STRATEGY SELECTION IN MENTAL ARITHMETIC PROBLEM SOLVING: A CASE FOR ADAPTIVE, NOT AUTOMATIC SELECTION

Alastair Mckenzie-Kerr

Thesis submitted to Cardiff University
for the degree of
Doctor of Philosophy

September, 2007

STRATEGY SELECTION IN MENTAL ARITHMETIC PROBLEM SOLVING: A CASE FOR ADAPTIVE, NOT AUTOMATIC SELECTION

Alastair Mckenzie-Kerr

Thesis submitted to Cardiff University
for the degree of
Doctor of Philosophy

September, 2007

UMI Number: U585047

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U585047
Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Page intentionally left blank

PUBLICATIONS AND PRESENTATIONS

1. *The findings reported in Chapter 2 of the thesis appear in the following:*

Mckenzie-Kerr, A., & Macken, W.J. (2006) The role of metacognitive mechanisms in mental arithmetic strategy selection. Submitted for publication to the *Journal of Experimental Psychology: Learning, Memory and Cognition.*

2. *Some of the findings reported in Chapter 2 of the thesis also appear in the following published conference proceedings article:*

Mckenzie-Kerr, A., & Macken, W.J. (2006) Metamemory in Strategy Selection: Evidence for a Dissociation in Familiarity-Based Judgments between Semantics and Item Activation Strength. *Proceedings of the 28th Annual Meeting of the Cognitive Science Society*, 1806-1811, Vancouver, BC.

Portions of the findings reported in Chapters 2 and 3 of this thesis were presented as a paper at a meeting of the BPS Cognitive Section, Leeds, September, 2005 and at a meeting of the BPS Cognitive Section, Lancaster, August 2006.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor Bill Macken for all of his guidance and support. For many of the opportunities I have had during the course of the PhD thanks must also go to Professor Dylan Jones.

I should also like to thank some of the people who made the last few years a pleasure. My office mates past and present, Amelia Woodward, Graph Lloyd and Sam Waldron thanks for not having too many strange habits. Also to the Dr.s in 64PP (Phil Morgan, Dr Nick, Rob Hughes, Helen Hodgetts and John Marsh) for many stimulating academic discussions, useful comments and a few beers every now and then. I'm very grateful for 'Irish' Steve, starting me off on an adventure in Visual Basic programming and friends in the Tower for a bit of football among many other things.

During the course of the PhD the support of my parents and grandparents has been invaluable and I need to thank my brother and sister in particular for reminding me of the simpler pleasures in life.

A very big thank you also goes to Bethan Whitston for putting up me with over the last few years. Your patience, love and support have been incredible. I finished it eventually and I promise that tomorrow I will definitely do the washing up!

This work was jointly funded by the EPSRC and QinetiQ.

CONTENTS

Declarations and Statements.. i
Publications and Presentations.. ii
Acknowledgements.. iii
Contents.. iv
Summary.. 1

CHAPTER 1 2
1.1 INTRODUCTION... 2
1.2 THEORETICAL APPROACHES.. 3
1.2.1 Approaches to strategy selection: The 'Automaticity' account......... 6
1.2.1.1 The Adaptive Control of Thought-Rational.......................... 7
1.2.1.2 The Instance theory of Attention and Memory...................... 10
1.2.1.3 The Distribution of Associations model.............................. 12
1.2.1.4 Summary of the Automaticity Accounts............................. 13
1.2.2 Approaches to strategy selection: The 'Adaptive' account.............. 14
1.2.2.1 The Source of Activation Confusion model.......................... 15
1.2.2.2 The Component Power Laws Theory................................. 19
1.2.2.3 The Adaptive Strategy Choice Model................................ 22
1.2.2.4 Summary of the Adaptive Account.................................... 24
1.2.2 Summary of theoretical approaches... 25
1.3 EMPIRICAL PHENOMENA... 27
1.3.1 Strategy selections can be made quickly and accurately................ 27
1.3.2 The relationship between strategy use and solution latencies........... 29
1.3.3 Problem familiarity correlates with reported strategy selections........ 32
1.3.4 Summary of key empirical phenomena....................................... 34
1.4 TASK OUTLINE... 35
1.5 INTRODUCTION SUMMARY.. 36
1.5.1 Empirical Series 1.. 37
1.5.2 Empirical Series 2.. 37

CHAPTER 2.. 39
2.0 ABSTRACT.. 39
2.1 INTRODUCTION... 40
2.2 EXPERIMENT 1a.. 41
2.2.1 Method.. 45
2.2.1.1 Participants... 45
2.2.1.2 Materials & Design... 46
2.2.1.3 Procedure.. 46

2.2.2 Results.. 47

2.2.2.1 Scoring Procedure.. 47

2.2.2.2 Strategy Selection.. 48

2.2.2.3 Selection Latency Analysis....................................... 50

2.2.2.4 Solution Latency Analysis.. 51

2.2.3 Discussion.. 52

2.3 EXPERIMENT 1b.. 56

2.3.1 Method.. 59

2.3.1.1 Participants.. 59

2.3.1.2 Materials & Design.. 59

2.3.1.3 Procedure.. 59

2.3.2 Results & Discussion.. 60

2.4 EXPERIMENT 2A.. 63

2.4.1 Method.. 65

2.4.1.1 Participants.. 65

2.4.1.2 Materials & Design.. 65

2.4.1.3 Procedure.. 65

2.4.2 Results.. 65

2.4.2.1 Scoring Procedure.. 65

2.4.2.2 Strategy Selection.. 66

2.4.2.3 Selection Latency Analysis....................................... 67

2.4.2.4 Solution Latency Analysis.. 67

2.4.2.5 Covariate Analysis.. 69

2.4.3 Discussion.. 70

2.5 EXPERIMENT 2b.. 72

2.5.1 Method.. 73

2.5.1.1 Participants.. 73

2.5.1.2 Materials & Design.. 73

2.5.1.3 Procedure.. 73

2.5.2 Results & Discussion.. 73

2.6 EXPERIMENT 3... 76

2.6.1 Method.. 77

2.6.1.1 Participants.. 77

2.6.1.2 Materials & Design.. 77

2.6.1.3 Procedure.. 78

2.6.2 Results & Discussion.. 78

2.6.1.1 Sum Familiarity and Answer Familiarity........................ 78

2.6.2.2 Sum Type... 79

2.8 GENERAL DISCUSSION... 82

CHAPTER 3 .... **87**
3.0 ABSTRACT 87
3.1 INTRODUCTION 88
3.2 EXPERIMENT 4 92
3.2.1 Method 95
3.2.1.1 Participants 95
3.2.1.2 Materials & Design 95
3.2.1.3 Procedure 95
3.2.2 Results 95
3.2.2.1 Scoring Procedure 95
3.2.2.2 Strategy Selection 96
3.2.2.3 Selection Latency Analysis 97
3.2.2.4 Solution Latency Analysis 98
3.2.3 Discussion 99
3.3 EXPERIMENT 5a 100
3.3.1 Method 107
3.3.1.1 Participants 107
3.3.1.2 Materials & Design 107
3.3.1.3 Procedure 108
3.3.2 Results 109
3.3.2.1 Scoring Procedure 109
3.3.2.2 Strategy Selection 110
3.3.2.3 Selection Latency Analysis 112
3.3.2.4 Solution Latency Analysis 113
3.3.3 Discussion 114
3.4 EXPERIMENT 5b 117
3.4.1 Method 118
3.4.1.1 Participants 118
3.4.1.2 Materials & Design 118
3.4.1.3 Procedure 119
3.4.2 Results 119
3.4.2.1 Scoring Procedure 119
3.4.2.2 Strategy Selection 119
3.4.2.3 Selection Latency Analysis 121
3.4.2.4 Solution Latency Analysis 122
3.4.3 Discussion 123
3.5 EXPERIMENT 6a 126
3.5.1 Method 130
3.5.1.1 Participants 130
3.5.1.2 Materials & Design 131
3.5.1.3 Procedure 131

3.5.2 Results 132
3.5.2.1 Scoring Procedure 132
3.5.2.2 Strategy Selection 132
3.5.2.3 Selection Latency Analysis 134
3.5.3 Discussion 135
3.6 EXPERIMENT 6b 136
3.6.1 Method 138
3.6.1.1 Participants 138
3.6.1.2 Materials, Design & Procedure 138
3.6.2 Results 138
3.6.2.1 Scoring Procedure 138
3.6.2.2 Strategy Selection 139
3.6.2.3 Selection Latency Analysis 140
3.6.3 Discussion 141
3.7 GENERAL DISCUSSION 142

CHAPTER 4 147
4.1 AIMS OF THE THESIS 147
4.2 SUMMARY OF FINDINGS: ESTABLISHING THE KEY EMPIRICAL PHENOMENA 147
4.2.1 Testing the Dual-Phase design 148
4.2.2 Problem Familiarity as a cue to strategy selection 150
4.2.3 The Selection-by-feature effect 152
4.2.4 Cue combination: When problem familiarity and problem features collide 153
4.2.5 Selection in context 154
4.2.5 Summary of key findings 156
4.3 THEORETICAL IMPLICATIONS 157
4.3.1 The Automaticity account of strategy selection 158
4.3.2 The Adaptive account of strategy selection 159
4.3.2.1 The Component Power Laws theory 160
4.3.2.2 The Strategy Choice and Discovery Simulation 160
4.3.2.3 The Source of Activation Confusion account 164
4.4 SUMMARY AND CONCLUSIONS 165

REFERENCES 167
APPENDIX A 179
APPENDIX B 181
APPENDIX C 185
APPENDIX D 186
APPENDIX E 193

SUMMARY

The present thesis examined the processes responsible for strategy selection in problem-solving tasks. Despite the salience of this mechanism there has been a dearth in empirical research in the paradigm. Existing accounts, primarily modelled upon simulations of data sets, propose that strategies are not selected per se, but problems are solved by an automatic attempt to retrieve a solution (the *Automaticity* account; Logan, 1988; 2002). Contrary to this account four studies presented in the first empirical series demonstrated that predicted retrieve/calculate selections could be made rapidly (within 850 ms) and accurately. This indicated that problem-solving comprises two dissociable phases, selection then solution. Selection was found to be sensitive to the familiarity of a problem and also specific problem features supporting an account in which selection may be determined by the type of problem and the context in which the problem is solved (the *Adaptive* account; Reder & Ritter, 1992; Siegler & Araya, 2005). Elucidating the mechanisms responsible for these effects, in the second empirical series, three issues representative of real-world problem-solving episodes were examined. When multiple cues to selection were available, the interplay between the cues either served to inhibit the effects of both cues, or facilitated the effect of one cue at the expense of the other. Problem familiarity effects were attributed to implicit procedures as these effects were apparently un-reliant upon conscious processes (Reder & Ritter, 1992; Schunn et al, 1997). However, the feature identification process, rather than the selection mechanism itself, was found to be reliant upon consciously directed processes (Siegler & Araya, 2005). The findings from these studies were used to evaluate existing accounts of strategy selection, and reflecting limitations in these models, candidate mechanisms are proposed to account for the key effects revealed in this thesis.

CHAPTER ONE

1.1 INTRODUCTION

Humans are faced with many problems to solve in life. For example, if you were walking through a forest and met a bear there are a host of responses you could make. The most reliable option would be to mentally simulate all of the potential outcomes for each candidate response, then based upon the success of each simulation, decide upon the appropriate action. But given the demands of the situation (i.e., a rapid response) and the limitations of human memory and attention, this approach is not always possible. A fundamental issue when choosing how to respond to a problem is determining whether its solution can be directly retrieved from memory (e.g., when I meet a bear I should run away) or whether the solution needs to be worked out (e.g., I have not met a bear before, but when I met a badger I ran away, so I should do the same with the bear). Research into rapid problem-solving has illustrated that humans are adept at choosing the appropriate action under are range of situational constraints including time pressures (Gigerenzer & Goldstein, 1996; Payne, Bettman & Johnson, 1988) and cognitive constraints such as concurrent memory loads (Broder & Schiffer, 2003). However, a large degree of controversy still remains as to how we rapidly select between different options and the types of factors which affect those selections.

Within the problem solving literature much is known about how problems are solved, in particular, how problem-solving strategies interact with long-term memory to produce an answer to a problem (e.g., Lebiere & Anderson, 1998; LeFevre, Sadesky & Bisanz, 1996). Relatively little is known about the factors which are responsible for the selection of problem-solving strategies. It is acknowledged that adult problem-solvers have a large battery of solution strategies available varying in sophistication (e.g., Siegler & Booth, 2005) and that selection may be influenced by a range of factors including individual differences (e.g., LeFevre & Kulak, 1994), the age of the problem-solver (e.g., Touron & Hertzog, 2004) and problem characteristics (e.g., Reder & Ritter, 1992; Lebiere & Anderson, 1998). Developing an understanding of these factors and the mechanism underpinning selection will contribute to the understanding not only of the problem-solving process but also the process of skill acquisition (e.g., Siegler & Araya, 2005; Touron & Hertzog, 2004)

For most problems there are a number of candidate strategies that can lead to a solution. Most models of strategy selection seek to account for the selection of the two principle classes of strategy; i) direct answer retrieval from long term memory and ii) calculation strategies (or *algorithms*). Indeed, aside from problem-solving, this procedural distinction is central to a range of tasks including vocabulary learning (Crutcher, 1989), spelling (Siegler, 1986), acquisition of linguistic rules (Bourne, Healy, Rickard & Parker, 1999) and visual numerosity judgements (Palmeri, 1997). To examine the decision process, in the thesis I shall identify the key issues highlighted within the strategy-selection literature using a mental arithmetic task designed to delineate between the selection of memory retrieval and algorithmic strategies. The utility of mental arithmetic tasks as a tool for examining strategy selection is widely acknowledged within cognition as "discoveries in mathematical cognition will have implications for general theories of skill acquisition and memory" (Rickard, Healy & Bourne, 1994, p. 1139). The aim of this research scheme is to contribute to the understanding of strategy selection in problem-solving mechanisms, examining existing theoretical frameworks, highlighting and investigating their shortcomings as a platform for further theoretical development. First I shall review the existing theoretical approaches that have been applied to the strategy-selection process, then detail the key empirical findings in the field to date and finally introduce the mental arithmetic task used in all the thesis experiments.

1.2 THEORETICAL APPROACHES

A large number of models have been developed to account for performance in arithmetic problem-solving tasks. The level of specificity to which these models are detailed varies. At one end of the continuum are examples of higher-level models which analyse performance in respect to the type of strategies applied to solve problems (i.e., the strategic-level of analysis). Conversely, more fine-grained, or low-level approaches decompose the strategic level of analyses into the building block procedures that together comprise each strategy (i.e., the procedural-level of analysis). Also such models commonly detail the selection and deployment of these procedures on a temporal dimension. As well as varying in specificity, the focal remits of the selection models also varies. Some are designed to model strategy acquisition and integration, some focus on strategy shifts as a function of learning, and others are presented as pure models of the strategy selection process. All of the models detailed in the following sections have been developed from computational simulations of empirical findings. It is important to note at this juncture that the absence of empirical tests of the predictions of these simulations is detrimental to the strategy–selection paradigm. I begin by providing a theoretical backdrop to the paradigm, then detailing the competing models of strategy selection, establishing shared features and points of contention between the models.

In line with the remit of this thesis I choose to limit the scope of this section to a theoretical review of models which detail the processes used to select particular strategies for particular problems, rather than detailing more general, higher-level models of arithmetic problem-solving. It should be noted that a number of models of arithmetic problem-solving were designed to explore issues tangential to this thesis, such as the types of code used to represent arithmetic problems and how arithmetic facts are represented in long term memory (e.g., Campbell, 1994; 2005; Clark & Campbell, 1991; Dehaene, 1992; Dehaene & Cohen, 1995; McCloskey, 1992; McCloskey & Macaruso, 1994). These models are particularly informative about issues of representation and encoding in arithmetic problem-solving, but do not address in detail the process by which strategies are chosen. Furthermore, to manage the scope of theoretical investigation, 4 conditions were established for 2 purposes. Firstly, it was necessary to exclude models with insufficient power to address the central issue of this thesis, i.e., the process by which a strategy is selected. Secondly, to redress limitations inherent in the generality of some of the models to real-world problem-solving scenarios the ecological validity of the models was deemed critical. Accordingly, the models detailed in the thesis are:

1. Applicable to a range of populations, rather than specialist populations such as children, adults or those with cognitive impairments.
2. Able to account for performance in selection tasks at a procedural (i.e., low-level of analysis) rather than strategic level (i.e., higher-level strategy identification).
3. Able to account for strategy selections in problems which have and have not been encountered previously.

4. Able to accurately predict solution latencies for given problems and strategies.

However, where useful, recourse to models which violate these criteria will be made.

Prior classifications of models within the literature, such as that proposed by Siegler (1997; Siegler & Shipley, 1995), delineated between models based upon *metacognitive* or *associative* mechanisms. The associative class of models (e.g., General Inductive Problem-Solving model; Jones & van Lehn, 1991) are reliant purely upon associative strength based mechanisms where the link between a problem, strategy and strategy success — reinforced by practise on a problem — dictates future strategy selections. In contrast, early metacognitive models, such as the Triarchic Theory of Mind (Sternberg, 1985) classically assume that selection is determined by consciously available knowledge, directed by a homunculean executive processor which decides what the cognitive system should do:

> The executive is aware of the system's capacity limits and strategies. The executive can analyse new problems and select appropriate strategies and attempt solutions. Very importantly, the executive monitors the success or failure of ongoing performance, deciding which strategies to continue and which to replace with potentially more effective and appropriate routines. In addition, the efficient executive knows when one knows and when one does not know, an important requirement for competent learning. (Schneider & Pressley, 1989, p. 91).

However, a wealth of more recent research has demonstrated that conscious processing is not necessarily required for accurate (i.e., where predicted strategy selection equates to actual strategy selection) strategy selection (e.g., Reder & Ritter, Schunn et al., 1997), and that associative strength-based factors do not solely determine strategy selection (Lebiere & Anderson, 1998; Reder 1987; Rickard, 1997).

Latter classifications have been based upon the type of information used to make selections. In one class of models, termed the strategy *base-rate* accounts by Schunn and colleagues (1997), strategies are selected in accordance with the relative proportion of successful applications of a strategy to a particular problem (e.g., Anderson, 1993; Lebiere & Anderson, 1997; Siegler & Jenkins, 198; Siegler & Shipley, 1995). This contrasts with a *familiarity*-based account in which the familiarity of the problem's terms dictates strategy selection (Reder & Ritter, 1992; Schunn et al., 1997). However, this distinction is flawed as the base-rate class of models cannot account for selections determined by the problem's terms (see Reder & Ritter, 1992; Schunn et al., 1997). Furthermore, both the familiarity and base rate accounts are undermined by the necessity of ad-hoc assumptions when modelling adult performance (see Reder & Ritter, 1992).

In this thesis, rather than distinguishing between the types of mechanism (associative or metacognitive) or the information (base-rate or problem familiarity) that may influence strategy selections models are classified upon the basis of a procedural distinction. Some models posit that strategies are selected within a dedicated strategy-selection phase. This operates prior to the deployment of solution strategies to solve the problem, affording an *adaptive* selection of strategies to fit the demands of the problem terms or situation. Other models specify that strategies are *automatically* applied to problems in a default order, thus circumventing the need for a dedicated selection phase. This distinction directly addresses the current challenge within the literature, namely to identify the processes by which memorial activations influence strategy selection.

