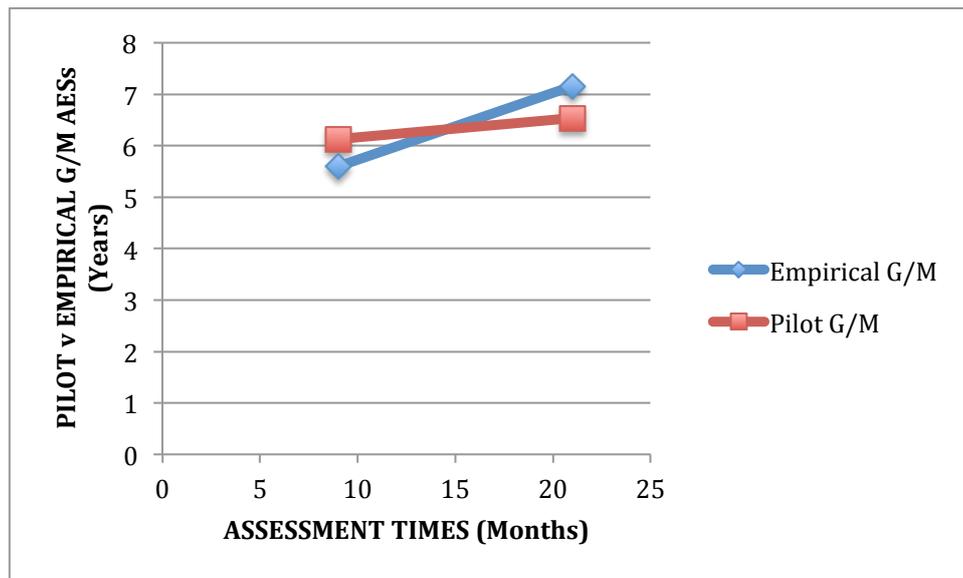


Figure 4.17: Comparisons of Edinburgh Pilot v Empirical Studies G/M AESs.

Interestingly, the Pilot Study mean AESs was significantly in front of the Empirical scores, but by month 21 the Empirical scores had gained highly significantly (at the 99% level) while the effect size was medium.

The next table shows the mean ages of the two cohorts.

Comparison of Empirical and Pilot Studies Mean Ages-at-Test

Month	Empirical Study A/T			Pilot Study A/T			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
9	5.922	70	0.323	5.969	90	0.324	-0.14	0.052	0.367
21	6.952	62	0.335	6.917	86	0.360	0.10	0.058	0.554
GAINS	1.030			0.948			0.24		

Table 4.33: Mean A/T Comparisons of the Edinburgh Cohorts.

The mean ages of the two cohorts were very similar, as would be expected, and this is confirmed by trivial effect sizes and insignificant **p** values.

These results are discussed more fully, (p.147).

4.7 Ability Categories

It was decided, in a spirit of curiosity and enquiry, to make a descriptive analysis to find if there might be supporting statistical evidence for the three anecdotal categories of ability generally found in schools, being those who:

- Will do well, almost regardless of how they are taught 10 - 15%
- Could and would do better, when taught appropriately 70 – 80%
- Will always struggle, regardless of how they are taught 10 – 15%

To explain the word anecdotal, the categories above are generally assumed in schools even though there may be no measurements to support these sizes. The simple selection rule was: all valid data for any pupil who had taken part in the initial assessments of either Study would be used. This gave a *pre-treatment* sample of 545 pupils that included the scores of the pupils from the original Edinburgh control schools to produce the following histograms and stem-and-leaf plots.

These histograms and stem-and-leaf plots are considered, together with all the other results, in the next Chapter.

4.8 Conclusion

The results that have now been presented are for a wider range of topics than had originally been planned. The main reason for this is actually being in schools to carry out the InCAS assessments and then to present the results, created communication influences – many of them non-verbal – that impinged on the original plans. As will become apparent, a characteristic of this Study has been to respond to such influences, as will be discussed, (p.178).

5. The Discussion

5. Introduction

The point has now been reached when the results of the Empirical Study, in particular, need to be considered and discussed before drawing some conclusions. The value of the Pilot Study lay in confirming that the InCAS computer adaptive assessment program would be appropriate for assessing the progress of the participating pupils and in evaluating the practicalities of working with schools. In all, the experiences from 12 primary schools, five being in Co. Durham and seven in Edinburgh made useful contributions, as did especially the three in Co. Durham and four in Edinburgh that took part in the Empirical Study itself.

This Chapter is in two parts:

- Discussion of the Empirical Study Results.
- Some general observations.

These reflect the main finding of the Empirical Study that:

- Mental arithmetic is not emphasised in the current curriculum.
- When given support, pupils can make significant age-related progress in this area.

Again, the main question is:

- Does the *Sumdials*' approach to learning number, based on the use of dedicated manipulatives (dials), produce statistically significant improvements in arithmetical automaticity?
- As secondary questions, are there statistically significant differences in the number attainments:
 - (vii) By gender (boys and girls)?
 - (viii) By location (between Co. Durham and Edinburgh pupils)?

The suitability of the main question for these findings is now clarified. It was assumed (by me) when it was set that the correct use of the *Sumdials*' manipulatives - the dials - would lead to arithmetical automaticity, but this actually never was the case, (p.207). To explain, repetition is used to memorise new facts (and procedures) and it was believed that the repetitions associated with using the dials to do sums would help in hardwiring new number facts. It is now known that this is not the case because such memorising must be done verbally and not visually/kinaesthetically. In the event, this became of little consequence because the lack of emphasis on mental arithmetic in the current taught curriculum resulted in virtually no observable automaticity. Inevitably, this meant that the main question could not be answered and it might then be assumed that the Study's results had no value. Such a conclusion would have missed the potentially very important discovery about the contribution of the approach apparently made to the age-related progress in arithmetic during the Study.

However, a case supporting our arguments for the importance of arithmetical automaticity was discovered after completion of the Empirical Study and will also be considered (p.94).

Part 5.2 Discussions of Empirical Study Results

5.2 Introduction

The key data have been consolidated from the Results tables to produce the following table and chart:

Summary of Empirical Study Gains by Session

Session (Subject)	Gain in AESs	Age Increase	Gain over Age (Years)	Rate of Gain (Months/year)
Mental Arithmetic	3.10	1.60	1.52	11.4
General Maths	2.31	1.60	0.71	5.3
Developed Ability	3.11	1.60	1.51	11.3

Table 5.34 Summary of Gains compared with increases in mean A/TS

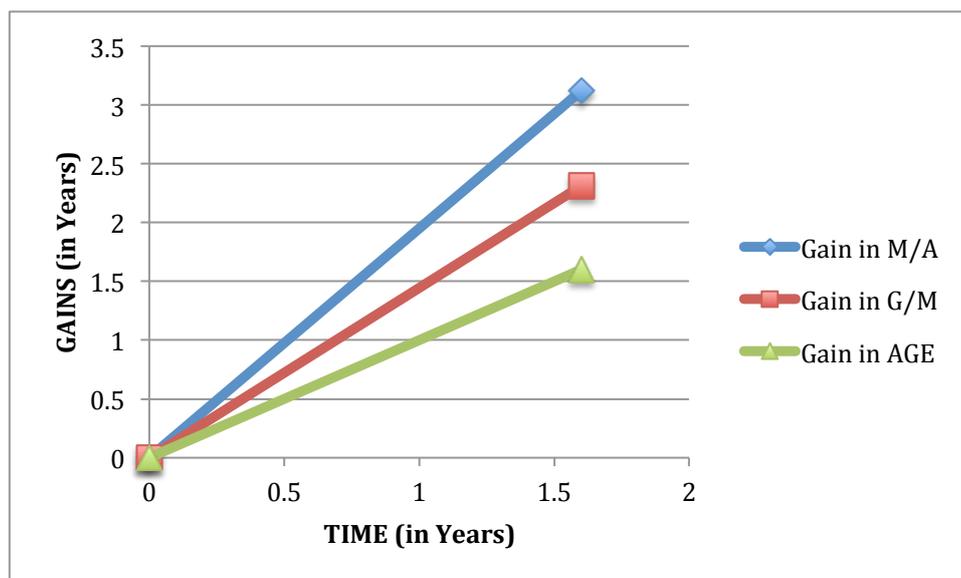


Figure 5.22: The Study gains in chart format (the D/A gain is not shown separately since it is virtually the same as the M/A gain).

These rates of gain are truly remarkable and this is particularly the case for Mental Arithmetic where the gain was the equivalent to almost one month for each month of the Study, while in the case of General Maths it was nearly half a month. The most likely explanation for the greater Mental Arithmetic rate of gain was the narrowing of the gap between its very low starting mean Age Equivalent Scores and that of General Maths. This suggests that mental arithmetic is generally being neglected.

5.2.1 Considerations

It would be reasonable to conclude that the use of dedicated manipulatives and specifically the adopted *Sumdials*' approach to learning number was the cause of these gains, since it was the main known change that had been effected. Confidence in this claim is justified because the gains were measured against the implicit national standard provided by the InCAS database, (p.93).

In support of this claim it can also be pointed out that such gains are the predicted outcomes of the *Sumdials*' approach with its use of manipulatives – the dials – and are consistent with the earlier suggestions of wider research into the effects of interventions being justified. For example, Bramald stated that his results merited further study while the conclusion of my longitudinal Study was the approach had

produced lasting improvements instead of typically “washing-out” within two or three years, Sylva (1994).

Thus, it can now be claimed that three independent studies have arrived at the same conclusion. Science teachers used to explain:

- One reading means nothing.
- Two readings the same are a coincidence.
- Three readings the same *are* proof.

It can be concluded on this principle that the *Sumdials*' approach to leaning number has been proved to be effective. The research team is convinced that this is because the dials satisfy the criteria for effective manipulatives. As a reminder, Threlfall asserts that ‘practical number apparatus has a role’ in learning number without specifying what it is. It has been proposed (by me) that their essential characteristics must:

- Model the basic arithmetic processes analogously.
- Be pupil friendly and easy to use.
- Be obvious – comprehensible – to pupils.
- The ways in which they are used should be easily demonstrable with the minimum use of words.
- Incorporate:
 - (i) The Arabic number symbols.
 - (ii) A number line (or other relational representation).
- Be constrained in the ways that they can be used so that they support effective practice.

It seems likely that the dials contributed to these positive results because they satisfied these criteria. Nevertheless, it cannot be ruled out that the teachers responded to taking part in a research project by becoming enthused about teaching number and making an extra effort. There was little obvious evidence of this apart from one experienced teacher who admitted that hitherto she had never looked forward to

number lessons, but immediately saw the potential of the *Sumdials*' approach and its dials and was delighted to follow the first 12 lessons of the teaching plan (learning how to use the dials and then apply them). Most of the other teachers were surprised and pleased by how easy it was to deliver the teaching plan – again, the first 12 lessons only – proved to be, but this does not necessarily signify extra effort being made. Another possibility is that the dials proved to be more effective than the teachers expected and this enabled the pupils to do more sums than usual in the time. Even if this was so, the results still confirm the outcome was very beneficial.

5.2.2 Alternative Analysis

A different line of analysis provides another possible explanation for the effectiveness of the approach and it is based on the Pearson's correlations summarised in the Results Chapter and now reproduced:

Empirical Study Pearson's Correlations

Pairings	N	Correlations
G/M v G/M (1&2)	148	0.953**
G/M v D/A	148	0.671**
G/M (1&2) v M/A	148	0.627**
G/M (1&2) v D/A	146	0.627**
G/M v M/A	146	0.610**
M/A v D/A	144	0.429**

** Correlation is significant at the 0.01 level (2-tailed).

Table 5.35: The Correlations between Sessions and Sub-sessions

It had been observed that the gains in the mean age equivalent scores of Mental Arithmetic and Developed Ability had moved in step during the Empirical Study (p.104) and an analysis between them found the earlier correlation: having done so, the additional ones were made to reveal stronger correlations. The real relevance of these is the experimental pupils would have been using the *Sumdial 10* (adding and subtracting up to 10) and the *Sumdial 20* (adding and subtracting up to 20) to develop their adding and subtracting skills.

Now, General Maths has the four modules, as previously described, (p.84):

- *Number 1.* This includes counting, informal arithmetic (i.e. a number problem presented as: ‘Here are 6 ice creams, if 3 are taken away how many will be left?’), partitioning and place value, fractions and decimals.
- *Number 2.* This deals with sorting, patterns, formal arithmetic, problem solving and algebra.
- *Measures, Shape and Space.*
- *Handling Data.*

In simple terms, further basic arithmetic skills need to be developed before the third and fourth modules could be attempted and the dials were developed to achieve this. This suggested the further analyses that used the combined mean Age Equivalent Scores of Numbers 1 and 2, as recorded in the table. All of them are now considered:

- The starting point is the need to ‘bear in mind the dictum that *correlation does not imply causation*’ (Kinnear & Gray 2011, p.8).
- All the correlations are strong and even very strong, but these could be due to coincidence of other unidentified causes, even though the most likely one is the effect of the dials.
- The very strong correlation between General Maths v General Maths (1&2) was to be expected and confirms that learning the basics of arithmetic prepared pupils for their Measures, Shapes and Space and also the Data questions.
- The other correlations were appreciably stronger than the Mental Arithmetic v Developed ability one even though it was analysed in the first place because of the chart showed them moving in step, (p.104). It had been produced in response to a superficial inspection of their results that suggested this might be the case. This prompted the conclusion that Mental Arithmetic must be the “driver” for the unexpected improvements in Developed Ability.
- It can be postulated that the rapid Mental Arithmetic gains were attributable to the better internal models of number that the dials helped the pupils to develop. Put another way, the dials allowed the pupils to acquire transferable competences independently of their teachers that enabled them to answer more Mental Arithmetic questions.

- More generally, it can be argued that the basics of arithmetic are equivalent to learning the alphabet as preparation for learning to read and write.
- The accumulating evidence suggests that the dials are effective even though this as yet cannot be *proved*, but there may soon be sufficient evidence to establish it beyond reasonable doubt.

However, when these observations are compared with the pre-study work of Bramald and myself a common underlying process can be detected and summarised as:

- Bramald’s Study was a short intervention on the entry-level *Sumdial 10* that focused on adding up to 10 and this could be described as manipulating the “counting” numbers (the integers) only.
- My follow-up study, as the original pupils were completing their primary school education, found that the treatment pupils’:
 - (i) National attainments were better than those of the control pupils.
 - (ii) Automaticity scores in *all* the basic arithmetic processes (not just addition) were better those of the control pupils, (p.35).
- It can be postulated from these findings that Bramald’s Study identified the role of the dials in establishing robust internal model of number for the treatment pupils that became the foundation for the other basic processes.
- It can be inferred from the correlations that a similar process started in the current study when the participating pupils were introduced to their *Sumdial 10s*.

If these points are valid then it can be concluded that a sympathetic introduction of the pupils to the *Sumdial 10* (once they have become number-ready) is likely to start the development of pupils’ robust internal models of number that will become beneficial to them for life. Analogies such as learning to start the engine of a car as the essential first step in driving or learning how to hold a golf club and address the ball is the essential first step in learning to play golf. When all three such activities are done sensitively and rewarded with success, then it is likely that confidence will grow to make them good performers for life with numbers, driving a car or playing golf.

It is acknowledged that further study still needs to be made before it can be claimed that the *Sumdials*' approach was the *sole* reason for the remarkable gains. Verification (or dismissal) of such an assertion is more likely to be made by cognitive development neuroscientists or by large scale randomised trials. Having said that, it is now believed that there is now sufficient evidence to justify such studies and, in the meantime, the research team plans to do further follow-up work to gather more evidence for examination.

5.3 Statistical Results

To return to the statistical results, the Study conclusions could have been:

- Significant statistical improvements in General Maths had been achieved.
- 'It does seem reasonable to suggest that [the *Sumdials*' approach] was the major reason' (Bramald, p.22) for the highly significant improvements in Mental Arithmetic as the result of improved automaticity.

This would have allowed the null hypothesis of the main question to be rejected.

However, the conclusion (of highly significant Mental Arithmetic gains) needs to be reconciled with the very limited observable evidence of *any* automaticity - answers at finger-clicking speeds - throughout all the InCAS assessments, as would certainly have been expected by the last assessments when the mean ages-at-test was 7.14 years. The reality was that most of the pupils were still *calculating* their answers in spite of apparently following the *Sumdials*' teaching plan because they had become their default methods and they still were not acquiring any traditional mental arithmetic skills. This is the almost inevitable outcome since:

- Mental arithmetic was not emphasised in the curriculum in either Co. Durham or Edinburgh, based on a total of 545 pupils who were originally assessed in the Pilot and Empirical Studies, in terms of their age related performance, and is a reasonable interpretation of the Mental Arithmetic histogram, (p.120).

This means some other explanation needs to be found for the outstanding gains in the pupils' Mental Arithmetic scores.

It was assumed at the planning stages for the Studies that Mental Arithmetic would be the obvious InCAS session to measure automaticity because its actual questions were consistent with this, (p.87). However, its suitability came into question when it was realised during the initial assessments that the time allowed for *each* Mental Arithmetic question was 30 seconds instead of a maximum of, say, five seconds that would be needed to assess automatic answers. Peter and I had not spotted this prior to the assessments proper when we were familiarising ourselves with InCAS by answering a few “test” Mental Arithmetic questions and we both completed them in less than five seconds.

Nevertheless, it can reasonably be concluded that *Sumdials*' approach, with the visual model of number and tactile emphases of the dials and the opportunities for practice that it afforded, was the cause of *both* the Mental Arithmetic and General Maths improvements. The evidence from this Study supports the conclusion that the use of the dials was more effective in helping pupils to acquire coherence and meaning to number than they had previously experienced. But that does not necessarily mean they had acquired automaticity.

In fact, the time taken to answer each Mental Arithmetic question (and General Maths also) is recorded, but it can only be retrieved manually and individually from the raw dataset of the InCAS scores: this line of enquiry has not been pursued because of the limited observable evidence of any automaticity.

By way of further explanation, there are four types of question in both General Maths (as just described) while for Mental Arithmetic they are:

- (i) Addition.
- (ii) Subtraction.
- (iii) Multiplication.
- (iv) Division.

It became apparent that most of the General Maths 1 and 2, and *all* the Mental Arithmetic questions were of the same *type* to such an extent the simplest way to

know whether they were Mental Arithmetic or General Maths questions was to look at the screen headers – hence, the decision to carry out the correlation analyses. Simply: both were measuring overlapping attributes.

This leads to two conclusions:

- The *Sumdials*' approach with its manipulatives supported gains in General Maths and especially so in Mental Arithmetic, but also quite unexpectedly in Developed Ability.
- Another method needed to be found to assess automaticity.

These are now considered starting with the effects of the *Sumdials*' approach. Having said that, the teachers found the diagnostic value of the score breakdowns between the basic arithmetic processes very helpful.

5.3.1 Contributions of the *Sumdials*' Approach

As the Summary Table shows the gains are remarkable by any standards and a supporting explanation for them is now suggested. The starting point is to consider the measurement units that are used, being Age Equivalent Scores. Again, these are derived from the Rasch equal interval linear scales used in all the InCAS results and indicate degrees of difficulty of questions; it will be recalled that a fuller explanation of Rasch scales is included in the Methodology Chapter, (p.89). The most important point now is that the scales are derived from very large databases that effectively become *national* scales. They are the number equivalents of Reading Ages.

This leads to the key point. It can be assumed that the effects of the *Sumdials*' approach are not included as constituent elements of the InCAS scales. This assumption is based on the limited take-up to date of the approach. Therefore, it can be concluded that these gains are attributable to the *Sumdials*' approach and the focused practice that it supported because it provided one commonality across the participating schools. It is again asserted that the dials develop robust internal models of the basic number processes *directly* through visual representation and tactile

affordances in place of the traditional decoding of *indirect* word-based explanations or use of less structured manipulatives.

There could, of course, have been other influences that led to the pupils' improved attainments such as their teachers allocating more *time* than usual to number work because they were taking part in a research study. This is unlikely with the general shortages of time in schools these days and the majority of primary school teachers are not looking for reasons to spend more time than necessary on number work, (p.141). Another possibility is that the dials proved to be more effective than the teachers expected and this enabled the pupils to accelerate and do more sums than normal in the usual time. The apparent benefits of this would have been confirmed by the InCAS results that were given to the schools promptly. The participating teachers appreciated them and it would be likely that they would be encouraged, especially when the results were becoming progressively better than expected; again, their diagnostic value was also welcome.

5.3.2 Developed Ability

To return to the above conclusions, three points are made about the Developed Ability scores and they are:

- As has been explained, it seeks to measure pupils' *acquired* (not taught) inter- and intra-personal skills and normally develops as a function of pupils' ages. This was found not to be the case and the correlation analysis was made on the hunch that Mental Arithmetic was leading to the unexpected improvements in Developed Ability; it seemingly supported such a conclusion.
- It was an obvious link to make that Mental Arithmetic was driving the Developed Ability gains (cf. the Summary Gains table above) when, in fact, the correlations suggest it was more likely to have been improvements in General Maths that were having a greater impact.
- However, it can also be argued that this is surely attributable to the *Sumdials'* approach since it would have been a new influence that had not been previously evaluated.

In short, the Pearson's correlations suggest that the three sessions of Mental Arithmetic, General Maths and Developed Ability became intimately interconnected *through the dials* and this is the real reason for the unexpected gains in Developed Ability. It cannot be ruled out that there is even a link between this point and Resnick's observation that learning number is essentially a socialising process.

In summary, the distinctive feature of the Studies was the use of the dials and it can be concluded on the available evidence that they contributed to the gains in, as yet, unexplained ways.

5.3.3 Assessing Automaticity

With regards to the second conclusion on the need for a way to measure automaticity, a simple solution could be to use the existing InCAS bank of Mental Arithmetic questions within a framework of answering as many as possible with a time limit of, say, six minutes (the time limit for the Developed Ability modules). It should be anticipated with automaticity that up to 100 questions would be answered in that time, on the assumption that both:

- Number fact recalls, and
- Mental calculation questions are asked.

However, such an apparently simple change is unlikely to be made solely on the convictions of two researchers who had learnt their arithmetic when memory work to achieve automaticity was *de rigueur*. The point is the InCAS questions measure degrees of difficulty whereas *all number facts have the same level of easiness/difficulty with automaticity*. In the meantime, pre-recorded pencil-and-paper automaticity tests will be used to make the arithmetical automaticity assessments in the Follow-on Study. Initially, there will be issues of being able to express the results as the equivalent of Age Equivalent Scores (AESs), but in time a database will be created that will allow an attempt to be made.

On reflection, I now realise that I first encountered the word 'automaticity' through Peter, but still assumed that memory work would be standard practice even though the

word itself was not part of the common parlance in schools; this was supported by the experiences of Bramald's Study. Hence, it was almost inevitable that I would assume the Mental Arithmetic measured automaticity attainments - a good example of the potential risks of unchecked assumptions?

5.4 Memory Work

Former teaching practices for memory work are now sometimes denigrated as rote learning or "learning parrot fashion", even though pupils could *recall* their number facts accurately, because they might not have understood what they meant.

The main influences against learning number facts are:

- The 'first principles/understanding' argument.
- Lack of time to achieve it.

These are now considered.

5.4.1 First Principles Argument

The first-principles/understanding argument has a very seductive appeal, especially in word-based cultures such as those of the developed world. Indeed, it is difficult to make a philosophical case against the concept that understanding a new topic must be better than simply learning the procedures for applying it. Achieving understanding in practice is not straightforward, as the following citation makes clear:

My concern with the question of understanding has its sources in the practical problems of teaching mathematics and such basic and naïve questions as: how to teach so that pupils understand? Why, in spite of all my efforts of good explanation they do not understand and make all those nonsensical errors? What exactly don't they understand? What do they understand and how?

Sierpiska (1994, p. xi)

Implicit in this is the possibility that understanding may not be definable in practicable terms. The reality with number is there is a limit to what pupils can understand at each stage of their experience. It is acknowledged, of course, that the actual levels of understanding achieved are influenced by the quality of teaching and

parental support together with the time made available; this is where the role of successful procedures usually comes to the rescue.

To give a personal example, it is only very recently that I understood (in the sense of being able to explain) why the simple *procedure* to divide by a fraction is: turn the divisor upside-down and multiply. I continue to use it because it works so well and, of course, there are many such arithmetical procedures that have emerged during the 3,000 years it has taken to develop numeracy. The point is it is unlikely to have made any *practical* difference to me if I had attained understanding in the first place in that I would have continued to use the tried-and-tested procedure.

Strategies, as they are known, are encouraged under the first principles-approach, but they are inherently less efficient than automaticity. For example, the steps of a strategy for 7×9 could be:

1. Add: 1 to $9 = 10$,
2. Multiply: $7 \times 10 = 70$,
3. Deduct 7 from 70 as: $70 - 7 = 63$.

As a comment, the knowledge of two number facts (the first two steps) is required to get the correct answer and then deducting 7 from 70 is likely to be error prone and even when it is carried out correctly and there is always the possibility that 9 is deducted instead of 7 . The automatic recall of the correct answer of 63 requires only one number fact to be known. This is faster and more efficient and contributes generally to number ability, as Krutetskii identifies, (1976 p.189). The contribution to conservation of memory is implied as is discussed later, (p.220).

Another example of the benefits of using procedures was given by a head teacher about his uncle who worked as a fitter in an engineering factory. When he needed to create a right angle he would use a compass to draw on a piece of scrap sheet metal the three sides of a triangle with the relative dimensions of three, four and five units. On a visit, his uncle then explained the theory behind his creating a Pythagorean triangle and his response was to give his uncle a piece of sheet metal and his compass and then asked him to make a right angle. He was simply making a practical

statement that all he needed to know was *what* to do and knowing *why* he did it, would not enable him to make better right angles!