1.2.1 Approaches to strategy selection: The 'Automaticity' account

Central to the Automaticity accounts of strategy selection is the premise that attention to a problem necessarily activates a memorial search for information pertaining to that problem, including its answer (the Obligatory Activation Assumption, Logan, 1988). This mechanism circumvents the problem of adjudicating between the competing classes of strategy (i.e., direct retrieval or answer calculation) as retrieval procedures are automatically initiated when a problem is attended to. Accordingly, this class of models does not strictly model the process by which strategies are selected as direct retrieval is always attempted first. However, models within this account can predict which strategy would be used to solve a particular problem; hence their predictions and specifications are directly comparable to those where strategies are chosen.

The Automaticity approach evolved from the classic response selection theory developed by Shepard (1957) and Luce (1959, 1963) which successfully predicted response probabilities in identification tasks from estimates of similarity and bias. However, these models were limited by their inability to predict response times (see Logan, 2004 for review). More recently, Bundesen (1993) among others successfully replaced similarity-choice models with independent race models, where different solution strategies 'race' to produce a solution to a problem. These models produced equivalent predictions of choice probabilities but also allowed response latencies as well as response probabilities to be derived, a key requirement of strategy selection models. Such an approach is successful in tasks purely requiring responses based upon direct memory retrievals, but for problems such as, "465 + 935 = ?", where it is highly unlikely that the problem can be solved by direct memory retrieval, specification of the mechanisms which select between different calculation algorithms (or *backup* strategies) is required. Accordingly, models of arithmetic strategy selection require a secondary mechanism to select between competing computational algorithms. In the following sub-sections, the specifications of models fitting the Automaticity classification will be presented in detail.

1.2.1.1 The Adaptive Control of Thought-Rational (ACT-R; Lebiere & Anderson, 1998)

Lebiere and Anderson's (1998) ACT-R 4.0 (hereafter referred to as ACT-R) model of arithmetic processing (see also Lebiere, 1999) provides perhaps the most comprehensive specification of strategy selection in mental arithmetic to date. Recent updates to the underlying ACT-R model have been made (Anderson et al., 2004; Anderson, 2005; Belavkin & Ritter, 2004), and where relevant, the implications of such updates will be applied to the Lebiere and Anderson (1998) model. The ACT-R architecture used to develop the arithmetic model is founded upon a distinction in memory between things we know (*declarative knowledge*) and the bank of processes and skills we have available to process memories (*procedural knowledge*). In respect to strategy selection, this dissociation neatly maps on to the two classes of solution pathway in arithmetic problem solving; retrieving the problem's answer directly from (declarative) memory and use of computational algorithms (i.e., procedural memory). When presented with a problem such as 17 + 13 = ?, the problem is represented in memory as a discrete chunk, containing three full slots, one for each operand (i.e., [17] and [13]), one for the operator (i.e., [+]) and a further empty slot ready for the problem's solution (Anderson, Reder & Lebiere, 1996; Lebiere & Anderson, 1998). Solution strategies are automatically applied to this representation in a prescribed order. As a consequence of the Obligatory Activation Assumption (Logan, 1988) an attempt to retrieve the answer initiates immediately upon presentation of a problem. If no answer is returned then a calculation strategy is selected and executed.

1.2.1.1.1 The direct retrieval procedure

To find a solution to the problem, the retrieval production (i.e., a strategy) searches memory for chunks which match the contents of the problem chunk (i.e., [17] [+] [13]) but also include the answer. The level of chunk activation, which is represented by base-level — or resting level of — activation and current activation levels, is used to determine which chunks are identified as candidate answers. To illustrate, the base-level activation of a chunk is derived using the following function:

$$Bi = \ln\sum_{j=1}^{n} t_j^{-d} \qquad \text{Base-Level Learning Equation (1)}$$

Where t_j is the time elapsed since the jth retrieval of chunk i, n is the total number of references to that chunk and d is the memory decay rate. Base-level activation is then added to the associative activation that has spread to the chunk to give the current level of activation:

$$A_i = B_i + \sum_j W_j S_{ji} \qquad \text{Activation Equation (2)}$$

Where B_i is the base-level activation (or strength) of the chunk i, W_j reflects the attentional weighting of the elements j that are slots of the current goal and the S_{ji} terms are the strengths of association from the elements j. To prevent incorrect matches between the problem chunk and candidate chunks, the current level of activation is converted into a match score, whereby a perfect match between the contents of the slots in the problem chunk and candidate chunk (excepting the answer slot) elicits a perfect match score. Match scores for imperfect (or partial) matches are penalised to prevent incorrect chunk retrievals. When candidate chunks are competing for retrieval, the chunk with the highest match score will be chosen, and the match score subjected to a preset threshold. If the score exceeds the threshold, the chunk, complete with the problem's solution, is copied into the problem chunk and made available for output (Lebiere & Anderson, 1998; Lebiere, 1999) thus completing the retrieval production.

1.2.1.1.2 The backup or calculation algorithm procedure

If the retrieval strategy fails to return a solution to the problem, the ACT-R simulation searches for appropriate calculation algorithms stored in procedural memory. For example, a common algorithm used by children to solve simple addition problems (i.e., 5 + 3 = ?) is the iterative counting produce (Lebiere & Anderson, 1998). Here the largest addend is taken as the anchor and incremented by 1 in a loop until the number of increments made equals the value of the other addend. Algorithms applicable to particular problems are ordered in a conflict resolution stack in accordance with their predicted likelihood of success, as indexed by the Expected Gain equation:

$$E = PG - C \qquad \text{Expected Gain Equation (3)}$$

Where G is the value of the current goal, set as a constant in the ACT-R model, reflecting the amount of time the model will allocate to pursuing a given goal. P reflects the probability that a production has been successful in the past (P = successes/[successes + failures]) and C (C = efforts/[successes + failures]) is the ratio between the sum of all past successes and failures of the production, where effort is the time from firing the strategy until the model achieves either a success or failure. Noise is also added to the expected gain value to prevent the model from settling prematurely upon strategies that gain early advantage from P and/or C.

The strategy with the greatest likelihood of success is executed first, if this fails to produce an answer, strategies in the stack are cycled though in a serial order until an answer to the problem is produced. Finally, once found, the answer is copied to the problem chunk, ready for output.

1.2.1.1.3 ACT-R summary

To summarise, in ACT-R (Lebiere & Anderson, 1998; Lebiere, 1999) when presented with a sum, the direct retrieval strategy automatically searches memory for a chunk matching the problem chunk. If a matching chunk of sufficient activation strength is found this chunk overwrites the problem chunk, including the empty slot ready for the answer. If however, the retrieval strategy fails to provide an answer, calculation algorithms applicable to the problem are arranged in a conflict stack according to their expected utility. Only one strategy can be employed at a time and they are executed serially until a solution to the problem is derived.

1.2.1.2 The Instance theory of Attention and Memory account (ITAM; Logan, 1988, 2002)

Similar to ACT-R, ITAM (Logan 1988; 2002) was conceived to be a comprehensive model of attention and memory. The 2002 version of ITAM has not been directly applied to strategy selection, however, selection in arithmetic processing has shaped the evolution of key mechanisms within the model (see Compton & Logan, 1991; Zbrodoff & Logan, 1986; Zbrodoff & Logan, 1990). Underpinning this model is an *Instance theory* of memorial representation (Logan, 2002). Dissimilar to traditional 'strength models' of mental representation such as ACT-R, where repeated presentations of a stimulus (or problem) serve to reinforce the associative link between stimulus and response, in ITAM each encounter with a problem and consequently the ensuing processing episode, creates a separate memorial representation or *instance*. Each instance contains all the information encoded during a processing episode, including a record of the stimulus encountered in pursuit of the goal, the participant's goal, the interpretation given to the stimulus with respect to the goal and the response made (i.e., the strategy used and its success). The numerosity of problem specific instances serves to 'strengthen' the link between stimulus and response. Specifically, as the number of accumulated instances for a particular problem increases, so does the likelihood that one of the instances will win the race against competing classes of instance (i.e., strategies). This serves to increase the rapidity and accuracy of responses in contrast to problems with relatively few accumulated instances.

1.2.1.2.1 Response selection in ITAM

As previously alluded, ITAM makes no formal distinction between the retrieval and calculation classes of strategy. On a procedural level, candidate answers are either derived by directly retrieving an instance which includes the problem's solution or from the output of a calculation algorithm. Consequently, when presented with a problem, ITAM does not choose between solution strategies, but, as an obligatory consequence of attention, automatically encodes and attempts to retrieve all instances pertaining to the problem complete with the answer (i.e., direct retrieval; Logan, 1988, 2002; Zbrodoff & Logan, 1986). Simultaneously it searches for and executes relevant calculation algorithms. Accordingly, both classes of solution strategy, retrieval and calculation, fire in parallel racing to return a candidate answer (see also Ashcraft, 1982; 1987; 1992).

The ITAM simulations can utilise three different models to monitor the accumulation of instances originating from the different solution pathways. The frequency and rapidity of instances activated in response to the presentation of the problem determines which of the three models is used (Logan, 2002). The simplest model, the *race model* terminates the race when it receives the first return from any of the different solution pathways. One counter monitors the accumulation of instances from the direct retrieval pathway and separate counters monitor the return from each of the calculation algorithms that are engaged. This model affords rapid solution latencies, however when there are returns to multiple counters (i.e., returns from competing strategies), the mechanism becomes error prone as the fastest return may not be the most accurate (Logan, 2002).

The second response selection mechanism, the *counter model* circumvents the limitations of the basic race model in situations with competing returns in different counters. As before, a counter monitors each solution pathway. Employing a decision rule, the counter only signals a return when *K* memory traces have been accumulated in that counter. If *K* is 1 (i.e., one complete instance has been returned to the counter) the model acts in an identical fashion to the simple race model. If *K* is set to 3, the counter waits for three instances to accumulate before signalling a response. This model serves to increase the accuracy of response selections by confirming the accuracy of the first instances returned, but at the expense of response latencies as a longer duration is required for solution pathways to return multiple instances.

Finally, the *random-walk model* monitors the solution pathway counters and waits until there are *K* more hits in one counter than the other. Similar to the counter model, the random-walk model affords greater accuracy than the race model but also accommodates the relative success of different solution pathways, resulting in protracted response times but greater accuracy when there is conflict between two strategies (Logan, 2002).

1.2.1.2.2 ITAM summary

When presented with a problem, ITAM automatically engages a memory search to recall all the instances, stored in memory, that pertain to the problem. This includes instances which can be classified as direct retrievals and instances comprising calculation algorithms. Instances are recalled in parallel, each racing to produce a solution to the problem. Three different types of counter can monitor the returns from each class of instance (or solution strategy; direct retrieval or calculation algorithms), determining which solution is accepted as the problem's solution.

As the latest ITAM conception has not modelled empirical data in the mental arithmetic paradigm, specific, quantifiable predictions cannot be derived from the model. However, the parallel execution of strategies in ITAM provides a useful counterpoint to the serial processing advocated within the ACT-R model and other models within the Automaticity account of selection.

1.2.1.3 The Distribution of Associations model (Siegler & Shrager, 1984)

The Distribution of Associations model (DOA; Siegler and Shrager, 1984) was one of the first comprehensive models of strategy selection and its specifications have been highly informative of a number of more recent models including ACT-R. The model was motivated by two remits; firstly to detail the process by which new strategies are developed and assimilated. Accordingly, the authors only simulated data from single-digit arithmetic sums (i.e., 3 + 4) in pre-school children which provide clearer evidence of strategy development and assimilation than data from adults. The authors also note that they suspect the same mechanisms are invoked by adults when making strategy selections (Siegler & Shrager, 1984). Secondly, the approach sought to debunk the notion that complex executive processes govern strategy selection (e.g., Sternberg, 1985) by illustrating that simple cognitive mechanisms can produce adaptive strategy selections.

Based upon a general semantic network view of knowledge representation in memory (see also Schunn et al., 1997), whereby knowledge is structured by associations between problems and their potential answers, the model picks the answer with the strongest association to the problem. For example, 3 + 5 is strongly associated with the correct answer 8, but also holds associations with the incorrect answers 2 (i.e., 5 – 3), 15 (i.e., 3 x 5) and other numbers neighbouring the correct answer, i.e., 7 or 9.

1.2.1.3.1 Strategy Selection in DOA

When presented with a problem, three processes, running in serial, are posited to determine strategy selections. First children attempt to retrieve the answer from memory; the success of this strategy is dependant upon the peakedness (i.e., the relative concentration of the associative strength distribution) of the association between problem and answer. A highly peaked distribution gives rise to greater confidence in the accuracy of an answer than a flat distribution. If confidence from the retrieval process falls bellows a preset threshold within the simulation, the next process taken by children is to elaborate the representation of the problem. For example, in a simple arithmetic problem such as 4 + 5 = ?, children may put up their fingers to represent the two addends. If this process fails to return an answer with sufficient confidence an algorithmic procedure would be used, perhaps counting the fingers the elaborated representation (Siegler & Shipley, 1995).

1.2.1.3.2 Summary of DOA

The DOA model invokes a strict serial processing order to strategy selection. Firstly, children attempt to retrieve the answer directly from memory, if this fails they attempt to elaborate the representation of the problem using fingers as visual/spatial cues to memory retrieval. Finally, if neither process is successful, a calculation algorithm is applied, in the case of single-digit addition problems, finger counting. Despite its apparent successes in modelling the pre-school children's data, demonstrating that the mechanism governing strategy selection was not reliant upon some homunculean executive the model was flawed in two key respects. First, it provided no specification of how competing calculation strategies were selected. Secondly, it did not learn from the experience it gleaned from the problem solving episode, no matter how many times the problem was encountered.

1.2.1.4 Summary of the Automaticity Accounts

Common to all the models in the Automaticity account is the fundamental assumption that humans do not have any control over the initial step taken when attempting to solve a problem. This is irrespective of the type of problem and the task conditions, following the Obligatory Activation Assumption (Logan, 1988). Furthermore, the suitability of candidate responses is based upon the prior success of a particular strategy, through an associative strength mechanism in ACT-R and DOA and the instance accumulator counters in ITAM. Beyond the automatic application of the retrieval strategy however, the models diverge in respect to the order in which strategies are applied to a problem. As Figure 1.1 illustrates, the ACT-R and DOA simulations only employ a backup or calculation strategy if the retrieval strategy fails to produce an acceptable solution, applying strategies in a strict serial processing order. In contrast, the ITAM model initiates retrieval and calculation classes of strategy in parallel, each strategy racing to provide an answer to the problem. In the following section the competing class of selection models, the Adaptive account, will be introduced.

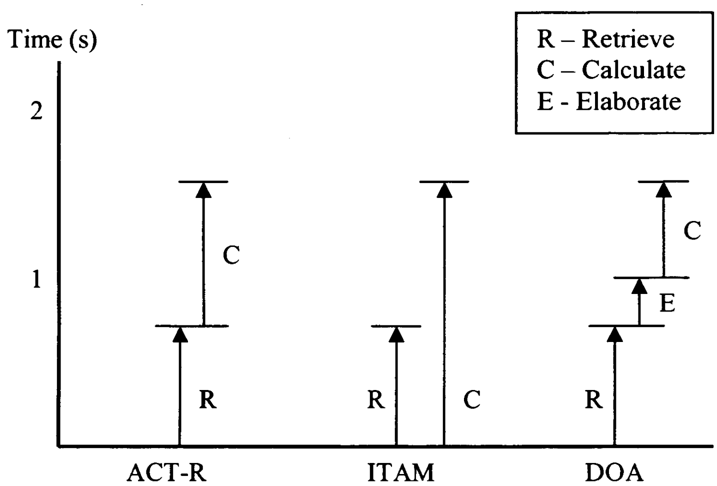

Figure 1.1: The order of processing proposed by each of the Automaticity models. Vertical arrows denote the action of a particular strategy (*retrieve*, *calculate* or *elaborate*), horizontal lines indicate the point at which an output can be returned from a particular strategy. Note that the time scale is for illustrative purposes and is not exact, although, a number of studies suggest that retrieval of a problems answer takes approximately 850 ms (see Reder & Ritter, 1992; Schunn et al., 1997; Staszewski, 1988).

1.2.2 Approaches to strategy selection: The 'Adaptive' account

The notion that problem-solving and thus strategy selection is adaptive is more commonly associated with complex problems solving tasks that afford consciously directed processing rather than tasks requiring rapid solutions. For example, Schunn and Reder (1998) reported evidence of adaptivity in the Building Sticks Task, an analogue of the classic Water Jars task (Luchins, 1942), where participants are required to add and subtract 3 sticks of different lengths to produce a stick of a target length. They also revealed adaptivity in the Kanfer-Ackerman Air Traffic Controller Task (Kanfer & Ackerman, 1989; Kanfer, Ackerman & Pearson, 1994), a complex simulation in which participants are required to queue and land planes abiding to a set of rules. However, burgeoning evidence indicates that adaptive responses, i.e., those which are influenced by the task demands and cognitive resources available (Payne, Bettman & Johnson, 1988), are also evident in problem-solving tasks that are not reliant upon conscious processing (Reder & Ritter, 1992, Schunn et al, 1997).

Central to the Adaptive account is the premise that humans have a range of solution strategies at their disposal and that they can be applied in a manner specific to the problem, producing rapid and accurate responses (Reder, 1987; 1988; Reder & Ritter, 1992; Schunn et al., 1997; Shrager & Shipley, 1998; Siegler, 1999; Siegler & Araya, 2005; Siegler & Shipley, 1995). Furthermore, due to the competition for resources, strategies cannot be executed in parallel, thus necessitating selection between competing strategies (Rickard, 1997; 2004). To circumvent this issue, a common feature of the adaptive models is that strategies are selected during a dedicated strategy-selection phase, which operates prior to the deployment of any solution strategies. As strategy selection must be able to operate very rapidly, the rational for this approach was largely founded upon the surge of interest in memorial mechanisms that operate before items can be retrieved from memory. In particular mechanisms responsible for monitoring and controlling the processes used to access memory (see Koriat, 2006 for review). In the following sub-sections three models, classifiable as 'adaptive' will be detailed.

1.2.2.1 The Source of Activation Confusion model (SAC; Reder & Ritter, 1992; Schunn et al., 1997)

Reder's approach was built upon her early research on question answering (1987). In strategy selection studies predating the publication of the SAC model, Reder's empirical approach to strategy selection paralleled the Recognition memory literature by investigating the distinction between intrinsic factors, such as problem familiarity (Reder, 1982; Experiments 4, 5, 6; Reder, 1987; Reder & Wible, 1984) and extrinsic factors, such as, task instruction manipulations (Experiments 1, 2 and 3; Reder, 1987). In direct contrast to the Automaticity accounts, Reder suggested that the first step undertaken when solving a problem is an evaluation of the problem terms (Reder, 1987). This evaluation was shown to operate very rapidly (i.e., within 850 ms) and with a high degree of accuracy suggesting that it could be a component of the problem solving process (Reder & Ritter, 1992; Schunn et al., 1997). This approach has been applied to retrieve and calculate strategy selections in mental arithmetic problems (Reder & Ritter, 1992; Schunn et al., 1997), specifically, testing double-digit multiplication sums (e.g., 26 x 12) and in one unsuccessful instance double-digit addition sums (Reder & Ritter, 1992; Experiment 1). Aside from strategy selection, the SAC model has been applied to recognition memory (Diana, Reder, Arndt & Park, 2006), specifically analysis of word frequency (Diana & Reder, 2006), contextual interference (Park, Arndt & Reder, 2006), list length and mirror effects (Cary & Reder, 2003).

In the SAC model, the problem solving process is split into two distinct phases. First, in a dedicated strategy selection phase, a solution strategy (retrieve or calculate) is selected, based upon a rapid analysis of the problem. In the second phase the selected strategy is then executed in an attempt to solve the problem. The model itself is founded upon a general semantic network model of memory (Schunn et al., 1997) in which linked, or associated nodes represent concepts within memory (see Figure 1.2). The mechanism responsible for selecting between strategies in the SAC model is highly dependant upon the manner in which problems are represented in the SAC model. In the next section I shall outline the features of the semantic model pertinent to strategy selection.

1.2.2.1.1 Problem representation in the SAC model

Whereas models such as ACT-R (Lebiere & Anderson, 1998) represent problems as chunks in memory, containing the whole problem within a discrete unit, the SAC model is based upon a different approach. Problems such as 17 + 23 are represented by nodes representing the whole problem (i.e., [17 + 23]), individual operands (i.e., [17], [23]) and the operator (i.e., [+]). The highly interconnected organisation of nodes within the semantic network means that numerical nodes may serve as operand nodes for one problem and the answer node for another problem (see Figure 1.2).

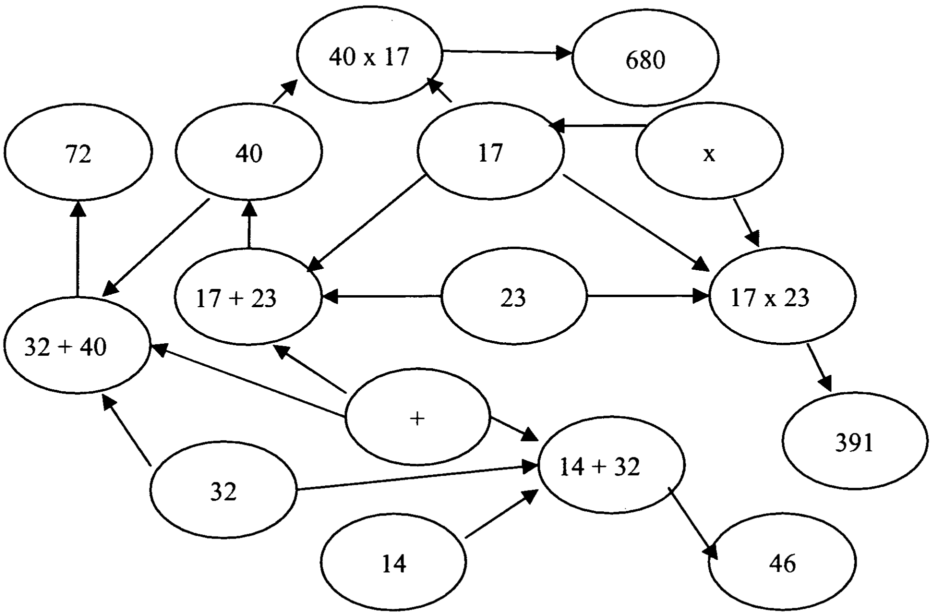

Figure 1.2: An example of the semantic network of representation employed within the SAC model, comprising separate nodes for the problem, problem components and the answer. Note that only a small section of the network is displayed.

Each node has a base level strength (or *resting strength*) which represents the history of exposure to that concept (indexed as *familiarity*). So more familiar problems, such as 2 + 4 will have greater base level strength than less common problems, such as 17 + 23. Increases and decreases in base level strength (B) are determined by the function below, where c and d are model constants and t_i is the time since the ith presentation.

$$B = c\sum t_i^{-d} \qquad \text{Base Level Activation Equation (4)}$$

This function accounts for the power-law decay of memories over time, where memories decay rapidly initially then more slowly after increasing delays. Also the power-law learning of memories with practise where the first exposure to a new item contributes more to its memorability than subsequent exposures (see Anderson & Schooler, 1991).

When a problem is presented, the three component nodes in this example [17], [+] and [23] are activated. In these *source* nodes, activation rises beyond the base level and spreads, via links running from the source nodes to nodes representing whole problems. The amount of activation spreading from a source node is dependant upon the level of activation in the source node itself and the strength of the links between the source nodes and receiver nodes. While link strength between nodes is dependant upon the number of times the two concepts represented by the node have been activated together, this strength is also subject to decay over time. So for the problem 17 + 23, in the local network illustrated in Figure 1.2, activation from the source node [17], would spread to the problem nodes; [17 x 23], [40 x 17] and [17 + 23]. From the second operand, 23, activation would spread to nodes representing [17 x 23] and [17 + 23], and from the operator, +, activation would spread to [14 + 32], [32 + 40] and [17 + 23]. The problem node with the greatest number of links back to the three source nodes (i.e., [17 + 23]) thus receives the greatest amount of activation which is added to the base level activation of that node. The SAC model takes the level of activation at the node with the greatest accumulation of activation as the basis for the strategy selection.

1.2.2.1.2 Strategy selection in the SAC model

When presented with a problem, the final level of activation (base level + spreading activation) at the most active problem node within the semantic network is used to determine a Feeling-of-Knowing (FoK). The FoK is then subject to a preset threshold criterion set in the simulation to determine whether to retrieve the answer or use a calculation strategy. Hence, if the most active problem node has a high level of activation, relative to other competing nodes, the FoK will be high and it is likely that the retrieve strategy will be chosen. Alternatively, if the level of activation is not much greater than other activated problem nodes, calculate is likely to be selected. Rather than making a binary decision between the retrieve and calculate strategies, the FoK threshold mechanism, which determines strategy selection, derives a probability of choosing to select the retrieve strategy. This is calculated by assuming a normal distribution of activation values with a fixed variance assumed and a fixed activation threshold for selecting retrieve. In the actual model runs, each 'participant' was randomly assigned different activation thresholds for retrieve selections taken from a set range. This was introduced to allow the model to replicate the between subjects variability demonstrated it empirical studies of arithmetic problem solving (e.g., Kirk

& Ashcraft, 2001). Furthermore, it must be noted that in SAC thresholds are fixed for the duration of a run.

1.2.2.1.3 Summary of the SAC model

In the SAC model, strategy selections are determined by a rapid analysis of a problem's terms. In this phase of the problem-solving process, which occurs prior to the deployment of any solution strategies, the patterns of activation elicited by the familiarity of a problem's terms produce a FoK which in turn is subject to a decision mechanism. Based upon the strength of the FoK either the retrieval or calculation class of strategies are selected.

1.2.2.2 The Component Power Laws Theory (CMPL; Rickard, 1997 and Rickard, 2004)

The primary purpose of CMPL is to model the transition from the selection of calculation to retrieval strategies as a function of learning. As a counterpoint to Logan's instance theory of automatic and parallel strategy execution, CMPL has been simulated on a range of tasks including alphabetic arithmetic (Experiment 2, Rickard, 1997; Rickard, 2004), pseudo-arithmetic (using a novel operator termed *pound*, Experiment 1, Rickard, 1997) and artificial classification learning tasks (Bourne, Healy, Parker & Rickard, 1999). Underpinning the model are three key assumptions, the first concerns the structural composition of calculation algorithms. CMPL treats calculation algorithms as a sequence of memory retrieval events (qualitatively identical to the direct retrieval strategy), where each step of the algorithm is a separate retrieval event. For example, when solving the problem 23 + 26 = ?, the first step in calculating the problem's solution is the memory retrieval 3 + 6. The second memory retrieval event being, 20 + 20, the third adds the products from the two intermediary sums (9 and 40) to get the final answer, 49. Secondly, in contrast to other models of serial strategy execution (e.g., ITAM) strategies cannot be executed in parallel; however candidate strategies can be activated in parallel ready for execution. Finally, unlike ITAM and ACT-R, CMPL assumes that the problem and its answer are represented separately in memory. Simple connectionist principles serve to specify the link between the problem and its answer, whereby practise on a particular problem reinforces the associative linkage.

The structure of the model is depicted in Figure 1.3. At the top level of the model is the overall goal for the task, which in this instance is to solve the problem. This node has excitatory connections to the two solution strategies at the subgoal level, retrieve and calculate, which serve to execute both the first step of the calculation algorithm (i.e., the first memory retrieval in the algorithm) and a direct retrieval from memory. These two nodes in turn have excitatory connections to the long-term memory nodes at the problem-level which are also fed by the stimuli and information contained within working memory (Rickard, 1997). The connections between the nodes at different levels in CMPL are reinforced by exposure to problem-solving episodes. Use of the retrieval strategy reinforces only the retrieval pathway connection, but use of the calculate pathway reinforces the connection in both pathways, eliciting a shift with learning from selection of the calculate strategy to retrieval.

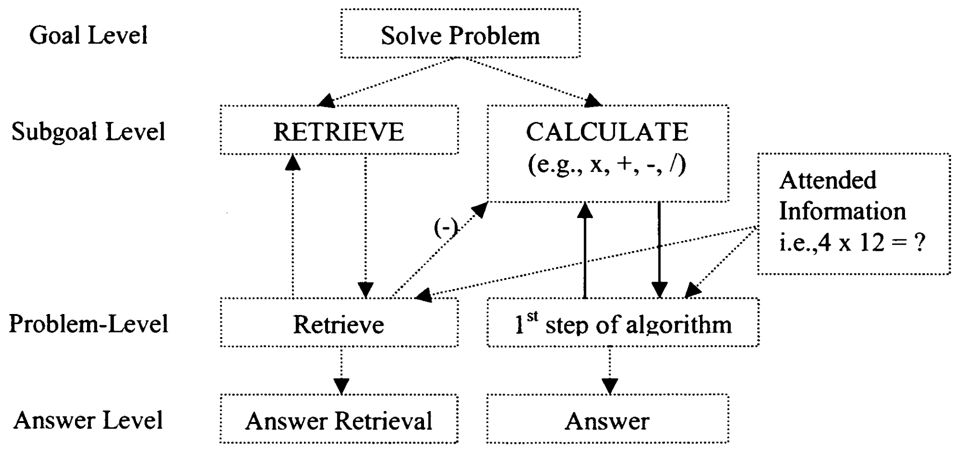

Figure 1.3: A schematic of the Component Power Laws theory of retrieve/calculate strategy selection in arithmetic problems. Note that '-' indicates an inhibitory pathway, all other links are excitatory. Adapted from Rickard, 1997.

1.2.2.2.1 Strategy selection in CMPL

Based upon Rickard's assumption that direct retrievals cannot be executed in parallel (Rickard, 2004) and that calculation algorithms are purely comprised of a series of memory retrievals, Rickard proposed a performance bottleneck where one strategy must be selected at the expense of another. When presented with a problem, candidate strategies are activated in memory in parallel during a dedicated strategy selection phase. Candidate selection is determined by the interactions between nodes at the subgoal and problem-levels for the direct retrieval strategy and the direct retrieval required for the first step of the algorithm. If activation levels at the subgoal and problem nodes in either strategy (i.e., retrieve or calculate) exceed a preset threshold, the pathway will be selected inhibiting the losing strategy. For example, if the subgoal for the first step of the algorithm and the corresponding problem-level node both reach the preset activation threshold, the direct retrieval strategy will be inhibited. Activation then is free to spread between the nodes in the algorithm strategy, eventually spreading to the answer node once the latter steps of the algorithm are completed. However, if the connection strengths between the subgoal and problem-level nodes within the direct retrieval strategy reach the threshold activation level before those in the algorithm strategy, the algorithm pathway will be inhibited. Activation then accumulates at the direct retrieval answer node without competition from the algorithm answer node. Contrary to serial models of strategy execution (e.g., ACT-R), where losing strategies are executed automatically upon failure of the original strategy to solve the problem, in CMPL the selection phase is reinitiated in light of the adjustments made to the associative connections elicited by the failed strategy execution.

1.2.2.2.2 Summary of CMPL

In CMPL candidate strategies are recalled from memory in parallel and compete for selection in parallel. The winning strategy, determined by the degree of activation elicited by an attempted retrieval of the answer (for the retrieval strategy) or retrieval of the answer to the first step of the calculation algorithm, then inhibits the losing strategy, concentrating all of the available activation upon the strategy with the greatest likelihood of success.

As Rickard acknowledges, subgoal (or strategy level), and problem-level (item level) factors are not the only influences upon strategy selection (Rickard, 1997). To date the model does not have any provision, or mechanisms which can account for contextual or any other task influences. Furthermore, the model is contingent upon the first step of a calculation algorithm being a direct retrieval. This may not always be the case, for example, when individuals choose to reorder a problem or break it down into smaller components which can be solved by direct fact retrieval, a series of complex processes are employed before the first retrieval in the algorithm is required.

However, CMPL does provide a useful and theoretically interesting structure within which to conceptualise the strategy selection process.

1.2.2.3 The Adaptive Strategy Choice model (ASCM; Siegler & Shipley, 1995), the Strategy Choice and Discovery Simulation (SCADS; Shrager & Siegler, 1998; Siegler, 1999) and the Strategy Choice and Discovery Simulation (SCADS*; Siegler & Araya, 2005)*

In light of criticisms of the DOA model Siegler and colleagues developed second (ASCM; Siegler & Shipley, 1995) and third generation models (SCADS; Shrager & Siegler, 1998 and SCADS*; Siegler & Araya, 2005) to fit the premise that people do not always attempt to retrieve the solution to a problem before applying other strategies (see Reder, 1982). Similar to the DOA model, ASCM, SCADS and SCADS* are computational models, applied to data from single-digit arithmetic problems (e.g., 4 + 3) and inversion problems (i.e., a + b – a, for example, 4 + 3 – 4) recorded from children aged from 4-5 (Shrager & Siegler, 1998; Siegler, 1999; Siegler & Araya, 2005; Siegler & Shipley, 1995; Siegler & Stern, 1998).

In all three models, strategy selections are determined by a common associative mechanism. When presented with a problem, four sources of historical data, derived from prior applications of each strategy, contribute to the predicted strength of each candidate strategy:

1. *Global data* comprises the speed and accuracy of a particular strategy's application to all problems.
2. *Featural data* comprises the speed and accuracy from prior applications of a strategy to problems containing specific features (e.g., both addends were presented in a larger font than average).
3. *Problem-Specific data* comprises speed and accuracy information pertaining to the application of specific strategies to specific problems.
4. *Novelty data* adds a weighting to new strategies, promoting selection of new strategies in the absence of a history of speed and accuracy data.

For each strategy, the relative weightings of each type of information differ in accord with the novelty of the problem. For example, when solving a commonly encountered problem, more information about the prior success of a particular strategy on that problem is available, in this instance problem-specific data is given a greater weighting than global data which provides information about the prior success of a strategy on all problems. Alternatively, when a strategy has not been used on a specific problem before, but on problems with similar features, only global and featural data are used when determining which strategy to apply. In the models, the weighted sources of information are then entered into a stepwise regression which computes the predicted strength of each strategy. The actual probability of choosing a strategy is proportional to the strategy's projected strength relative to the strength of all the strategies combined. This is calculated for each candidate strategy, the strategy with the greatest probability of success is chosen. SCADS*, unlike ASCM and SCADS, also includes feature detectors which are active during the encoding of a problem. These are sensitive to features such as the magnitudes of the operands, the type of operators in a problem, whether any operands are the same, whether all operands are odd or even and the colour and size of the operands (Siegler & Araya, 2005). When encoded, if any of these features are present, information pertaining to the success of each candidate strategy, in light of the features is also activated influencing strategy selection.

If a calculation algorithm is chosen, the procedure is run until an answer is returned. Alternatively, if retrieval is selected an identical procedure is employed as detailed in the DOA model, where an answer is retrieved when the associative strength between the answer and problem exceeds a confidence criterion. If however, the first choice pathway fails to return an answer within the allotted timeframe, a second round of selection occurs but with the failed strategy removed.

*1.2.2.3.1 Summary of ASCM, SCADS and SCADS**

In these three models, the suitability of all candidate strategies is assessed in parallel and based upon knowledge of the prior success of each candidate strategy. The winning strategy is then executed; if it fails to produce a solution then the selection process is repeated. Dissimilar to any of the other models previously detailed, in addition to the associative strength mechanism which operates as a default determinant of strategy selections, SCADS* includes a feature detection mechanism which influences strategy selection when particular features are detected in the problem terms. Other simulations such as ACT-R and the SAC model deal with this problem by invoking ad-hoc assumptions. However, the tractability of complex information processing requirements of the ASCM, SCADS and SCADS* simulations, where the value of each of the four types of information (global, featural, problem-specific and novelty) are entered into a stepwise regression to determine the potential of each candidate strategy is questionable when strategy selection is required under time constraints. This complexity also makes it very hard to extract concrete predictions of performance on specific stimuli in adult populations.

1.2.2.4 Summary of the Adaptive Account

The models detailed in this section all posit that, when given a problem to solve, a strategy is selected before any attempt to derive an answer to the problem is initiated. This is not to say that selection is a controlled process in the tradition of the early metacognitive models (e.g., Sternberg, 1985) where an executive homunculean processor dictates selection, but that selection is adaptive. It is responsive to the type of problem presented, the context within which it is presented and specific characteristics of that problem. As Figure 1.4 illustrates, in the SAC model, during a selection phase (represented by the dashed arrows), before a strategy is deployed the decision is made whether to deploy the retrieve strategy or not. In CMPL, retrieve and calculate strategies compete for selection during this phase. If the retrieve strategy wins, the answer is output, however if the calculate strategy wins the remainder of the algorithm is processed until an answer is produced. SCADS*, uses a more complex mechanism whereby strategies compete for selection during a selection phase, the decision process is also augmented by a feature detector.

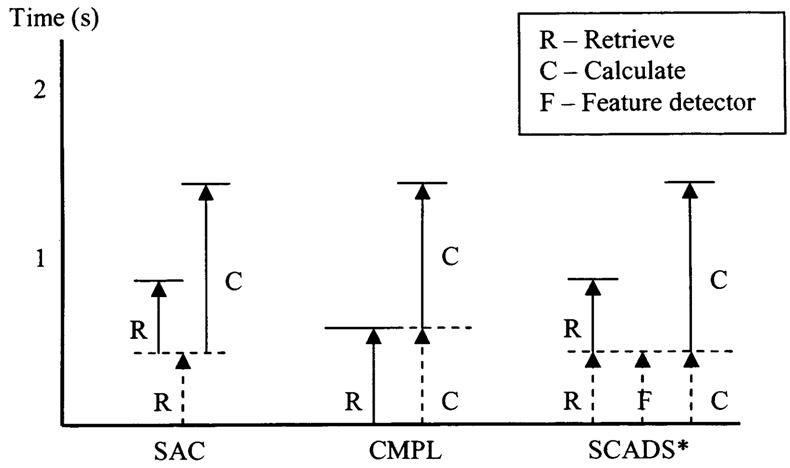

Figure 1.4: The order of processing advocated in each of the adaptive models. Dashed vertical arrows denote instances where the potential application of a strategy is being assessed during a selection phase. Dashed horizontal lines represent the point at which a strategy selection is made. Solid vertical arrows represent the execution of a strategy and solid horizontal lines the optimal point at which an answer can be produced. Note that the most recent of Siegler and colleagues models, SCADS*, is depicted here. Furthermore, for simplicity only two strategies, out of the host of potential candidate strategies considered by SCADS* during the strategy selection phase are included in this figure.