Enlisting help from my grandsons has previously been mentioned and this is now illustrated by reproducing the answer to a question given by one of them from an exam paper:

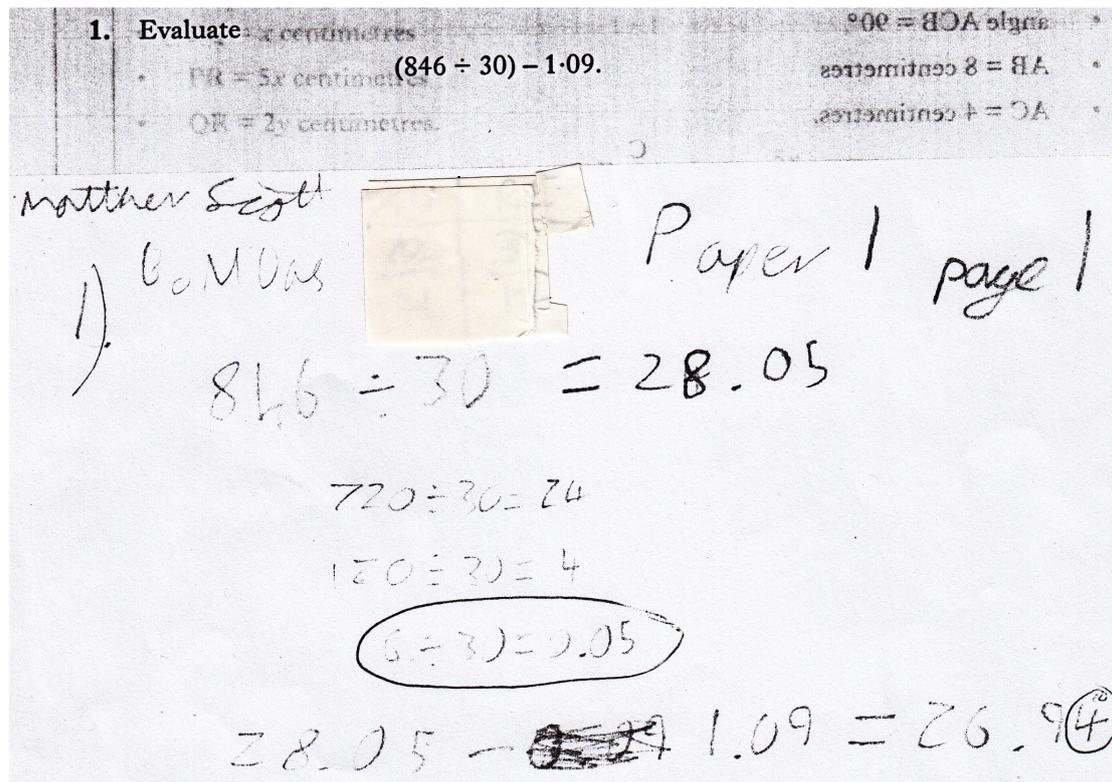


Illustration 5.6: Copy of the Answer.

In view of the poor quality of reproduction, his answer is now transcribed:

BoMDas

1)

$$846 \div 30 = 28.05 \text{ (This answer had been entered after completing his calculations, as below)}$$

$$\begin{aligned} 720 \div 30 &= 24 \\ 120 \div 30 &= 4 \\ 6 \div 30 &= 0.05 \end{aligned}$$

$$28.05 - (0.09 \text{ crossed out}) 1.09 = 26.94 \text{ (the 4 had a very small 10 above it and both were enclosed within a circle).}$$

5.4.1.1 Comments:

- ‘BoMDas’ is the acronym for: brackets, of, multiply, divide, add and subtract, being the sequence in which the calculation steps are made – and was shown because he had been taught marks would be awarded for including it.
- ‘846’ had been decomposed into 720, 120 and 6 before each was divided by 30 and then added together (no workings shown) to arrive at an incorrect answer because of a mistake made when dividing 6 by 30. It suggested that he had arrived at the fraction $\frac{1}{5}$ and ‘converted’ it into the decimal 0.05 (a common error).
- The final subtraction was incorrect, presumably because the number fact of $15 - 9 = 6$ had not been hardwired into his long-term memory and his unknown method of calculating the derived fact was faulty.
- The overall comment is a well-designed question had revealed gaps in his knowledge of the basics of arithmetic and, being a STEM student, is consistent with Gibson’s workshop observations, (p.216).

My answer was:

$$(846 \div 30) - 1.09 = 27.11$$

No workings are shown because the evaluation was done entirely ‘in my head’.

My metacognitive account is:

- I inspected 846 and realised it was divisible by 3 (because the sum of its three digits = 18 and is divisible by 3 – a procedure learnt at school); this would have been done in my short-term memory.
- 846 was then divided by 3 to arrive at 282 and the decimal point was then moved one place to the left to divide by 10 to get 28.2 (another procedure retrieved from long-term memory). Thus a combination of number facts and procedures were retrieved from long-term memory and manipulated in my short-term memory.

- I then deconstructed 28.2 into 27 and 1.20 and stored 27 in my short-term memory while 1.09 was subtracted from 1.20 to give 0.11 and stored it in my short-term memory. Finally, 27 was combined with 0.11 to give the answer of 27.11.
- My estimated time to evaluate the question was about 15 to 20 seconds.
- It is very likely that the short-term memory registers would be over-written once their original items had been used (as would happen with computers) and I had probably forgotten what the question was by the time I had answered it.

What was the point of this exercise? The main one was my grandson and I learned number (arithmetic in my case) on opposite sides of the New Maths divide and it provided a good opportunity to find out if teaching and learning methods had changed. They had!

Some more background concerning my grandson reinforces another point. He is generally perceived to be very bright as Peter, wearing the two hats of his maths teacher and his tutor, confirmed. The tutoring arose because my grandson had developed a medical condition that severely reduced his stamina to the extent that he was only able to attend two or three lessons each day for the last two years at secondary school. Special tutoring arrangements were made and these included Peter tutoring him for maths that led to a successful entry to university. This was in spite of starting with a very uneven knowledge of number, as evidenced in the above example. There might have been three influences at play:

- He had fallen between the two stools of
 - (i) Not mastering the basic arithmetic procedures supported by a good knowledge of his number facts.
 - (ii) Not understanding what he was doing.
- He had “discovery” teachers at primary school that encouraged ownership of his own complicated methods instead of using them as starting points to show him the more efficient algorithms, as a ‘connector’ teacher would have done (p.59).
- He was still operating on the insecure foundation laid at primary school.

My own recollection is teaching and learning was much simpler, albeit within much less ambitious curricula, and the results were probably better than now - because the foundations were being taught more thoroughly. And to keep matters in perspective, my own self-assessment is that I would have been at the top end of the middle group when at school. Peter agreed and then added that I would now have been at the top of the top group. This suggests that the performance of number skills have declined considerably during my lifetime and his time as a teacher. I am convinced that a proper study of grand parents v grandchildren would arrive at the same conclusion.

It is believed these examples demonstrate that using tried-and-tested procedures can be very effective even though they may not satisfy the intellectual comfort that comes with working from first principles and developing understanding, as these are time and mental-energy consuming (and may not actually be successful).

5.4.1.2 Lack of time Argument

The lack of time to achieve fluency in number facts is an obvious explanation for teachers to use for their pupils not having automatic recall of their number facts, but the realities are much more complicated; they are discussed more fully when teacher qualifications are being considered, (p.211). Suffice it to say now, the issues include lack of conviction of its importance, not knowing how to achieve it, acceptance of the first principles argument, fear that the pupils would find it boring or too challenging, to name some of them.

It must be pointed out that achieving automaticity in primary schools – when young brains are much more “sponge like” than when they reach secondary school – leads to a much more productive use of time. To delay the development of fluency or ‘hardwiring’ of number facts until secondary schools is a false economy because it takes much longer to achieve then and reduces the time available for higher subject teaching and learning.

5.5 Initial Conclusions

The main conclusion must be that the *Sumdials*’ approach to learning number during the Study was very effective for the pupils in their earlier years at primary schools.

The gains for Mental Arithmetic, in particular, and General Maths confirmed the effectiveness of the approach.

However, it was not possible to assess the approach's effectiveness in 'hardwiring' number facts to achieve automaticity because it is no longer being emphasised and assessed. The observations made during the InCAS assessments confirmed that *true* mental arithmetic, in terms of a focus on instant recall, was no longer a key arithmetical goal. It can be anticipated that even greater gains would be achieved once mental arithmetic, based on automaticity and the development of fluency in mental calculation, is reinstated as an essential capability for mastery and fluency with number. Thankfully, apart from not being able to provide a complete answer to the main question, the Study was able to demonstrate considerable arithmetical gains. Identifying scope for further improvement (of InCAS) allows consideration to be given in anticipation of when the need to measure automaticity becomes accepted.

In character, my response on discovering that mental arithmetic was no longer being emphasised was to find the reasons for such a very unexpected development. To me, there had to be a reason(s) and it (they) had to be discovered, (p.208).

5.6 Secondary Questions

The secondary questions were:

- Are there statistically significant differences in the number attainments:
 - (i) Gender (boys and girls)?
 - (ii) By Location (between Co. Durham and Edinburgh pupils)?

5.6.1 By Gender

The results confirmed that there were no statistically significant differences between the boys' and the girls' assessments and this was as predicted for pupils of their ages, Dowker (2005, p.7). Also, consistent with her observation 'Males are more likely to

be *extremely* good at mathematics' was the two extreme cases in the Empirical Study were both boys.

To make a personal comment, my working life has mainly been in manufacturing industry where arithmetical abilities were essential and male colleagues dominated. Thus, these findings surprised me - hence my clinging to the finding that the mean scores of the boys were higher, even if not significantly so, than those of the girls. The data were available for analysis without having to make any arrangements to collect them and arriving at the widely predicted conclusions at least confirm that the dataset is reliable.

The other reassuring point is that Dowker, (2005, p.3) treats *arithmetic* as a stand-alone subject that is quite separate from mathematics, as is the context for Threlfall's paper. The relevance of this point was explained (p.126).

5.6.2 By Location

The main reason for comparing the Co. Durham and Edinburgh pupils is they learn in two different national regimes. The opportunity to assess the differences *had* to be taken, especially as the data would have already been collected during the InCAS assessments without making any special arrangements.

The starting ages of the English and Scottish Schools are also different. It was thought initially that the simplest method to compensate for this would be to deduct the Ages-at-Test from subject Age Equivalent Scores (AESs) of Mental Arithmetic, General Maths and Developed Ability before calculating the mean scores. However, this would produce negative net scores in many instances that could have caused conceptual difficulties with the participating teachers and was replaced by straight comparisons of AESs; this was consistent with all the other results. The Co. Durham schools did better than the Edinburgh ones with effect sizes of medium/large and **p** values that were significant at the 95/99% levels. These findings are now considered.

Firstly, as background, it was always intended that all data would only be analysed at *group* level for both practical and ethical reasons. Thus, the Co. Durham and

Edinburgh pupils would be two separate groups, the boy and girls would be another pair and the Mental Arithmetic, General Maths and Developed Ability scores became other groups paired with their corresponding ages-at-test (once the control schools were no longer participating in the Studies). This also ensured the performances of individual teachers would not be implicitly assessed and the teachers, recognising this, willingly co-operated in the research. Moreover, neither Peter nor I would be competent to assess teachers and the focus of the research could have become blurred if it had been included.

Again, the research aim was to assess the effectiveness of the *Sumdials*' approach and applying it to both groups ensured a degree of uniformity in teaching and learning within the two groups. Thus, the need became one of trying to identify *why* the Co. Durham pupils had scored better than their Edinburgh counterparts. This is now considered and as far as could be judged, both groups were:

- Following very similar programmes in very similar settings.
- In apparently similar socio-economic catchment areas.

Four possible explanations are considered:

- The settings.
- Ages at Test.
- InCAS data.
- The Teachers.

5.6.2.1 The Settings

The research team's superficial impression was the Co. Durham and Edinburgh schools were using similar approaches in spite of the NNS curriculum in the former and the 5 to 14 Years Guidelines in the latter. This, of course, was the reason for making the comparisons in the first place. However, there were many similarities between the two groups such as both worked in small groups and finger-counting nose/head-tapping being widespread in both locations to give only two examples. In practice, the pupils' accents were the constant reminders of the classroom locations.

Having made that point, the resulting general ethos of the two groups may well have been different and we had not discerned this.

5.6.2.2 Ages at Test

Ages at test may have contributed to the better results achieved by the Co. Durham pupils and the different class nomenclatures used in England and Scotland may have indirectly exacerbated this.

To explain, the starting dates of the Co. Durham pupils are six month earlier than the Edinburgh ones giving classes with birthdays in the following year groups:

- Co. Durham: 1st September to 31st August the following year.
- Edinburgh: 1st March to 28/29th February the following year.

The effect of this is the Co. Durham pupils would have had six more months classroom experience and this would have rendered them *more likely to be number-ready* by the time they were introduced to the Sumdials' approach. The point here is that a higher proportion of the Co. Durham pupils would be likely to benefit from their dials to develop robust internal models of number than would be the case with the younger Edinburgh pupils.

This is consistent with Piaget's developmental stages and highlights the possible difficulties that could arise if pupils do not have appropriate mathematical experience as a result of starting their formal subject learning too soon. This is supported by the results at the end of this Study that showed:

- The Co. Durham pupils' mean Mental Arithmetic scores were 0.83 years (10 months) ahead of the Edinburgh pupils even though they had become only 0.32 years (4 months) older.
- Similarly, their mean General Maths scores were 0.60 years (7 months) ahead while being only four months older.

These conclusions are clearly evident in the Mental Arithmetic and General Maths charts, (pp.112/113). Nevertheless, there was always the possibility that the mean age equivalent scores used throughout the Study itself concealed skewed results and, as a check, boxplots were produced to confirm this was not the case:

Comparative Boxplots of the Co. Durham and Edinburgh Pupils

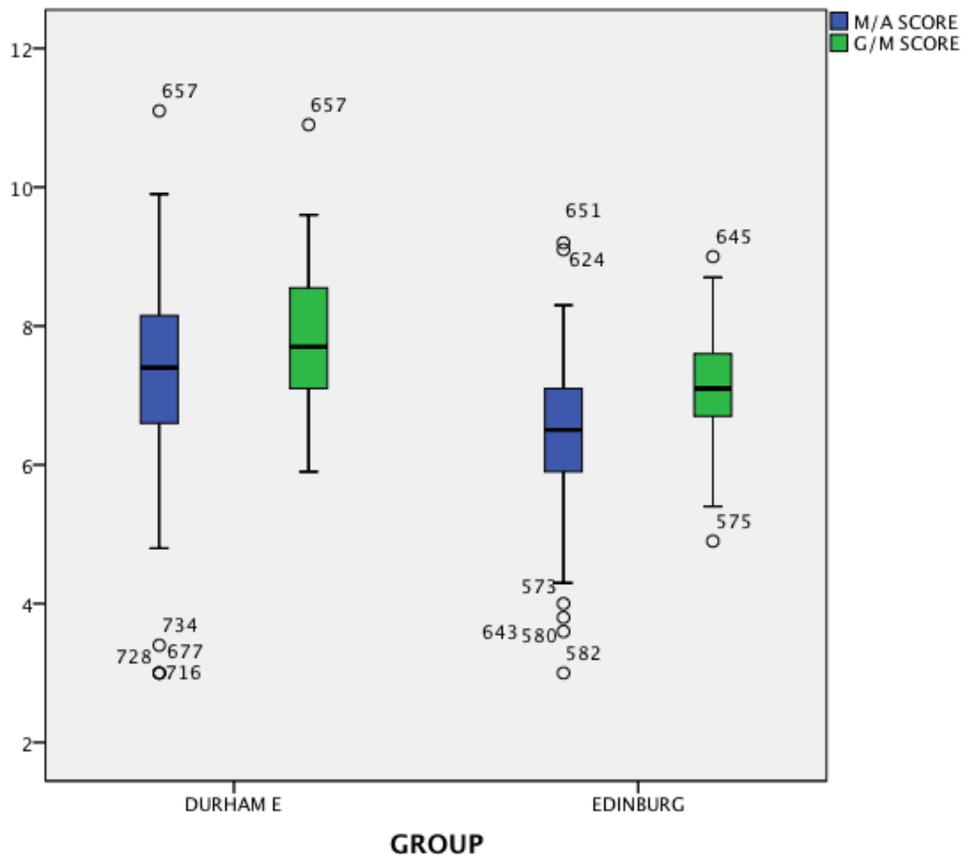


Figure 5.23: Co. Durham and Edinburgh M/A and G/M AESs Boxplots for Month 21

In short, the relevant Results charts show the gaps had only narrowed for Mental Arithmetic during the first year and there was no overall gap reduction for General Maths throughout the Study, while these boxplots are consistent with the Results Tables based on mean Age Equivalent Scores, (pp.110/1). However, more evidence is required to support the hypothesis that starting formal number learning before pupils are number-ready may have a lasting effect and it is now intended to reassess the surviving pupils at the end of their Y4/P4 years. Starting age possible effects are discussed later, (p.148).

5.6.2.3 InCAS

The versions of InCAS used in this Study were compiled from assessments made in English schools and this could have put the Edinburgh pupils at a disadvantage because of differences between the two curricula. However, the time to confirm or refute this could have been difficult to justify for what was a secondary question.

5.6.2.4 Teachers

Another possible explanation for the Co. Durham pupils' success is their teachers were more effective in mathematics teaching than the Edinburgh ones. Again, no attempt was made throughout the Study to assess the abilities of individual teachers, but the *subjective* impressions of the Co. Durham teachers were they seemed to have more "presence" or authority than the Scottish teachers did. Consistent with this, all school arrangements in Scotland had to be made through the head teachers whereas they were made directly with the class teachers in Co. Durham. Indeed, we never met the current head teachers in four of its schools while the other two schools were so small that it would have been difficult to avoid them!

It can be noted that the standard deviations of month 21 for both the Edinburgh Mental Arithmetic and General Maths were smaller than the Co. Durham ones and that suggests that the Edinburgh teachers could be more effective at keeping the class progressing together. Thus, further investigation would be required to resolve this aspect, but it would be outside our competencies or the scope of this research.

5.6.2.5 Conclusion

In conclusion, the statistical evidence allows the null hypothesis to be rejected because, for Mental Arithmetic, the effect size was medium and the **p** value was highly significant at the 99% level, while for General Maths the corresponding values were large and significant at the 95% level. In the event, any or all of the above considerations could have contributed to the result.

The most likely explanation is the Co. Durham pupils had become more number-ready because they were six months older and had been in classrooms for 12 months longer. It is now planned to assess the surviving pupils at the end of the current session to determine whether or not the gaps between the two groups have changed.

However, it cannot be ruled out at this stage that the Co. Durham teachers were more effective in supporting progress in mathematics.

5.7 Other Results

There are two other Results that are presented even though they were not part of the original Empirical Study plan:

- Ability classifications.
- Starting Ages.

They are now discussed.

5.7.1 Ability Categories

The anecdotal evidence in schools suggests that pupils may be classified as those who:

- Will do well, almost regardless of how they are taught 10 - 15%
- Could and would do better, when taught appropriately 70 – 80%
- Will always struggle, regardless of how they are taught 10 – 15%

It seemed likely that such a breakdown had always been generally accepted without any measurements ever having been made. In light of this a descriptive analysis was made and confirmed that the three classifications are reasonable (pp.120/1). Two points arise out of this widely accepted classification and they are:

- These classifications are self-selecting and seldom disputed in the sense that:
 - (i) The first group have a different order of number ability from the others that is both very apparent and is not usually disputed. Such pupils typically *retain* new learning after only *one* explanation.
 - (ii) The second group is likely to need several different explanations before its pupils apparently ‘grasp’ some new topic, but still remain dependent on learning appropriate procedures.
 - (iii) The third group struggle to learn.
 - (iv) These groupings tend to persist throughout school and into adult life.

- The same groupings are just as applicable to teachers, head teachers and indeed to all occupations.

As a general observation, the policy makers favour the top and bottom groups by allocating more resources at the expense of the middle group. Interestingly, when the follow-up Study was being set-up the teachers were very pleased that they were asked to select their typical ‘middle-of-the-road’ pupils (from the middle group). Our reasons were:

- Those in the top group, especially, learn readily and do not need extra support.
- Raising the mean number attainments of the large middle group by even as little as 5% will enhance the national competitiveness in due course.

As a general comment, it would be helpful if there were a greater awareness and acceptance that such distributions are facts of life and the challenge is to respond to them constructively on the evidence improved mean scores of both groups.

5.7.2 Starting Ages

Starting ages has become something of an issue in that pupils here are often only four years old when they start school while in Europe they usually start when they are six or seven years old; that is more in line with Piaget’s development stages. They are considered now because, as previously mentioned, (p.116), an experienced teacher, who took part in both the Pilot and Empirical Studies, wondered if her Empirical pupils’ answers using their dials actually *meant* anything to them. As further background, she made this comment during the informal “chat” at the end of the lesson we had observed her pupils using their dials. It was so well delivered that our regret was we had not videoed it to use it for teacher training generally! However, these pupils were five months younger than those in the Pilot Study when they were introduced to their dials.

The comparisons between the two cohorts show that her concerns were justified for General Maths only at Month 9, but by Month 21 the Empirical pupils had made

better progress in both subjects, but especially so in General Maths with an effect size of medium and a **p** value that was highly significant (at the 99% level), (p.106). These Results are for the whole cohorts of both Studies and not just her classes. A possible explanation for the General Maths gains could be the teachers were on their second iterations by the time of the Empirical Study and so had become more proficient and confident in delivering the *Sumdials*' approach to the benefit of the pupils.

Overall, the Results suggest that the actual starting ages for primary schools *per se* are not the issue. The real issue, in the case of number, is ensuring that the pupils are number-ready and have appropriate early arithmetical experiences before the formal number teaching and learning commences; it is up to the teachers' professional judgments to determine when that is. As a guide, it is suggested that the dials could be used as the bridge between becoming number-ready and starting formal number learning.

It can be hypothesised that this is what had actually happened with the experimental cohort and is consistent with an experienced teacher's concerns (p.240). To explain, her observation that the answers her pupils were getting by using their dials may not have meant anything to them was valid, but the dials had started the process of establishing robust internal models of number. The structure of the *Sumdials*' teaching plan used in Bramald's study in 2001 provides a useful background for the current studies. It was based on traditional arithmetic practice – derived from our own childhood experiences - for learning addition up to 10 and had two sections. Again, they were:

- Section one consisted of 12 lessons in two parts:
 - (iii) Instructing pupils on how to use and apply their dials.
 - (iv) Answering adding question on worksheets.
- Section two consisted of 15 memory-work lessons.

It was expected that by the end of the six weeks course the majority of the pupils would be able to add up to 10 and have good automatic recall of their number facts.

The point now is: *the treatment teachers of the first Study found the approach unexceptional and willingly followed it.* It is relevant to this Study that one of the teachers, who took part in Bramald's Study (13 years earlier), confirmed the validity of this point. The experiences gained in his Study provided the foundations that were used for the Empirical Study and led to the very good gains by Month 21 when it was completed.

One observation from the Follow-on Study suggest that ensuring the pupils become number-ready *before* they start their formal number learning is a move in the right direction: this is more important than the rules that determine when pupils start their formal education.

5.8 Conclusions

The overall conclusions that have been reached after consideration of the Results are:

- The outstanding results can be attributed to the effectiveness of the *Sumdials'* approach to learning number with the manipulative dials that are its distinctive feature.
- It was not possible to assess automaticity scores because automaticity is no longer being taught – being insufficient time to achieve it.

Possible causes for the lack of automaticity are now discussed starting with the consequences of ever increasing government control of education from top to bottom. The implication of this change is that the original micro study acquired a macro focus.

6. The Swiss Study

6. Introduction

The primary role of this chapter is to provide a bridge between the accounts of the intended RCT chapters (Part 1) and the reflections on the originality and learning inherent in the action research cycles (Part 2). Its structure will be the same as the overall structure of the thesis.

It will be helpful to describe how two Swiss schools become part of this Study, albeit indirectly, before considering its parts of:

- The Swiss results and some comments on them.
- Reflecting upon some wider issues within Swiss education.

Part 1. The Swiss Results

6.1.1 Background

The RCT Pilot Study was completed shortly after the sad death of my wife and friends in Germany and Switzerland became concerned when they discovered that I was not planning to take any holiday. In response, I agreed that I would visit both sets of friends in one trip provided visits to local kindergartens could be arranged. This was to satisfy myself that their later school starting-ages was not due to formal education taking place in the kindergartens.

In the event, this was confirmed in Germany, but I discovered to my (suppressed) annoyance that such a visit could not be arranged in Switzerland: instead a small rural school was visited with a composite class for its first four years. This was very successful and led to another school participating in the following years. The pupils of both Swiss schools were assessed using InCAS, as in the Empirical Study.

These school involvements were not part of the original plan and could not become part of the experimental design mainly for logistical reasons, but it was sensed their assessments could provide useful contextual comparisons – as proved to be the case.

6.1.2 The Results

The same assessment procedures were used, but with translating help available as necessary. The first school results are summarised below, starting with the mean Mental Arithmetic Age Equivalent Scores compared with the mean Ages-at-Test.

Charts are not produced because the Swiss pupils were neither a direct part of the Study, nor strictly comparable. Also the samples were very small and spread over four age groups. However, their scores are considered below in terms of contextualising our understanding.

The Swiss School (1) Mental Arithmetic Results

Assessment	Mental Arithmetic			Ages at Test			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
First	9.046	14	2.93	8.612	17	1.42	0.19	0.550	0.423
Second	10.306	17	1.87	9.541	17	1.55	0.49	0.380	0.061
GAINS	1.260			0.929			0.30		

Table 6.36: The comparison between the Swiss School (1) mean M/A AESs and A/Ts.

It was immediately apparent during the actual assessment sessions that these pupils had a different order of mental arithmetic skills, as confirmed by these results, compared with those of the UK pupils.

The Swiss School (1) General Maths Results

Assessment	General Maths			Ages at Test			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
First	8.822	17	1.85	8.612	17	1.42	0.12	0.450	0.500
Second	9.935	17	1.09	9.541	17	1.55	0.29	0.299	0.206
GAINS	1.113			0.929			0.17		

Table 6.37: The comparison between the Swiss School (1) mean G/M AESs and A/Ts.

The Swiss School (1) Developed Ability Results

Assessment	Developed Ability			Ages at Test			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
First	11.498	16	2.59	8.612	17	1.42	1.71	0.303	0.000
Second	11.853	17	1.81	9.541	17	1.55	0.29	0.390	0.000
GAINS	0.355			0.929			-1.42		

Table 6.38: The comparison between the Swiss School (1) mean D/A AESs and A/Ts.

Once again, these results are of a different order than the UK ones and are considered shortly with the Mental Arithmetic and General Maths ones.

The results of the second school are now summarised and they will also be considered shortly along with those of the other school.

The Swiss School (2) Mental Arithmetic Results

Assessment	Mental Arithmetic			Ages at Test			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
First	8.837	5	4.073	8.887	5	1.43	0.03	1.285	0.971
Second	9.480	5	1.87	9.000	5	1.44	0.13	0.503	0.394
	0.643			0.506			0.10		

Table 6.39: The comparison between the Swiss School (2) mean M/A AESs and A/Ts.

The Swiss School (2) General Maths Results

Assessment	Treatment Group G/M			Ages at Test			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
First	9.152	6	1.92	8.744	6	1.43	0.22	0.510	0.461
Second	9.550	6	1.47	9.250	6	1.44	0.19	0.450	0.535
GAINS	0.398			0.506			-0.03		

Table 6.40: The comparison between the Swiss School (2) mean G/M AESs and A/Ts.

The Swiss School (2) Developed Ability Results

Assessment	Treatment Group D/A			Ages at Test			Effect Size	T Test	
	Mean	N	S/D	Mean	N	S/D		S/E	p
First	11.197	6	2.34	8.744	6	1.43	1.17	0.652	0.013
Second	12.600	6	2.12	9.250	6	1.44	1.71	0.772	0.007
GAINS	1.403			0.506			0.54		

Table 6.41: The comparison between the Swiss School (2) mean D/A AESs and A/Ts

6.1.3 Discussion of the Swiss Results

Again, being able to assess two very small Swiss rural primary schools was quite fortuitous and was not part of the original Study plan. However, it was obviously too good a *research* opportunity to pass up when the opening was there. It proved to be relevant because it was immediately apparent that *the Swiss children were all applying automaticity* when they were being assessed – using the same InCAS versions that the Study children had used.

Some general observations about the Swiss results are made now:

- The mean AESs of all results were higher than the mean Ages-at-Test or the same (once) in contrast with the UK results.
- The Mental Arithmetic scores were higher than the General Maths ones also in contrast with the UK ones. One possible explanation is the ‘translators’ found the General Maths questions more difficult to translate than the Mental Arithmetic ones and so required more time; the children who did their own translating may have encountered the same difficulties.
- The Developed Ability scores were much higher than their mean Ages-at-Test scores and this is also in marked contrast with the UK Study scores.
- The ways in which the pupils did their InCAS assessments showed they:
 - (i) Could subitise.
 - (ii) Could do mental arithmetic.
 - (iii) Had hardwired their number facts.
 - (iv) *Never* counted on their fingers!

- The usual explanation for lack of emphasis on hardwiring number facts given by the teachers of the Study schools was insufficient time, but also it could have concealed a lack of conviction about its importance. Again in contrast, one of the Swiss teachers has a minute-glass on her desk to consolidate number facts by asking pupils 20 random number facts in one minute, equivalent to three seconds to answer each question. In contrast, Study teachers thought this did not give enough time to *calculate* the answers – confirming the impression that they were unconvinced of the need to hardwire number facts.

Some other more general impressions are now made:

- Some rural schools probably do better than the urban ones in Switzerland, as is anecdotally the case in *some parts* of the UK and possibly many developed countries. This is another reason why direct comparisons with the UK urban scores would probably be of only limited value.
- However, the villagers’ commitment to education was convincingly shown at a workshop on the *Sumdials*’ approach that was given for the pupil’s parents of the larger school. The pupils’ assessments started at 0715hrs while the workshop started at 1930hrs and yet all the children came back with their parents - and even some grandparents - together with teachers from other schools leading to “standing-room only”. The workshop was a joy because of the interest and willingness of everyone to participate including the children demonstrating correctly how the dials should be used. Towards the end, one pupil (ten years old) found himself giving an impromptu demonstration on the beamer of adding pairs of three digit numbers together (from InCAS). Not only was he getting the correct answers, he was calculating them “in his head” to give a convincing demonstration on the contribution automaticity makes in traditional mental arithmetic.
- Their very high Developed Ability scores were almost certainly the direct product of their local environments in which the benefits of good education are a part of their culture and expectation; they were motivated to learn.

Such outcomes may also be attributed to a long established Swiss tradition of bottom-up education that is responsive to *local* needs. Education policy is determined at canton level (there are 26 cantons in Switzerland) while inspectors are appointed at area level and will typically visit a class once or twice each year, or more frequently if there are problems or complaints. Such visits are arranged directly with the teachers themselves, usually about one week in advance, and may only be for a morning or afternoon at the most - another striking contrast with inspections in the UK.

Interestingly, during an informal comparing of notes with one of the Swiss teachers revealed a fundamental difference in attitudes summed up by an implicit notice on the outside of his classroom door that states: “I am my own boss”! One effect of this is school head teachers do not and cannot influence their teachers’ pedagogies or how they run their classes. The policies of individual schools are determined by the votes of teachers at meetings chaired by the head teacher whose main responsibilities are to ensure that the school buildings are properly maintained and that there is an effective administration. The contrast with schooling in the UK could hardly be greater!

6.1.4 Initial reflections

It was quite providential that these two Swiss schools became involved during my Study even though their assessment timings precluded them from being part of the experimental design. However, the way the Swiss pupils conducted themselves throughout was how I had expected the pupils in the Pilot and Empirical Studies would perform. This observation greatly influenced my subsequent approach in that they convinced me that my own learning experiences must have been similar to those of these Swiss children. My reaction to Switzerland being the highest placed non-Asian country in the OECD (2013) results was that there must be a connection between this and my observations.

The impression should be well established by now that all is not well in the world of education. This can be attributed to the excessive political control at national level and, then, having set the strategic/macro aims, attempting to implement them through top-down tactical/micro initiatives. Many of these have not been tested by prior research to confirm that they will be effective and a ‘one-size-fits-all’ would be relevant to *all* actual local needs. In contrast, the German-speaking Swiss cantons

started work in 2011 to establish a common curriculum by 2021. There could be two changes of UK governments in that time that could lead to further changes of direction for education – hardly conducive to stability. Again, it can reasonably be postulated that there is must be a link between such a measured pace in Switzerland and the OECD (2013) results showing that it is now the top non-Pacific rim country (p.209).

6.1.5 Initial Conclusion

The Swiss results were very impressive, albeit based on a very small sample and, therefore, caution must be used in before drawing any conclusions. However, they do point to the need for further investigation at the national (educational) policy level and will be carried out as an action research enquiry. The findings will be reported in the second part of this chapter.

Part 2. Reflections

6.2.1 Inquiry

Action research was used to inquire into what influences educational attainments in Switzerland. The usual structure was used consisting of:

- **Intent:** Identify the main influences on educational outcomes in Switzerland.

- **Process:** The main steps included:
 - (i) Observing activities and forming impressions during school visits.
 - (ii) Gathering, analysing and reviewing relevant reports that provided information on policies, structures and outcomes.
 - (iii) Reaching provisional conclusions and discussing them with teachers and my Swiss friend, as my mentor, while my Supervisor continued to act as my sceptical colleague.
 - (iv) Finalising conclusions.

- **Audience:** The Examiners as mediated by the Supervisor and the wider mathematics education community after completion.

My Swiss friend acted as mentor on Swiss education. He was a primary teacher in the early stages of his career before becoming the chief executive of an organisation caring for the severely handicapped from infancy to beyond retirement ages. He is now nominally retired.

The reminder is now made that the original focus of my Study was a student research micro RCT that developed into macro reflections on education in the UK. This was one outcome of my experiences of visits to two Swiss primary schools that led to it becoming a bridge between the two parts of my thesis. It is essentially personal and something of an overview while being sufficient to fulfill its role: it is believed that within its acknowledged limitations it is competent, but does not pretend to be exhaustive.

Process:

The four steps are now described:

- **School Visits:** The primary reasons for the visits were to carry out the InCAS assessment and the impressive results have been presented and comment upon in the first part of this chapter. However, it was inevitable that ‘comparing notes’ took place and it was through this that the following conclusions were made:
 - (i) Two of the teachers (one from each school) were outstanding, very dedicated and acting virtually autonomously. However, there are increasing difficulties in recruiting candidates for teaching with the right attitudes.
 - (ii) The overall Swiss results were high (see above) and this can be attributed to high standards of teaching and pupil motivation, backed-up by parental support.
 - (iii) Head teachers are not expected to interfere in individual classrooms in acceptance of the “I am my own boss” syndrome (p.156).

- **History and Politics:** To appreciate Swiss education, a brief summary of Switzerland's structure and divisions of responsibility need to be known. It is a land-locked country with no natural resources that explains its conviction on the value of education, as was indicated in the workshop (p.155). It was established in 1848 in its current form as a Confederation of 26 Cantons (20 full and six "half") and 2,362 Communities. It is essentially a bottom-up democracy working within three levels of responsibility where citizens have several rights over the final decisions. The philosophy is public services should always be delivered from the lowest sensible level and all levels have their own budget approved by the relevant citizens. Five cantons are French-speaking, one Italian-speaking while the remaining ones are German or bilingual. Cantons are responsible for education and have been very effective in satisfying local needs while contributing to the overall national qualifications and their contribution to national employment rates, as an OECD Country Note confirms (2013).

However, Federal politics is unexpectedly becoming involved in education in response to increasing mobility and the anguish through children encountering different curricula from canton to canton – the most obvious one being no standard practice on when language teaching should start – or if it should be French or English. This has given rise to a federal requirement for 'harmonisation' in education throughout the land.

- **Provisional Conclusions:** In response, the five French-speaking cantons promptly agreed and implemented their new curriculum. Meanwhile, the 21 German and bilingual-speaking cantons appointed a specialist group of experts to draw-up Curriculum 21 covering the first nine years of education for implementation in 2021. Its work was carried out behind closed doors and without consulting teachers or schools. It identified 4,753 items of competence to achieve after 9 years.

The concept of 'Competence' represents a radical change because it would measure pupils' outputs in contrast with the inputs of what teachers have to teach. This is the **key** concept behind this discussion and claims to be the

modern approach in education and it is highly criticised in principle and, in particular, its insufficient lack of natural science and number. The most important concern is that it represents a fundamental change in education from bottom-up to top-down or in other words, a doctrinal initiative to bring teachers into line with the loss of their highly respected didactic freedoms. In addition, are the unknown costs and effectiveness of the resulting bureaucracy, (adapted from the ‘550 against 550’ memorandum, 2014). Not surprisingly, it was not well received when published and the canton ministers requested that the number of items be reduced to less than 1,000 and to report back within months. In the meantime, responses by individual cantons threaten to derail the whole Curriculum 21 project. However, if matters are not amicably resolved, the Federal Government, following a plebiscite, has the powers to step in to ensure ‘conformity’ of curricula aims is reached throughout the cantons. However, there is growing opposition that could lead to a referendum against the whole process.

On reflection, this is a needless tragedy in the making in that overall Swiss education is very effective and the envy of most other countries, as the recent OECD PISA results confirm. The colloquialism “If it ain’t broke, don’t fix it!” seems to be very relevant. Swiss education’s strength derives through responding to local needs even though all is not perfect, as the curricula inconsistencies show. With goodwill, it should be possible to resolve them at cantonal level instead of, to use another colloquialism, “taking a sledgehammer to crack a nut!” being appropriate for the fundamental change that is being envisaged. It can reasonably be asked: Is there no awareness of education experiences and outcomes in the UK or the USA?

- **Final Conclusions**

These conclusions now include corrections from Switzerland and can be seen as being final.

Audience:

I discussed this presentation of my Swiss experiences as an action research inquiry with my Supervisor and the Examiners and therefore added this bridging Chapter according to their suggestions.

6.2.2 Conclusion

The unexpected findings and developments in Swiss education at cantonal and federal levels provide a good lead in preparation for the reflective second part of this thesis.

7. Action Research

7. Introduction

The data collection for both the Pilot and Empirical Studies, together with their analyses, had been completed when, in consultation with my Supervisor, it became apparent that my approach had many of the characteristics of Action Research. It is now briefly considered because of its relevance to this research.

7.1.1 Considerations

Primary schools were the shared environments with teachers in which this Study took place and action research is generally associated with teacher self-improvement initiatives. A brief overview of the history of action research, its principles and applications now follows.

Its origins date back to the early twentieth century in response to perceived political and social injustices at grass-root levels, mainly in former colonial countries, according to Somekh and Lewin (2011, pp.94/5). It can be surmised that teachers were encountering such injustices in their daily lives and individually took emancipatory actions to address them. Subsequently, the teachers realised that their initiatives were effective and, in turn, could be adapted to help them in their personal self-improvement aims.

This links in with Bell (2005, p.8) who describes action research as an *approach* that can be applied when:

‘... specific knowledge is required for a specific problem in a specific situation or when a new approach is to be grafted on to an existing system’,

Cohen and Manion (1994, p.194).

It would be perfectly reasonable to see the *Sumdials*' approach to learning number as belonging to the second part of their definition and to classify it as action research.

Bell also asserts that action research is:

- Neither a method nor a technique
- Applied research.

It is difficult to reconcile her first point with the *plan-do-review* cycles described by Baumfield, Hall and Wall (2008 pp.4/5) or *planning-action-monitoring-reflection* cycles that can form “spirals” as successive cycles are completed, described by Waters-Adams (2006, p.5) are neither a method nor a technique.

The terminology of action research has still to become standardised and the terms that will now be adopted (in **bold**) with their equivalents are:

- **Teacher-researcher:** practitioner-initiated.
- **Sceptical Colleague:** critical friend or outside facilitator.

The self-improvement situations within which action research can be applied are virtually endless. To illustrate this, two different examples are now given starting with a colleague’s response to the cohort who could not understand his established verbal pedagogy:

- **Intention:** To develop a new pedagogy that:
 - (i) Modelled well the rules for positive and negative integers.
 - (ii) Used visual (seeing) and tactile (doing) methods.
 - (iii) Kept the use of verbal explanations/instructions to a minimum.
- **Process:** Trial-and-error methods to create a manipulative using safe material in response to the **Intention**.
- **Audience:** The cohort.

The same framework is now applied to this Study to illustrate that action research is not limited to self-improvement in classrooms/schools:

- **Intention:** Answer the research question:
 Does the *Sumdials*’ approach to learning number, based on the use of dedicated manipulatives (dials), produce statistically significant improvements in arithmetical automaticity?

- **Process:** Successfully complete a part-time distance learning PhD at the School of Education, Durham University.
- **Audience:** The Examiners as mediated by the Supervisor and the wider mathematics education community after completion.

The contrast between my colleague's circumstances, accountable to himself and completing a short intuitive cycle by himself in three weeks, and my study accountable to the University, working within the University and involving several schools in both Co. Durham and Edinburgh for seven years illustrates the scope of this broad nature of action research. Neither of us was addressing political or social injustices, though we were both addressing what can be perceived as cultural or societal challenges that confirm action research as a systematic approach to non-arithmetical problem solving

The role of a sceptical colleague is briefly mentioned because of its considerable potential value. My colleague, Peter, did not have one because he was combining the two roles in himself, as head of his department. It can be noted that he had completed a successful project without being aware that it was action research. However, being a sceptical colleague was *one* of several roles of a Supervisor; others included mentoring, based on greater subject knowledge and experience of schools, and was very helpful as was his acting as a sounding board.

7.1.2 Classification

The next step is to determine a classification for action research. In general, the aim of research is to advance knowledge and the scientific method (with its eight stages) is widely used to achieve this:

- Hypotheses, hunches and guesses.
- Experiment designed; samples taken; variables isolated.
- Correlations observed; patterns identified.
- Hypotheses formed to explain regularities.
- Explanations and predictions tested; falsifiability.
- Laws developed or disconfirmation (hypothesis rejected).

- Generalisations made.
- New theories.

Hitchcock and Hughes (1995, p.23)

New theories imply *proof* and this is usually achievable in the natural sciences, including physics. Since all research is not amenable to such rigorous methodologies, research can be classified as either:

- Pure
- Applied.

Their distinguishing features can be expressed as:

Pure		Applied
Proof	v	Sufficiency of evidence
(Physics)		(Engineering)

A good example of pure research is the confirmation by CERN of the theoretical equations that postulated the existence of the Higgs' boson. It can now be replicated anywhere – a key feature of the scientific method - by any properly qualified team with a suitable laboratory.

In contrast, much engineering design is based on “what works” sufficiency of evidence that has stood the test of time. For example, most applications in civil engineering and structural projects are unique and cannot be generalised as theorems. London's Southbank Pedestrian Bridge is a good example in that its design is structurally sound, but it had not been checked for wind resonance problems before it was built. The problem was easily overcome using further tried-and-tested practices.

Another example is the rule used by builders to calculate the depths of beam required to span voids. It is: one inch for each two feet of span plus two inches. Thus, for a 12' span the depth of beam would be:

$$12/2 \times 1'' + 2'' = 8''.$$

This worked well even if it led to over-specified beams. Relevantly, it illustrates a good practical application of arithmetic – the point of this comment will also become clear later.

Both examples used “formulae” – often little more than rules-of-thumb - derived from accumulated evidence and their value is enhanced when different researchers produce similar results and so increase confidence in them. In time, sufficient evidence is accumulated to establish such outcomes as scientific facts, according to Malofeeva (2009, p.34). This explanation confirms Bell’s classification (above) that action research is applied research.

7.1.3 Interim Conclusion

The influence of action research on this Study is considered (p.172), but the closing reminder is that most of the Study had taken place while seemingly applying action research without being aware that this was the case.

7.2 Research and Action Research

7.2 Introduction

The possible contribution of action research to this Study is now considered even though the Study was well advanced before there was any awareness of action research; it is considered within the following Sections:

- Planning with action research.
- The contributions of action research.
- The limitations of the Study seen through action research.
- Educational inquiry supported by action research.
- Influences of action research on personal learning.
- A personal reflection on action research

The usual action research framework will be applied with its three stages of:

- Intention,
- Process,
- Audience.

They are applied to the relevant research experiences to illustrate how action research might have been applied to this research.

Each Section concludes with personal reflection, when appropriate.

7.3 Planning with Action Research

7.3 Introduction

The outlines of research plans can be relatively straightforward to prepare in that they are the first steps of converting ideas into actions and are usually drawn-up in isolation. Once prepared, the implementation of any plan is likely to encounter unanticipated developments and the ‘spirals’ associated with action research with their three stages of:

- Intent.
- Process.
- Audience.

Using such stages should ensure that a systematic structure is followed and helps to retain stability as the accommodating changes are made. In the case of this Study, this would contribute to the impression that its implementation was under control. This would help to retain the confidence of the participating schools involved in such a study for the first time as any sense of panic was avoided.

7.3.1 Background

The *Sumdials*' approach to learning number has two components, as was described. (p.32):

- Dedicated manipulatives or 'dials' that model well the basic arithmetic processes for the following progression:
 - (i) Adding and subtracting up to 10.
 - (ii) Adding and subtracting up to 20.
 - (iii) Multiplying and dividing up to 50.
 - (iv) Multiplying and dividing up to 100.
- Hardwiring the associated number facts into pupils' long-term memories.

It is now believed that the main role of the dials is to help pupils to develop robust internal models of these processes, whereas the other discrete role is the properly delivered memory work is **automaticity** - the *instant and accurate recall without any conscious mental activity* of the previously memorised number facts.

Again, the three action research stages are now applied:

- **Intent:** To conduct a typical RCT involving 16 primary schools.
- **Process:** The original plan was to conduct an RCT empirical study involving 16 schools to arrive at eight pairs with four each in Co. Durham and Edinburgh. The InCAS computer adaptive program of CEM would be used to collect data from the pupils for analysis using SPSS 19 to determine whether or not they were statistically significant (a minimum level of 95%), supported by effect sizes of medium or greater.
- **Audience:** The Examiners as mediated by the Supervisor and the wider mathematics education community after completion.

The two previous Studies and both the Pilot and Empirical Studies may have used implicit *plan-do-review* cycles similar to those of action research. Their Intents and Audiences remained the same throughout the current Study, but the Processes were

changing in response to the developments that were encountered within the participating schools. The actual cycles were:

- Cycle 1: Bramald's Study, being a pre-study cycle.
- Cycle 2: The researcher's follow-on Study of Cycle 1.
- Cycle 3a: Initial assessments of the Pilot Study pupils (in Edinburgh).
- 3b: End of session assessments of Pilot Study treatment pupils.
- Cycle 4a: Initial assessments of the Empirical Study pupils (in both Co. Durham and Edinburgh).
- Cycle 4b: End of session assessments of Empirical Study pupils together with the Pilot Study pupils.
- Cycle 4c: End of session assessments of all participating pupils.
- Cycle 5a: Initial Follow-on Study assessments of pupils in their first year of primary school at the same seven Empirical Study schools.
- Cycle 5b: End of session assessments of the Follow-on Study pupils.
- Cycle 5c: Planned next end of session assessments of the Follow-on Study pupils.

The analogy of a funnel (for transferring liquids from large to smaller containers) explains well the basic refining processes that were behind the developments that took place as the Studies progressed. The common feature may have been the repeated application of *plan-do-review* procedures starting with Cycle 3a and for each subsequent cycle. Most of these were initiated by the researcher with the aim of improving further the *Sumdials'* approach to learning number that emphasises the importance of the two separate processes of:

- Developing robust internal models of the basic arithmetic processes.
- Hardwiring number facts into long-term memories.

Comments are now made on these cycles.

Cycle 1

Bramald's Study in 2001 was the first cycle to collect independent data in the belief it would support the conviction that the *Sumdial's* approach and its

resources were sound. It was hoped this would be confirmed by statistically significant results that, in turn, would encourage teachers to adopt it. In the event, his conclusions were less clear-cut than had been hoped, but the teacher feedback was constructive and applied to improve the approach.

Cycle 2

This took place five years later to identify whether or not there was any measurable evidence of enduring improvements in number attainments of the treatment pupils as they were completing their primary education. Their improvements were highly significant, but teachers still did not take-up the approach, (p.35). However, one surprising outcome was that this Cycle indirectly led to this Study.

Cycle 3a

The original plan was to do an RCT Pilot Study in Edinburgh with eight schools making four matched pairs. Only seven schools were willing to take part with four of them wanting to provide treatment classes while the remaining three were happy to be controls. All first year cohorts were assessed (N = 200) to provide a Study baseline.

Cycle 3b

It was only possible to reassess the treatment pupils in the time available before the end of the year. The manual statistical analyses of the results suggested that the treatment pupils' gains justified proceeding with the Empirical Study at the start of the new session.

Cycle 4a

The Empirical Study started by assessing 12 new cohorts at the original seven Edinburgh schools together with cohorts from five schools in Co. Durham.

Cycle 4b

The end-of-session assessments were carried out on all the remaining treatment schools – by then, most of the control schools had fallen by the wayside, including those from the Pilot study.

Cycle 4c

The assessments at the end of the next session marked the end of the data collection phase of the Empirical Study. Only seven schools were now involved, being four in Co. Durham and three in Edinburgh. The results showed that all was not well in learning number and this applied especially to mental arithmetic.

In light of this it was decided to do a Follow-up Study with the same Edinburgh schools and three of the Co. Durham ones (N = 190). The plan was that they would follow a simplified number curriculum with the aim of establishing secure number bases that included all the participating pupils acquiring automaticity.

Cycle 5a

The beginning of session assessments were carried out for the planned six new classes, but one of the schools was going to be involved with a major re-organisation and one of the original control schools took its place at very short notice.

Cycle 5b

The assessments at the end of session were carried out at the six participating schools and the results indicated *tokenism*. By this is meant continuing with their usual curricula while making token attempts to accommodate the requirements of the Follow-on Study.

Cycle 5c

The next assessments were planned for the end of session and the hope was that the results would show that commitment had replaced tokenism. To achieve this would include much closer monitoring to ensure this happened.

- **Process:** The point of these summaries is to show that the underpinning structure can be understood as that of action research to provide a consistent

framework that facilitates the comparison of results from each cycle. Maintaining the same structure of each case (pupil's record) throughout allowed them all to be part of the same dataset (N = 545). The same structure of records would also be used in the Follow-on Study and that will allow the performances of the two Studies to be compared. This should indicate whether or not commitment has replaced tokenism.

- **Audience:** The Examiners as mediated by the Supervisor and the wider mathematics education community after completion.

7.3.2 Conclusion

The striking feature of the Empirical Study in particular was the on-going and unplanned changes that took place, mostly attributable to the general messiness of primary schools. The framework is that an implicit action research approach could have contributed to bringing the Study to a clear conclusion and led to the Follow-on Study.

7.3.3 Reflections

However, it cannot be claimed solely that action research, *per se*, brought about such an outcome. It could also be attributed to the persistence of an unchanged research team with their understanding of the InCAS assessment program while working with very supportive treatment schools.

The possible contribution of action research to the Study will now be considered and it is concluded that it has been appreciable in unexpected ways that are now discussed.