However, the models differ in respect to the types of information they use to resolve selection. In the SAC model selection is based upon an assessment of the familiarity of a problem's terms. CMPL employs knowledge of the prior success of a strategy derived from base-level and associative strengths. While in ASCM and SCADS four classes of information are used to derive strategy selection, all based upon the prior success of the problem-solving episode. Each source of information is then weighted and combined to determine which strategy has the greatest chance of success. Augmenting that mechanism SCADS* also employs a secondary mechanism, a feature detector, which is sensitive to specified features inherent within a problem.

1.2.3 Summary of theoretical approaches

As I have shown in the previous section a significant degree of controversy still remains as to whether strategies are applied automatically to a problem, or whether they are selected adaptively in accordance with the demands of a problem. Some models propose that selection occurs before any attempt to solve the problem is made (SAC, CMPL, ASCM, SCADS and SCADS*), alternatively that selection only occurs after the first strategy deployed fails to return an answer (DOA, ACT-R), or not at all (ITAM). As some of the models are embedded within general architectures of cognition (e.g., ACT-R and ITAM) or are used to model performance in a variety of tasks (e.g., SAC and CMPL) this is a core issue in cognition which requires resolution as the ramifications of this issue are far-reaching. The models do however converge upon the notion that selection can be resolved by simple cognitive mechanisms that monitor and/or control access to knowledge stored in long term memory, rather than opaque conscious or 'executive' mechanisms. Very little empirical work has been undertaken to test the predictions of these models and only three attempts have been made to directly contrast the predictions of the competing models (Nino & Rickard, 2003; Reder & Ritter, 1992 & Rickard, 1997).

It may be possible that inconsistencies between the models are a by-product of the different methodologies and stimuli used to produce the datasets upon which the simulations are founded. The datasets simulated by ITAM and CMPL utilised either alphabetic or pseudo-arithmetic paradigms (Compton & Logan, 1991; Rickard, 2004; Rickard & Bourne, 1996; Rickard, Healy & Bourne, 1994) rather than conventional arithmetic tasks. The SAC model is based upon a paradigm where participants are repeatedly exposed to conventional, but difficult, arithmetic problems (e.g., 23 x 46 = ?) and predict what strategy they would use to solve the problem before attempting to solving it (Reder & Ritter, 1992; Schunn et al., 1997). As mentioned previously, no ITAM simulations have been run on arithmetic tasks to my knowledge and the component mechanisms of this model, similar to CMPL, were tested using artificial arithmetic tasks (i.e., alphabetic- or pseudo-arithmetic) rather than conventional arithmetic problems. Furthermore, all of the models are simulations tested upon either one or a small number of datasets, for example, the ACT-R, ASCM and DOA models were all tested upon the same dataset, each model differing in respect to the order in which strategies are applied to a problem.

Following Bourne, Healy, Parker and Rickard (1999) it appears likely that "learners can and do change the basis for performance in a learning or memory task as experience with or practice on the task proceeds" (Bourne, Healy, Parker & Rickard, 1999, p.224). To identify the actual strategy selection process used by adults it is necessary to test adult performance on conventional mental arithmetic tasks. Whilst this thesis acknowledges the contribution of the models based upon artificial arithmetic tasks and experimentally induced knowledge, a thorough investigation of the factors that actually impinge upon actual adult performance is clearly required (Siegler, 1999). To evaluate the reliability of the competing models the empirical work in the thesis uses experimental designs which directly contrast the predictions of the models, testing adult performance on conventional arithmetic tasks.

1.3 EMPIRICAL PHENOMENA

Whilst a number of theoretical models of strategy selection exist, there has been comparatively very little empirical research testing their predictions. In light of the shortfall of empirical research within this paradigm, in the next section I shall review the key findings from the mental arithmetic and strategy selection paradigms which stand as features that a comprehensive model of strategy selection must fit. The first phenomenon detailed pertains to the order of processing in the problem-solving episode (1.3.1), the other two sections (1.3.2 and 1.3.3) to the factors that have been shown to influence strategy selection.

1.3.1 Strategy selections can be made quickly and accurately

The most compelling evidence that strategy selections can be made rapidly and accurately is from a series of experiments by Reder and colleagues (Reder & Ritter, 1992; Schunn et al., 1997) designed to contrast the predictions of the Automaticity and Adaptive accounts of selection. To recapitulate, the Automaticity class of models stipulate that strategies are applied in an automatic fashion, as an obligatory consequence of encoding, conversely, the Adaptive accounts assert that strategies are selected during a discrete phase which occurs prior to strategy execution. In their dual-phase design, on each trial participants first indicated what strategy (retrieve or calculate) they would use to solve the presented problem within a time limit of 850 ms and then solved the problem. The time limit was imposed to prevent participants from solving the problem, identifying what strategy they used, or should have used, and using this hindsight report as a basis for their predicted strategy selection. Previous studies have illustrated that 850 ms is required to execute to completion a direct retrieval (Staszewski, 1988), accordingly responses exceeding this time limit were excluded from the analysis.

In line with the predictions of the SAC model, participants were able to make strategy selections well within the 850 ms time limit. Furthermore, the percentage of *late* responses (i.e., those exceeding the time limit) was below 34% in all experimental conditions (see Table 1.1) indicating that selections could be made rapidly.

Table 1.1

All figures are means, rounded to whole integers. Sources 1 and 2 were taken from Reder and Ritter (1992) Experiments 1 and 2 respectively, source 3 from Experiment 1, Schunn et al (1997).

		Strategy selection latency (ms)		Late to choose strategy (%)		Solution Latency (ms)	
Source	Stimuli	Calculate	Retrieve	Calculate	Retrieve	Calculate	Retrieve
1	Multiplication	685	760	19	34	8930	3660
	Addition	750	625	33	12	1910	780
2	Multiplication	640	695	18	26	9130	1280
	Sharp	645	580	16	26	7850	1700
3	Multiplication	647	607	9	11	7787	1415
	Sharp	645	596	12	11	7235	1376

To test the accuracy of the responses it was reasoned that the chosen strategy in each trial should correspond to the amount of time required to solve the problem, i.e., the chosen strategy was actually used to solve the problem (Lebiere & Anderson, 1998; Reder & Ritter, 1992; Schunn et al., 1997). For example, retrieve selections made in the selection-phase of each trial should be accompanied by short solution latencies in the solution phase, calculate selections by longer solution latencies. As per the predictions of the SAC model it was found that rapid strategy selections could be made accurately: Retrieve selections were followed by significantly shorter solution latencies than when calculate was chosen in the selection phase, indicting that the rapid strategy selections were corroborated by actual strategy use in each trial.

This finding was evident in three separate experiments using arithmetic stimuli (Reder & Ritter, 1992; Schunn et al., 1997) but can only be accounted for by the SAC model, directly contradicting the predictions of the Automaticity models. However, this finding has to be treated with a degree of caution as it has not been replicated in any other papers than those authored by Reder and colleagues, nor has the reliability of this effect been tested using different experimental designs. In other paradigms which measure judgements of predicted future performance, such as Judgements-of-Learning (Koriat & Ma'ayan, 2005; Koriat, Ma'ayan & Nussinson, 2006; Nelson, 1996; Nelson & Dunlosky, 1991) or Ease-of-Learning (Koriat, Ma'ayan & Nussinson, 2006; Nelson & Narens, 1990), it is accepted that the accuracy and consistency of these judgements may be questionable (Koriat & Ma'ayan, 2005), one problem being that the mechanisms that underlie such responses are not fully understood (Koriat, 2006; Koriat, Ma'ayan, Nussinson, 2006; Smith, Shields & Washburn, 2003). Another possibility is that the finding is contingent upon the design of the experiment, in which participants are asked to predict their subsequent strategy selection, before solving a problem. Following the assumptions of the Automaticity models this would suggest that order of processing in this design (i.e., strategy selection then strategy execution) is artefactual, rather than being the natural processing order in problem solving assumed by Reder and colleagues. Furthermore, analysis of the strategy selection latencies was largely overlooked in all studies (Reder & Ritter, 1992; Schunn et al., 1997). Such an analysis may be highly useful in determining exact the relationship between FoK and strategy selection (e.g., Lachman & Lachman, 1980) and identifying whether Reder and colleagues dual-phase design is representative of the manner in which adults normally solve problems in the real world.

1.3.2 The relationship between strategy use and solution latencies

This relationship is exemplified by a key phenomenon in the mental arithmetic literature, the *problem-size effect* (Groen & Parkman, 1972). As Figure 1.5 illustrates, there is a positive correlation between solution latencies and the magnitude of a problem's solution (see also Penner-Wilger, Leth-Steensen & LeFevre, 2002). Problems with a larger solution, and hence larger operands, elicit longer solution latencies than problems with smaller operands and solution. Furthermore, for larger problems, a greater frequency of incorrect answers is given than in smaller problems (Zbrodoff & Logan, 2005). This effect has consistently been shown to be present in multiplication, addition, division and subtraction problems, in production problems (i.e., 9 + 7 = ?) and verification problems (i.e., 9 + 7 = 14, true or false?). Furthermore, the problem size effect is found in individuals of different ages, cultures and languages and applies to problems with single- and multiple-digit operands (see Zbrodoff & Logan, 2005 for review).

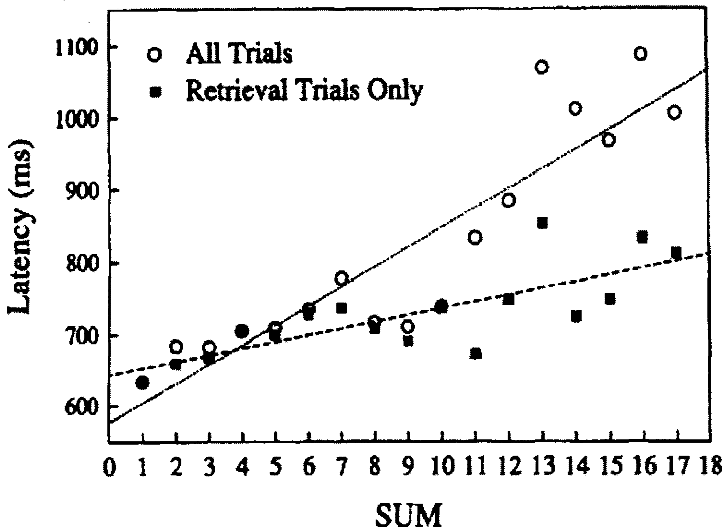

Figure 1.5: Median latencies for single-digit addition sums, collapsed across participants for all trials (open circles) and for trials in which the retrieval strategy was used (closed squares). Tied problems (i.e., 6 + 6 = ?) were omitted. The dotted line represents the best-fitting regression line, where the problems sum was used as a predictor. The dashed line only represents trials where the retrieval strategy was selected (adapted from LeFevre, Sadesky & Bisanz, 1996).

Recent interpretations of this effect assert that the positive correlation between solution latencies and the magnitude of a problem's solution is due to a principled variation in the strategies applied to problems (Zbrodoff & Logan, 2005). Problems with larger operands require more complex solution strategies than smaller problems. This effect is borne out by the assumption that particular problems are consistently solved by specific strategies (Campbell & Fugelsang, 2001; Campbell & Timm, 2000; Campbell & Xue, 2001; Hecht, 1999; 2002; LeFevre et al., 1996; LeFevre, Sadesky & Bisanz, 1996). For example, LeFevre, Sadesky and Bisanz (1996) reported that participants used retrieval more frequently with simple problems (i.e., those which summed to less than 10) and a calculation strategy, termed *transformation*[1], more frequently for problems summing to more than 10 (see Figure 1.6). This suggests that conventional approaches to reporting latency data, which are insensitive to strategy use (i.e., averaging across trials then participants), may be responsible for this positive

[1] The calculation strategy, transformation (LeFevre, Sadesky & Bisanz, 1996) is used to break a problem into smaller sums which then can be solved using retrieval procedures. The answers from these intermediary sums are then recombined to solve the original problem. Referred to elsewhere in the literature as a *decomposition* strategy (Hecht, 1999; 2002; Siegler, 1987).

correlation (see the dotted line in Figure 1.5). However, as the dashed line in Figure 1.5 illustrates, problem size effects are also evident within strategies (in this instance the retrieval strategy) as well as between strategies (but see Hecht, 1999; Lee & Kang, 2002; Seyler, Kirk & Ashcraft, 2002; Wheeler, 1939). But it must also be noted that calculation strategies (e.g., transformation and counting) typically produce greater problem size effects than the retrieval strategy (Campbell & Timm, 2000; Campbell & Xue, 2001; Compton & Logan, 1991; LeFevre, Sadesky & Bisanz, 1996; Siegler, 1987; Siegler & Shrager, 1984).

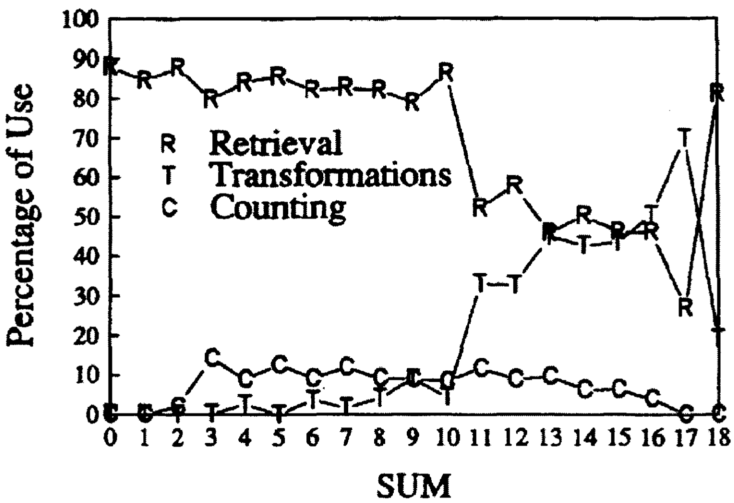

Figure 1.6: The percentage of single-digit addition problems solved by reported strategy use; retrieve (R) and the transformation strategy (T). Also, the counting strategy (C) where participants reported incrementing one operand by the magnitude of the other (adapted from LeFevre, Sadesky & Bisanz, 1996).

It is however the exceptions to the problem size effect that are most revealing of the mechanisms used to solve problems. For example, problems comprising *decade* numbers (10, 20, 30… 90) or *fives* (5, 15, 25…95) may be solved rapidly by direct fact retrieval where problems comprising other numbers of similar magnitude (e.g., +/- 1) require calculation procedures. Most adults would rapidly solve 20 + 30 = ? using the retrieval strategy, whereas 19 + 29 = ? would normally require the application of a calculation procedure eliciting a longer solution latency. This effect violates the positively skewed distribution of responses which normally fit solution latencies to answer size (Penner-Wilger, Leth-Steensen & LeFevre, 2002) representing a switch in selection from calculation to retrieval procedures (LeFevre, Sadesky & Bisanz, 1996; Zbrodoff & Logan, 2005).

In summary, the reliability of the problem size effect has established the magnitude of a problem's answer as a key predictor of strategy use in models of strategy selection including ACT-R, ITAM, CMPL, ASCM, SCADS and SCADS*. However, all the models barring SCADS* have struggled to account for exceptions to the effect without invoking ad-hoc mechanisms which are taught to recognise the exception problems and what strategy to apply to the problem. The feature detection overlay in SCADS* differs only from an ad-hoc mechanism in that it has the ability to learn new exception problems. Perhaps more importantly however, the reliable relation between solution latencies and strategy use acts as an accepted analytical tool used to identify strategy selections in arithmetic tasks (e.g., Lebiere & Anderson, 1998; Reder & Ritter, 1992; Schunn et al., 1997). Specifically, shorter solution latencies are expected if the retrieval strategy is used, whereas longer solution latencies are expected if a calculation strategy is employed. However, this principle is unable to distinguish between the different types of calculation strategy employed.

1.3.3 Problem familiarity correlates with reported strategy selections

A number of studies have suggested that problem solving is mediated by a mechanism which rapidly monitors the contents of memory (Koriat & Lieblich, 1977; Metcalfe, 1986; Metcalfe & Weibe, 1987; Reder, 1987). In a series of experiments modelled by the SAC account, Reder and colleagues examined the sensitivity of a monitoring mechanism, which elicits FoK judgements, to the familiarity of mental arithmetic problems (Reder & Ritter, 1992; Schunn et al., 1997). They investigated what processes underlie FoK judgements and whether these judgements (and hence mechanisms) directly influence strategy selection. Previously Reder (1987) had demonstrated in non-arithmetic problem solving studies that priming a problem's terms induced stronger FoKs, but without accompanying benefits to the recall or recognition the problem's answer. From this it was inferred that judgements of problem answerability may have been based upon the familiarity of a problem's terms — and by extension FoKs — rather than, as all the other models advocate, the magnitude of a problem's solution.

In Experiment 1 (Reder & Ritter, 1992), reported that greater exposure (and hence familiarity) to a problem confers a greater likelihood of choosing the retrieve strategy in selection tasks. Using the same dual-phase experimental design as detailed in 1.3.2, participants were repeatedly exposed — with systematically varying frequency — to double-digit addition and multiplication problems (e.g., 16 + 31 or 17 x 23) in order to provide their participants with opportunity to learn answers. For each presentation of a problem they were required to make a rapid retrieve or calculate strategy selection and then solve the problem. Frequently encountered problems elicited greater frequencies of retrieve strategy selections than rarely encountered problems.

To isolate which components of a problem influence strategy selection, novel problems were presented towards the end of the experiment which looked similar to the previously learnt problems, but were reconfigured with operator switches (17 x 23 to 17 + 23), operand reversals (17 x 23 to 23 x 17) or novel combinations of operands (17 x 23 to 17 x 24). Results indicated that the exposure effect was controlled by the familiarity of the problem terms, rather than the familiarity of the problem's answer, in particular by the familiarity of the pair of operands in the problem. However, these effects were only reported for multiplication problems. As the authors acknowledge, the incentive scheme designed to promote accurate strategy selection and rapid solution latencies flawed the results from the study using addition problems (Reder & Ritter, 1992). In a second experiment, a novel operator with a longer solution algorithm, named *sharp*[2] (#), replaced the addition problems used in the first experiment. As before, greater familiarity with a problem (i.e., exposure to a problem) elicited a greater frequency of retrieval strategy selections, irrespective of the operator in a problem.

In keeping with the tradition set by ACT-R, ITAM, SCADS* and CMPL which asserts that strategy selections are determined by the prior success of a strategy in providing a solution to a problem, a major criticism of the finding is that accurate strategy selections under speeded conditions could be informed by an *early read* or partial retrieval of the answer (e.g., Blake, 1973), rather than the familiarity of the problem's terms. An early read account assumes that the progress of a solution strategy in solving the problem is monitored online and a progress report can be accessed to inform responses demanded prior to the completion of the strategy. To rule out this possibility, Schunn et al (1997; Experiment 1) using double-digit

[2] Sharp (#) problems were calculated as follows, the solution 40, to the problem 23 # 67 = ?, would be derived by completing three steps; [(2 + 6) * (3 + 7) * 3] modulo 100, then [8 * 10 * 3] modulo 100 which equals 240 modulo 100 = 40.

multiplication problems, employed a similar priming technique as in Reder and Ritter (1992) but split the procedure into two distinct phases. In the first block of trials, participants were repeatedly presented with specific problems and requested to select a strategy (priming only the problem terms), in the second block, participants were required to select strategies and then answer the problem (priming both the problem terms and problem's answer). In contrast to the ACT-R, ITAM, ASCM, SCADS, SCADS* and CMPL accounts, it was found that strategy selection was not influenced by priming of the problem's solution as well as the terms of the problem. This suggests that the familiarity of a problem's terms is the key determinant of strategy selection as per the predictions of the SAC account.