7.4 Contribution of Action Research

7.4.1 Background

One aspect of undertaking a part-time distance-learning doctoral study is that neither the supervisor nor students come to know one another well; this is exacerbated when a student has only limited academic experience. The upshot is that their respective expectations may not be well matched and there is only limited opportunity to correct them since the optimum tutorial frequency is generally considered to be one hour per

month. This point resonates with Watson's concept of affordances, constraints and attunements Watson, (2004). Moreover, the limited availability of time for the informal interactions that allow them to crystallise is not available to absorb the academic ways of exchange, (p.188).

Several factors contributed to this and they include:

- Busy diaries making meeting difficult to arrange.
- No pressing deadlines, particularly during the first five years of a nominal six years course.
- Allowing the student to solve unexpected problems, as is appropriate for doctoral courses.

As explained earlier, a relevant influence in this Study was I have been 'my own boss' for over 35 years (running small family businesses) and made decisions to deal with the unexpected *without referring to anyone*. In short, I made things 'happen'. Thus, when I found it was not going to be possible to answer my research question (because mental arithmetic was no longer emphasised in the curriculum and number facts were not being memorised), it was a natural and spontaneous response to try to discover why this had happened.

Now, the crucial point was I continued with my Empirical Study signalling that everything was proceeding apparently to plan as interim results were discussed. There was no cover-up on my part hoping "it will be alright on the night". However, it did not worry me once it became clear that my research question could not be answered in the way that I intended, because this seemed not to be an unusual outcome for PhD studies.

7.4.2 Action Research

Action research is generally associated with teacher-researcher (practitioner-initiated) self-improvement programmes such as seeking better personal pedagogies in specific areas, according to Baumfield, Hall and Wall (2008, p.4)). To achieve this, *plan-do-review* cycles are implemented with its structure of:

- Intent (improve pedagogy).
- Process (find new resources/methods that would improve the pedagogy).
- Audience (classes and, indirectly, pupils' parents and head teachers).

It is good practice to enlist individually the support of a 'sceptical colleague', typically a more experienced colleague or an outsider, such as a university researcher or retired teacher, to act as a mentor or sounding board, as discussed below.

Now, narrowly defined, I did not meet any of these specific educational criteria in that:

- **Intent:** The intent of the Study was to advance knowledge (in learning number) and there was no self-improvement dimension in that the aim was to gather good data, as would be expected of research, and there was no possibility of enhancing career prospects, in my case, through a successful completion.
- **Process:** School involvements were in two parts of:
 - (i) Assessing pupils number performances through analysing the collected data.
 - (ii) Liaising with the schools and their teachers.
 Success in these would provide schools with independent and diagnostic results. The data would allow the research itself to progress.
- **Audience:** The Examiners as mediated by the Supervisor and the wider mathematics education community after completion.

However, there is another way of looking at the research process and it focuses on the contribution of the sceptical colleague.

7.4.3 Sceptical Colleague

The traditional roles of the sceptical colleague are essentially acting as:

- A monitor keeping a watchful eye on the progress of a self-improvement project.
- A mentor making suggestions, based on either experience or theory, when appropriate.
- A sounding board for the researcher to test out new ideas or tentative conclusions.
- An encourager responding to setbacks or unexpected difficulties who has already been “round the block” several times and probable had experienced most of them in one form or another.

The reality in this Study is that the participating teachers had become quasi-teacher-researchers while the research team (Peter and me) was the real initiator, but both parties benefitted because:

- The pedagogies of the participating teachers improved:
- The opportunities to observe pupils actually doing their sums while answering their InCAS questions proved to be very informative because their teachers were not present and could not influence. This was probably the greatest influence on the direction of the research because the pupils were observed as they answered the questions *their* “doing it their ways”. (The relevance of their questions is explained under the explanation of the InCAS program, (p.130).
- Data were obtained that allowed the statistical analyses to be made and conclusions to be drawn:
- A genuine two-way channel was established between teachers and researchers that improved the quality of teachers’ pedagogies and the authenticity of the research.

However, this need not be the end of a link with action research in that it can be argued that my Supervisor was also the *de facto* sceptical colleague fulfilling the same four roles above over a greatly extended elapsed time than would be usual.

7.4.4 Conclusion

It was not intended to use action research in this Study, but that something very similar to it took place contributed to improvements in the participating schools while also enhancing the authenticity of the research suggests the genuine two-way street was a bonus!

7.4.5 Reflections

The outcomes suggest that the plan-do-review structure provides a good framework for self-improvement and practical improvement – that have been well tried-and-tested in engineering environments.

Thus, it could be claimed that action research is a meeting point between the more philosophical or interactive approaches of teaching and the impersonal problem-solving instincts naturally adopted by engineers.

7.5 Study Limitations

7.5 Introduction

It is likely that most research studies have their limitations and this one is no exception. Its limitations are now considered through an action research perspective and it is concluded that some of them could have been mitigated if there had been a suitable experienced sceptical colleague. However, it is believed that good has come out of the Study in spite of the way in which it was conducted and there is the prospect of more benefit to come.

7.5.1 Background

It could be said that this Study just “happened” in the sense that it was not a staging post on a well-planned academic career. It was more an outcome of several coincidences that include:

- Peter asking me to help in the development of his *Sumdials*' approach to learning number. This coincided with reaching normal retiring age and my son

taking over the day-to-day running of the family packaging business. I accepted because I believed:

- (i) I could contribute to the development of his system and address some of the issues associated with the long-term decline in number skills.
 - (ii) It was in my “DNA” to become involved in such a project – and still is!
- My acceptance seemingly being justified by the approach’s resources such as the production of the dials, teaching plans and sales/marketing activities becoming better established.
 - Bramald’s Study being commissioned in 2000 to carry out his short intervention in the hope it would help sales, but with hindsight this did not happen because it was not realised that the *real* issue was a reluctance to change pedagogies (p.67).
 - My follow-on/longitudinal Study five years later showing that his treatment pupils had achieved highly significant *enduring* benefits from Bramald’s six weeks intervention (p.31). This was unexpected in that such interventions normally “wash-out” within two or three years.
 - However, my report only interested his former colleague who was to become my Supervisor on taking-up his Chair at the University.

He suggested that I could do a part-time distance-learning PhD to start after he had moved. This was very surprising and I accepted without giving it any serious thought, believing I would be able to cope.

Again, my background was an unusual preparation in that my main experience of education was as a customer (when a pupil) while my first degree was in electrical engineering, but my Master’s in Business Administration (by dissertation) involved some exposure to university practices. Nevertheless, completing a doctoral degree did not seem to be out of the question other than that my perception was my written ‘academic’ English would not be good enough.

However, one retrospective conclusion after I had started is I would have benefitted from closer supervision that included my being confronted more firmly, as befits a loose cannon. It can readily be argued that I allow myself to become side-tracked and such changes of direction would have been less likely if I had seen my Supervisor more frequently as a ‘trusted colleague’, according to Baumfield, Hall and Wall (2008, p.68). Alternatively, it could be held that I needed to develop my self-monitoring skills as part of the process of acquiring independence of mind as a researcher.

The limitations of the Study are considered now that part of its background history has been re-summarised. This will be carried out using an action research approach.

7.5.2 Study Limitations

The headings for consideration are:

- The main question itself.
- The structures of the Pilot and Empirical Studies.
- The InCAS mental arithmetic module.
- Teacher capabilities.

7.5.2.1 The Main Question

The flaw in the main question, as has been discussed more fully (p.124), is that the repetitive use of the dials only contributes to developing robust internal models of the basic arithmetic processes and *their contribution to acquiring automaticity are likely to be limited*. This was not appreciated (by me) at the start of the research and even if it had been it is unclear how the question would have been reworded because of the strong perceptual link between the *Sumdials*’ approach to learning number - based on its use of dials – and developing automaticity.

With the benefit of hindsight, it is very unlikely that any schools would have signed-up to research into the effectiveness of good memory work in that the majority of them seemed to be unconvinced of its importance and unwilling to make sufficient time available to achieve it.

The question now becomes:

- Would a conscious action research approach have avoided this problem?

Its three stages are now used to answer this:

- **Intent:** Consideration of the research question itself was needed at the beginning, but not even Peter would have been able to query its soundness if he had been consulted as a sceptical colleague. He believed the repetitions associated with using the dials contributed to memorising number facts, as also did Bramald. However, that was essentially the role my Supervisor was discharging and he could only have spotted the issue if he had had the relevant detailed knowledge of the *Sumdials*' approach and the InCAS program.
- **Process:** In all likelihood, it would have been confirmed as being appropriate when the RCT was planned.
- **Audience:** The Examiners as mediated by the Supervisor and the wider mathematics education community after completion.

In point of fact, my initial three years of the Study were devoted to the Literature Review, including the uses of manipulatives in learning number, and consideration of the Methodology to be used. Hence, the research question was finalised without any detailed questioning of its suitability immediately prior to the Pilot Study and at that point urgency of a new term starting acted as focus on starting the Pilot Study before the 'window' in that academic year closed.

On reflection, it is now clear the whole approach to my research was perhaps too flexible and responsive to schools' needs and pressures in terms of the use of the *Sumdials*, as has already been acknowledged and is consistent with the way I have operated over most of my life in adapting to contingencies and demands!

7.5.2.2 The Study Structures

The structure of both the Pilot and Empirical Studies illustrated well the dictum: a little knowledge is dangerous. It was known that a minimum number of schools would be needed if the Study were to have sufficient statistical power. I based my

assumptions on a calculation formula that produced the answer of 14 schools. Thus 16 was chosen because that would:

- Provide eight treatment/control pairs, four in Co. Durham and four in Edinburgh:
- Preserve the statistical power if up to two schools dropped out.

In the event, it was realised during the Pilot Study that 16 schools would require far greater resources than Peter and I could provide. In fact, replacing the control schools by the participating pupils' Ages-at-Test (A/Ts) as the control produced more robust samples even though it was an expedient choice. This resulted in the Study changing from being a typical RCT to a within-subject study that made it much more manageable.

Again, the three action research stages are now applied:

- **Intent:** To conduct a typical RCT involving 16 primary schools.
- **Process:** A sceptical colleague would have queried the need to have 16 schools for a student study and might have suggested using pupils' A/Ts to provide a better control in place of depending on messy primary schools. This change would have been entirely appropriate for a within-subject study and have led to a greatly reduced logistical load. It would have been accepted as a very sensible change (by me).
- **Audience:** The Examiners as mediated by the Supervisor and the wider mathematics education community after completion.

A sceptical colleague with a detailed knowledge of the *Sumdial's* approach would have identified the key issue of over-commitment and pointed out that a smaller sample would have been appropriate for a student study.

7.5.2.3 Mental Arithmetic Module of InCAS

The fundamental problem for this Study lay in the time allowed to answer each Mental Arithmetic question during the assessments. It was found to be 30 seconds,

when much shorter times would have been more appropriate, as is explained again below. However, the InCAS approach, even with this major limitation was accepted because:

- It would be simple to administer and, importantly, has a very large database.
- The schools would prefer the pupils' scores for both modules to be available at the same time and in the same formats.
- It was assumed the actual times taken would be readily accessed and more appropriate cut-off times could be applied post-assessment.

This last point is now considered more fully, in view of its importance. It may be recalled that two types of mental arithmetic question are asked (with suggested answering times in brackets):

- Number-fact recalls (five seconds to see/hear the question, select the correct answer, being one of four options displayed on the laptop screens, and manipulate the keypad to select it).
- Mental calculation questions (ten seconds to calculate the answer the single step required to arrive at the correct answer and then select it, as before).

It was assumed, on the basis of the demonstrated ease of accessing the times taken to answer individual questions that they would be readily available with the results. It turned out that this was not the case and the actual times would need to be extracted manually. This would have been an immense and time-consuming task that, in the event, would have added little to the results because it was concluded:

- True mental arithmetic, based on automatic *recall* of previously learned number facts, was no longer being emphasised.
- The concept of automaticity was virtually unknown in the primary schools.

Again, the three action research stages are now applied:

- **Intent:** To assess the participating pupils' automaticity.
- **Process:** A sceptical colleague with the requisite knowledge and expertise to highlight this specific limitation of InCAS.
- **Audience:** The Examiners as mediated by the Supervisor and the wider mathematics education community after completion.

In practice, teachers would normally be best placed to assess their pupils' automaticity, provided they were convinced of its importance. This was the case with one of the Swiss teacher who used a one-minute glass to time 20 number-fact questions (p.155). A variation on this is using pre-recorded question tests and answer sheets allowing, say, five seconds per question (to provide time for the questions to be asked and then to write the answers).

In fact, this apparent set-back became helpful in identifying a crucial explanation for the declining number skills and in pointing the way forward for further research once the most likely major reason for the decline had been confirmed – lack of automaticity and a lack of a general appreciation of its importance in the context of arithmetic as the *foundational* subject of mathematics..

On reflection, having made that point, it is helpful to re-emphasise that the greatest benefit in the whole research came through observing how the pupils answered their InCAS questions without their teachers being present. This was unexpected, but hugely helpful to this Study.

7.5.2.4 Teacher Capabilities

The *unspoken* assumption of this Study was that the 'seeing and doing' (constrained discovery) approach would make a greater contribution towards becoming fluent with number than the effectiveness of the participating teachers. Thus, there would be no need to assess the teachers' abilities. However, with hindsight that position is not tenable, even though it may be widely held. The reality is that teachers should be developing their pupils' responsibility for learning and their ability to discharge this must be a function of their own number knowledge, according to Seeger, Voigt and Waschescio (1998, p.15-16).

Again, the three action research stages are now applied:

- **Intent:** To assess the participating pupils' automaticity without assessing the abilities of the participating teachers.
- **Process:** Again, it would have been considered if a sceptical colleague had pointed out the limitation this would be imposing on the value of the Study.
- **Audience:** The Examiners as mediated by the Supervisor and the wider mathematics education community after completion.

Thus, the explicit avoidance of *any* teacher assessment during the Pilot and Empirical Studies limited its overall value, (p.146). However, primary teachers' general number abilities are considered more fully, while noting now that they may have influenced the pupils' scores differentially, (p.211). It can be added that the majority of the teachers seemed to be part of the 70-80% middle category and would have similar abilities (p.119).

7.5.3 Conclusion

The first two limitations of this Study may have been avoided with greater foresight in terms of understanding the demands and complexities of schools or with greater resources to manage a large-scale project.

On reflection, my conviction has grown that the Study was worthwhile in that it provided reliable results at the micro level while adding a very helpful perspective on the overall learning of number at the macro level. This has led to the recommendation that arithmetic and automaticity are accorded their former places of importance in the curricula as soon as it is practicable.

Much has been learned during the seven years of the Study and the aspects that probably had the greatest impact on personal learning are now highlighted within a context of action research. The hope is that this provides some indications of the potential benefits of action research.

7.6 Action Research and Educational Inquiry

7.6 Introduction

Number attainments for school leavers continue to decline (SSLN, 2014) and there are many causes. One such cause is many primary school teachers had become convinced that ‘maths-is-a-hard-subject’ while they were still at school. This mind-set needs to be changed and it is likely that an action research approach would be a very effective first move.

However, this should be seen as a holding operation while better-qualified students are being recruited to become primary teachers. Fluent automaticity - instant *recall* of number facts – would be an essential requirement for selection. At the same time, educational policy must ensure that arithmetic is accepted as being *at least* as important as reading and writing.

7.6.1 Background

It is recommended that arithmetic be established as a subject in its own right since it and geometry were the original foundation subjects for all mathematics since the ancient Greek times. Unfortunately, the idea of ‘maths-is-a-hard-subject’ and cultural attitudes towards mathematics have become deeply embedded in the national epistemology with the result that the majority of primary school teachers, along with the parents, have become unwitting transmitters of this belief to pupils. It is now asserted that all of the developed world’s cultures are word-based, but a number of Asian cultures place greater emphasis on the role of number: one consequence is the more mathematically confident cultures are gaining a competitive advantage in an increasingly technological age.

7.6.2 Action Research

Action research has been adopted in schools as a practitioner (teacher)-initiated self-improvement approach that should contribute to the improvement of number skills generally. As a reminder, Bell (2005, p.8) describes action research as an *approach* that can be applied when ‘specific knowledge is required for a specific problem in a specific situation’, from Cohen and Marshal (1994, p.194). It would seem to be tailor-made for the need to raise number attainments when:

- Specific knowledge is required for:
 - (i) A specific problem – improving teacher pedagogies.
 - (ii) In a specific situation – the need to improve number skills generally.
- A new approach is to be grafted on to an existing system – the current failing number learning system.

It is recommended that teachers are provided with appropriate scaffolding to help them to improve their own pedagogies and, in turn, their pupils' number skills (p.235). An outline of how this might be achieved by applying an action research learning approach is now given under its main headings.

7.6.2.1. Intent: The teachers' intents must be along the lines of:

Their pupils will become better with number than they were themselves at the same stage/age.

This addresses the declining emphasis in confidence and competence with number that needs to be broken. This means the teachers themselves need to take the first step by committing themselves to their *own* improvements: it must be something they *want* to achieve because they have become convinced it is essential.

Action research would be a very appropriate approach to deliver the required improvements and many teachers could already have discovered its value and there should be relevant experience available in schools on the effectiveness of action research. Based on this point, the next supportive step is to enlist a suitable sceptical colleague with relevant experience to act as a mentor and an encourager. Ideally, it would be the school's numeracy co-ordinator, provided her number skills were sufficient, a better (with number) colleague, an external person such as a university researcher or a retired teacher. My own

experience confirmed this principle through the help that Peter has given me throughout this Study, as both a technical adviser and a sounding board.

7.6.2.2 Process: The majority of teachers embarking on such a course is unlikely to know *precisely* their number strengths and weaknesses; identifying them at the beginning will be beneficial (Peter has written specific texts that could be used to achieve this). The point is each teacher is likely to have different needs and action research, being individually driven, would accommodate this applying the personalised *plan-do-review* cycles as her ‘treatment’, following her diagnosis. (Suitable self-improvement modules are also available for this purpose.)

7.6.2.3 Audience: Individual teachers.

It must be believed that improving all pupils’ number skills will become an accepted key priority of all schools - even though this cannot be assumed in a word-based culture.

7.6.3 Comment

Such a programme may be very ambitious, but:

- Desperate situations require desperate actions – while avoiding confusing activity with action.
- It is a bootstrap move to breakout of the current decline in number skills using existing resources while better ones are prepared.
- Some of the Follow-on Study schools will be approached to pilot it.

If there is no response, the fall-back position is to encourage them to follow a proper automaticity programme using the Swiss roll approach for memorising approach together with the former practices of daily short sessions first thing every morning that include activities such as “10 a day”. It is believed that acquiring good automaticity will transform teachers’ confidence in their own number abilities.

This is proposed as a holding operation until proper remedial actions are in place. It has been suggested that arithmetic has always been the poor relative of the 3 Rs in this country and the consequential conviction that maths-is-a-hard subject should not come as a surprise. One pragmatic but effective practice in primary schools is “to look the other way” when a “good” maths teacher takes her colleagues’ number lessons while they take her English lessons!

However, the long-term need now is to recruit better-qualified arithmetic (not maths) teachers immediately to be used in schools as specialists (as in the above practice). At the same time, the entry qualifications for new entrants need to be raised *permanently* so that only those with demonstrable automaticity are admitted. It is essential to establish, as a matter of policy, that arithmetic is accepted as being *at least* as important as reading and writing (p.183). Then, and only then, will the national competitiveness be regained.

7.6.4 Conclusion

It can never be known how this Study would have progressed if I had applied an action research approach from the outset. However, this consideration of its benefits convinced me that an action research approach would contribute in disseminating my findings.

It is appropriate to include some personal reflection on action research, bearing in mind that I only became aware of it towards the end of the Study. These are based on personal metacognition.

7.7 Action Research and Personal Learning

7.7.1 Action Research

Overall, the action research framework with its three stages will again be applied:

- Intention,
- Process,
- Audience.

These are now considered.

- **7.7.1.1 Intent:** The overall intent remained, under the guidance of Supervision (also in part acting as a sceptical colleague), to carry out and complete a successful research study culminating in satisfying educational research criteria that the *Sumdials*' approach to learning number is effective based on:
 - Use of its dials:
 - Acquiring true automaticity (of number facts).

- **7.7.1.2 Process:** The process would be divided into the following activities:

Carrying out a relevant literature review that included:

- (i) A general overview of the main elements of learning number.
- (ii) Specific observations about the contributions made by manipulatives in learning number.
- (iii) Writing sections, as appropriate, to demonstrate learning and for inclusion in the thesis in whole or part.

Designing the Study methodology and then to:

- (i) Select primary schools in Co. Durham and Edinburgh that were willing to take part in the research.
- (ii) Collect and analyse data and then present the Results.

Drafting and revising a thesis incorporating the relevant sections described above.

Examination.

7.7.1.3 Audience: The Examiners as mediated by the Supervisor.

Some specific learning outcomes of this process are now considered.

7.7.2 Learning Curves

The majority of my everyday applications of number were learnt long before the digital age. Thus I have been confronted by a series of steep learning curves that include:

- 'Office' and specifically: Entourage, Excel and Word.

- InCAS, the computer adaptive program for assessing the pupils.
- SPSS 19 for the statistical analyses of the data.

Perseverance has been rewarded, albeit with an unwelcome expenditure of nervous and emotional energy, but even that would not have been sufficient in the case of SPSS 19; it was the support from the University almost beyond the call of duty that saved the day.

I still have to learn how to make searches for journals and academic papers and my experience has a direct read-across to pupils learning number. In my case, I had to start before I was “digital ready” and from that moment searches-are-difficult became part of my personal epistemology and, like pupils with number, I have developed “strategies” to conceal this. It will be interesting (to me) to discover if the wounds can be healed.

On the basis of my own experiences, the isolation associated with being a distance-learning student may not be adequately appreciated and it arises through there being no one ‘along the corridor’ to seek advice or generally to have a chat (p.174). In making this point, I admit to being a fully paid-up member of the “When all else fails, try reading the instruction manual” society and when I repented of my ways, I struggled to make sense - back to Resnick and Schoenfeld – of what was written.

This experience points strongly to the benefits from sceptical colleagues being *local* impracticable though it would be in the majority of cases.

7.7.3 A Practical Example

The stage for dissemination of my findings had almost been reached, the importance of which is rightly emphasised by Baumfield, Hall and Wall (2008, p.122). An action research approach was adopted to ensure that the findings (assuming they stood up to critical examination) would be effectively disseminated. The first step was to develop my dissemination skills and to achieve this is now illustrated using the action research stages:

- **Intent:** To ‘make the transition between a thesis and a publication’ by attending an appropriate workshop.
- **Process:** To be trained in the relevant skills to enable me to:
 - (i) Identify the appropriate publishers to disseminate my findings.
 - (ii) Present my findings in ways that will appeal to them.
- **Audience:** To be identified through the Process, as above.

This would be the first *plan-do-review* cycle and it anticipated that many more such action research cycles would follow. In the event, it was discovered that direct approaches to education authorities with suitably tailored presentations were likely to be more effective than indirect approaches through journal articles.

7.7.4 Conclusion

What has been learned must be personal, but it has been written to help others by showing that an action research context should be helpful in research. In particular, it is now accepted that the role of the sceptical colleague in providing help and encouragement had not been fully appreciated.

However, on reflection the fact is that a successful outcome for doctoral research courses depends primarily on students’ efforts under the guidance of their supervision. The right action research support structure should contribute to such efforts being better focused.

7.8 Reflection on Action Research

7.8.1 Background

My career must be the main source for this reflection on a comparison between schools, especially primary schools, and business generally of the applicability of action research. To me, the prevailing school ethos and sub-epistemology can be traced back to the appointment of the Newcastle Commission in 1858 that led to a conservative teaching tradition being established and maintained – as befits Brown’s

observation about ‘the combined good sense and inertia of the teaching profession’, Thompson (1999, p.15).

In contrast, towards the end of my electrical engineering studies, our professor (the only one in the electrical engineering department!) advised the final year class to keep informed on the electrical properties of silicon dioxide (that would lead to the transistor). I suspect he would have been very surprised by how prescient his advice was to become with all its truly transformational developments! One consequence in particular was the very rapid and continuing changes in business, being most relevant to my own experience. Much of it, including my assignments in management consultancy, has been with “family” businesses rather than with big organisations and the relevance of this is now described.

7.8.2 Classifications

Two labels are now used to classify schools and small businesses and they are:

- Spread sheets:
- ‘Seat-of-the-pants’ as colloquially applied to customer/business responsiveness, a term of both denigration and sneaking admiration!

Simply, ‘spread sheet’ organisations are managed-by-procedure with little discretion being allowed and limited entrepreneurialism evident. Usually, their management systems are reliable and everyone is ‘playing it by the book so that such organisations’ banks have little anxiety and, in turn, become disappointed when they cannot persuade them to take out bigger loans. Typically, they will have staffs of over 200 people and have a national customer base. In contrast, seat-of-the-pant companies are run by a founder boss or are family businesses with ‘hands-on’ flexible practices relying on gut-feel or hunches; their aims are to keep going by meeting their customers’ needs and with whom they usually have long-standing personal relationships. This tends to take priority over improving their financial returns and they typically have staffs of less than 25 while serving a local market.

It will have been obvious, based on the way I have carried out my study that I am a typical small businessman driven by reaction and/or reflection. Reaction often involved helping customers who have run-out of stock and urgently need replacement packaging and, indeed, there is much satisfaction from helping them. However, reflection is the key “24/7” activity and consists of a never-ending mulling about mainly trivial incidents or remarks that can lead to anticipatory action. For example, another machine was ordered once it was realised it would be needed to cope with the likely increase in demand for presentational packaging once Edinburgh became a capital city following devolution. It can be added that only a back-of-envelope calculation *might* have been made before ordering the machine and a spreadsheet would not have existed to be consulted.

7.8.3 Comparisons with Action Research

It will be evident from this brief analysis that the business development process is *continuous*, being driven by reaction and reflection: it is essentially forward-looking while reviews are limited. In contrast, the teacher-researcher self-improvement approach that Peter used when devising his first dials would be unusual in business. Developing a better method usually involves several members of staff working together as they shared in the overall – and unspoken - aim of improving their business.