1.3.4 Summary of key empirical phenomena

To date none of the models of strategy selection have been able to account in full for all of the key empirical findings. The most controversial finding, that strategies can be selected quickly and accurately (Reder & Ritter, 1992; Schunn et al., 1997), suggests that when solving a problem, a strategy is selected before any attempt to derive a solution to the problem is initiated. However, as there has been no direct empirical investigation to my knowledge designed to question the contrary position (i.e., the predictions of the Automaticity models) this finding has to be treated with some caution.

In addition to the distinction between the Automaticity and Adaptive selection models, based upon the manner in which strategies are applied to a problem, a second issue in the literature is highlighted by the empirical work undertaken within the paradigm. A degree of controversy exists in respect to the factors invoked to determine strategy selection. Most models posit that associative strength based factors are the key determinants of selection (for the Adaptive selection models, i.e., ASCM, CMPL, ASCM, SCADS and SCADS*) or predictors of actual strategy usage (for the Automaticity models, i.e., ITAM, ACT-R and DOA): Specifically, the strength of the link between a representation of the problem and the problem's solution. On a superficial level this approach meshes with the problem-size effect (Groen & Parkman, 1972; Zbrodoff & Logan, 2005). Smaller problems are solved more rapidly and accurately than larger problems, which are less common in day-to-day life or schooling. In such accounts, the frequency of prior exposure to a problem and its solution serves to reinforce the associative link between the problem and its solution, thus increasing the accuracy and rapidity of problem solutions. In contrast the SAC model asserts that strategy selections are determined by a rapid evaluation of a problem's terms, specifically an analysis of the familiarity of the operands in a problem. This approach asserts that a metacognitive mechanism monitors the familiarity of a problem's terms and bases strategy selection upon the magnitude of the FoK elicited by the familiarity of a problem's terms.

In light of the dearth of empirical research in the paradigm, in this thesis two key controversies will be further investigated. Firstly, that of whether strategies are selected before they are deployed to solve a problem, and secondly, the factors which determine strategy selection will be identified. In the following section, the general methodology used to investigate these issues will be detailed.

1.4 TASK OUTLINE

In order to investigate the two key issues outlined in the previous section, a task was required which afforded a fine-grained analysis of the strategy selection process. A secondary consideration was that the two key factors may interact and thus would require investigation in a design which allows both issues to be examined in tandem and in isolation. One option was to simply present participants with a sum and ask them to solve it. In this design, it is possible to infer strategy use on each particular problem from solution latencies, whereby short solution latencies would indicate a retrieve strategy selection, long solution latencies a calculation strategy selection (e.g., Hecht, 1999; 2002). This design would demonstrate the stimulus driven factors that may influence strategy selections (e.g., the magnitude of a problem's solution or the familiarity of the problems terms) but would not provide any information about the manner in which the solution strategy was selected (i.e., automatically or adaptively). A second possibility was to record strategy selections in a pseudo (e.g., Experiment 1, Rickard, 1997; Schunn et al., 1997) or algebraic arithmetic task (e.g., Compton & Logan, 1991; Experiment 4, Logan, 1988; Experiment 2, Rickard, 1997; Rickard, 2004). Here participants would be taught arithmetic facts through repeated exposure to specific problems, a design normally used to chart the transition from algorithmic problem solving to direct memory retrieval as a function of learning. However, this design does not afford an examination of how strategy selections are made for commonly encountered problems, one of the goals of this thesis.

The third option was to employ the Game Show design (see Reder, 1987; Reder & Ritter, 1992; Schunn et al., 1997), a methodology with the power to identify the factors that determine strategy selection and to examine the mechanics of the selection process. Each trial in the methodology has two distinct phases, the first immediately followed by the second. In the first phase of each trial, the selection-phase, participants are presented with a problem and asked 'when solving the problem would you retrieve the answer directly from memory or use a calculation algorithm?' Here, a time limit of 850 ms is set for the retrieve or calculate strategy selection, the predicted strategy selection and the time required to make the selection are recorded. In the second phase of each trial, the solution-phase, participants are asked to solve the problem as quickly and accurately as possible. The actual response and the time taken to make the response are recorded.

This design affords an insight into the manner in which strategies are applied to a problem. Accurate and rapid predicted strategy selections can only be made in the selection-phase if a distinct strategy selection phase precedes strategy deployment (Reder & Ritter, 1992; Schunn et al., 1997). Furthermore, the solution phase of each design attests to the accuracy of the predicted strategy selection. Accurate retrieve selections made in the selection-phase should be followed by short solution latencies in the solution phase. Calculate selections made in the selection-phase should be followed by longer solution latencies in the solution phase. Lastly, by manipulating the types of problems presented, or the frequency with which they are presented, it is possible to identify these problem-driven factors that influence strategy selection.

1.5 INTRODUCTION SUMMARY

In this chapter I have reviewed the computational models of strategy selection and the key empirical findings pertinent to the paradigm. In doing so, key areas of contention have been highlighted in respect to the structure and the predictions of these models. To illustrate the degree of controversy within the literature, all eight models detailed previously make different predictions and are based upon different assumptions. Accordingly, a program of systematic and comprehensive experimental research is required to test the veracity of these models.

The empirical work in the thesis will investigate the two key issues; the manner in which strategies are applied during a problem-solving episode (i.e., automatically or adaptively) and the identification of factors which determine strategy selection. Predictions from each of the computational models will be contrasted in a number of experiments which employ a methodology designed by Reder and colleagues (see Reder, 1987; Reder & Ritter, 1992; Schunn et al., 1997) to develop the SAC model.

1.5.1 Empirical Series 1

This series was designed to test the competing Adaptive and Automaticity accounts of strategy selection. To do so, predictions derived from one of the Adaptive accounts, the SAC model (Reder & Ritter, 1992; Schunn et al, 1997), were used as a focal point for the investigation as out of all the models detailed in Chapter 1 this model is detailed to the highest resolution. Not only was it possible to investigate the operation of the selection mechanism itself, but also the suitability and robustness of the dual-phase design for subsequent experiments. The series goes on to adopt a systematic approach to identifying the key factors responsible for strategy selections. While in the SAC account selection is determined by the familiarity of a problem's terms (Reder & Ritter, 1992; Schunn et al, 1997) all of the other accounts propose that the association between a problem and its solution would determine whether the retrieve or calculate strategy was selected. Manipulations were employed to examine this issue also investigating whether selection is influenced by particular problem features. Reder and Ritter (1992) noted that in some instances participants appeared to base strategy selections purely upon the type of operator in a problems (i.e., x or +). Examining this contention, problem features were manipulated and evaluated within the framework of the Adaptive and Automaticity accounts.

1.5.2 Empirical Series 2

The studies reported within the second empirical series were designed to refine the problem familiarity and problem feature effects revealed in Chapter 2. Of the existing accounts only the SAC model predicted that the familiarity of a problem would influence selection while none of the accounts were able to detail in full the empirical signature of the problem feature effects. Accordingly, the experiments presented in this series sought to establish the mechanisms through which their influence upon selection is realised. In particular whether strategy selections are driven by consciously directed procedures, or whether unconscious processes are responsible for problem familiarity and feature effects. Furthermore, the investigation also considered how the influence of these factors is realised in real-world problem-solving episodes. Following Cary and Reder's (2002) proposal, the notion that in real-world problem-solving episodes a range of factors, not just problem familiarity and problem features, may influence selection was considered. In particular, two issues were explored, that of how the selection mechanism accommodates multiple cues to retrieve/calculate selection and how selection is influenced by biases inherent in the wider processing context.

CHAPTER TWO

2.0 ABSTRACT

Predictions derived from the SAC model were used as a basis for examining the selection mechanism. In all of the experiments it was evident that retrieve/calculate selections could be made rapidly, and with a degree of accuracy, contrary to the predictions of the Automaticity accounts. Accordingly, following the Adaptive account, it was proposed that selections are made in a distinct phase before an attempt to execute a problem-solving strategy is engaged. Confirming the procedural demarcation between selection- and solution-phase it was evident that the dual-phase design is a sufficiently robust methodology for subsequent investigation (Experiments 1b, 2b and 3). Furthermore, the familiarity of a problem, rather than its answer was shown to influence selection (Experiments 1a and 2a) supporting the predictions of the SAC model (Reder & Ritter, 1992; Schunn et al, 1997). However, problem feature effects upon selection could not be easily assimilated into that model (Experiments 2a and 2b) suggesting that the selection mechanism is sensitive to more factors than previously conceived.

2.1 INTRODUCTION

The experiments presented in this Chapter 2 test predictions derived from the models detailed in the Chapter 1. A methodology was used to investigate two key facets of the models. Firstly, the issue of whether strategies are chosen adaptively in a distinct selection phase was addressed; specifically, whether strategy selection precedes an attempt to solve the problem (i.e., the Adaptive account) or whether strategies are applied automatically, in a default order, to solve the problem (i.e., the Automaticity account). Secondly, the types of arithmetic problems used within this design were manipulated to identify the particular qualities of a problem that influence selection. These two ends will be achieved by focusing on the methodology underpinning the SAC model, the Game Show design (Reder & Ritter, 1992; Schunn et al., 1997). This stands as the most comprehensive example of experimental work in the arithmetic strategy-selection paradigm to date and as an appropriate starting point for the empirical work in this thesis.

Experiments 1a and 2a in the present chapter employed the dual-phase Game Show design. In this methodology, participants are first presented with a problem and are required to make a rapid strategy selection, choosing either the retrieve or calculate strategy (i.e., the *selection-phase*). In the *solution-phase,* which immediately follows the selection-phase participants are required to solve the problem. Experiment 1a examines the influence exerted upon selection by the familiarity of a problem's terms and the familiarity of the problem's answer. The problems presented in the current study were designed to follow the predictions of the Problem Size effect (Groen & Parkman, 1972), whereby solution latencies increase in accordance with the magnitude of the addends in a problem. For example, solving 26 + 29 = ? takes longer than solving 16 + 19 = ?. Experiment 2a, using the same methodology, examined the selection process in three types of problem that violate the positive correlation between addend magnitude and solution latencies predicted by the Problem Size effect (Groen & Parkman, 1972). For example, 20 + 30 = ? can be solved much more rapidly than 19 + 29 = ?, despite the similarity in addend magnitude. Three types of problem, *decades* (i.e., 20 +30), *fives* (i.e., 25 +35) and *mixed* (i.e., 25 + 30) were used to investigate whether specific features common to a particular problem, for example that *both addends are multiples of 10*, influenced the selection process. Experiments 1b and 2b stand as replications of each respective experiment in which only the first phase, the selection-phase, is tested. Results from these experiments indicated whether the act of solving a problem in the solution-phase of each trial influences subsequent performance in the selection-phase of each trial. This will indicate whether the dual-phase Game Show design used in this thesis is validated or undermined. Adopting the reverse rationale, Experiment 3 examined whether crosstalk between the selection-phase influences performance in the solution-phase of each trial by testing the solution-phase in isolation.

2.2 EXPERIMENT 1a

Results from Reder and colleagues' (Reder & Ritter, 1992; Schunn et al., 1997) studies, hereafter referred to as the *Game Show* studies, suggest that strategy selections can be made rapidly, accurately and are influenced by the frequency of exposure to specific problem constituents (Reder & Ritter, 1992; Schunn et al., 1997). The more frequently an individual is exposed to a problem the more likely they are to choose to retrieve the solution to that problem. Together these findings suggest that strategy selections are made in a distinct phase. This operates before an attempt to solve the problem is initiated and is determined by a rapid evaluation of the frequency of exposure to, or familiarity of, a problem's terms. The current experiment was conducted for two purposes. Firstly, to extend the core findings of the Game Show studies, examining how they relate to the competing models of strategy selection. Secondly, as a general test of the dual-phase Game Show design and how the adaptations made to that design in this study influence selection.

As the findings from the Game Show studies act as the foundation for the SAC model, the current study also stands as a direct test of the predictions of this account. First I shall detail the two key modifications made in the current experiment to Reder and Ritter's (1992) design. Both of these key changes were introduced to establish the robustness of familiarity effects upon selection once problematic features of the classic design were removed. In the Game Show studies, problem familiarity was manipulated by varying the frequency with which specific problems and specific problem components were presented within the experiment. This approach to familiarisation served to control — and was used to manipulate — the strength of long-term memorial representations. For example, specific problems or specific pairings of problem components (e.g., '23 +') partnered with other operands (i.e., 23

+ *38*, 23 + *49* or 23 + *16*), were presented on multiple occasions during the experiment to induce familiarity. More frequent exposure to a particular problem, or pair of problem components presented together served to increase their familiarity.

A further aspect of the Game Show design was that, to encourage participants to actively engage in learning these problems and their answers, an incentive scheme was used. This sought to establish experimentally induced familiarity differentials between problems (or components of the problem) that were presented frequently or infrequently. The scheme utilised four payoff situations designed to discourage participants from defaulting to the 'safer' option in the selection task and selecting calculate for every problem. Furthermore, it encouraged participants to learn the arithmetic facts they derived from processing the problems. This promoted the selection of the retrieve strategy in the selection-phase and the use of retrieval procedures during the solution-phase of the experiment. Fifty points were awarded on a trial when participants selected retrieve during the selection-phase and produced a correct answer in the solution-phase within 1050 ms or 1400 ms (depending upon the number of digits in the answer, 3 or 4 respectively). If the calculate strategy was chosen in the selection-phase, followed by a correct answer produced within 24 s or 30 s (again dependant upon the number of digits in the answer), the participant was rewarded with 5 points. Alternatively, if participants produced a correct answer but violated the deadlines in either phase, 1 point was awarded. No points were given when subjects failed to meet the deadline in both phases, or failed to solve the problem correctly. Participants were given .005 cents per point and awarded a bonus of $1 if their point total exceeded the current top score (see Reder & Ritter, 1992).

Over a series of experiments, this scheme proved successful, promoting retrieve selections for difficult problems (i.e., double-digit multiplication problems and double-digit problems with a novel operator termed as *sharp*). However, for double-digit addition problems (Reder & Ritter, 1992, Experiment 1) scores were unexpectedly low. Many participants reported that they chose retrieve purely in attempt to beat the incentive scheme and were unable to solve the problems within the time limit allowed for a correct retrieve selection. Accordingly, the authors presented only a superficial analysis of the data from the addition problems, focusing primarily upon multiplication problems. To circumvent this flaw, the approach taken to problem familiarity in the current experiment was to derive measures of problem familiarity from conventional measures of long-term familiarity. This approach is widely used in tests of familiarity in a number of related paradigms such as tests of recognition memory (see Yonelinas, 2002 for review). Problem familiarity was derived by adding the familiarity rating (see Gielen, Brysbaert & Dhondt, 1991) of each addend in a sum. The relationship between independent observations of numbers x and y and concordant observations of x and y stipulates that more frequent independent observations equate to a greater frequency of concordant observations over time (see Church & Hanks, 1990 for similar argument). It was predicted that sums classed as high in familiarity are likely to elicit a greater percentage of retrieve selections in the selection-phase of each trial than unfamiliar sums.

Using this approach, three key findings underpinning the predictions derived from the Game Show design and consequently the SAC model of selection, were examined. Firstly, the Game Show studies demonstrated that strategy selections can be made rapidly, indicating that selection occurs before a deployed solution strategy is executed to completion. In this study, similar to the previous Game Show studies, a time limit of 850 ms (see also Reder & Ritter, 1992; Staszewski, 1988; Vernon & Usher, 2003) was set to ensure that participants were unable to base retrieve or calculate selections made in the selection-phase upon feedback derived from completed solution procedures. Campbell and Austin (2002) have shown when solving single-digit addition problems that 57% of answers were initiated within 750 ms and 73% within 900 ms. Accordingly, the additional encoding time necessary for processing double-digit addend sums — as opposed to single-digit sums (Reder & Ritter, 1992) — suggests that more than 900 ms will be required before a response derived retrospectively from a completed solution procedure can be returned. Strategy selections made within 850 ms are therefore argued to be determined by a strategy selection mechanism rather than feedback from completed solution procedures (Reder & Ritter, 1992; Schunn et al, 1997)

The second key finding from the Game Show studies is that these rapid strategy selections are largely accurate, i.e., statistically, they predict actual strategy use, a finding testified to by the time taken to solve the problem in each trial. When retrieve was chosen in the selection-phase it was found that solution latencies in that trial were relatively short. Conversely, protracted solution latencies were evident in the solution-phase if calculate was selected in selection-phase (Reder & Ritter, 1992; Schunn et al., 1997). A similar approach to accuracy was adopted in the current study to determine the accuracy of selections. Finally, Reder and Ritter (Experiment 1;

1992) have shown that when participants were frequently (as opposed to infrequently) presented with either the first operand, first operand and operator pairing or second operand, they were more likely to choose retrieve in the selection-phase and produce shorter solution latencies when solving the problem. From this it was tentatively suggested (because of the aforementioned confounding influence of the incentive scheme employed) that for double-digit addition problems, the familiarity of a problem's terms tempted participants to respond with erroneous retrieve selections. More importantly, this would indicate that strategy selection was indeed sensitive to the familiarity of the problem. To test the problem familiarity effect, in the present experiment problems were either low or high in pre-experimental familiarity, rather than using a technique which experimentally induced familiarity. If selection is influenced by the familiarity of a problem's terms it was predicted that a greater percentage of retrieve selections, and consequently a lower percentage of calculate selections, would be made for problems with relatively familiar as opposed to unfamiliar terms.

Contradicting the predictions of the SAC model, the Automaticity class of models (i.e., ACT-R, ITAM and DOA) and one model from the Adaptive class of models, the CMPL account, assert that apparently rapid and accurate strategy selection is determined by a competing account, the *early read* account [3]. Common to these models is the notion that selection is determined by the strength of the associative link in memory between a problem and its solution. Generally speaking, the act of encoding a problem automatically initiates the search for the problem's answer (Logan, 1988). Accordingly, on an operational level, the familiarity — as familiarity and associative strength are largely inseparable in the models of strategy selection — of the answer will determine (in the case of CMPL where strategies are selected before deployment) or predict (in the case of the Automaticity models where strategies are applied in an automatic manner) strategy selection. It may be that when making a strategy selection under time pressure, individuals can utilise feedback from incomplete solution procedures as a cue to selection. For example, in *tip-of-the-tongue* states the answer itself cannot be recalled but information pertaining to the answer such as the first letter of the answer, the location of a primary stress, or

[3] It should be noted here that the complex array of factors that are posited to guide adaptive strategy selections in the ASCM, SCADS, SCADS* simulations make the models virtually unfalsifiable on an experimental level. However, findings from this experiment (and the remainder in this thesis) where possible will be applied to these models.

knowledge of the number of syllables in an answer may be known (Brown & McNeill, 1966). The strength of the tip-of-the-tongue state acts as an indicator of the probability that the actual answer can be recalled.

In line with the difficulty of the problems presented in the current study and in the absence of an incentive scheme, it was deemed unlikely that any of the problems could be solved using direct retrieval procedures and that the calculate strategy would be selected in a higher percentage of trials than the retrieve strategy. Accordingly, effects of problem and answer familiarity upon strategy selection would be evident in the percentage of calculate strategy selections reported in the selection phase. Although Reder and Ritter (1992; Experiment 1) demonstrated that answer familiarity did not influence strategy selection, the present experiment includes a direct test of the early read account (cf. Blake, 1973) in light of the modifications made to the original Game Show design. For problems with answers rated as low or high in familiarity it was anticipated that a greater percentage of calculate selections would be expected in the low answer familiarity condition than high answer familiarity condition if an early read mechanism influences selection. However, in the absence of answer familiarity effects, if a greater percentage of calculate selections were evident in the low problem familiarity condition then this would support the SAC account. This would indicate that pre-existing familiarity, rather than experimentally induced familiarity influences strategy selection. To ensure that strategy selections could be made within the 850 ms time limit, the strategy selection latencies were recorded, affording an opportunity to examine for the first time whether the time taken to select a strategy is influenced by either familiarity measure. Furthermore, problem solution latencies were recorded as a measure of the accuracy of the strategy selections made during the selection-phase of each trial, whereby shorter latencies were anticipated if retrieve was chosen in the selection-phase rather than if calculate was selected.

2.2.1 Method

2.2.1.1 Participants

Twenty-four undergraduates from the School of Psychology at Cardiff University were given course credit for their participation. All were native English speakers reporting normal hearing and corrected or normal vision.

2.2.1.2 Materials & Design

The addends in each problem summed to less than 100 and were drawn from a sample ranging from 12 to 49. Each addend pair was from the same decade class (e.g., 23 + 29), none of the addends were divisible by 5 or 10 and there were no tied addends (i.e., 23 + 23). Two variables were investigated; sum familiarity (i.e., the summed familiarity of the problem terms) and answer familiarity. The four experimental conditions were contrasted in a repeated measures design; low sum familiarity/low answer familiarity, low sum familiarity/high answer familiarity, high sum familiarity/low answer familiarity, and high sum familiarity/high answer familiarity (see Appendix E, table E1, for stimuli). Participants completed 16 practise trials followed by 64 experimental trials.

Familiarity scores were derived from the number frequency measures developed by Gielen, Brysbaert and Dhondt (1991). In that study, twenty participants rated the numbers between — and including — 0 to 99 on a scale ranging from 1 (extremely rare) to 6 (extremely frequent) in respect to their frequency of occurrence in everyday life. Ratings were then linearly rescaled (by subtracting 1 and multiplying the results by 100) to give a familiarity rating of between 0 and 500 for each number. Sum familiarity ratings in this study were obtained by simply adding the familiarity ratings of the two addends in each problem. The answer familiarity was purely the familiarity rating of each problem's answer. Two repeated measures ANOVAs were run, confirming that there was a significant difference between the low and high levels of sum familiarity, $F(1, 15) = 105.34$, $MSE = 4142.89$, $p < .001$ and answer familiarity; $F(1, 15) = 412.63$, $MSE = 427.50$, $p < .001$.

2.2.1.3 Procedure

The experiment was programmed and compiled in Visual Basic 6.0. Participants were instructed that they would be presented with a series of arithmetic problems, each trial comprising two phases. Each trial commenced when the participant pressed the enter key to start the program, after which a fixation mark ("X + X") appeared in the centre of the screen, it flashed three times, each flash interleaved by a blank screen for 850 ms. On what would have been the fourth appearance of the fixation mark, the problem appeared in its place and the selection-phase commenced immediately. An on-screen caption prompted participants to make a retrieve/calculate selection by pressing the *r* button for retrieve or *c* for calculate.

The *r* button was located on the *z* key and the *c* button on the *m* key of a standard qwerty keyboard. Participants were advised before the experiment that their selection had to be made within 850 ms and that the time limit did not allow for much consideration of the problem. Furthermore, that retrieve should be selected for trials where they "would be able to instantly provide a solution if they were solving the sum, for example 10 + 10". Calculate should be chosen "when the sum may need to be broken down, such as 14 + 17. Here, the sum may be broken into 4 + 7 and 10 + 10, the products of those recombined to answer the original sum". The same examples were relayed to each participant verbally by the experimenter and also as part of the written instructions. Care was taken to provide the exact same instructions to each participant. As Kirk and Ashcraft (2001) previously demonstrated, instructions emphasising the usage of the retrieve strategy increased self-reported use of retrieval procedures relative to a non-biased condition. Similar effects were found when task instructions emphasised the use of the calculate strategy in problem-solving tasks (see also Blöte, Van der Burg & Klein, 2001; Gardner & Rogoff, 1990; Siegler & Lemaire, 1997). Accordingly, use of neither strategy was emphasised and equivalent time and description were provided for the retrieve and calculate strategies.

If a strategy selection was made within 850 ms participants immediately proceeded to the solution-phase. The runtime program automatically recorded the strategy selection or marked the response as *late* if a selection was not made within the time limit. In the event a strategy selection was late, the solution-phase of each trial started immediately after the expiry of the 850 ms time limit. In this phase participants were instructed to "solve the sum as quickly and accurately as possible". The answer was entered by the participant using the numerical keypad on a standard qwerty keyboard, and 'enter' was pressed to confirm the answer. The runtime program recorded the participant's given answer and the time taken to enter that answer. As the experiment was self-paced, participants were required to click on a 'continue' button to proceed to the next trial.

2.2.2 Results

2.2.2.1 Scoring Procedure

Two measures were taken during the selection phase; the *strategy selection* (retrieve or calculate) and the *strategy selection latency* which was recorded from the initial point at which the problem was presented up to when the participant selected retrieve or calculate. If a selection was not made within the time limit no selection latency was recorded for that particular trial and the response was marked as *late*. In the solution-phase, the participant's *given answer* and accuracy of the answer (in respect to the correct answer) was recorded. The *solution latency* was automatically recorded by the program, either from the entry of a strategy selection, or if the selection was late, at the expiry of the 850 ms time limit until the participant pressed the 'enter' key to confirm his/her given answer. Both latency measures were accurate to approximately ±2 ms and for analysis, split according to the retrieve/calculate selection made in selection-phase. So, on each trial a selection latency, solution latency and sum solution were tagged to the retrieve or calculate selection made during the selection-phase. If a strategy selection was late the data from that trial were removed from the final analysis.

As a consequence of this experimental design, participants could only select the retrieve or calculate strategy. Accordingly, if a participant chose to select calculate for all the problems in one of the four experimental conditions there would not be any entries for the other measures (strategy selection latency, sum solution, or solution latency) for retrieve in that condition as they are tagged to the strategy selection made in the selection-phase. This serves to reduce the degrees of freedom in some statistical analyses reported in the thesis. Participants with missing values are automatically removed from the repeated measures analysis used in the thesis by SPSS. This however only applies to analyses where the two alternative fixed choice is required, such as selecting between retrieve or calculate. In instances where missing values are justified, such as when considering the percentage of incorrect sum solutions where a missing value indicates a correct response, missing values are replaced with a zero. Throughout the thesis, statistically non-significant main effects and interactions are not reported unless theoretically warranted.

2.2.2.2 Strategy Selection

As predicted, mean strategy selections reported in Table 2.1 illustrate that calculate was selected in a considerably higher percentage of trials than retrieve. Accordingly, the results analysis in the current study will focus upon the influence exerted by the experimental variables upon the percentage of calculate selections rather than retrieve selections. This approach is adopted in light of the dependant relationship between retrieve and calculate selection such that the factor/s which influence calculate selections will have an inverse effect on retrieve selections. One corollary with such an analysis is that although retrieve and calculate are dependent measures, the percentage of late responses also need to be considered. If the tendency to produce a late response is influenced by either familiarity measure then the dependant relationship between retrieve and calculate selections may be compromised. For example, in the low sum familiarity/low answer familiarity condition, on average the retrieve strategy was selected in 16.1% of trials, the calculate strategy in 73.7% of trials. In the remaining 10.2% of trials selections were not made within the time limit and were hence marked as late. If in another condition 20% of trials had late responses and this was caused by one or more of the experimental manipulations then the hypothesised dependant relationship between retrieve and calculate would need to be reconsidered. In this study, a repeated measures ANOVA demonstrated that neither sum nor answer familiarity, influenced late selections, nor was there a significant interaction between those two variables (all Fs < 3.72, all ps $> .08$). Accordingly, it is surmised that the dependant relationship between the percentages of retrieve and calculate selections was not differentially influenced by condition type.

Turning to the influence exerted by the two experimental variables upon the percentage of calculate selections, a 2 (sum familiarity; low vs. high) x 2 (answer familiarity; low vs. high) repeated measures ANOVA revealed that a greater percentage of calculate selections were made for unfamiliar than familiar problems, indicating for the first time that pre-experimental familiarity — rather than experimentally induced familiarity — influences strategy selection, $F(1, 23) = 13.2$, $MSE = .28$, $p = .001$. Furthermore, null effects of answer familiarity upon the percentage of calculate selections were revealed, $F(1, 23) = .38$, $MSE = 68.8$, $p = .54$, demonstrating that the familiarity of a problem's solution does not influence the selection process.

Table 2.1.

Summary by condition of mean retrieve (Ret) and calculate (Calc) strategy selections (in %), strategy selection latencies (in ms) and solution latencies (in ms). Corresponding standard deviations are in parentheses.

	Low Answer Familiarity				High Answer Familiarity			
	Low Sum Familiarity		High Sum familiarity		Low Sum Familiarity		High Sum familiarity	
Measure	Ret	Calc	Ret	Calc	Ret	Calc	Ret	Calc
Strategy selected (%)	16.1 (16.2)	73.7 (17.6)	24.7 (17.4)	59.9 (20)	16.9 (17.5)	70.6 (22)	32 (23.6)	59.6 (25)
Selection latency (ms)	568 (102)	510 (83)	599 (84)	510 (77)	547 (121)	522 (96)	589 (84)	506 (94)
Solution latency (ms)	3904 (1490)	4369 (1210)	3979 (1540)	4475 (1270)	3708 (1360)	4401 (1040)	3627 (1290)	4000 (1250)

2.2.2.3 Selection Latency Analysis

In total, only 12.2% of strategy selections were not made within the time limit of 850 ms. Similar to results evident in the Game Show studies (Reder & Ritter, 1992; Schunn et al, 1997) this indicates that participants were comfortably able to make strategy selections within the time limit. In respect to the experimental manipulations, a 2 (sum familiarity; high and low) x 2 (answer familiarity; high and low) repeated measures ANOVA revealed that of calculate selection latencies were insensitive to sum familiarity, $F(1, 23) = .76$, $MSE = .002$, $p = .39$, and answer familiarity, $F(1, 23) = 0.28$, $MSE = .001$, $p = .6$. Similarly, it was found that the time taken to select the retrieve strategy was insensitive to the familiarity of the problem's answer, $F(1, 15) = 1.55$, MSE = .01, $p = .23$. However the time taken to select the retrieve strategy was influenced by the familiarity of the problem, $F(1, 15) = 7.66$, $MSE = .002$, $p = .014$. Post-hoc comparisons indicate that retrieve selections were made more slowly for familiar problems than unfamiliar problems (*M* difference = 31 ms, $p = .014$). Furthermore, when comparing the mean selection latencies for the two strategies (retrieve or calculate), a repeated measures ANOVA, including sum and answer familiarity, revealed that calculate selections were made more rapidly than retrieve selections, $F(1, 38) = 6.24$, $MSE = .02$, $p = .02$.

2.2.2.4 Solution Latency Analysis

Only 4.7% of solution latencies were excluded from the analysis for exceeding 10 s. Incorrect sum solutions accounted for 7.4% of responses demonstrating that participants were able to complete the solution-phase with a high degree of accuracy. A 2 (sum familiarity; high and low) x 2 (answer familiarity; high and low) repeated measures ANOVA run on solution latencies tagged to calculate selections made during the selection-phase revealed a significant interaction between sum familiarity and answer familiarity, $F(1, 23) = 4.95$, $MSE = 1.54$, $p = .04$. Simple effects indicated that answer familiarity effects were only evident for familiar problems. Familiar problems with more familiar answers were solved more rapidly than familiar problems with unfamiliar answers, $F(1, 23) = 6.97$, $p = .02$. Solution latencies tagged to retrieve selections made in the selection-phase were immune to effects of sum familiarity, $F(1, 16) = .22$, $MSE = .83$, $p = .65$ and effects of answer familiarity were marginally significant, $F(1, 16) = 4.36$, $MSE = .53$, $p = .05$. They followed the same pattern as solution latencies tagged to calculate selections in that familiar problems with familiar answers exhibited shorter solution latencies than familiar problems with unfamiliar answers.

To examine the accuracy of strategy selections made in the selection-phase, the strategy selection was compared to the solution latency produced in the solution phase of each trial (see also Reder & Ritter, 1992; Schunn et al., 1997; Siegler & Lemaire, 1997). To recapitulate, it was predicted that if retrieve was chosen in the selection-phase and participants could, to some degree at least, predict which strategy they would use, shorter solution latencies should be evident in the solution phase of each trial than if calculate was selected in the selection phase. As Table 2.1 indicates, problems where calculate was selected in the selection-phase were partnered by longer solution latencies than when retrieve was chosen in the selection-phase. However, this trend failed to reach statistical significance, $F(1, 39) = 1.78$, $p = .19$. In the present experiment however, the solution latencies evident in Table 2.1 were befitting of the type of sums presented to the participants, none of which were expected to be solved by direct retrieval procedures. To corroborate this relation, Figure 2.1 illustrates that the percentage of calculate selections made increased in line with the length of time required to solve the problems ($r = .85$, $p = .003$) indicating that the percentage of calculate selections made was befitting of the actual difficulty (as indexed by solution latencies) of the problems tested.

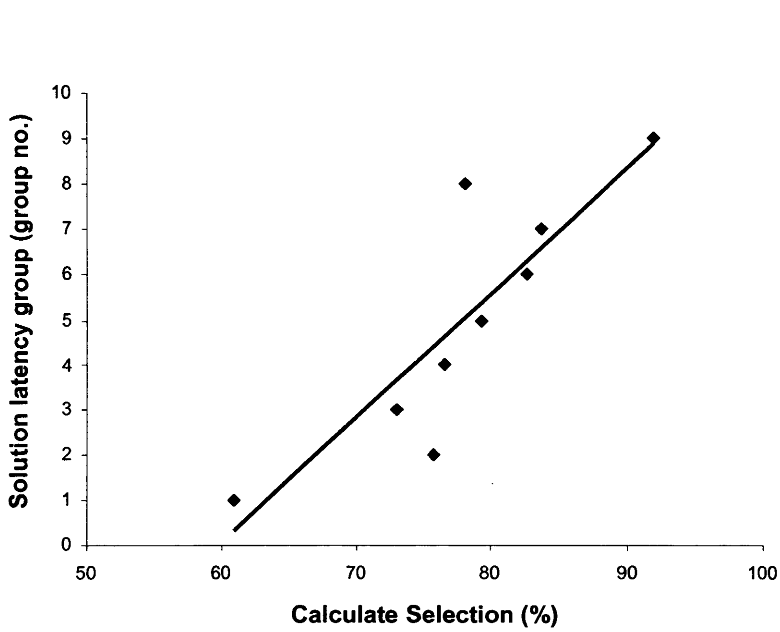

Figure 2.1: The percentage of calculate selections made within each solution latency group collapsed across sum and answer familiarity conditions. Solution latency groups run from 1-2 s, 2-3 s, 3-4 s, 4-5 s, 5-6 s, 7-8 s, 9+ s. The trend line represents a strictly linear relation.

2.2.3 Discussion

The present experiment was designed for two specific purposes, firstly to test the suitability of the dual-phase methodology as a basis for further empirical investigation of the selection process. Secondly, to test the core findings derived from the Game Show studies which act as the foundation for the SAC model. In respect to the first aim, the design used in the present experiment produced comparable responses to the Game Show studies. Participants were able to make rapid strategy selections with a high degree of accuracy (i.e., which predict subsequent solution latencies). In a significant departure from the Game Show design, strategy selection was shown for the first time to be influenced by pre-experimental familiarity (rather than experimentally induced familiarity). This suggests that strategy selections in conventional arithmetic tasks, not only in the dual-phase design, are sensitive to the familiarity of the problem rather than being contingent upon task learning. Furthermore, an incentive scheme, as used in the Game Show studies, is not required to obtain accurate indices of performance in this methodology and task. Accordingly, it seems apparent that the experimental design employed here is an effective means by which to obtain accurate indices of performance in strategy selection.

When considering the underlying mechanism responsible for selection the current study contrasted predictions derived from the familiarity based Adaptive account proposed by the SAC model and the early read account supported by the predictions of the Automaticity models and CMPL. To recapitulate, in the early read account selection is determined by the strength of activation elicited by an obligatory attempt to solve the problem (see Logan, 1988), determined by the strength of the association between the problem and its solution. In this experimental design, the time limit imposed upon the selection-phase of each trial precludes the participant from solving the problem and basing the predicted strategy selection upon a retrospective interpretation of the problem-solving episode. Accordingly, the progress made in solving the problem, at the point where this process is necessarily truncated so a selection can be returned within the time limit, may serve to determine selection. Here the strength of the association between the problem and its answer will predict the progress made by retrieval procedures in solving the problem. For example, in problems with more familiar answers, and hence a stronger association between problem and answer, greater progress will have been made in finding the answer than in problems with unfamiliar answers. Accordingly, from this account it was predicted that more retrieve selections would be made for problems with familiar answers than problems with unfamiliar answers. However, the findings from the current study fail to support the early read account as strategy selections were insensitive to the familiarity of the problem's answer. Furthermore, the speed with which selections were made, consistently within 850 ms and the evident accuracy of strategy selections demonstrated by solution latencies serves to question the predictions of models reliant upon the early read account.

In contrast, the findings from the present study fit the predictions of the SAC model (Reder & Ritter, 1992; Schunn et al., 1997) where a rapid assessment of the familiarity of the problem's terms determines strategy selection. To rule out the possibility that other factors contributed to the problem familiarity effect a covariate analysis was conducted (see Appendix A). This demonstrated that there were no other problem-level factors which covaried significantly with calculate selections. A series of regressions were also conducted to examine whether any other characteristics of the problems (i.e., answer magnitude, answer familiarity, whether a carry would be required when solving the problem) influenced selection (see Appendix B). It was found that both answer magnitude and problem familiarity accounted for a highly significant but comparable amount of the variance in calculate selections, 46% and 44% respectively. Accordingly, the possibility remains that either answer magnitude or problem familiarity could account for strategy selections in this experiment. Two lines of converging evidence, although failing to rule out the possibility that answer magnitude influences selection, support the problem familiarity-based account of selection. Firstly, for answer magnitude to influence strategy selection a model of selection akin to the early read account would be required. However, the implausibility of an early read account has already been outlined in this section. Secondly, Reder and colleagues have previously demonstrated by dissociating answer magnitude and problem familiarity that selection is determined by problem familiarity. In their priming methodology they systematically manipulated the frequency with which problems with a range of answer magnitudes, or components of those problems, were presented to participants in a pre-test phase. In a test phase they demonstrated that increased exposure to a problem (or its components) serves to increase its familiarity and the likelihood that the retrieve strategy is chosen in subsequent problem solving episodes (Reder & Ritter, 1992; Schunn et al, 1997). No effects of answer magnitude were reported demonstrating using their methodology that problem familiarity rather than answer magnitude determines selection.