In retrospect, I was applying my business experience to my study as, for example, when I concluded that collecting data from control schools was likely to be problematical. My response was a problem-solving one similar to that used by the faculty member who considered and rejected several possibilities quite quickly to arrive at a good solution (Grouws p.356)). In much the same way, I decided, without consulting my Supervisor, that pupils’ Ages-at-Test would make as least as good a control as the originally planned control schools.

It can be concluded that action research is:

- Appropriate for *individual* teacher-researcher initiated self-improvement in the nominally structured and traditional environments of schools.

- Less appropriate in the small business flexible environments where *collective* actions lead to spontaneous improvements often in response to unpredicted external events.

To return to this Study, it would be natural for my Supervisor, with his teaching and academic career, to assume the changes that took place during the Empirical Study, in particular, were the outcomes of action research. However, it should now be apparent that this was unlikely to be the case for someone who had spent most of his working life in the reactive problem-solving world of small businesses.

7.8.4 Conclusion

Would it have helped me if I had been aware of action research at the beginning of my Study? Possibly, but I could have fallen between two stools in that the more structured application of action research would have curtailed my reactive spontaneity, while leading to a more rigorous study (cf. Kilpatrick as cited by Schoenfeld, p.347). By doing so, the opportunity to include the Swiss schools would not have been taken and this would have impoverished this Study and the breadth of its investigations.

In conclusion, action research has much to offer in the right circumstances and these are more likely to include primary schools than last minute or unsystematic enterprises.

8.1 Reflections on the Literature Review

8.1 Introduction

The literature review (Chapter 2) is in two sections. To recap, they are based on:

- A synoptic or wide-ranging review focused on pupils' learning of mathematics and arithmetic in particular appropriate for answering the research questions and providing a general background for the issues under investigation. It covered the main issues associated with learning number and the possible benefits of using manipulatives.
- An additional focused review when it was thought the main question could not be answered and is relevance to current practices associated with the teaching and learning of number that needed specific consideration having been prompted by experiences in schools.

These reflections are now made on the initial review that provided a good background for this Study with its introduction to the benefits of automaticity. However, it was discovered during the Pilot and Empirical Studies that automaticity was no longer being explicitly taught. Possible causes of this are considered in this Part. In spite of this, some very relevant points were highlighted and are the focus of these reflections.

8.1.1 Background

The approach now used to reflect on the literature (in the first Part) is to compare my initial responses with my current conclusions. My first Annual Review Essay, written after six months of study, summarised my initial conclusions:

- The priority then being given to rigour in the widespread use of *quantitative* statistical “treatment A versus treatment B” comparison studies predominated in the scientific study of thinking, learning, and problem-solving had led to increasing frustration in the United States when compared with the Russian *qualitative* approach, according to Kilpatrick (1978). The consequence in the search for experimental rigour, researchers had lost touch with truly mathematical behaviour, Grouws (1992, p.347). I persisted with a primarily

quantitative study in spite of this clear warning, for reasons that will be explained even though it widened into a qualitative study in response to the findings and my reflections on these, (p.74).

- I had not grappled with the epistemology as I thought it was the sort of work that more abstract cognitive/developmental researchers were involved in and that it had little *practical* to offer in terms of taking action to solve the immediate need to improve number skills. Nevertheless, the following quotation ‘registered’:

“Understanding is both developmentally and culturally bound. What a person understands and how he or she understands is not independent from his or her development stage, from the language in which he or she communicates, from the culture into which he or she has been socialized”. (Sierpiska, 1994. P.138).

This ‘made sense’ and was, indeed, ‘an aha’ moment because it explained so much of what had been experienced during school visits during this Study and during earlier workshops.

- This ‘mind-opening’ process was continued by the following citations:

“For Pólya, mathematical epistemology and mathematical pedagogy are deeply entwined. Pólya takes it as a given that for students to gain a sense of the mathematical enterprise, their experience of mathematics must be consistent with the way mathematics is done. The linkage of epistemology and pedagogy is a major part ... and elaborates a particular view of mathematical thinking – discussing mathematics as an act of sense making that is socially transmitted. It argues that students develop their sense of mathematics – and thus how they use mathematics – from their experiences with mathematics (largely in the classroom). It follows that classroom mathematics must mirror this sense of mathematics as a sense-making activity, if students are to come to understand and use mathematics in meaningful ways”, by Schoenfeld.

Grouws, (1992, p.339).

- Resnick, tracing contemporary work to antecedents in the work of George Herbert Mead (1934) and Lev Vygotsky (1978), states that:

“... we may do well to conceive of mathematics education less as an instructional process (in the traditional sense of teaching specific, well-defined skills or items of knowledge), than as a socialisation process”.

Resnick, (1988, p.58).

- It was very reassuring that both Krutetskii (1976) and Schoenfeld (1992), representing the communist and capitalist systems, attached the same high importance to maths as the foundational subject that was the driver for increasing competitiveness. In spite of this, the word-based culture of the developed countries continues to deny its importance.

However, Resnick’s comment, as cited by Schoenfeld, (Grouws 1992 p.340) was the one that made the greatest impact by *dispelling my initial attitude of:*

‘...all that would be necessary to raise standards in learning arithmetic and mathematics was better instruction’ (Brought about through the use of a particular type of learning resource!)

There were many other detailed learning points that only made sense once this fundamental change of mind-set had taken place.

8.2. Reflections

The reality is that most of these initial learning points had receded in my mind as the arrangements were made for the Empirical Study, its implementation including carrying out the InCAS assessments in schools, analysing and writing-up the results, discussing them and generally progressing the Study. However, very recently, I wanted to locate a particular citation and thinking it was in Schoenfeld’s chapter, I skim-read it without finding it and then read it properly. This re-reading provides the driver for these reflections.

The title of his chapter is:

**LEARNING TO THINK MATHEMATICALLY:
PROBLEM SOLVING, METACOGNITION,
AND SENSE MAKING IN MATHEMATICS.**

It is something of a useful *tour de horizon* with 171 references and was published in 1992, so it is not necessarily up-to-date. Nevertheless, it lacked a definition of ‘mathematics’: does it have the same meaning as in the UK and, if so, is there a subject equivalent to ‘arithmetic’ – the number focus of this study? The relevance of this is considered shortly. While considering definitions, Schoenfeld pointed out that agreed definitions of ‘problem solving’ and ‘metacognition’ were not in place when he wrote his chapter, Grouws (p.337 and 347).

His main points are:

- A résumé of the changing decade themes following the then USSR successful launch of the *Sputnik* in 1957 that included in fairly quick succession:
 - (i) The 1960s: “New math” – that met a similar fate as its counterpart in the UK.
 - (ii) The 1970s: Back-to-basics that proved to be equally ineffective.
 - (iii) The 1980s: Problem solving and metacognition that led to mathematics being seen ‘as an act of sense-making, (p.337)
 - (iv) The 1990s: The combining of ‘what might be called the cognitive and social perspectives in human behaviour, in the theme of enculturation’ (p.347).

Enculturation can be seen as the outcome of Resnick’s insight, as above.

However, my own learning of number conditioned me to expect that an instructional process must be the only one, but that changed while watching pupils playing with their learning dominoes. They had changed their teacher’s rules and were helping one another: this example convinced me of the validity of Resnick’s insight that learning arithmetic is more a social process than an instructional one. Again, Watson’s concept of affordances, constraints and attunements extends his insight as pupils’ progress to more formal learning of arithmetic, (p.49).

This experience confirmed Schoenfeld's closing point that collaboration between teachers and their pupils and also amongst pupils was the most promising way to develop the all-important mathematical thinking that enables pupils to become members of the mathematical enterprise. It is added that these pupils had not reached the number-ready stage, when they would be introduced to their dials, but had already experienced collaborative learning and there should be no inherent limitation in the dials that would restrict this in the future. In the belief that this will be the case, Watson's pupils would have expected all their future arithmetic learning to be collaborative, (p.49). The particular episode was in one of the Follow-on Study classes and encouraging the teachers to become *guides* instead of *instructors* would become a key feature of the new Study. For example, a *guiding* teacher should be able to help her pupils to discover *collaboratively* how to use their dials. Relevantly, it is likely to confirm the need for the teachers to:

- Be a connector.
- Have good subject knowledge.
- Be enthusiastic about their subject.

This has already been recommended, (p.68).

8.2.1 Comments

Some comments are now made about Schoenfeld's chapter that were relevant to this Study:

- It is difficult to avoid the conclusion that the policy makers were guilty of confusing activity with action with their resultant decade themes, having explained that stability is almost essential for effective teaching. Were they responding to political imperatives to "do something"? The Finns have consistently demonstrated the benefits of stability.

- No mention is made of the importance of automaticity that must surely enhance arithmetic and mathematics performances. In fairness, Schoenfeld's review was strategic and not tactical.

It is now readily acknowledged that his overall contribution eventually became the most helpful and relevant part of the original literature review: it was convincing that learning mathematics is a very much more complicated process than had hitherto been appreciated. However, it can reasonably be pointed out that maths teachers and their pupils will have a much better chance of prospering if the pupils arrive in secondary school with a secure foundation of arithmetic that includes the inestimable contribution provided by fluency with number gained through acquiring automatic recall of number facts.

8.2.2 Interim Conclusion

The concluding point is made on ability categories (p.118). Simply, it is likely that in any cohort only 10-15% have the natural inclinations and potentials to become mathematicians (and will benefit from good automaticity). The middle group (70-80%) will become much better at arithmetic through good automaticity and, in doing so, will enhance their general skill levels to become members of the mathematical enterprise even though they will never become mathematicians (p.45). However, they will become better able to support the top group. It is this that will improve the competitiveness of the country. And, of course, the groups are largely self-selecting, but automaticity raises the attainments of *both* groups.

8.3 Le Dénouement

These reflections were being completed when Peter made a web search unconnected with this Study and came across the paper: *Developing Automaticity* by Crawford. It was truly a "Eureka!" moment for him and confirmed his conviction that automaticity provides the key to becoming fluent with number, based on over 40 years experience as a maths teacher. It was the missing piece of the jigsaw puzzle and clinched the case for the essential importance of automaticity.

8.3.1 Developing Automaticity

The provenance of the paper was unknown (to us), but there is evidence that it has been peer-reviewed by other authors who cite the date as 2003 (and this convention is followed). It is admitted that no attempt was made to discover why it was written on the principle: “Do not look a gift-horse in its mouth”! Reading between the lines conveys the possibility that it was Crawford’s valedictory and, as such, he was summing up his life’s work as a man who was at peace with himself. I believe our picking up his baton would have pleased him.

His thesis (2003 pp.9/10) can be summed-up as: there is a three-stage process through which pupils become fluent with number:

- **Concrete** as pupils develop their numerosities to connect quantities, symbols and words. These are achieved through counting, finger counting, counting-on, counting-back, number learning activities and games that develop the basic concepts of number as they become “number-ready”.
- **Strategic** as pupils start to learn number facts and strategies to calculate or derive them when needed.
- **Automatic** as pupils develop the retrieval of previously learnt number facts *without any conscious mental activity*.

As he cited (p.10):

‘When facts have been well practiced, they are “remembered” quickly and automatically – which frees up other mental processes to use the facts in more complex problems (Ashcraft, 1992; Campbell, 1987b; Logan, 1991a)’.

There has been considerable evidence of the application of the concrete and strategic stages – mainly variants of finger counting throughout the Study and also the Follow-on Study. Indeed, the main reason for quality improvement officers (QIOs) attending

their local authority workshops during the late 1990s was they hoped to learn some strategies. They were nonplussed when they discovered I had only been taught my number facts and did not know any strategies in the way they understood them.

However, I have probably developed many during my life that are personal to me and would only be usable by people with *identical* learning and background experiences. Importantly, they are based on my automatic recall of number facts, again, *without any conscious mental activity*. Or as it is described:

- ‘Modern theories argue that the process underlying automaticity is memory retrieval: According to these theories, performance is automatic when it is based on direct-access, single-step retrieval of solutions from memory rather than some algorithmic computation (Logan and Klapp, 1991a, p. 179)’, from Crawford, (2003, p.10).

However, not everyone accepted Crawford’s thesis, as he wrote (p. 10):

- ‘Baroody made one of the last forceful defenses of the alternative model that adults continue to use, albeit very quickly, “rules, procedures, or principles from which a whole range of combinations could be reconstructed” (1985, p. 95)’.

It can be observed that Baroody’s defence was seemingly effective in that automaticity is still virtually unknown as a word or a concept in the world of education. Unhelpfully, many promote strategies as a means of demphasising memorisation (Crawford 2003, p.12).

To return to automaticity, Crawford cites the large body of research that describes the effective contribution that automaticity makes in helping pupils to become fluent with number. In particular, measurements of direct retrieval response times are typically less than one second compared with three or more seconds when strategies are used to produce answers with higher error rates (Crawford, pp.12/16). In some ways the

current widespread avoidance in developing automaticity is truly remarkable when it was observed:

“Automaticity is not genius, but it is the hands and feet of genius” (Bryan & Harter, **1899**; as cited in Bloom, 1986 – Crawford, p. 14, emphasis added).

The value of automaticity lies in its contribution to *conservation of memory*, (p.221). Again, in simple terms, memory resides in the brain that has a finite capacity, like any other organ. Thus, direct retrieval of number facts when calculating requires less mental effort and this allows more – and higher order – calculations to be made. Peter’s repeated experience (as a maths tutor) is that able pupils who understand the mathematical principles involved fall down because they do not know their number facts; this saps their confidence as they inefficiently try to calculate them *each* time they are needed. Having done so, hopefully correctly, they have become “tired” and lost their calculating momentum to conclude, once again: maths *is* a hard subject.

Helpfully, and possibly as a response to overcome the general (teacher?) antipathy to memory work, Crawford devotes the largest part of his paper (pp. 19 to 32) to the practical steps that lead to the embedding of number facts into pupils’ long-term memories. Reassuringly, he uses the Swiss roll model approach that we advocate of:

- Repetition.
- Revisiting.
- Consolidation (p.228).

Moreover, he cites tried-and-tested procedures, confirmed through much research, that use a series of small, but logical, memory steps to construct a complete repertoire of number facts. In many respects, these are likely to provide the biggest practical contribution in ensuring that automaticity is desirable and attainable. This will be evaluated as part of the Follow-on Study.

8.3.2 Comment

To explain this last observation, it seems after all that teachers have a latent wish to do systematic and effective memory work provided it is properly structured and not

rote learning. Crawford has addressed this point by summarising the research findings and Peter has broken down the 693 basic arithmetic number facts into small *logical* groups ready to be built into the Swiss roll model. Previously, the teachers had convinced themselves that there was little point in attempting memory work once they had discovered how ineffective the linear memory model is – especially, when they were so short of time (p178).

8.3.3 Confirmation

The discovery of Crawford's paper was perfectly timed to support our conviction about the vital contribution automaticity would make in improving pupil's arithmetic attainments. This independent confirmation was a very welcome reassurance after the apparent setback, when the main question could not be answered, because it has greatly strengthened our arguments on the importance of automaticity.

Encouragingly, the decision-makers and staffs of the Follow-on Study schools became utterly convinced of automaticity's paramount importance. They committed themselves to achieving it as a means of increasing their pupils' arithmetic attainments.

8.4 Conclusions

The two main conclusions that can be drawn from these reflections are:

- Learning number becomes more effective as a socialising process rather than an instructional one.
- Independent support from Crawford's paper supports the case for automaticity that has been argued throughout this thesis is completely justified.

However, the reality is that there is still much to be done before the important contribution that automaticity would make towards acquisition of number fluency and mastery becomes generally accepted by the policy-makers and then implemented in schools.

9.1 General Review

9.1 Introduction

It was realised that the general situation in schools had become very fluid and it seemed unlikely that any research into its causes would have been carried out or explanations available for such a development. This coincided with its increasing awareness through the pages of, for example, the Times Educational Supplement Scotland (TESS) and the media generally. This included the Open EYE campaign and its publication: *Too Much, Too Soon?* House (2011) that included topics such as ages for starting formal learning and potential damage arising from starting formal learning before pupils' cognitive development had progressed sufficiently.

9.2 History

As is so often the case, there were many different influences in play that had led to the current lack of emphasis on traditional mental arithmetic. Some of these are now considered starting at the time *Sumdials Ltd* was established in 1997. This was before the introduction of the '5 to 14 Guidelines' in Scotland and the 'National Numeracy Strategy Framework' in England and Wales in 1999 that marked:

‘... the tightest ever control by government on primary mathematics, with central prescription not only of national curriculum and national test, but also of teaching style’, according to Brown (Thompson 1999, p.15).

Two comments are now made:

- Clearly, these developments brought about a huge change in education as government involvement moved from its previous relatively laissez faire approach to one of total control.
- The full effects of this new control had not fully taken effect by the time of Bramald's Study – hence the earlier observation that the treatment teachers in both England and Scotland found the detailed *Sumdials'* teaching plan unexceptional and were not surprised that memory work was included, (p.31).

It was after that time that the phrase ‘initiative fatigue’ was being heard in schools with ever-increasing frequency following the relentless top-down requests – shorthand for government initiatives. These have led to the current overloaded – cluttered – curricula with the resulting shortages of time. The time available for each learning topic had unwittingly been reduced either to seven minutes, as calculated by one head teacher, or anecdotally, the inspectors’ figure of four minutes confirms the lack of time to teach pupils their number facts.

It was hinted that there might be other reasons for not undertaking memory work and one possible explanation is teaching plans (from on-high) were becoming increasingly prescriptive and did not specifically require pupils to hardwire number facts into their long-term memories. Pupils are required to ‘learn’ or to ‘know’ them and may, indeed, have achieved this by the end of a lesson or end of a week, but then most are quickly forgotten. Another possible reason for the lack of memory work is that many primary teachers may never have learned their number facts themselves to be confident about their recall. In turn, they were not convinced of its over-riding importance or fully understood how to achieve number fact automaticity for their pupils. This is just one example of the top-down syndrome that is now briefly considered.

9.2.1 Top-Down

Top-down is shorthand for the process whereby the appropriate government departments establish education policy and objectives and then micro-manage their implementation. This was an integral part of both the National Numeracy Strategy Framework (England and Wales) and the 5 to 14 Guidelines (Scotland). This continues to make an ever-increasing impact – not necessarily for the better - and this needs to be considered, even if it is a subject that may not be based on any academic research so far on its effectiveness and likely consequences.

Such outcomes are almost inevitable, as explained by Harford (2011, pp. 37–79). His argument is that when policy implementation involves many layers of management, the feedback loop becomes very ineffective because of the complexity and contradictions of the information being transmitted back to the top. This is

exacerbated by the instincts of the upward transmitters to filter out adverse data that they sense would be unwelcome to their superiors. He illustrates this principle with an account of the second Gulf War when it was going badly and how it improved once the middle ranking officers *on the ground* started to ignore their instructions from the Pentagon. It was remarkable how many of these officers had social science PhDs and could identify, relate and respond to the *real* concerns of the local people that were quite unconnected with what the Pentagon believed them to be. The parallels with education could be very close.

9.2.2 Starting Ages for Formal Learning

Having already mentioned the issue of starting ages for formal subject learning, it is now considered. The earlier citation provides a good introduction:

‘Exposing pupils to mathematics instruction early on seems to be a natural step in addressing this difficulty’ [of under-achievement in the United States], as asserted by Malofeeva (2009, p.75).

This could be consistent with research into the innate number abilities of the pupils under the three to six years age of the pupils in her group, (pp.41/76).

Such research shows these abilities are much greater than Thorndyke’s *empiricist* framework that pupils’ minds are “blank slates” at birth and they only become ready to learn arithmetic by the ages of six to seven years old, as cited by Sarama & Clements (2009, pp.3-19). One outcome has led to the development of pupils’ learning trajectories to capitalise on such research findings in support of the argument that it is the ‘limitations of the society and its schools’ rather than the limitations of the pupils themselves. This is well summed-up by:

‘What pupils are capable of at a particular age is the result of a complex interplay among maturation, experience, and instruction. What is developmentally appropriate is not a simple function of age or grade, but rather is largely contingent upon prior opportunities to learn’.

(Sarama and Clements, 2009, p.25).

This may explain the reported general findings of the *73 Education at a Glance: OECD Indicators* (2013) that reveal the inferior United States’ (and England’s) attainments relative to most other developed nations being the result of such research

findings not contributing to more effective pedagogies. (Scotland did not participate on the grounds of cost.)

In contrast, pupils' indicative starting ages in European countries, apart from the UK, are seven years old, but the possibility of reducing this by one year was considered in North Rhineland-Westphalia, Germany in the 1970s. However, it was decided to test the benefits, if any, from such a change (in contrast with the usual UK approach?) and a study by Schmerkotte (1978) was carried out with 50 kindergartens. 25 of them took part as a treatment group following an early academic programme emphasising maths and reading while the other 25 followed the traditional child- and play-centred approach. The conclusion was that the treatment group showed a slight initial gain at the expense of reduced social skills, but by the third year there was no measurable gain and the starting age remained unchanged, according to Suggate (House 2011, p.239).

It could be that the later starting ages had been arrived at through experience and that Piaget's findings on pupils' development confirmed they were appropriate. To enlarge on the point above on arriving at effective pedagogies, the research findings mentioned by Sarama and Clements (2009) are likely to have been made by gifted researchers working in near laboratory conditions that are very different to those found in messy primary schools with typical teachers. If this is the case, then the first step should be to encourage local improvements, possibly incremental, to existing pedagogies while new research-based pedagogies are developed.

9.2.3 Conclusion

This extension to the Literature Review has developed in response to the changes in the direction this research took following the initial findings of the Pilot and Empirical Studies. In particular, was the discovery that mental arithmetic was no longer being emphasised and the resulting need to consider its consequences on this Study.

The overriding need now is to regain stability in primary schools through discontinuing further top-down initiatives immediately while the teachers regain time

to reflect on and review what is actually happening in their own classrooms. Once they have achieved stability again, it needs to be borne in mind that it takes three iterations/years to implement properly a change in teaching methods, (p.68).

9.3 Some General Observations

9.3 Introduction

It was suggested earlier that the extent of government direct involvement in education provides a good barometer of national self-confidence. Simply, there seems to be an inverse relationship whereby the greater the governmental involvement (and control) in education, the lower the national self-confidence. It was only about 20 years ago that it was being proclaimed that “Education, Education, Education!” would solve all the country’s ills. This was another step along the road to the ever-increasing government control that now is almost total – a process described by Brown (Thompson 1999, pp.3-17). It can be asked if there is a link between the recent PISA international comparisons with the moderate UK scores, (OECD, 2013)?

9.3.1 Government Control

The issues are very complex, but a simplistic explanation is governments have been establishing and amending education policies and increasingly micro-managing their implementation. This has not produced the desired results because a one-size-fits-all approach has been used. The organisational Achilles’ heel of this approach is the low quality feedback from school classrooms to the policy makers that are leading to ever more irrelevant initiatives. The effects of establishing academies or free schools are not considered because it is still early days and effective education is more likely to come through long-term stability.

The different time scales of governments and schools is a further source of tension in that governments are always preparing for the next election, which will never be more than five years away, while schools work within much longer time horizons. Teaching is a very conservative (with a small ‘c’) profession and this is appropriate since they are preparing pupils for life. This makes governments impatient, as Baker demonstrated, while teachers resent their becoming political pawns along with their

pupils, as Brown pointed out, (Thompson 1999, p.11). Thus tension is inevitable especially when teachers are naturally dedicated to meeting the needs of their pupils and need sufficient time to evaluate and then plan properly how they will implement new initiatives and methods.

The impression should be well established by now that all is not well in the world of education and this can be attributed to the excessive political control at national level and, then, having set the strategic/macro aims, implementing them through top-down tactical/micro initiatives. Many of these have not been tested by prior research to confirm that they will be effective and relevant to actual local needs. In contrast, the German-speaking Swiss cantons started work in 2011 to establish a common curriculum by 2021. In that time there could be two changes of UK governments that could lead to further changes of direction for education – hardly conducive to stability. It can reasonably be postulated that there is likely to be a link between such a measured pace in Switzerland and the OECD (2013) results showing that it is now the top non-Pacific rim country. However, the measured pace in drawing-up a new curriculum has been overtaken by the Federal wish for ‘harmonisation’ of all cantonal curriculums (p.158).

It is as an outsider that all these points are made and it is concluded that there is now an urgent need to achieve stability within UK education. This will be best achieved by avoiding any further initiatives or changes until schools and teachers get their bearings; it is acknowledged that the current levels of change are going to make this very difficult to achieve. However, it will have to happen and the sooner the better it will be in the long run. Might Finland, where its strategic education policies were agreed about 30 years ago, provide a good example of the benefits of stability?

In an ideal world, the role of governments would be limited to establishing educational policies and priorities while delegating to schools via their local authorities responsibility for tactical adaptations in response to *local* priorities and needs. However, this will only work if all political parties can agree long-term *strategic* objectives and then engrave them in stone – preferably granite! That is essentially the Finnish model and one important effect is its education became

depoliticised to everyone's benefit and, in turn, the essential long-term stability was established.

Again, time is required to carry out appropriate pilot studies to assess the soundness of proposed changes and only then can the detailed preparations for their implementation be made. The introduction of New Maths illustrates the point of insufficient time being allowed in that not enough textbooks were available in time and, more importantly, the capabilities of the teachers to deliver it effectively had not been adequately considered. Visionary mathematicians promoted its introduction during a relatively *laissez-faire* period towards education by governments. However, education departments drew the conclusion that establishing control over schools would greatly simplify the implementation of new initiatives.

Realistically, external circumstances are always changing and many of these will require *tactical* adjustments to be made. However, governments' impatience over recent times betrays a lack of trust in schools and teachers as they increasingly micro-manage their own policies by telling schools and teachers what they have to do and then ensuring that they do it. This has led to endless initiatives – hence, “initiative fatigue” – followed by dutiful implementation even though the teachers know the initiatives are not working. This provides a good illustration of the pitfalls of such an approach by those on high get their inspection agencies to investigate and discover that the teachers are doing exactly what is required and, so, they cannot be blamed for the disappointing results!

Worse still, as teachers implement the top-down requirements of overloaded (cluttered?) curricula they are becoming “de-professionalised” as they have to concentrate on implementation rather than on the needs of their pupils, (Hughes, 2007). This is well illustrated by an observation about a probationer made by a head teacher that she would not be retaining her: ‘because she only did what she was told to do’. In other words she was not analysing and reflecting on her pupil's individual difficulties to come up with alternative pedagogies to meet their actual needs. In fairness to the probationer, she did not have the experience to know how to deal with the micro-management implicit in the detailed prescriptiveness of what had to be

taught - even though it may have been quite irrelevant to her *local* needs. In short, she and many others like her had become “obedient messengers”.

There are other drawbacks to such prescriptiveness and these include the development of “box-ticking” approaches to teaching and learning (as has already been mentioned) that includes trying to stick to the allowed times to deliver each topic that make no allowances for the inevitable “strops” of young pupils. In the case of the lack of hardwiring of number facts, the inevitable response has been that there was not enough time.

It is believed that most teachers would identify with these points and this raises the question of how it has come about? To deal with the cluttered curricula feature, it is probably the result of responses to *transient* political needs, leading to the new topics becoming *permanent* features of the curriculum. Another possibility is those who draw up curriculums are not part of the mathematical enterprise. If they had been, they would have instinctively realised that insufficient time would be available for typical middle-of-the-road teachers to cover the ground working with typical middle-of-the-road pupils. Again, the reminder is made that there are only four/seven minutes available for each topic. Thus, a pupil responding to a call of nature would miss a topic!

9.3.2 Teaching Number

It would be natural to attribute the difficulties pupils have with number to teachers’ inadequate subject knowledge and undoubtedly this is a factor. However, the reality is that the teachers are more the *victims* of administrative expediency. The main contributory factors were:

- Projections about 20 years ago indicated that there would be a shortage of teachers.
- This coincided with the emerging shortage of jobs for new graduates.
- The training colleges had to accept much higher quotas for their PGCE courses. For example one had its quotas increased from 350 graduates to

500 and then 600 in successive years (and knew the numbers were excessive).

- Such courses attracted many of the new graduates who had been unsuccessful in finding employment in their original career choice and started the de-professionalisation of teaching (p.210).
- The entry qualification of a ‘C’ in GCSE (England and Wales) or Standard Grade (Scotland) in maths is not high enough to ensure candidates have the necessary *in-depth* subject knowledge.
- The first degree of many primary school teachers is likely to have been in word-based subjects, maths having been dropped at school at the earliest opportunity.
- There became insufficient teaching posts available for the new teachers and one consequence was a growing feeling of insecurity, especially amongst primary teachers, leading to their doing what was *prescribed* even though they knew it was not working – the situation encountered in the Study schools.

To explain the link between C pass qualification level and in-depth subject knowledge, the unwritten custom and practices in secondary schools are:

- A pass: Enrol, almost automatically, for the higher maths course and in-depth subject knowledge taken for granted.
- B pass: ‘Let’s have a go and see how you get on’ – reasonable subject knowledge assumed.
- C pass: Relief - having demonstrated only limited subject knowledge and aptitude. Drop maths by mutual agreement.

This summary may seem cynical, but it is the reality and C passes usually would typically be insufficient for admission to the mathematical enterprise. The reality is that too many primary school teachers have entered the profession under the mistaken belief: “any graduate can teach” – adapted from “Anyone can be a teacher”! Neither is true because teaching is a *vocation* and not a port-in-a-storm for unemployed graduates. To give other examples, career mathematicians would be unlikely to make

good schoolteachers or tournament (golf) professionals are unlikely to make good teaching professionals because very different outlooks and skills are required. Having said that, some graduates-become-teachers have discovered that they have the instinctive aptitudes and intuitions to become excellent teachers (Peter being an outstanding example), but they are the exceptions.

The need for good subject knowledge and enthusiasm for number is to enable teachers to make the all-important *connections* between what pupils are doing as they learn number and to the next steps in their learning. It requires good subject knowledge to spot quickly the causes of mistakes and then to correct them before anxiety sets-in. These assume special importance, as pupils are becoming number-ready through the socialising activities such as number games - Resnick's conclusion – to encourage learning rather than just letting the pupils play.

It has already been suggested that the decline in number skills actually started following the introduction of New Maths. Based on the very small sample of automaticity tests that some teachers took, it was found that those who were over 50 years old (in 2001) had no difficulties, while the younger the rest were, the more mistakes they made (p.42). One effect, as the William's Review (2008) acknowledged, was the UK had become one of the few developed countries where it was quite chic to be proud of poor maths skills (Para 4) – and this would likely include many primary school teachers. The official belief was such shortcomings would be rectified through CPD delivered by maths specialists; this was unrealistic (to me) and is supported by two personal experiences.

The first one occurred while leading a workshop for school numeracy co-ordinators – presumably the best arithmetic teachers in their respective schools – and I was asked very early on: 'How many number facts do you know'? I had never thought about it and could not answer it (subsequently, I estimate it would be more than 750 and that is far fewer than the 2,000 words - with their spellings - of a basic vocabulary). The point is they quickly perceived my level of arithmetic proficiency to be of a totally different order to theirs and, yet, they would be expected to coach their colleagues. A second issue emerged and it was my lack of "strategies" because I was never taught any when I was at school, presumably because the teachers were unquestionably

convinced that recalling number facts *instantly* was much more efficient than calculating them.

The second workshop was for local authority QIOs (quality improvement officers who advise *schools*) where I found myself in a small group with two of them. We were asked to solve five straightforward word-based problems working together, but I finished them by myself in about three minutes – due to not hearing the instructions properly – and discovered my new colleagues were still trying to choose a strategy to use for the first question. Three points emerged:

- They were very impressed by my speed of calculation and admitted they knew very few number facts.
- They felt they had to use “official” strategies in their problem solving and the concept of using spontaneous methods (that come naturally once number fluency and mastery have been acquired) was quite foreign to them.
- Their number skills, to me, were very weak and yet they were expected to advise schools on how to improve their teaching of number – the point made in my submission to the William’s review.

These experiences confirm that very few people are likely to be part of the mathematics enterprise because innumeracy is almost endemic in the UK. Hence, my earlier remark that something is amiss in the world of number education is justified and supported by Ma (1999, p. 26).

That teaching and learning number in primary schools has not collapsed is almost entirely due to the commitment and dedication of the teachers as they try to do their best for their pupils. The price the teachers are to pay for this is working up to 65 hours each week preparing ‘innovative’ presentations for learning topics (see below). This cannot be allowed to continue for many reasons and one is teachers need time to think and reflect about what is happening in their classrooms and to consider ways of improving their own pedagogies. In fact, Peter and I were doing it for them by making our comments and suggestions – a service that has been welcome and put to

good use. To give one example, they had not realised until we pointed it out that parents were no longer helping their pupils to learn number, as used to be the norm.

A confirmation of teachers' lack of time to think and reflect is borne out by my practice of writing briefing comments in response to specific InCAS results or school experiences. Ten such notes had been written and could be regarded as separate mini-studies that were not a direct part of the Study. When they were analysed 14 themes had been mentioned, some of them more than once, and a summary of the analysis is given in Appendix 5.3. However, the most striking point is four themes only:

- Accounted for 60% of the mentions with Lack of Time being the most frequent - six times – (p.203).
- They all dealt with the general teaching and learning environments of schools.
- *None* of them are subject specific.

An explanation of how this situation could have arisen is now suggested. Its root lies in the career progressions of civil servants who are likely to be word-based administrators with limited number skills, as has already been mentioned. They transfer from department to department such as agriculture, social services and education as their careers progress, but without ever acquiring the in-depth subject knowledge to become competent in assessing the technical advice they receive. To mention a rule-of-thumb from business, 95% about a business can be learned in six months, but it takes another 20 years to acquire the remaining 5% that is essential for success.

Such civil servants become like birds eating breadcrumbs without ever realising what a loaf of bread is. These comments on civil servants' lack of technical competence seem to be supported by a TESS article (19 July 2013 p.7) citing the work of Campbell, LHAE, Toronto University.

9.3.3 Interim Conclusion

This section can be summed-up by concluding that politics now drives education with activity becoming confused with action. Having said that, describing the life skill contribution of arithmetic should be seen as a constructive response.

9.4 The Contribution of Arithmetic

Arithmetic and geometry were the original foundation subjects for all number. Only a very limited amount of geometry is now included in school curricula and arithmetic could be going the same way, as a quick trawl through the subject indexes of the literature confirms. It was one of the pillars of the 3Rs (of Reading, 'Riting and 'Rithmetic) in British primary/elementary education from 1862. However, it ceased to be a stand-alone subject in the 1980s in Scotland, probably in response to timetabling pressures rather than any policy decision, and possibly as a fall-out from New Maths in England.

It is now argued that these were unfortunate steps because, again, arithmetic (with geometry) is the *foundational* subject of all mathematics and not a part of it, as confirmed by Ma (1999 p. 116-8). The analogy of building a house without foundations, while using up their materials as and when the house is built, is apt. Some of the consequences of these were summarised at a workshop when the main findings from 'a database of over 2,000 errors taken from first-year university maths exam scripts' by STEM students (science, technology, engineering and mathematics) were described. One conclusion was that the origins of many of the errors could be traced back to inadequate or incomplete teaching and learning in *primary* schools, according to Gibson, Goldman and Grimfeld, (2005).

The context for their point is maths is a highly structured and hierarchical subject that makes it difficult to correct unsound teaching or to fill any gaps at secondary schools. One such example of limited subject knowledge is: "Subtracting *always* reduces the size of a number". In short, the arithmetic foundations must be securely laid once pupils become number-ready, (p.212). Passing driving tests is a good analogy because success confirms that the basics of driving a car have been learned as the essential first step in becoming good drivers, be it as driving taxis, ambulances, buses, fire engines, lorries or even F1 racing cars.

The parable of the wise and foolish builders (Mt 7:14-17) illustrates well the point that is being made on arithmetic's essential role as the foundational subject for mathematics.

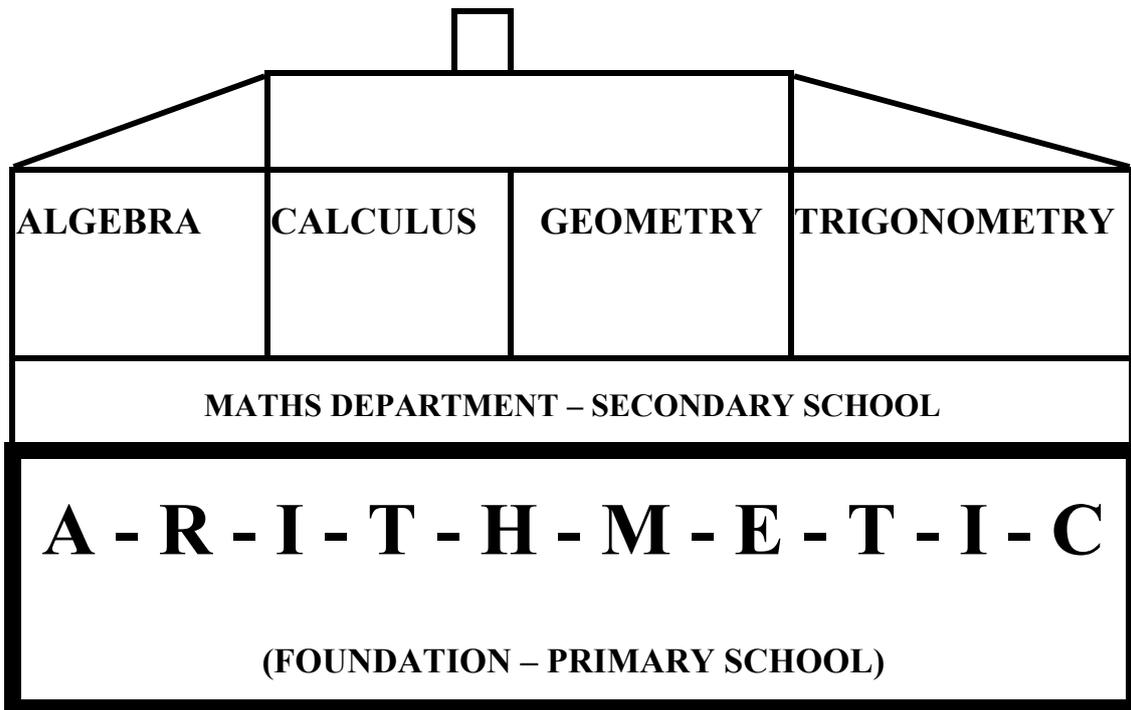


Figure 9.24: The Arithmetic House with a Foundation (built on rock) after the storm.

In contrast, the current situation can be likened to the house that collapsed in the storm.

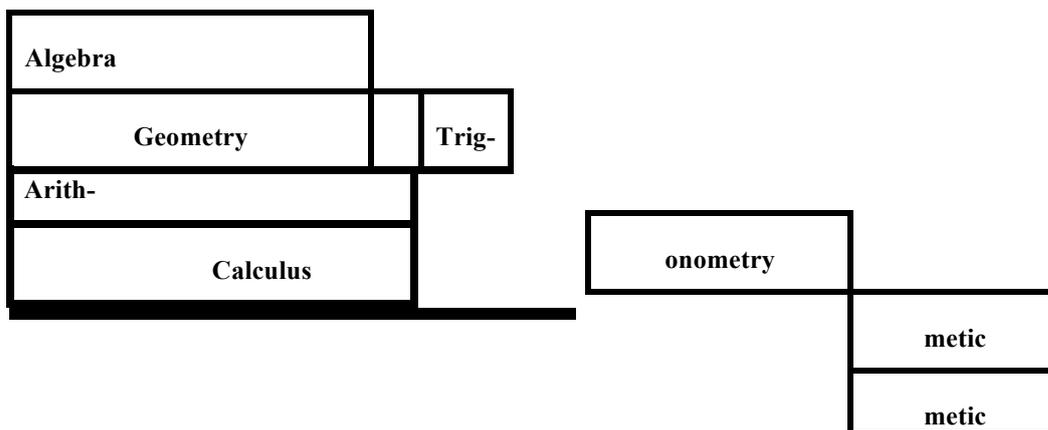


Figure 9.25: The Number House with no foundation (built on sand) after the storm.

Experience suggests that as much as 90% of the adult population only needs arithmetic and, in support of this, Peter's experience is he has hardly used any branch

of mathematics outside his classroom during his adult life. Those who do need mathematics will always be grateful for the secure foundation of arithmetic.

As a reminder, the teaching plan that was provided for Bramald's Study and was used in the current Study had 27 lessons in two sections:

- 12 lessons explaining how to operate the dials, followed by worksheet-based exercises to become confident in using the dials for addition up to 10 (and to make a start in laying the foundations for addition, as one of the basic processes of arithmetic).
- 15 lessons to hardwire the associated adding number facts in the pupils' long-term memories.

It is now clear that these are two separate – even unconnected – activities. However, the memory lessons were not delivered in this Study, as the observations during the InCAS assessments confirmed (the Swiss pupils were the unconnected exceptions). Hence, the conclusion was that true mental arithmetic had effectively been abandoned as a taught subject. Possible reasons for not carrying out the memory work include some or all of:

- Using understanding/first-principles strategies to calculate number facts is better practice than hardwiring them.
- Insufficient time for proper memory work.
- The pupils would find it boring.
- A general lack of understanding and appreciation of its importance.
- Lack of capability of the teachers.

These are now considered.

9.4.1 Understanding/First-Principles

This aspect has already been discussed, when the seductiveness of this approach was emphasised (p.134).

9.4.2 Insufficient Time

This is a plausible explanation within a context of overloaded curricula, but it probably conceals the more fundamental reasons that have already been discussed (p.139).

9.4.3 Boring

The primary school ethos until very recently was pupils had to enjoy themselves at school and that took precedence over learning. This was queried in the participating schools and subsequently modified when it was pointed out that:

- The *raison d'être* of schools used to be teaching and learning.
- We had no recollection of actually learning our own number facts.
- It was good preparation for work in that 95% of all jobs are boring!
- The use of *Sumdials* has not been considered to be 'boring' by the pupils

9.4.4 Lack of Understanding

One of the aims of the Follow-on Study is to collect evidence of the considerable benefits that automaticity will make on improving number skills even though it is not necessarily accompanied by good understanding. It is hoped that such an outcome would lead to a bigger and properly resourced study to be undertaken to confirm the importance of automaticity

9.4.5 Teacher Capabilities

This has already been discussed, (p.211), but in essence it was observed that the teachers' total commitments to satisfying their pupils' general learning needs was hindered by the teachers' relative lack of confidence in their own number capabilities and, in turn, to deliver memory work effectively.

9.4.6 Interim Summary

It should now be apparent why it was concluded that 'all is not well in the world of (mathematics) education' (p.208). The way forward is to identify a constructive way to make the essential improvements in number skills that will accrue through good automaticity and the consequential improvements in international competitiveness.

9.5 The Contribution of Automaticity

To return to the main question of the research, it was set in good faith based on both my childhood experiences and achieving automaticity – the hardwiring of number facts into long-term memories - was one of the aims of Bramald's Study. It has already been concluded that the success of the *Sumdials*' approach (with its use of dials that model the basic arithmetic processes well) helped the pupils to develop robust internal models of these processes through visual and kinaesthetic – seeing and doing – methods. However, the repetitions pupils make in operating their dials only contribute towards developing robust internal models of number: they are unlikely to contribute to hardwiring of number facts, as is now explained.

The first step is to clarify what automaticity's actual role is in becoming fluent with number. Simply, good automaticity has the effect of apparently increasing the capacity of brains by making them more efficient when carrying out the arithmetical calculations that are such a day-to-day feature of life. Two analogies are now used to illustrate one very important benefit of automaticity and it is: conservation of short-term memory by reducing the number of calculation steps it has to make. Memory, as part of the brain tires with use, like any other organ, and this is where automaticity contributes by increasing its stamina, as is now explained.

9.5.1 Memory Conservation

It is commonplace experience that the brain tires as it makes calculations and there are a finite number of steps it can make in one session in much the same way that people can only make a finite number of steps on a long walk. Walking stamina can be progressively increased through good training and by using walking-poles while reducing the weights of clothing, boots and backpacks that can be likened to number strategies. However, the real gains come through good route planning that allow shortcuts to be made. Automaticity is the equivalent of this because it improves the efficiency of the brain when it is engaged in number work.

This is now illustrated using the analogy of a simple computer that has:

- An input device such as a keyboard (that is equivalent to eyes or ears),
- A central processing unit (CPU, equivalent to short-term memory) consisting of:
 - (i) An executive controller responding to a program, and
 - (ii) Between five and nine working registers for calculations,
- Read-only-memory (ROM - equivalent to long-term memory),
- A printer that gives answers (to the eyes).

Computers use pre-written programs (lines of instruction or code) to control all their operations of reading input data, processing them and then outputting the answers, following the program's instructions. It is pointed out that it can be reasonably be assumed the early computer designers used their own calculation methods as models for their new computers.

A simple problem illustrates the steps in evaluating "4 + 3 =?" The lines of code to do this, when there is *no* ROM, could be:

1. Set x and y registers to 0,
2. Read-in question (4 + 3 =?) from input device,
3. Store 4 in x register,
4. Store 3 in y register,
5. Add 1 to x register,
6. Subtract 1 from y register,
7. Test if y register = 0 and jump to 9 if it is,
8. Jump back to line 5 (if y ≠ 0),
9. Send answer (7) to output device (when y = 0),
10. Jump back to line 1 (for next question).

In this example, 11 calculating steps (5, 6, and 7 three times and 8 twice) are required to do the calculation itself (and there would be another four for each number when the addend is > 3 or four fewer for each number that it is < 3).

However, computer programmers realised that having a ROM holding all the number facts (in a look-up table) would reduce the number of calculating steps in solving a mathematical problem:

Line	Instruction
1.	Set x and y registers to 0,
2.	Read in question from input device,
3.	Store 4 in x register,
4.	Store 3 in y register,
5.	Call-up from the ROM the value for $4 + 3 (= 7)$,
6.	Enter 7 into the x register,
7.	Send answer (7) to the output device,
8.	Jump back to 1 (for next question)

This program now only requires 2 retrieval steps (5 and 6) regardless of the size of the second addend to achieve much greater processing efficiency (over 500%) by installing a ROM look-up table. Look-up tables simply greatly reduce the amount of repetitive calculation needing to be made by short-term memories in answering such questions: other examples include log tables, distances between cities in road atlases or the pigeonholes in staffrooms for papers, messages, etc.

The parallels with memory will be obvious, but two points need to be made in connection with automaticity:

- Once an item of information has been stored in long-term memory (as has just been described) it is there for life unless trauma or a degenerative condition is sustained. To give a good example, the elderly mother of a Study schoolteacher is suffering from extreme memory loss, but still knows and can accurately *recall* all her number facts.
- In contrast, short-term memory used for number tasks can only hold such information for short durations (measured in minutes) and has only sufficient registers (between five and nine) to store information if none of them are

required for executive functions such as carrying out computer operations as in the example above.

To return to the walking analogy, automaticity is equivalent to good route planning that greatly reduces the number of steps needed to complete a calculation and so allows *many more calculations to be made before tiredness takes over*. And, importantly, the level of accuracy of recalled number facts will be higher than calculated ones.

To illustrate this principle, Peter's experience as a maths tutor for secondary pupils preparing for exams is relevant here. All of them, *without exception*, understand the processes involved, but encounter difficulties because they do not know their number facts. Thus they are repeatedly diverted from evaluating their questions while they apply their strategies to calculate, usually, the addition or subtraction facts. The outcome, even when they carry out such secondary calculations correctly, is a loss of momentum to become "tired" – like walkers – and discouraged because they realise they should have known the answers to the secondary calculations. This is further exacerbated when they get incorrect overall answers, as my grandson's example illustrated (p.136).

These analogies and explanations have described automaticity's role in helping pupils to develop fluency and mastery with arithmetic. Now its contribution to mental arithmetic is considered.

9.5.2 Automaticity and Mental Arithmetic

The contributions made by automaticity in acquiring mastery with number are now described. A good starting point was the observations made during the InCAS assessments. For example, the majority of pupils were using their finger-counting procedures when answering their adding and subtracting questions. To illustrate this, take the example $4 + 3 = ?$ Typically, the steps were:

- Read/hear the question.
- Put up four fingers (almost invariably of the left hand) and then three fingers on the other hand.

- Count the extended fingers on both hands by either nose or head tapping.
- Select the correct answer on their laptops, being one of four possible options.
- Depress the laptop touch pad.

The main comments are:

- This procedure was first observed in Edinburgh and it was assumed that this was a local practice (in response to the need to improve attainments quickly), but it was concluded that this practice was being *taught* nationally because exactly the same procedures were observed in the Co. Durham schools.
- One consequence is the pupils were learning to count-on (or count-back) starting with the larger addend first instead of being taught to add or subtract (up to 20). The considerable limitation of such practices were confirmed by the anguish caused by the question: $8 + 3 = ?$
- The bigger concern is the pupils were being taught to use *external* models and substituting counting-on and counting-back for adding and subtracting. To explain its limitation, how can multiplication - repeated addition - be taught if pupils do not know how to add?
- Is this yet another example of a top-down initiative that has been introduced without prior research to establish its long-term effectiveness in a highly hierarchical subject like arithmetic? The point made by Coe (2013 p.2.3).

Now it can be argued that if pupils are getting the right answers (albeit very inefficiently) that is what matters. However, that does not allow for the effects of finger counting becoming the default method for life. Most secondary maths teacher will testify that their pupils are using their default methods (finger counting) – usually under their desks - and especially so when under pressure.

It has already been pointed out that mental arithmetic involves manipulating data – making calculations ‘in the head’ - without using pencil and paper methods. This is achieved through:

- Analysing the problem and deciding how to solve it.
- Recalling the appropriate number facts.

- Applying them correctly to solve the problem.

The Mental Arithmetic module of InCAS assesses these activities and is now illustrated with two actual number fact questions in fairly adjacent sequence:

- $4 + 3 = ?$

In word-based format:

- “Four rockets are flying together (as the caption for the picture) and are joined by another three rockets. How many are there now”?

The majority of pupils used their finger counting procedures, as just described for the first question and then used it again when answering the second one. Such practices were repeatedly observed in both the Mental Arithmetic and General Maths modules and signified that they were seen as independent questions with no commonalities. It was the repeated observation of these practices that led the conclusion that traditional mental arithmetic, as a subject, was no longer being developed. Such an outcome is almost inevitable because the use of *external* models is likely to lead pupils to regard each question as isolated and unconnected problems. This is inefficient even though the pupils were getting the correct answers – the point made by Krutetskii (1976). Schoenfeld (1992) would have endorsed Krutetskii by not admitting them into his mathematical enterprise. The analogy of a rowing boat with only one oar would be apt – erratic progress would just be possible, but very inefficiently.

9.5.3 Possible Explanations

How has this happened? Possible explanations include:

- Overloaded curricula (accompanied by unrealistic targets) that leave insufficient time to teach all topics properly – the four/seven minute restriction.

- The limited availability of time provides a good pretext for teachers to skip number work in a maths-is-a-hard-subject culture, especially when they are not confident about their own subject knowledge.
- Memory work is perceived as being boring and this is unwelcome in an ethos where pupils must enjoy themselves at school. Thus pupils are being denied the von Glasersfeld's self-generated satisfaction that comes through achieving progress as they answer increasingly difficult questions (p.53). The subject of boredom is also outside the scope of this Study, but in passing it is generally accepted that 95% of *every* job is boring – even if not admitted.
- The breakdown of the unofficial compact between parents and schools in sharing the teaching of number that has been accelerated by the arrival of the digital age (p.48).