If problem familiarity rather than answer magnitude does influence selection the question still remains as to why these two factors accounted for an almost identical proportion of the variance in calculate selections. One explanation can be found in an argument proposed by Siegler and Araya (2005). They stated that the frequency of exposure to a problem (and consequently the familiarity of a problem) is predicted by the magnitude of a problem's answer. Studying learning materials they discovered that individuals are exposed to smaller problems more often than larger problems. This is reflected in the significant negative linear relationship between exposure to a problem (and therefore the familiarity of the problem) and the size of the problem which can be indexed by the magnitude of a problem's answer. Accordingly, as found in the present experiment, a significant negative correlation between problem familiarity and answer magnitude would be expected.

One finding that could not be easily reconciled by the SAC model was that calculate selection latencies were made more rapidly than retrieve selection latencies. Schunn et al (1997) propose that retrieve/calculate selections are resolved by a single threshold against which FoK responses are subjected. To recapitulate, the activation level elicited by the problem terms of the most active node in memory determines the strength of the FoK. Low FoKs stem from low levels of relative activation and propagate a low probability of selecting retrieve. Conversely with a high FoK, where the relative spread of activation between nodes is disparate, the probability of selecting retrieve will be greater (Schunn et al., 1997). When considering that the SAC model posits a single threshold it would be predicted that retrieve and calculate selections (i.e., the decision not to retrieve) would be made at the same juncture in time. Accordingly, further specification of the selection criterion is required.

It may be possible that the SAC's single threshold mechanism acts in tandem with a conditional rule forcing an extended search to find further information validating the selection (including the problem's solution) if the activation level of the most active node is relatively, but not especially high. Evidence from related experimental paradigms supports this notion, suggesting that individuals search longer for answers to questions that they already know (Gruneberg, Monks & Sykes, 1977; Lachman & Lachman, 1980; Reder, 1987). Affording additional time to confirm whether retrieve should be selected allows the individual to obtain more information to verify the veridicality of the candidate response. However, an alternative explanation can be derived from the recognition memory literature. Explanations of the Fan effect (Anderson, 1983) suggest that the number of facts linked to a particular item regulate old (i.e., the item was evident in a study list) and new (i.e., the item was not seen in the study list) responses. Within this paradigm, effects of strength and interference have opposite effects on the time taken to retrieve items from memory. Generally it is established that by practising memory retrievals on particular items subsequent responses on the same item become faster (Yonelinas, 2002). However, the retrieval of facts rich in associates is slower than retrieval of items with few associates (Reder, 1988). This is because of the competition for the available, but limited, activation is greater than when there are few associates. Consequently the activation value required for one node to breach the threshold for retrieval is not reached as quickly when relative levels of activation at more than one node is equable. In the case of strategy selections, it is likely that there are fewer associates, or that the associates will have less strength for problems which are solved by calculate rather than retrieve selections. Accordingly, this interference account may be responsible for calculate selections being made more rapidly than retrieve selections. An alternative proposition is that separate threshold mechanisms may be responsible for retrieve and calculate selections. The SAC model proposes a single-counter mechanism in which the FoK determines whether the retrieve strategy should be selected. Here calculate selections occur by default, as a function of not choosing the retrieve strategy, conversely a dual-counter account (see Nelson & Narens, 1990) proposes that separate counters (and hence thresholds) determine selection. Accordingly, as separate thresholds determine retrieve and calculate selection responses can be returned at different points in time.

In summary, the present study has served to confirm the three key predictions of the SAC model. Firstly, that strategy selection is made very rapidly, within a timescale which indicates that selections are not determined by completed solution procedures. Secondly, that selections are made with a degree of accuracy and thirdly, that the familiarity of the problem, rather than the familiarity of the problem's answer influences selection. The analysis presented in this experiment fails to rule out the possibility that answer magnitude influences selection (as examined in Appendix B), however, lines of converging evidence support a problem familiarity-based account as the most likely description of the selection process. The SAC model fails to account for the finding that calculate selections are made more rapidly than retrieve selections. Furthermore, the findings illustrate for the first time that pre-experimental familiarity influences selection and that an incentive scheme is not required to stimulate accurate responses. In Experiment 1b a potential corollary of this experimental design is examined. It may be that the act of solving a problem after selecting a strategy serves to bias subsequent strategy selections. Specifically, by solving a problem after selecting a strategy, participants may be able to learn which strategy to correctly select for latter problems in the experiment (i.e., within-task learning). Furthermore, the following experiment affords an opportunity to test the veracity of the results revealed in the present study.

2.3 EXPERIMENT 1b

Reder and colleagues' Game Show design was originally based upon the assumption that problem-solving is characterised by two distinct phases; the strategy selection phase during which a strategy is chosen, followed by a solution phase in which the chosen strategy is executed in an attempt to solve the problem. Evidence from Experiment 1a serves to affirm this postulated order of processing. It was evident that selections can be made quickly (i.e., before the answer can be retrieved from memory) and that they predict which strategy is actually used to solve a problem. As the focal remit of this thesis is the strategy selection mechanism, rather than the process by which a solution is derived from a deployed strategy, it is necessary to lesion the selection- and solution-phases to ensure that in this experimental series the two phases operate in isolation. In the Game Show studies the coupling of these phases was used for two purposes; firstly, as a means of testing the accuracy of selections, and secondly, to encourage within-task learning of the presented arithmetic problems. The stimuli they used were all difficult problems which at the start of the experiment could not be solved by retrieval procedures (e.g., double-digit operand multiplication problems). Repeated exposure to these stimuli and their components served to familiarise participants with these problems. Once familiarised, a greater percentage of retrieve selections were elicited in the selection-phase and in the solution-phase direct retrieval procedures were employed more often. However, Experiment 1a indicated that pre-experimental familiarity, as opposed to familiarity artificially induced within the experiment, can also influence selection. Accordingly, it is necessary to ensure that familiarity induced during the course of the experiment does not influencing selection in this experimental design which, as in Experiment 1a, was attributed to pre-experimental familiarity.

To examine this potential confound, in Experiment 1b the selection-phase was run in isolation (the *selection-phase design*), where the strategy selected (retrieve or calculate) and the time taken to select a strategy in each trial were recorded. The principle aim of the current study was to examine whether the act of solving a problem influences future strategy selections. This was done by contrasting performance in the dual-phase design to the selection-phase design. Specifically, examining whether feedback from the solution phase, which may reveal the accuracy of prior strategy selections, influences subsequent strategy selection. In the dual-phase design the solution phase acts as a probe of the strategy selected as it provides the participant with the opportunity to evaluate the accuracy of their strategy selections. This can be achieved by comparing the strategy chosen in the selection-phase to the strategy successfully used to solve the problem, providing a useful source of information which may direct future strategy selection. One possibility is that the presence of a probe may influence the strategies selected in latter problems. However, contrary to this position, Green, Cerella & Hoyer (2000), using a design which probed only half of their participants, found that the presence or absence of a probe did not influence selections. If feedback between the selection- and solution-phase does influence performance it would be predicted that the percentage of calculate (and hence retrieve) selections would differ significantly (see Rickard, 2004 for similar argument) in respect to the influence exerted by problem and answer familiarity between the two experimental designs. Therefore, if effects of sum but not answer familiarity are reported in the selection-phase design (as in the dual-phase design) this would indicate that feedback from the solution-phase does not modulate the effect of these variables upon the strategy selections reported in Experiment 1a. More importantly, that pre-experimental rather than experimentally induced familiarity can influence strategy selection. This confirms that the dual-phase design is representative of the normal processing order used in arithmetic problem-solving and provides an accurate measure of selection.

In the eventuality that the influence exerted by problem and answer familiarity does not differ between experimental designs the secondary aim of this study was to analyse of the validity of the strategy selection latencies reported in Experiment 1a. To recapitulate, mean selection latencies ranged between 500 ms and 600 ms, well within the allotted 850 ms time frame. This indicated that participants did not require the full time allocation to make accurate selections. Of greater interest is the finding — contrary to all the models of strategy selection detailed in Chapter 1 of this thesis — that calculate selections were made more rapidly than retrieve selections. Furthermore, only retrieve selection latencies were influenced by the familiarity manipulations, whereby latencies were longer for more familiar problems.

Before it is possible to draw further conclusions from these results, the present experiment was designed to replicate these three findings using the selection-phase design. In the dual-phase design, as soon as a strategy selection is made, participants were required to solve the problem. One possibility is that in addition to the selection-phase time limit, the immediate onset of the solution-phase acts as a further time pressure, forcing participants to rush their responses. Whereas in the experimental design employed in the present study, the secondary time pressure is circumvented as after each strategy selection (or expiry of the 850 ms time limit) the runtime program paused, prompting participants to click a 'continue' button to start the next trial. The same stimuli were presented in the current experiment as in Experiment 1a, testing the influence exerted by problem and answer familiarity upon the selection process. It was predicted that if the feedback gleaned from the act of solving a problem (i.e., the probe) influences subsequent strategy selections the pattern of influence exerted by sum and/or answer familiarity in the present experiment would differ significantly from that evident in Experiment 1a. Furthermore, if the solution-phase in the dual-phase design influenced the amount of time required to make strategy selections, mean durations in the current experiment would differ from those evident in Experiment 1a.

2.3.1 Method

2.3.1.1 Participants

Twenty-four undergraduates from the School of Psychology at Cardiff University were given course credit for their participation. All were native English speakers reporting normal hearing and corrected or normal vision and had not participated in any of the other experiments in this series.

2.3.1.2 Materials & Design

See corresponding section in Experiment 1a, the same set of materials and experimental design was employed in the present study.

2.3.1.3 Procedure

Exactly the same procedure and instructions were given to participants as in Experiment 1a, with three exceptions. Firstly, whereas participants in Experiment 1a were informed that there would be two phases in each trial (i.e., a strategy selection-phase and a solution phase), here participants were told that they had to select a strategy in each trial. Secondly, upon completion of the selection-phase (i.e., immediately after a strategy selection was made, or upon expiry of the 850 ms time limit), a 'continue' button appeared on the screen which once clicked initiated the lead-in to the next trial. Finally, participants were advised that they would not be required to solve any of the problems at any point in the experiment.

2.3.2 Results & Discussion

In this design the two measures taken were recorded in an identical manner to Experiment 1a, the actual *strategy selection* and the *strategy selection latency.* In line with the difficulty of the double-digit addition problems the calculate strategy was selected most often (see Figure 2.2). In respect to the principle aim of this study, there were main effects of sum familiarity upon calculate selections, $F(1, 23) = 34.58$, $MSE = 73.4$, $p < 0.01$, replicating the pattern of results found in the dual-phase design in Experiment 1a. The calculate strategy was selected less often (and hence retrieve more often) for more familiar problems, and null effects of answer familiarity, $F(1, 23) = 1.88$, $MSE = 135.4$, $p = .18$, were evident. Therefore, it seems reasonable to conclude that in the dual-phase design, the feedback that can be extracted from the act of solving a problem did not contribute to the action of problem familiarity during the course of this experiment.

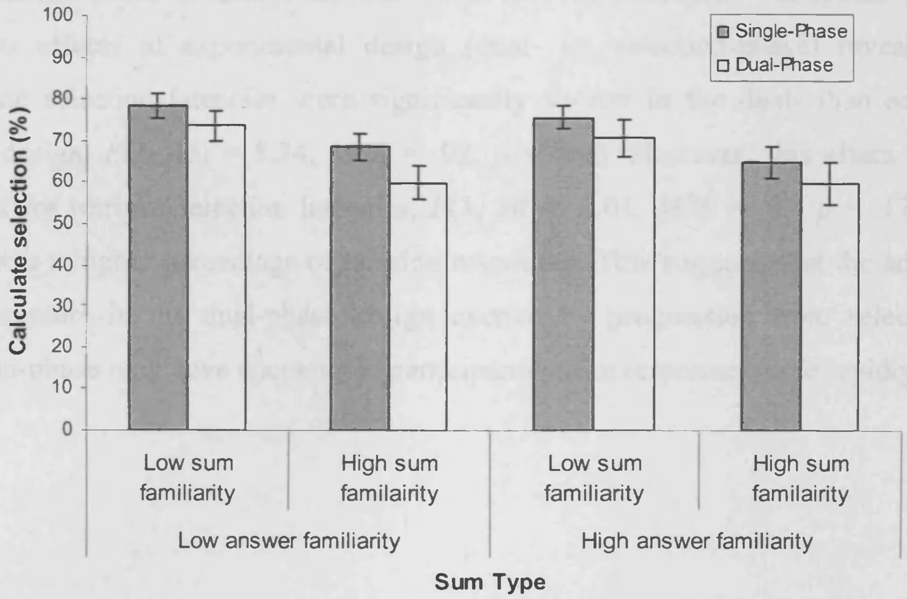

Figure 2.2. The percentage of calculate selections by condition. Grey bars represent data from the selection-phase design employed in Experiment 1b, white bars the data from the dual-phase design Experiment 1a. Error bars represent the standard error of the mean.

In total, 8.14% of strategy selections were not made within the 850 ms time limit compared to 12.2% in the dual-phase design. The mean duration of calculate selections were insensitive to answer familiarity, $F(1, 15) = 3.81$, $MSE = .001$, $p =$

.07, and sum familiarity; $F(1, 15) = 2.55$, $MSE = .001$, $p = .13$[4], replicating the key effects observed in the dual-phase design used in Experiment 1a. Analysis of retrieve selection latencies revealed a minor divergence from Experiment 1a. In the dual-phase design retrieve selections were made more slowly on average for familiar problems than for unfamiliar problems. In this design however null effects of both sum and answer familiarity, $F(1, 15) = .25$, $MSE = .006$, $p = .62$ and $F(1, 15) = 2.87$, $MSE = .005$, $p = .11$, were evident, thus failing to replicate this effect.

The secondary aim of the current study was to assess the validity of the selection latency measure in respect to the average time taken to select both strategies. As Table 2.2 demonstrates, there was a significant interaction between answer familiarity and the strategy selected (i.e., retrieve or calculate), $F(1, 38) = 5.23$, $MSE = .003$, $p = .03$. Simple effects reveal marginally significant differences between the mean latencies of retrieve and calculate selections but only in problems with familiar answers $F(1, 38) = 3.4$, $p =.07$. This finding partially replicates the effects evident in Experiment 1a where a significant difference between the duration of retrieve and calculate selections in each of the four within subjects conditions was found. Between subjects effects of experimental design (dual- vs. selection-phase) revealed that calculate selection latencies were significantly shorter in the dual- than selection-phase design, $F(1, 46) = 8.34$, $MSE = .02$, $p = .006$. However, this effect was not evident for retrieve selection latencies, $F(1, 30) = 2.01$, $MSE = .03$, $p = .17$, where there was a higher percentage of missing responses. This suggests that the additional time pressure in the dual-phase design exerted by progression from selection- to solution-phase may have encouraged participants make responses more rapidly.

[4] As in Experiment 1a, the degrees of freedom for this analysis, and subsequent analysis in this section, are lower than 23 as only 16 out of the 24 participants tested made retrieve selections in each of the four conditions.

Table 2.2.

Summary by condition of mean retrieve (Ret) and calculate (Calc) strategy selection latencies (ms) for dual- and selection-phase design. Standard deviations are in parentheses.

	Low Answer Familiarity				High Answer Familiarity			
	Low Sum Familiarity		High Sum familiarity		Low Sum Familiarity		High Sum familiarity	
Selection Latency	Ret	Calc	Ret	Calc	Ret	Calc	Ret	Calc
Selection-phase	609 (117)	568 (79)	603 (102)	594 (74)	638 (130)	562 (84)	608 (94)	582 (80)
Dual-phase	568 (102)	510 (83)	599 (84)	510 (77)	547 (121)	522 (96)	589 (84)	506 (94)

In summary, the findings from this and Experiment 1a reveal that the act of solving a problem and the feedback that can be derived from this act did not impact the influence exerted upon selection by problem and answer familiarity for these types of problems and in this experimental paradigm. That participants were not explicitly required to solve the problems after making a predicted strategy selection, a technique used by Reder and colleagues to experimentally induce familiarity, suggests that pre-experimental familiarity influenced selection in the current study experiment. Similar to Experiment 1a, strategy selection in this study was influenced by problem familiarity but not by answer familiarity such that more familiar problems elicited a greater percentage of retrieve strategy selections. Replication of the problem familiarity effect confirms that the dual-phase design provides a useful method to analyse the strategy selection process. However, it was found that selections were made more rapidly in the dual-phase design than in the selection-phase design. This suggests that in the dual-phase design the additional demand of solving the problem after making a selection forces participants to make more rapid responses in the selection-phase. As the influence of sum and answer familiarity upon selection was similar across experimental designs there is no reason to suspect that the added time pressure has a negative affect on the accuracy of strategy selections.

In respect to the secondary aim of the study, to replicate the other key effects illustrated in Experiment 1a, the mixed findings were revealing. From the analysis of selection latencies, inconsistent effects of sum and answer familiarity between experimental designs suggests that familiarity effects upon selection latencies should be treated with caution and may as such be an unstable measure. Furthermore, as in Experiment 1a, it was evident that calculate selections were made more rapidly than retrieve selections. Contradicting the predictions of the SAC model (Reder & Ritter, 1992; Schunn et al, 1997) this provides some supports Nelson and Narens (1990) dual-counter account of rapid memorial searches, as well as the interference account described within the Fan effect (Anderson, 1983). This finding also fits with an account in which a conditional rule dependant upon the strength of the FoK, affords the individual more time to search for information validating the proposed strategy selection. Further consideration of these hypotheses will be made throughout the course of the experiments presented in this chapter and in more detail in Experiments 5a, 5b, 6a and 6b of Chapter 2.

The present experiment has served to replicate the key findings from Experiment 1a, also demonstrating that the dual-phase design provides an appropriate methodology for investigating the influence of pre-experimental familiarity upon the selection process. Experiment 2a uses the dual-phase design to test a different class of double-digit addition stimuli that a number of strategy selection models have failed to model successfully.

2.4 EXPERIMENT 2A

Experiments 1a and 1b demonstrate that strategy selection in arithmetic problems is influenced by a rapid assessment of the familiarity of a problem's terms in line with the predictions of the SAC model. However, the specifications of the SAC model, built primarily upon results from double-digit multiplication stimuli, failed to account comprehensively for strategy selections in double-digit addition sums. In their Experiment 1, Reder and Ritter (1992) presented a mixture of addition and multiplication problems where participants consistently chose retrieve for double-digit addition — as opposed to multiplication — sums. Solution latencies indicated that a high percentage of retrieve selections were inaccurate: Participants reported they were made in an attempt to beat the incentive system used to stimulate rapid and accurate responses. Accordingly, the authors argued that in this design, the operator type (+ or x) was used as a cue to strategy selection and that the apparent lack of accuracy in participants' strategy selections was motivated by the incentive scheme used in their design. Such a finding promotes the notion that the presence of specific problem features in a task, such as operator type, irrespective of the familiarity of the problem terms, may also influence rapid strategy selections.