There is some light at the end of the tunnel in that pupils generally know their multiplication/division tables better than their adding/subtracting number facts, based on anecdotal evidence from secondary schools. Thus some memory work is still being taught in primary schools, but the harmful consequences of not knowing the adding/subtracting facts remains considerable.

The challenge is to get general acceptance that hardwiring the number facts of

- Adding/ subtracting up to 10
- Adding/ subtracting up to 20
- Multiplying/dividing tables up to 50
- Multiplying/dividing tables up to 100

must be achieved while pupils are still at primary school, if they are going to be successful with the secondary mathematics curriculum. There are only 693 number facts and, as has already been explained, long-term memory becomes less receptive to new facts after those ages. Again, leaving any gap filling of missing number facts to secondary school is inefficient use of scarce time.

In conclusion, it is now acknowledged that memory is a very complex subject and that this is a very over-simplified explanation, but it is hoped has allowed a good case to be made for the reintroduction of teaching and learning of automaticity.

9.6 Memory Work

Repetition is the key to all memory work, be it training for sport, learning to play a musical instrument, driving a car and evaluating sums. It was believed at the time of Bramald's Study that the repetitions associated with using the dials (physical activity) would develop automaticity for adding number fact (up to 10). Indeed, this was one of his conclusions. However, that appears to be incorrect and the apparent automaticity that the treatment pupils achieved through completing the 15 memory lessons were *making a start* in hardwiring adding number facts up to 10 using verbal/auditory methods. The crux of effective memory work is the repetitions must use the same medium as the one that is being hardwired.

The above activities, except memorising, are physical ones that require physical activities to hard-wire them into "muscle" memories. However, *number facts are word-based* and that explains why repeated use of dials – a physical activity – is largely ineffective in hardwiring number facts; their hardwiring must come through appropriate *verbal/auditory* repetitions, as are now considered.

9.6.1 Hardwiring Number Facts

Currently, such memory work that is being carried out in schools can be represented schematically as:

Current Practice

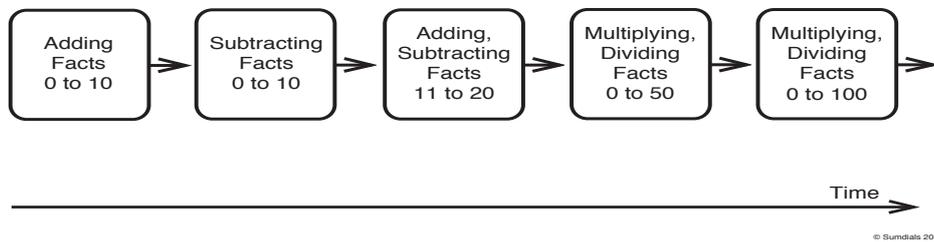


Figure 9.26: Current practice using a linear approach to memorising number facts.

This can be described as a linear box-ticking approach in which each stage is learned in turn. However, it is only effective over the short-term and this is, in effect, what students do as they revise for exams only to discover shortly after that most of the memorised information is progressively forgotten. In many cases, this is of little consequence – provided they passed their exams! Its attraction to schools is it does not require much time as pupils are prepared for a class test at, say, the end of the week. The box can be ticked once the test confirms that the requisite memory work has been successfully carried out, and the next item on the list is taught.

Unfortunately, it has usually been forgotten as many teachers have discovered when they revisit past items. To describe an actual experience of this point, Peter was preparing a class for a test at the end of a week, but realised they would not be ready by then, so he told the class he was postponing the test until the following week. The response was: ‘But we will have *forgotten* it by then!’ ‘Out of the mouths ...’!

The following schematic diagram, usually referred to as the “Swiss roll” model, illustrates the effective hardwiring process with its three components:

- Memorisation.
- Revisiting.
- Consolidation.

"Hardwiring" Number Facts

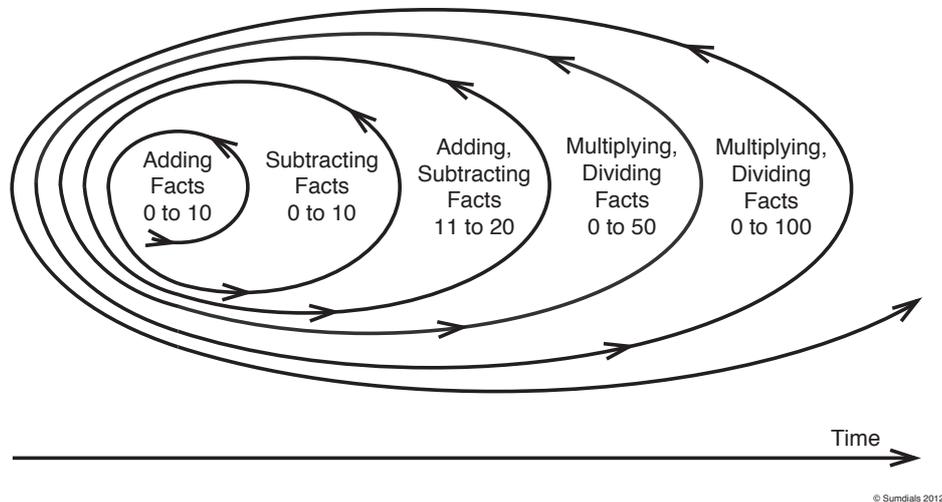


Figure.9.27: The Swiss roll model for hardwiring Number Facts.

The memory work associated with the first stage is the same repetitive learning that would be used in the linear model and is the initial hardwiring step of the consolidation process, as would have taken place in Bramald's Study. It becomes almost continuous and may best be achieved through a little-and-often approach such as a short session (10/15 minutes) at the beginning of *each* day. This was the usual practice in former times and involved *verbal* repetitions that remain an essential element of hardwiring number facts. The diagram starts with memorising the adding number facts up to 10 to be followed by the subtracting facts up to 10. However, the crucially different step is the subtracting ones *are merged or intermingled* with the adding ones unlike the linear approach where they are kept in their separate boxes.

This is repeated for each successive new tranche of number facts and continues until all 693 number facts have been hardwired into long-term memories. Thus during the hardwiring phase *any and every* number fact is in play until the task is completed when all the adding/subtracting facts up to 20 and the multiplying/dividing number facts up to 100 are hardwired. Having achieved that, it will still be necessary to

revisit all the number facts from time to time. Such times are when new facts are being added, such as the value of π or the imperial/metric conversion factors.

Recent research findings confirm hardwiring involves permanent brain changes and number facts will continue to be forgotten until the changes have been consolidated and embedded through completion of the appropriate neurological processes (Menon, V., 2014). It is up to the professional judgment of teachers to determine the contents and activities of each daily session, but, again, the aim is to achieve the *accurate and instant recall* (of all number facts) *without any conscious mental effort*. That implies finger-clicking speeds of response and it can safely be assumed that even short delays in answering indicate that thinking and mental effort - calculation - is taking place.

Because of its importance it is mentioned again that one Swiss teacher uses a minute-glass and her aim is that each pupil will be able to answer 20 random number questions in one minute or three seconds per question that includes the time taken to ask the question. An instructive experience now is that after some InCAS assessments the pupils were given a memory test that entailed answering 10 pre-recorded number fact questions on each of:

- Adding up to 10
- Subtracting within 10
- Adding up to 20
- Subtracting within 20.

The pupils enjoyed their new “quizzes”, but the teachers felt not enough time had been allowed for each question; it was 6.5 seconds compared with the 3 seconds in Switzerland. Clearly, the teachers believed the only way to answer such questions was to *calculate* the answers and the concept of recalling previously embedded number facts was unknown to them.

This seems to be generally accepted as, for example, 30 seconds being allowed to answer each question in the Mental Arithmetic module (of InCAS). When this length was queried, it was explained that it should be sufficient for pupils to answer enough

questions to avoid them becoming discouraged. Thus the module should more correctly be known as: “Mental Strategies”, being the second type of applying strategies, as identified by Crawford (p.200).

30 Second Challenge, Lock (2008) is relevant because it is made-up of 30 seconds mental arithmetic questions that have ten steps, as the following examples shows:

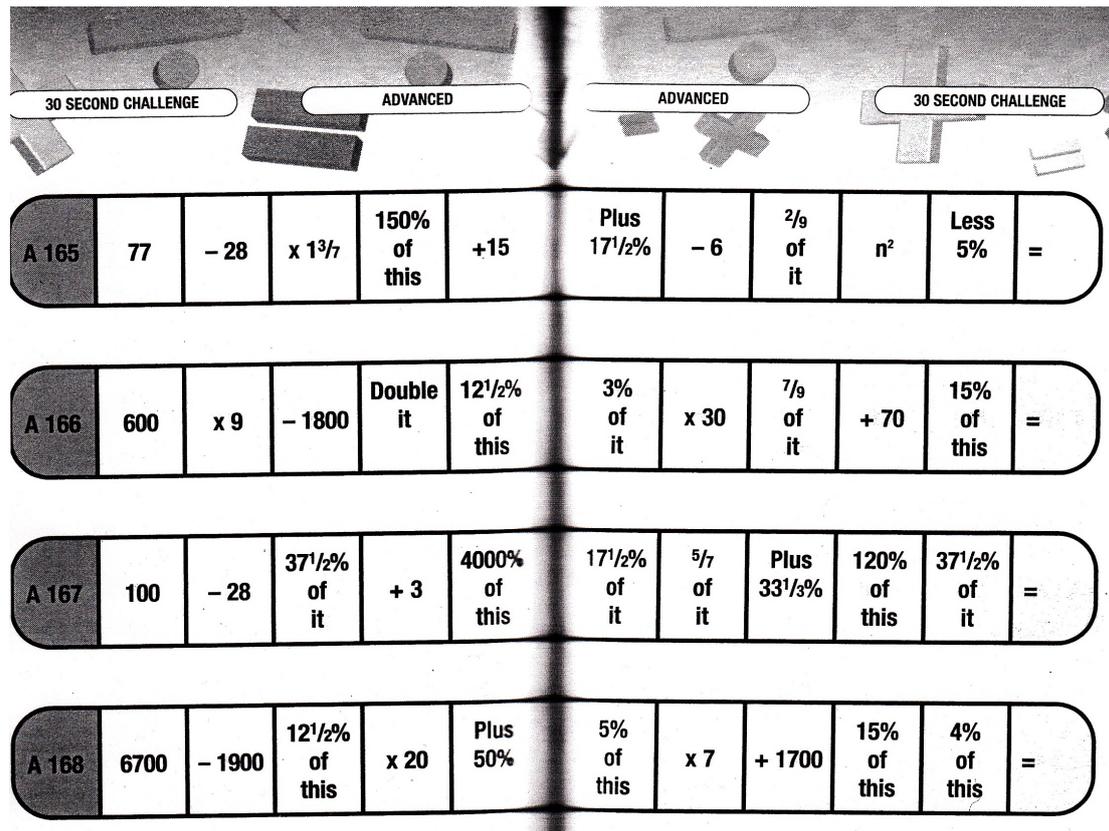


Illustration 9.7: The reproduction of a page of *30 Second Challenge*, by Lock.

Only three seconds per step are allowed on average and most of them involve calculations and not just number-fact recalls. In short, it requires finger-clicking speeds to answer each step and being able to do so would almost certainly gain entry to Schoenfeld’s mathematical enterprise!

9.6.2 Recollections

Now is an appropriate time to return to the before-and-after New Maths comparisons based on my own recollections of a junior school staffed exclusively by transmitters operating in environments where:

- Good education, based on acquiring a good working knowledge of the 3Rs, was generally seen as the road to advancement. Parents expected their children to learn and supported the schools in achieving this. One consequence was that classes with over 50 pupils (not uncommon) needed transmitter teachers to cope.
- The number emphasis was on teaching (a) factual knowledge – number facts - and (b) procedural knowledge - learning the ‘tried and tested’ algorithms being the distillation of 3,000 years’ development and experience. On a personal note, I have no recollection whatsoever of learning my number facts and would not be in the least surprised if it involved drill or rote learning. (I believe I knew them by the time I was nine years old.)
- Subject teachers were the norm once we were about nine years old and many had been retrained having become redundant from manufacturing during the Great Depression; their arithmetic was real because they used experience-based practical examples instead of ones from textbooks. This also meant they were natural connectors (behind their transmission styles).
- Conceptual knowledge and, hence, understanding seemed not to have been an issue. Simply, applying the appropriate procedures, retrieving the relevant number facts and getting the right answers made everyone happy.
- Lastly, head teachers had almost total autonomy and political oversight was virtually unknown, according to Brown, (Thompson, 1999 p.4).

This section is concluded with some personal examples based on experiences when I was conducting workshops demonstrating the paramount importance of automaticity. For example, to show teachers what automaticity is, they were asked if they could recall from when they were at primary school, as appropriate:

- The registration number of the family car.
- Their home telephone number.
- Their mother’s ‘divi’ number.

To *their* surprise they answered the relevant questions without hesitation, so they were then asked the corresponding numbers when they were at training college and hardly any of them could give an answer. My omission then was to ask them if they knew the registration numbers their current cars. However, the point had already been made that automaticity is real - even though it may never have been mentioned during their training or subsequently.

9.6.3 Observation

It is trusted that a convincing case has been made for automaticity's potential in enabling pupils to develop fluency and mastery with number. However, automaticity only comes into play to expedite solving sums once the basic arithmetic processes can be applied confidently; that is by no means the case. A means to achieve this is now proposed.

9.7 A Remedy

In an ideal world, there would be sufficient properly qualified arithmetic teachers for all primary schools to put its teaching and learning on a sound long-term footing. Importantly, establishing formally such arrangements would signal that arithmetic is an essential life skill. Again, the germ of such an arrangement was encountered in my longitudinal Study when one teacher, who was good at arithmetic, took her colleagues' number lessons while they took her other subject lessons and it worked well for everyone - colleagues and pupils (p.186)!

And to avoid any doubt, good personal automaticity would be an essential qualification to become such a teacher.

9.7.1 Scaffolding

Such an ideal does not exist and the need is to create suitable scaffolding in conjunction with primary teachers themselves to enhance their capabilities to become more confident about teaching number. The first step could be to adopt a narrow definition of arithmetic, such as:

- Arithmetic includes the manipulation of integer numbers (0 to 9) in counting, adding, subtracting, multiplying and dividing, and applying

the results to answer everyday situations. These activities would be enabled by the *recall* of previously hardwired number facts from long-term memories.

This definition is deliberately constrained so that teaching of rational numbers (fractions, decimals and percentages) will only be attempted once confidence in the basic processes and their applications has been achieved. Simply, this is a limited objective definition that should be ‘fit for purpose’, but subject to modification by the teachers as they build up their own self-confidences through a series of action research spirals (p.162). Progress would be confirmed by more positive attitudes to number and better results by their pupils.

It has been mentioned several times that arithmetic (and mathematics) is a very hierarchical subject as is illustrated by the following abbreviated table:

Number Hierarchy Table

Type of Number	Examples	Descriptions	Years
Counting/Natural	1, 2, 3, 4, etc.		1
Whole Numbers	0, 1, 2, 3, 4, etc.		2,3,4
Rational Numbers	$\frac{3}{4}, -5 = \frac{-10}{2}$	Ratios of two integers – fractions, decimals and percentages.	5,6
Real Numbers	$\sqrt{2}, \pi, e$, etc.		Secondary
Complex Numbers	$a + bi$	Where ‘i’ = $\sqrt{-1}$ while ‘a’ and ‘b’ are real numbers.	Secondary/ University

Table 9.42: A simplified Number Hierarchy table.

It is stressed this hierarchy table is only an indicative subject background, but it should be sufficient to show what is meant by emphasising that arithmetic is the foundation of mathematics and both are hierarchical. Also, there should be sufficient

time for pupils in their first four years to achieve the limited objective of mastering the basic processes of arithmetic and acquiring *complete* automaticity.

9.7.2 Framework

This section shares the ground above (pp.184-186) from a different perspective. Both are included because of the importance in raising teachers' number capabilities. The aim is to establish a macro setting for the scaffolding that gains full teacher acceptance that:

- Raising the overall attainments of the large middle group will be most easily achieved through improving number skills. This will increase the country's overall competitiveness – something that is urgently needed.
- Research for the Basic Skills Agency by Bynner and Parsons (2000) suggests:
 - (i) Improving number skills led to improved word skills, but it was a one-way street in that improved word skills have no impact on number skills.
 - (ii) Good number skills *do* enhance employment prospects more than reading and writing do according to Dowker (2005, p.14).

In support of the first point, one teacher using the *Sumdials'* approach realised that pupils spontaneously started to write sentences for the first time as though they had acquired a new skill or improved self-confidence.

As further background, it was striking that the participating teachers in the Follow-on Study became enthusiastic once they realised that it would involve their middle pupils – because they are neglected in the sense that extra resources are almost invariably made available for the top and bottom groups. The hope must be that participating teachers will convince themselves that:

- Becoming number-ready first (before any formal number learning) does lead to future benefit.
- Their involvements in drawing-up their own teaching programmes in response to their pupils' needs will generate greater confidence with number.

This Study has now become located in teaching and learning basic arithmetic and the comments that follow relate only to arithmetic (as defined above):

- This Study's definition of arithmetic is deliberately restrictive, aimed at helping pupils to learn the basics of number and arithmetic. It is equivalent to learning the alphabet and how to apply its letters to write simple sentences using appropriate grammar of "The cat sat on the mat" genre.
- The near universal base 10 number system and its standard procedures took more than 3,000 years to develop and refine to become completely reliable. Furthermore, it had become familiar to all generations and this allows the older ones to help the younger ones.
- Much has been learned recently about learning methods for arithmetic and especially on embedding number facts in long-term memories that contributes so much in helping pupils to become more effective in their uses of arithmetic – and increase their confidence.
- The contribution that a socialising process makes in learning number is likely to be considerable, especially during the earlier years. To use the dominoes example, it cannot be said precisely how or when it contributed to the pupils becoming number-ready, but it is likely that peer-to-peer learning played a part, (p.196).
- It is very probable that pupils subconsciously think and reflect about their active learning games without realising it. This is probably a normal part of the learning processes.

It is believed that these points can be used in preparing pupils to learn arithmetic in a new way and in support of this the teaching plans of the *Sumdials'* approach are already being modified in readiness to include the teachers' suggestions and requirements.

However, the scaffolding should only be seen as a holding operation.

9.8 Contribution of Number Resources

Reflection about an apparently trivial episode subsequent to the Pilot and Empirical Studies revealed the key that make dials effective in helping pupils to learn number. To describe the episode, the *Sumboard 10* is another active learning resource – but not a manipulative - that has been devised (by Peter) to help pupils in becoming number-ready, as illustrated:

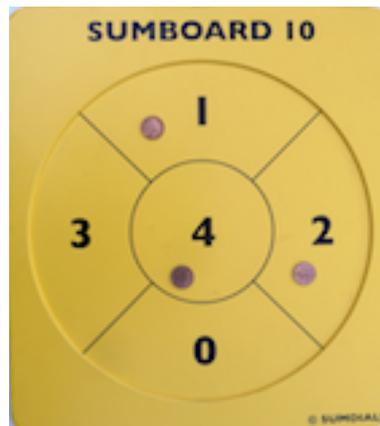


Illustration 9.8: A Sumboard 10 is a classroom adaptation of a dartboard

A *Sumboard 10* uses three coins or counters (instead of darts) being flicked onto it by pupils taking turns. Several pupils were playing together – socialising - and one player scored a 1, 2 and a 4, but for all that it meant to her she could have described it as a red, a blue and a green. However, one of the others said it was: ‘1 + 2 + 4’ signifying she realised they could be added together, even though she could not actually do it. However, if a dial had been produced at that point – because the connection had been made – then the pupil would have discovered addition because *she was number-ready* whereas her classmate was not. More importantly, it illustrated the contribution of Resnick’s socialising insight towards learning number.

What other evidence supports this? Bramald’s study took place when the pupils were likely to have been *number-ready* because they were one year older than those of the current Studies. It can be postulated that the improved attainments found in his treatment pupils in the longitudinal follow-up five years later were the result of his six weeks intervention making a start in establishing robust internal models of the adding process that were subsequently adapted to the other arithmetic processes, (p.35/36).

In a different vein, it is probable that a Swiss pupil who was using an abacus was number-ready, but an abacus is a calculating tool and not a manipulative for developing number skills. Thus, an abacus needs the appropriate number skills before it can be used, whereas the dials are a means of acquiring skills and having done so – that is established a robust internal model of number? – to progress rapidly. It can also be suggested that the pupil in the Qualitative Study (p.40) who described a dial as ‘cool’ was number-ready and her response indicated she had instantly acquired an improved internal model of number – and it made sense to her in the same way that happened with this Swiss pupil.

One conclusion can be drawn and it is:

- Pupils *must* be number-ready before they will benefit from using dedicated manipulatives to prepare them for learning new arithmetical experiences.

That is the prerequisite for learning number and the introduction of manipulatives before pupils are number-ready may explain their mixed results. This finding is now being emphasised during the implementation visits to the schools taking part in the Follow-up Study.

9.9 Developing Mastery and Automaticity with Number

9.9 Introduction

Even though it would not be possible to tackle the issue of developing fluency and automaticity in number directly, it was decided to continue with the main Empirical Study while explanations were sought for the lack of automaticity in pupils’ responses and their reliance on counting ‘strategies’.

9.9.1 Explanation and Personal Reflection

The Pilot Study showed that the Mental Arithmetic (M/A) scores were appreciably lower than the General Maths (G/A) scores and also for the initial assessments of the Empirical Study. The cause, as already explained, was that mental arithmetic was no

longer being emphasised. In particular, it became clear while watching the pupils during their assessments that the vast majority had no automatic recall of their number facts. The teachers confirmed that this was to be expected as there was not enough time to teach it, but it was sensed that if the time had been available it would not have been used for memory work to develop instant *recall* of the required number facts.

It was apparent that the teachers were happy to use the dials because they found the teaching plans easy to follow and the pupils mostly enjoyed using their dials. Moreover, using them correctly helped the pupils to meet the requirements of their normal learning aims and so became welcome resources for the teachers. This confirmed that the manipulatives – the dials - were effective, but it is pointed out that using them was a consequence of the research; the participating teachers had not *chosen* to use them. However, it became clear that not insisting on carrying out the memory work needed to achieve automaticity avoided a needless impasse.

The origin of the impasse was my assumption that automaticity would be taught as standard practice as it was when we were at school and, indeed, during Bramald's (2001) study. Thus, it was never discussed when the arrangements were being made with the schools for the studies and the relevant point about manipulatives became the focus for the participating teachers who became happy to use the dials because their potentials could be seen following their preparatory training for this Study. Effectively, the teachers saw the correct use of the manipulatives as an end in itself, rather than contributing to arithmetical fluency and mastery.

It was clear that it would be easier to find explanations for the lack of emphasis on mental arithmetic if contact with the schools was maintained. Thus, a virtue was made out of continuing with the assessments and, in any event, that was desirable if the credibility of school research in general was to be maintained. Fortuitously, at much about the same time of the potential impasse, the head teacher of one of the participating schools wondered if the pupils were starting their formal subject learning (of arithmetic) before they were number-ready. To make a personal observation, I was almost certainly 'fed' the point, but I was more than happy to run with it because an experienced teacher (at another school), who took part in both Studies, had previously made a related point, albeit based on her experiences.

Her Empirical Study pupils had started five months earlier in the session than her Pilot Study ones had and, therefore, were five months younger. However, she wondered if the correct answers of the Empirical pupils, using their dials, actually *meant* anything to them. This was little more than a throwaway remark during some informal chat at the end of the lesson, but she was making the same point that the head teacher had: her pupils were not number-ready. Her remark was taken seriously because we had just observed her using the dials and we regretted we had not videoed it to become part of a training pack for teachers new to the dials!

9.9.2 Conclusions

This analysis explains how true mental arithmetic ceased to receive sufficient emphasis in the curriculum is the consequence of political imperatives that led indirectly to arithmetic being subsumed into the broader subject of mathematics. The need now is to restore stability by establishing the education's long-term strategic objectives and then trusting schools and teachers to implement them. As part of this, it is recommended for number that its importance is acknowledged by:

- Re-instating arithmetic as the foundational subject for mathematics and, as a reminder, up to 90% of the population will only ever need arithmetic.
- Emphasising automaticity will enhance number attainments and help pupils to become fluent with number putting them well on the road to mastering it.
- Developing suitable scaffolding for those teachers who may lack sufficient in-depth number knowledge to teach it effectively.
- Appointing suitably qualified and trained arithmetic teachers who are:
 - (i) Members of the 'mathematical enterprise'
 - (ii) 'With good subject knowledge, and
 - (iii) Enthusiastic about their subject' (Schoenfeld in Grouws 1992, p.349).

In conclusion, the illustration below was produced for primary schools as an alternative to word-based explanations about what was actually happening in their schools and all teachers, without exception, confirmed it is "spot-on"!

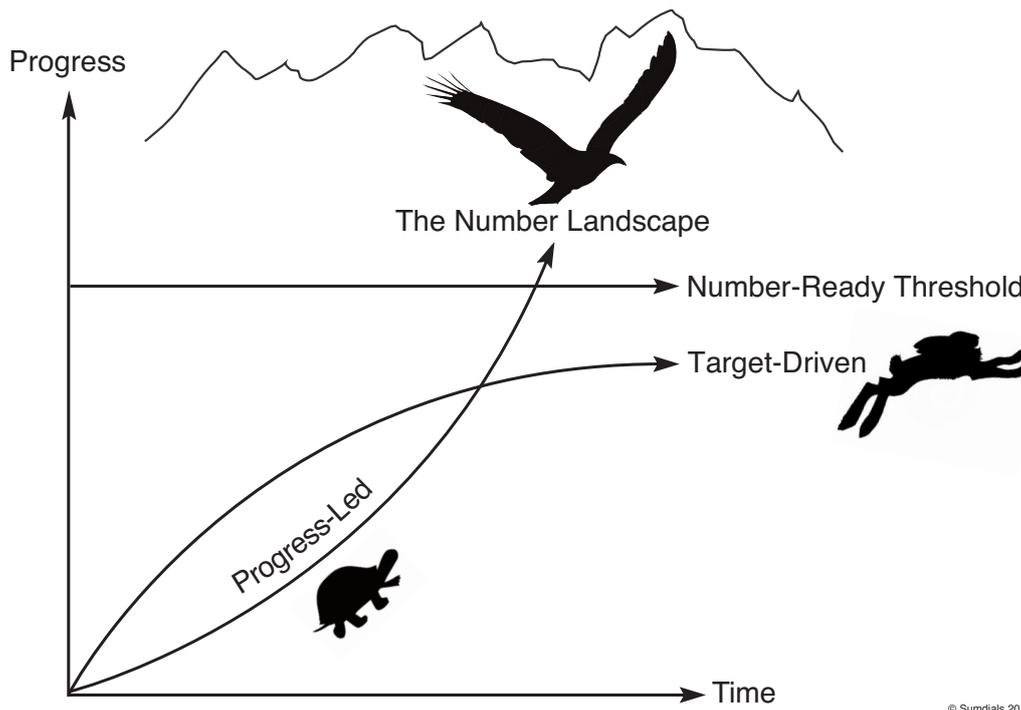


Illustration 9.9: A Schematic Representation of current number learning outcomes.

Its origins can be traced back to a discussion with a secondary school maths teacher and a QIO several years *before* this Study started. To emphasise a point, it originated instinctively when the two straight line axes and the two graph lines were drawn on the back of an envelop as a *representation* that allowed me to make a point very effectively *without using words*. This came naturally to me and is widely used by mathematicians and, in my case, engineers to communicate with one another. Relevantly, Peter realised that his “clients” no longer draw diagrams now and his former colleagues’ experiences with their pupils is similar. In my engineering days, our usual medium of communication was through chalk-drawings on the floor!

Subsequently, the original graph has been embellished to arrive at a “picture” to represent the influence that the dials make to learning number once pupils have become number-ready – *through the establishment of robust internal models of number*. It is now believed that the finding of the Studies should allow many more pupils to be tortoises - instead of becoming demotivated hares – and then with implementation of the above recommendations to become high-flying number eagles!

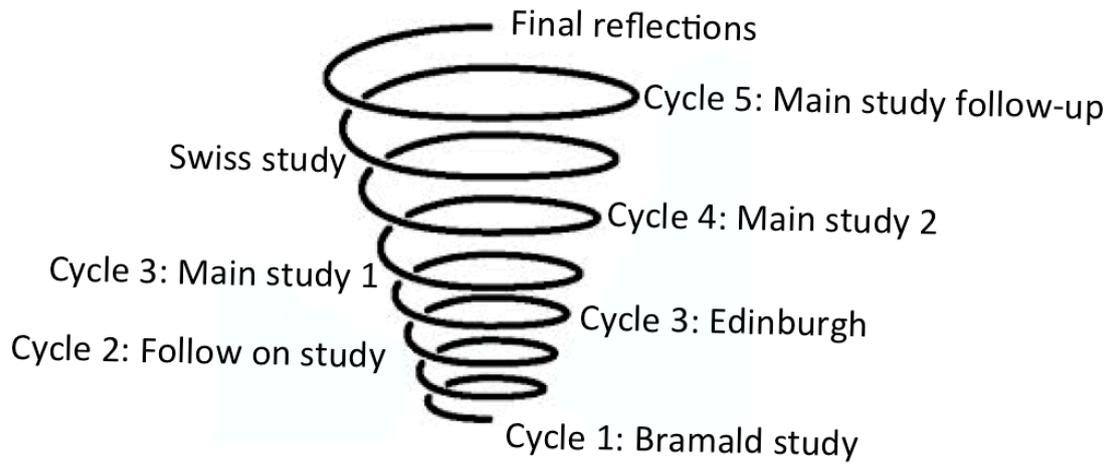
9.10 Concluding Reflection

Having reached this stage, a simple fable explains the long-term effect of abandoning automaticity – the hardwiring of number facts into long-term memories. Not having automaticity is like having cars with only a first and a reverse gear. They are sufficient for learning to drive and indeed to pass driving tests in that hill starts, three-point turns and reversing into gaps to park alongside pavements together with all the other driving maneuvers that can be successfully completed, albeit not very efficiently.

However, driving such cars on busy roads and motorways would be extremely fraught if not impossible. Ultimately, such cars would only be seen on quiet country lanes – because driving on all the other roads would become so hard. In the fullness of time, their drivers came to terms with their limitations while acknowledging that they were falling behind. One day a car with an elderly driver and passenger was lost on such a lane. They asked for directions and the local who helped them was then given a lift because he was going on their way.

He observed that the car was going much faster than his own car and the engine was much quieter. He eventually realised that this was because it had six forward gears, but try as they would the elderly driver and passenger could not get him to understand the reasons and he remained unconvinced of the benefits. In due course, he gratefully went on his way still wondering how he and his fellow drivers were being left behind.

And the point? The large majority of current educationalists, be they policy makers or teachers, attended school after New Maths was introduced and accepted the seductive argument that understanding and applying first principles – strategies – was the best way to master number in spite of the accumulating evidence to the contrary. The reality is they cannot comprehend the transforming contribution that accrues through automaticity in mastering number and so do not instinctively appreciate the need to make priority time available for it in the curriculum. They are also being left behind.

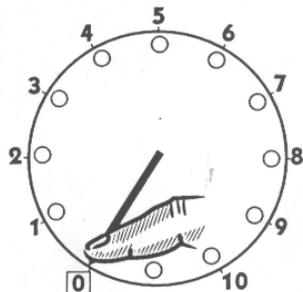
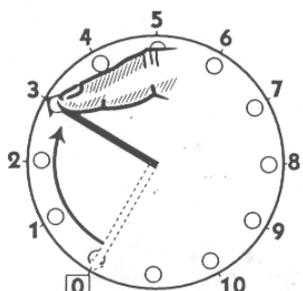
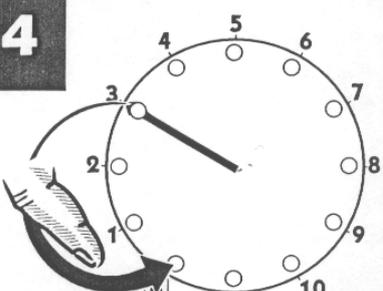
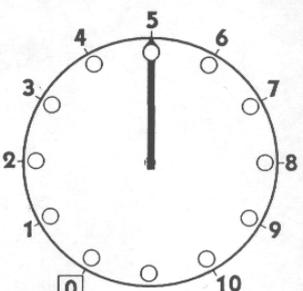


Appendix 1.1

Entry-level Instruction Sheet

To add $3 + 2$ using your **SUNDIAL 10**, this is what you do

The dial has holes (like an old fashioned telephone dial) of a size suitable for a child's index finger.

<p>1</p> 	<p>2</p> 
<p>Set arrow to 0 by dialing with your finger until the arrow points to 0</p>	<p>The arrow should now be pointing to 0 This is the start position</p>
<p>3</p> 	<p>4</p> 
<p>Dial in 3, by moving the dial with your finger from 0 to 3</p>	<p>Leaving the arrow pointing to 3, remove your finger and replace it at 0</p>
<p>5</p> 	<p>6</p> 
<p>Dial in 2 by moving the dial with your finger from 0 to 2</p>	<p>Remove your finger from the dial. Arrow now points to the answer : 5</p>

Instruction Sheet for adding two single digit numbers together.

Appendix 1.2

Quality Study interview Questionnaire

Factual Questions

What is Your Name?

What is Your Date of Birth?

What was Your Primary School?

What was the Name of Your P2/Y1 Teacher?

Did you ever use a SUMDIAL 10?

If 'yes', can you remember how to use it? If 'no', would you like to see how it works?

Did you like using it?

Would you have liked to use it?

Ability in Mathematics

Are you: Very good/ quite good/ average/ not very good/ poor at it?

Why did you give your answer?

Attitude to Mathematics

Do you: Really like it/ quite like it/ not bothered/ not like it very much/ not like it?
What is your reason (for your choice)?

Do you think mathematics is: Very important/ quite important/ 'so so'/ not very important/unimportant? What are your reasons (for your choice)?

What are your favourite subjects? Why do you like them?

Does your ability in mathematics affect your success in your favourite subjects? If so, in what ways does it affect them?

“Minitest” (mental)

1. $15 + 9 = ?$

2. $46 - 8 = ?$

3. $7 \times 9 = ?$

4. $600/20 = ?$

5. $1/3$ of 21 =?

6. $3/4$ of 24 =?

7. $1/2$ of $4/5 = ?$

8. 20% of £20.00 =?

Did you enjoy your “Minitest”?

Do you have any questions or comments?

(Acknowledge appreciation for helping in the research)

Appendix 5.3

School Activity Analysis Table

Title	Main Theme(s)	Repeat/Other Themes
Rationale of the Sundials System for Learning Number	Cultures of Developed World are word based.	
Automaticity Study	Defining/explaining automaticity and its contribution to number.	Pupil-led/target-driven chart.
Study Reflections	Lack of time for consolidation due to overloaded curricula.	The importance of automaticity. Pupil-lead/target-driven chart.
Arithmetic	Starting formal learning too early	Insufficient time to consolidate learning - especially number facts.
Forward to Arithmetic?	Target-setting issues. Finger counting become <i>external</i> models. No subitizing. M/A no-go area. Arithmetic as foundation subject for maths.	Busyness, time pressures, new initiatives (top-down). Starting formal learning too early. Lack of time for consolidation of learning. Implications of three ability groups.
Behind the Curve	Re-instating arithmetic as a stand-alone subject, being the foundation of maths.	Importance of automaticity. Starting formal learning too early. Breakdown of compact between homes and schools in sharing learning of number. Implications of three ability groups.
Points for Learning Number		Repeating/ consolidating many of the above themes.
“If it ain’t broke . . .”	Need to “blitz” memory work.	Implications of three ability groups. Reinstating arithmetic as a stand-alone subject. Importance of consolidation. Recapping earlier themes.
The Number Landscape	Linear v “Swiss Roll” models for memory work	Three ability groups. Waiting until children are number-ready. Consolidation. Hare/tortoise chart.
Acquiring Number Sense	Learning arithmetic as a socialising process.	Only 10/15% of population will ever need maths in their lives.

APPENDIX 7.4

NUMBER-READY CHECKLIST

Pupil Name:

Class:

Experience & Outcome	Date Achieved	Initials
Can count forward 1 to 10		
Can count forward from a any number within 1 to 10		
Can recognise numerals 1 to 10		
Can write numerals 1 to 10		
Can count backwards 10 to 1		
Can count backwards from any number within 10 to 1		
Understands 0 (Zero)		
Can subitize (recognising the number of items in a set - up to 5 or 6 - by just glancing at it – no counting)		
Understands the equivalence of quantities, numbers and numerals		
Can play “Numerosity Snap”		
Can use “more”/”less” correctly		
Can use “greater”/”smaller” correctly		
Can use “same”/”equal” correctly		
Can use “enough”/”not enough” correctly		

“COMPENDIUM OF GAMES”

1. Numerosity Snap
2. Dominoes (Standard)
3. Ludo
4. Snakes & Ladders
5. School-based Board Games, e.g. 4 or more.

© Sundials, 09.01.14

Bibliography

- ‘Assessment of Achievement Programme Fifth Survey of Attainment in Mathematics’ (1997), Edinburgh: The Scottish Office.
- Ainsworth, S., Bibby, P. and Wood, D. (2002) ‘Examining the effects of different multiple representational systems in learning primary mathematics’, *The Journal of the Learning Sciences* Vol. 11(1) (pp.25-61).
- Askew, M. (1999) ‘It ain’t (just) what you do: effective teachers of mathematics’ in Thompson, I. (ed.) *Issues in teaching numeracy in primary schools* (pp.91-102). Buckingham: Open University Press.
- Barmby, P., Harries, A.V., Higgins, S.E. and Suggate, J. (In Press) ‘The array representation and primary children’s understanding and reasoning in multiplication’: *Education Studies in Mathematics*.
- Baumfield, V., Hall, E. and Wall, K. (2008) *Action Research in the Classroom*, London: Sage Publications.
- Beishuizen, M. (1999) ‘The empty number line as a new model’, in Thompson, I. (ed.) *Issues in teaching numeracy in primary schools* (pp.157–168). Buckingham: Open University Press.
- Bell, J. (2005) *Doing your Research Project*, Maidenhead: Open University Press.
- Blöte, A.W., Klein, A.S., and Beishuizen, M. (2000) ‘Mental computation and conceptual understanding’. *Learning and Instruction* Vol. 10 (pp.221–247).
- Bond, T.G. and Fox, C.M. (2001) *Applying the Rasch Model*. Mahwah: Laurence Erlbaum.
- Bramald, R. (2001) *An evaluation report into the effectiveness of Sumdials’ entry level materials for young children learning mathematics* Newcastle: University of Newcastle.
- Brown, M. (1999) ‘Swings of the Pendulum’ in Thompson, I. (ed.) *Issues in teaching numeracy in primary schools* (pp.3-16). Buckingham: Open University Press.
- Bruner, J.S. (1966) *Towards a Theory of Instruction*. Boston: Harvard University Press.
- Bynner, J. and Parsons, S. (2001) ‘Qualifications, basic skills and accelerating social exclusion’ *Journal of Education and Work*, 14(3) (pp.279-291).
- Butterworth, B. (1999) *What Counts*, New York: The free Press

Carpenter, T.P., Fennema, E. and Franke, M.L. (1996) 'Cognitively Guided Instruction: A Knowledge Base for Reform in Primary mathematics Instruction', *The Elementary School Journal*, Vol. 97 (1), (pp. 3–20).

Carpenter, T.P., Fennema, E., Peterson, P.L., Chiang, C.P. and Loeff, M. (1989) 'Using knowledge of children's mathematics thinking in classroom teaching: An experimental study', *American Educational Research Journal*, Vol. 26 (4), (pp.499–552).'

Carbonneau, K.J., Marley, S.C. and Selig, J.P. (2013) 'A Meta-Analysis of the Efficacy of Teaching Mathematics with Concrete Materials', *Journal of Educational Psychology*, Vol. 105, (2) (pp.380–400).

Chandler, D. and Torbert, B. (2003) 'Transforming inquiry and action interweaving 27 flavors of action research' *Action Research*. Vol. 1(2), (pp.133-152).

Coe, R. (2013) 'Improving Education: A triumph of hope over experience' *Inaugural Lecture* Durham University: Durham.

Cohen, L. and Manion, L. (1994) 'Research methods in education' *Occasions when action research as a method is appropriate*, London: Routledge.

Cohen, L., Manion, L. and Morrison, K. (2007) *Research Methods in Education*, New York: Routledge.

Davenport, P. and Howe, C. (1999) 'Conceptual Gain and Successful Problem-solving in Primary School Mathematics', *Educational Studies*. Vol. 25, (1).

Davis, P.J. and Hersh, R., (1980) 'Nonanalytic Aspects of Mathematics' *The Mathematical Experience* (pp.301–316) Boston: Birkhäuser.

Dehaene, S. (1997) *The Number Sense*, London, Penguin Books.

Demetriou, A., Christou, C., Spanoudis, G. and Platsidou, M. (2002) 'The Development of Mental Processing: Efficiency, Working Memory and Thinking'. Boston: Blackwell Press.

diSessa, A.A. (1987) cited by Goldin, G.A. in Janvier, C. (ed.) *Problems of Representation in the Teaching and Learning of Mathematics*, (p.131) Erlbaum: Hillsdale, N.J.

diSessa, A.A. (1993) 'Towards an epistemology of physics', *Cognition and Instruction*, Vol. 10(2-3) (pp.105-225).

Dowker, A. (2005). *Individual Differences in Arithmetic*, Hove: Psychological Press.

Dowker, A. (2007) 'What can intervention tell us about the development of arithmetic?' *Educational and child psychology*, Vol. (24) (pp. 64-82).

- Duncan, A., (2006) 'Should we bin our Dienes (and the Decomposition Algorithm with it)?' *Primary Mathematics*. Vol. 10 (2), (pp. 18–22),
- Ellis, P.D. (2011) *The Essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge: Cambridge University Press.
- Else-Quest, N.M., Hyde, J.S. and Linn, M.C. (2010) 'Cross-national patterns of gender differences in mathematics: a meta-analysis'. *Psychological Bulletin*, 136(1), (p.103).
- Ernest, P. (1983) 'Thinking of A Funny Line', *Mathematical Education For Teaching*, Vol. 4 (2) (pp.30 – 40).
- Ernest, P. (1985) 'The Number Line as a Teaching Aid', *Educational Studies in Mathematics*, Vol. 16 (4) (pp. 411-424).
- Ernest, P. (1986) Mental Number Line Images, *Teaching Mathematics and its Applications*, Vol. 5 (1) (pp. 1 – 5).
- Ernest, P. (1988) The Mental Number Lines of a Sample of Adults, *International Journal of Mathematical Education in Science and Technology*, Vol. 15 (3) (pp.389–395).
- Gardner, H., (2006) 'The Development and Education of the Mind' Oxon: Routledge.
- Gibson, I.R., Goldman, F.H. and Grinfeld, M. (2004) 'Mathematics at the School-University Interface in Scotland' Unpublished.
- Goldin, G.A. (1987) in Janvier, C. (ed.) *Problems of Representation in the Teaching and Learning of Mathematics*, (p.131) Erlbaum: Hillsdale, N.J.
- Greeno, J.G. (1998) 'Thinking practices in mathematics and science learning' Mahaw: Lawrence Erlbaum Associates.
- Gross, J. (2007) 'Supporting children with gaps in their mathematical understanding: the impact of the National Numeracy Strategy (NNS) on children who find mathematics difficult.' *Educational and Child Psychology*, Vol. 24 (2), (pp.146-156).
- Gunderson, E.A., Ramirez, G., Levine, S.C. and Beilock, S.L. (2012) 'The role of parents and teachers in the development of gender-related attitudes.' *Sex Roles*, Vol. 66 (3-4) (pp.153-166).
- Goldin, G.A. (1998) 'Representational systems, learning, and problem mathematics': *Journal of Mathematical Behaviour*, Vol. 17 (2), (pp. 137–165).
- Harford, T. (2011) 'Adapt' London: Little, Brown.
- Harries, A.V. (undated) 'Representing mathematical ideas' Durham University.

Harries, A.V., Barmby, P. (undated) 'Representing and Understanding Multiplication' Durham University.

Hembree, R. and Dessart, D.J. (1986) 'Effects of Hand-Held Calculators in Precollege Mathematics Education: a Meta-Analysis' *Journal for Research in Mathematics Education*, Vol.17 (2), (pp. 83–99).

Hughes, M. (1986) 'Children and Number', Oxford: Basil Blackwell.

Hughes, M. (2007) 'Enhancing primary literacy and mathematics through home-school knowledge exchange' London: Economic and Social Research Council.

Hollis, S. and Campbell, F. (1999) 'What is meant by intention to treat analysis? Survey of published randomised control trials'. *British Medical Journal* Vol. 319 (7211) (pp.670-674).

Jensen, R.J. (1993) (ed.) *Research Ideas for the Classroom: Early Childhood Mathematics*. New York: Macmillan.

Jordan, N.C., Kaplan, D., Ramineni, C. and Locuniak, M.N. (2009) 'Early math matters: kindergarten number competence and later mathematics outcomes' *Developmental Psychology*, Vol. 45(3), (p.850).

Kandel, E.R. (2006) 'In Search of Memory' New York: W.W. Norton & Co.

Kaput, J.J. (1987) in Janvier, C., (ed.) *Problems of Representation in the Teaching and Learning of Mathematics*, (p.131) Erlbaum: Hillsdale, N.J.

Kinnear, P.R. and Grey, C.D. (2010) *IBM SPSS Statistics 18 made simple*. Hove: Psychology Press.

Klein, A.S., Beishuizen, M. and Traffers, A. (1998). The empty number line in Dutch second grade: realistic v gradual program design *Journal for Research in Mathematics Education*, Vol. 29 (4), (pp.443–464).

Kroesbergen, E.H. and Van Luit, J.E. (2003) 'Mathematics Interventions for Children with Special Educational Needs A Meta-Analysis' *Remedial and Special Education*, 24(2), Vol. (pp.97-114).

Krutetskii, V.A. (1976) 'The Psychology of Mathematical Abilities in Schoolchildren' Chicago: The University of Chicago Press.

Leash, R., Post, T. and Behr, M. (1987) in Janvier, C. (ed.) *Problems of Representation in the Teaching and Learning of Mathematics*, (p.131) Erlbaum: Hillsdale, N.J.

Lock, D.L (2008) *Daily Mail 30 Second Challenge*, Chester: Marks & Spencer.

Long, C., Wendt, H. and Dunne, T. (2011) Applying Rasch measurement in mathematics education research: steps towards a triangulated investigation in the

multiplicative conceptual field. *Educational Research and Evaluation*, Vol. 17. (pp.387-407).

Ma, L. (1999) 'Knowing and Teaching Elementary Mathematics' Mahaw: Lawrence Erlbaum Associates.

Malofeeva, E. (2009) 'Meta-Analysis of Math Instruction with Young Children' Saarbrucken: VDM Verlag Dr. Muller.

Menon, V. (2014) 'New research sheds light on how children's brains memorize facts' *News Center, Stanford Medicine*.

Montague-Smith, A. (2002) 'Mathematics in Nursery Education' London: David Fulton.

Morris, S.B. (2003) 'Estimating Effect Size from the Pretest-Posttest-Control Design' *Society for Industrial and Organizational Psychology*, Orlando: Conference.

Munn, P. and Reason, R. (2007) 'Arithmetical difficulties: developmental and instructional perspectives.' *Educational and child psychology*, Vol. 24 (2) (pp.5-15).

Nisbet, I. (2013) 'Is there a place for China's wise *laoshi*?' *Times Education Supplement Scotland*, 14th April, (p.25).

Noffke, S. and Somekh, B. (2005) 'Action Research' in Somekh, B and Lewin, C. (eds.) *Theory and Methods in Social Research*, (pp.94-101) London: Sage Publications.

Nunes, T. (1992) 'Ethnomathematics and everyday cognition' in Grouws, D.A. (ed.) *Handbook of Research On Mathematical Teaching and Learning*, (pp. 557-574). New York: Macmillan,

OECD (2007) '*Understanding the Brain: The Birth of a Learning Science*', OECD Publishing.

OECD (2013), *Education at a Glance 2013*'' *OECD Indicators*, OECD Publishing.

Resnick, I. B. and Ford, W.W. (1981) 'The Psychology of Mathematics Instruction' Hillsdale: Laurence Erlbaum Associates.

Reynolds, D. and Muijs, D. (1999) 'Numeracy matters: contemporary policy issues in the teaching of mathematics' in Thompson, I. (ed.) *Issues in teaching numeracy in primary schools*, (pp.17-26). Buckingham: Open University Press.

Sarama, J. and Clements, D.H. (2009) *Early 'Childhood Mathematics Education Research'*, New York: Routledge.

Schoenfeld, A.H. (1992) 'Learning to think mathematically: Problem solving, metacognition, and sense making in Mathematics' in Grouws, D.A. (ed.) *Handbook of Research On Mathematical Teaching and Learning*, (pp.334-370). New York: Macmillan.

Scottish Survey of Literacy and Numeracy (SSLN) (2014), Edinburgh: Scottish Government.

Seeger, F. (1993) 'Representations in the mathematics classroom: Reflections and constructions.' in Seeger, F., Voigt, J. & Waschescio, U. (eds.) *The Culture of the Mathematics Classroom*, (pp. 308–343). Cambridge: University Press.

Sierpinska, A. (1994) '*On Understanding in Mathematics*', London: Falmer Press.

Smith, I. (1996) 'Supporting teaching for effective learning' Dundee: Scottish Consultative Committee on the Curriculum.

Suggate, S.S. (2011) 'Viewing the Long-Term Effects of Early Reading with an Open Eye' in House, R. (ed.) *Too Much, Too Soon?*, (pp.236-246). Stroud: Hawthorn Press.

Sylva, K. (1994) 'School Influences on Children's Development' *Child Psychology and Psychiatry*, Vol35 (1). (pp.135-170).

Szendrei, J. (1996) 'Concrete Materials in the Classroom' in Bishop *et al* (eds.) *International Handbook of Mathematics Education*, (pp.411-434).

Tariq, V.N. (2002) 'A decline in numeracy skills among bioscience undergraduates' *Journal of Biological Education*, Vol. 36(2), (pp.76-83).

Terzis, V., and Economides, A.A. (2011). 'The acceptance and use of computer based assessment'. *Computers and Education*, Vol. 56(4), (pp.1032-1044).

Threlfall, J. (1996) 'The Role of Practical Apparatus in the Teaching and Learning of Arithmetic'. *Educational Review*, Vol. 48(1), (pp.1-13).

Tymms, P. (2004) 'Are standards rising in English primary schools?' *British Educational Research Journal*, Vol. 30(4), (pp.477-494).

von Glasersfeld, E. (1987) in Janvier, C. (ed.) *Problems of Representation in the Teaching and Learning of Mathematics*, (p.131) Erlbaum: Hillsdale, N.J.

Watson, A. (2003) 'Affordances, Constraints and Attunements in Mathematical Activity' *Proceedings of the British Society for Research into Learning Mathematics* Vol. 23(2), (pp.103–108).

Watson, A. (2007) 'The nature of participation afforded by tasks, questions and prompts in mathematics classrooms' *Research in Mathematics Education* Vol. 9(1), (pp.111-126).