In the present experiment, for the first time in the paradigm, a systematic examination of the influence exerted by specific problem features upon the selection process was undertaken. Rather than using an incentive scheme and presenting problems with different operators (i.e., + and x) similar to the Game Show studies, a problem addend manipulation was employed. Problems were comprised of integers which were multiples of 5 (*decades,* i.e., 10, 20, 30, 40... and *fives,* i.e., 15, 25, 35, 45...). A number of sources of evidence converge upon the proposition that decade and fives numbers are processed in a different manner to other numbers of similar magnitude violating the assumptions of the problem size effect (Groen & Parkman, 1972). For example, in arithmetic tasks problems comprising these types of addends are solved more rapidly and with lower error rates than problems comprising other numbers of a similar magnitude (Campbell, 1994; Campbell & Graham, 1985; Campbell & Oliphant, 1992; LeFevre, Sadesky & Bisanz, 1996). Furthermore, decades are ranked as highly familiar numbers (Gielen, Brysbaert & Dhondt, 1991) and are used very frequently in a range of calculation procedures to simplify complex problems. Another possibility is that the primacy of decades numbers in arithmetic tasks may arise from the organisational principles used by the verbal naming code of numerical quantities (e.g. *thirty*-one or *seventy*-nine) and consequently verbal encoding processes (Fayol & Seron, 2005, cited in Campbell, 2005; McCloskey, 1992; McCloskey & Macaruso, 1995). English, German, Spanish, Italian and French verbal numerical systems are all constructed around a base of 10, e.g., *twenty*-one, *thirty-four*. To illustrate, a memorial representation for the number 5030 may be {5}10EXP3 {3}10EXP1, the digits in braces (e.g., {3}) indicate quantity representations and the 10EXP*n* indicates an exponent of 10. For example, 10EXP3 represents 1000 (McCloskey, 1992; McCloskey & Macaruso, 1995).

By manipulating the types of addend pairs in a problem, the present experiment was designed to examine whether the strategy selection process is affected by particular problem features; specifically, whether addends in a problem are both decades, fives or contain one decade and one fives addend. Furthermore, Experiment 2a sought to replicate two findings from Experiment 1a while testing problems which were likely to be solved by a mixture of retrieval and calculate procedures, rather than just calculate procedures: firstly, that solution latencies correspond to the retrieve/calculate selection made in the selection-phase of each trial which testifies to the general accuracy of strategy selections, and secondly, that calculate selections are made more rapidly than retrieve selections, a finding that pertains to the underlying mechanics of the strategy selection mechanism.

2.4.1 Method

2.4.1.1 Participants

Twenty-four undergraduates from the School of Psychology at Cardiff University were given course credit for their participation. All were native English speakers reporting normal hearing and corrected or normal vision and had not participated in any of the other experiments in this series.

2.4.1.2 Materials & Design

The sum of each problem amounted to less than 100 and all addends were drawn from a sample of integers divisible by 10 (20, 30, 40, 50, 60, 70) or 5 (15, 25, 35, 45, 55, 65, 75). Three levels of *sum type* were prepared, each level comprising 12 novel problems: *decades* sums comprised two decade addends (e.g., 20 + 50), *mixed* sums contained one decade and one fives addend (e.g., 20 + 55) and the addends in *fives* sums (e.g., 25 + 55) were both fives. All problems were addition sums, each comprising two addends presented in the centre of the screen on one line, e.g., "40 + 50". The font used to present the sum was Arial, size 48 and none of the sums comprised tied addends (e.g., 40 + 40). Questions were arranged into different pseudo-random orderings for each participant. Sixteen practise problems were presented at the beginning of the experiment (see Appendix E, table E2, for stimuli).

2.4.1.3 Procedure

Procedural details were exactly the same as those employed in Experiment 1a.

2.4.2 Results

2.4.2.1 Scoring Procedure

The same four measures were recorded as in Experiment 1a; the *strategy selection* (retrieve or calculate), the *strategy selection latency*, the *given answer* and the *solution latency* (see corresponding section in Experiment 1a for more detail). To briefly recapitulate, both latency measures were tagged to the preceding strategy selection, so in each trial a strategy selection latency and solution latency was recorded for either the retrieve or calculate strategy.

2.4.2.2 Strategy Selection

In accordance with the difficulty of the problems, the retrieve strategy was chosen in a higher percentage of trials than the calculate strategy. As Table 2.3 illustrates, the percentage of retrieve selections was highest in the decades condition and lowest in the fives condition. A repeated measures ANOVA with three levels of sum type (decades, mixed and fives) revealed a significant main effect of sum type, $F(2, 22) = 9.07$, $MSE = 291.74$, $p = .001$. Planned pairwise comparisons indicate that the percentage of retrieve selections was sensitive to the three levels of sum type. Retrieve was selected more often — hence calculate less often — in the decades condition than fives ($p = .001$), decades than mixed ($p = .02$) and mixed than fives ($p = .03$).

As in previous experiments, to ensure that the dependant relationship between retrieve and calculate selections is not confounded by the percentage of late selections, a repeated measures ANOVA, run on the percentage of late responses in each sum type condition was conducted. It was found that the number of late selections was not influenced by sum type, $F(2, 22) = .9$, $MSE = 1.16$, $p = .41$. This indicates that the percentage of late responses in the current experiment did not compromise the dependant relationship between retrieve and calculate selections.

Table 2.3.

Summary by condition of mean retrieve and calculate strategy selections (in %), strategy selection latencies (in ms) and solution latencies (in ms). Standard deviations in parentheses.

	Decades		Mixed		Fives	
Measure	Retrieve	Calculate	Retrieve	Calculate	Retrieve	Calculate
Strategy selected	82.6	12.5	69.8	26.4	56.9	43.4
(%)	(30.9)	(22.9)	(43.5)	(39.1)	(43.7)	(41.8)
Selection latency	509	416	556	515	567	505
(ms)	(52)	(77)	(45)	(48)	(89)	(62)
Solution latency	1426	1971	1829	2793	2904	3294
(ms)	(218)	(387)	(380)	(474)	(459)	(805)

2.4.2.3 Selection Latency Analysis

Only in 9.26% of trials were strategy selections not made within the 850 ms time frame, a figure which compares to the 12.2% of selections excluded from Experiment 1a for the same justification. This indicates that despite the difference in dominant strategy selections between experiments (calculate in Experiment 1a, retrieve in the present experiment) participants were equally able to make rapid strategy selections within the time limit. Due to the low percentage of calculate selections returned in this study repeated measures ANOVA was not conducted on calculate selection latencies. Significant effects of sum type were revealed when considering the time taken to select the retrieve strategy, $F(2, 16) = 9.49$, $MSE < 0.01$, $p = .002$. Pairwise comparisons indicated that retrieve selection latencies were shortest in the decades condition, on average 47 ms faster than the mixed condition ($p = .003$) and 58 ms faster than the fives condition ($p = .03$). This effect may be more accurately recast as demonstrating the difference between the encoding demands of decades and the mixed and fives conditions, as only a 9 ms mean difference in retrieve selection latencies was evident between the mixed and fives condition. This indicates that selections were made most rapidly when both addends were divisible by 10. It is proposed that this finding is largely contingent upon the time taken to encode the three different types of problems. Where decades addends are encoded more rapidly than addends in mixed and fives problems. As Table 2.3 illustrates, and similar to the findings from Experiments 1a and 1b, retrieve selection latencies were consistently longer than calculate selection latencies $F(1, 23) = 13.71$, $MSE = .004$, $p = .001$.

2.4.2.4 Solution Latency Analysis

Error rates were negligible, 0.34% of sum solutions were incorrect in the decades condition, 2.08% in the fives and 2.43% in the mixed condition and only 5.2% of solution latencies were excluded for exceeding 10 s. A repeated measures ANOVA revealed that solution latencies tagged to retrieve selections were influenced by the type of sum presented (decades, mixed or fives), $F(2, 16) = 177.81$, $MSE = .11$, $p < .001$. Pairwise comparisons revealed that solution latencies for retrieve selections were longer in the fives than decade conditions, longer in the mixed than decade and fives than mixed conditions (all ps $< .05$). Therefore, decades problems were solved most rapidly while fives problems evoked the longest solution latencies.

When examining the accuracy of selections, solution latencies tagged to retrieve selections were on average shorter than solution latencies tagged to calculate selections in each condition, $F(1, 23) = 49.22$, $MSE = .15$, $p < .001$. This demonstrates, as predicted, that when retrieve was chosen in the selection-phase participants went on to solve the problem more rapidly in the solution-phase than when calculate was chosen in the selection-phase. Further analysis of the relationship between solution latencies and retrieve selections is illustrated in Figure 2.3. In the decades ($r = -.73$, $p = .03$), mixed ($r = -.85$, $p = .004$) and fives ($r = -.72$, $p = .01$) conditions, significant negative correlations between solution latencies (arranged into in 500 ms groupings) and the percentage of retrieve selections illustrates that shorter latencies were evident where there was a higher percentage of retrieve selections confirming the accuracy of selections.

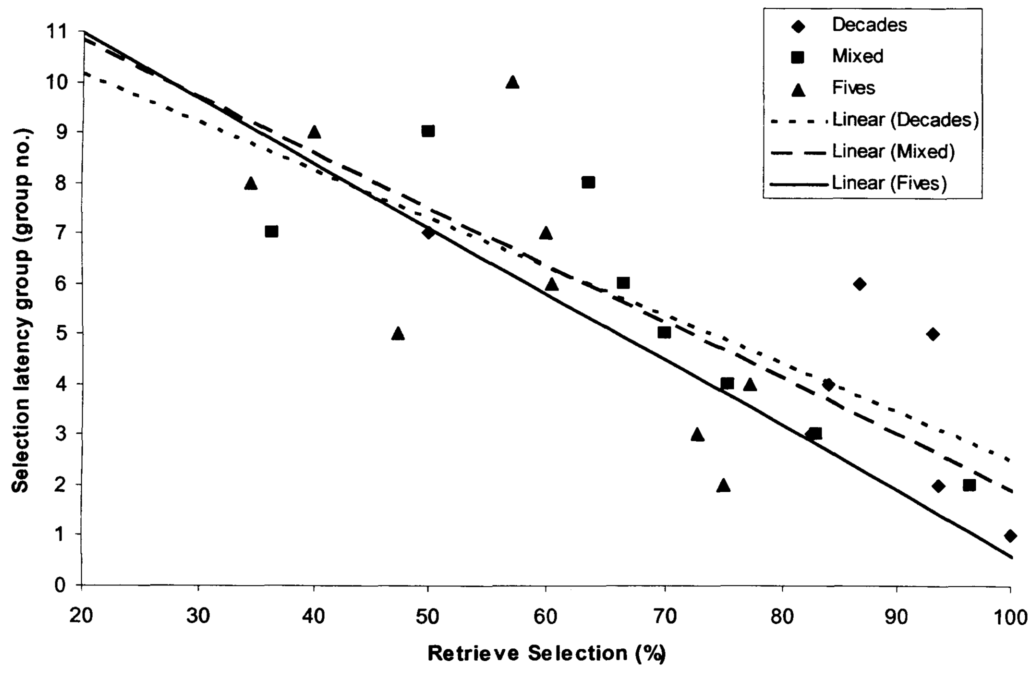

Figure 2.3: Correlation between the percentage of retrieve selections and solution latency group. Solution latency groups run from 0 s at 500 ms intervals up to 5 s, group 11 comprises all remaining responses from 5 s to 10 s latencies.

However, solution latencies in neither the mixed nor fives conditions met the criteria set at 1.4 s by other authors for a true direct retrieval (see Reder & Ritter, 1992; Schunn et al., 1997; also Staszewski, 1988). In the Game Show studies solution latencies were recorded from the point of sum presentation to the initiation of a vocalised answer. Here, allowing for a typed response and recording up to the point of answer confirmation, latencies under 2 s may be more representative of direct retrievals. By this reasoning, as Table 2.3 illustrates, solution latencies tagged to retrieve selections in the decades and mixed conditions may represent accurate strategy selections, but in the fives condition longer mean latencies reveal that a mixture of retrieval and calculation procedures may have been used to solve the problems. Interestingly, when questioned at the end of the experiment, a number of participants admitted to selecting retrieve for fives problems even though with hindsight they realised that they actually solved some problems, such as 35 + 45 = ? using three — albeit very rapid — addition sums; for example, 30 + 40, 5 + 5 then 70 + 10. Accordingly, apparently incorrect retrieval selections in the fives condition may be attributed to an inaccurate definition of retrieval procedures rather than incorrect selection *per se*.

2.4.2.5 Covariate Analysis

Despite the influence of sum type upon the percentage of retrieve selections in this experiment which supports the notion that specific problem features influence strategy selection, it was necessary to examine whether item-level variables, such as the familiarity of the problems' terms covaried significantly with sum type. If this was the case it may be that sum familiarity is responsible for the effects primarily attributed to sum type in the preceding sections of this analysis. Accordingly, two linear mixed models were run upon the percentage of retrieve selections in each condition in order to identify any potential covariates. Three covariates were considered, each of which shared a linear relationship to the percentage of retrieve selections; sum familiarity (the familiarity rating of both addends in a problem summed), the familiarity of the first addend and the familiarity of the second addend. Two models were tested, the first comprising sum familiarity as a covariate and sum type as a repeated measures factor (df = 6, AIC = 152.79) which provided a significantly better fit to the data than the second model (p = .01) which was comprised of sum type as a repeated effect and the familiarity ratings of the first and second addends as covariates (df = 6, AIC = 160.57). Notably, in the sum familiarity model the repeated effect of sum type, $F(2, 22.4) = 15.32, p < .001$, and the covariate sum familiarity, $F(1, 13.07) = 5.14, p = .04$, both reached significant levels suggesting that sum type was not solely responsible for the variation in the percentage of retrieve selections but that sum familiarity also played a role. To delineate between these two factors, effects of sum familiarity were analysed within the three sum type conditions (decades, mixed and fives). If the familiarity of a problem was responsible for the percentage of retrieve selections returned in each condition then a positive correlation between problem familiarity and the percentage of retrieve selections recorded would be expected. However, in each condition non-significant correlations were reported (all *r*s < .53, all *p*s > .07) indicating that problem familiarity effects were not masquerading as sum type effects. A series of regressions presented in Appendix C also confirm this finding, demonstrating that sum type is the greatest predictor of the retrieve strategy selections reported in this study.

2.4.3 Discussion

The results from the present experiment clearly replicate certain aspects of Experiment 1a demonstrating that strategies can be selected rapidly and accurately, favouring the case forwarded by the SAC model. However, contrary to the SAC models predictions, and replicating the findings from Experiments 1a and 1b, calculate selections were made more rapidly than retrieve selections. Of greater interest is the finding that retrieve strategy selections reported in this experiment appeared to be determined by specific problem features, specifically whether the addends in this experiment were decades or fives numbers. To illustrate, problems in the decades condition which comprised two addends divisible by 10 into integers elicited a greater percentage of retrieve selections than problems with either one addend divisible by 10 (i.e., mixed condition) or none (i.e., fives). This finding stems from the first systematic investigation of the influence exerted by specific problem features upon the selection process, illustrating that problem familiarity is not the sole determinant of strategy selection. Where problem features are shown to influence selection, in future experiments such effects will be termed *selection-by-feature* effects.

The Game Show studies have previously shown that selection in problems which violate the problem size appears to be based upon specific features inherent in a problem. Indeed the SAC model simulations include a conditional parameter "*does participant decide to never retrieve for one of the operators*" (Reder & Ritter, 1992; Schunn et al., 1997, p.12). Similarly, ACT-R has invoked additional assumptions in an ad-hoc fashion to model selection in problems divisible by 10 (Lebiere & Anderson, 1998). However, the specificity of such parameters stipulates that any additional parameters are identified in an ad-hoc fashion. The only model to include a mechanism with the potential to account for selection-by-feature effects is SCADS* (Siegler & Araya, 2005). In this model a feature detection mechanism, running in parallel with encoding processes, identifies both task relevant (e.g., the magnitude of the operands, whether the operands are ties) and irrelevant problem features (e.g., the size and colour of the operands). For example, in the current study a task relevant feature in the decades condition would be *both addends end in a zero* (i.e., a decades problem, 20 + 40 = ?). For each feature identified, SCADS* keeps track of two proportions; one proportion tracks the number of trials in which a particular feature was detected and in which a specific strategy (e.g., retrieve) produced good performance (i.e., a correct response and a faster solution latency than normal) relative to the total number of trials in which the feature was present. The second counter monitors the number of trials in which a particular feature was absent and the strategy produced unusually good performance relative to the number of trials in which the feature was absent. If a large enough difference between these two proportions exists over the course of several trials, the presence or absence of a particular feature is used to compute the potential strength of a strategy thus influencing the selection process (Sigler & Araya, 2005).

In isolation, the mechanism seems capable of accounting for the effects of sum type in the current study. However, to recapitulate, SCADS* not only bases strategy selections — which occur in a distinct selection-phase prior to strategy execution — upon feature detections, but also historical data concerning the prior success of strategy applications to specific problems, and the prior success of the strategy in general. The latter two measures are determined by the strength of the link between the representation of a problem and its answer, and hence on an experimental level would be indexed by the answer familiarity measure employed in Experiments 1a and 1b. As null effects of answer familiarity were reported, the model as a whole is undermined. However, the structure of the feature detection mechanism may be informative of the manner in which specific problem features influence selection. Further examination of the functionality of this mechanism will be presented in the studies presented in the following chapter.

To test the veracity of the findings reported in the present experiment, Experiment 2b seeks to replicate the key findings employing the selection-phase design and testing the same stimuli to identify whether within-task learning is responsible for the selection-by-feature effect. More importantly, it affords an opportunity to examine how potentially useful problem features are adopted by the selection process.

2.5 EXPERIMENT 2b

The current experiment was conducted for two purposes: Firstly, to replicate the selection-by-feature evident in Experiment 2a and secondly to examine whether the selection-by-feature effect is influenced by solving similar problems in previous trials. As in Experiment 1b, the selection-phase design was used and the same stimuli presented in Experiment 2a were deployed.

In Experiment 1b it was shown that pre-experimental problem familiarity influenced selection rather than experimentally induced familiarity as shown in the Game Show studies. However, as Experiment 2a illustrated, strategy selections for decade, mixed and fives problems were not influenced solely by problem familiarity but crucially by specific problem features. By lesioning the selection- and solution-phases of the dual-phase design, in the current experiment it was be possible to identify whether the problem features that were used to determine selection in Experiment 2a are identified as a consequence of selecting a strategy, then solving the problem. Or whether they were identified from arithmetic problem-solving episodes prior to the experiment and had been employed in instances prior to the experiment. As in Experiment 1b where the pattern of influence exerted by the two familiarity manipulations (problem and answer) did not differ between the selection-phase and dual-phase designs the same rationale was adopted here. Null effects of sum type upon the percentage of retrieve selections (unlike Experiment 2a) would indicate that problem features are only learnt through the act of selecting a strategy then solving the problem within the experiment. However, if the percentage of retrieve selections made do vary by sum type this will indicate that the problem features that do influence selection have been derived from processing episodes prior to the experiment and thus the selection- and dual-phase designs are representative of the selection process used in real-world problem solving.

Based upon the finding that sum type influenced the percentage of retrieve selections made, the secondary aim of the study was to examine some key comparisons derived from the previous three studies in this chapter. Firstly, that given the relative difficulty of these problems, in comparison to those presented in Experiments 1a and 1b, the retrieve strategy should be selected more often than the calculate strategy. Secondly, in respect to the duration of strategy selections, calculate selections should be made more rapidly than retrieve selections, and finally, effects of sum type should be evident upon the duration of retrieve selections, whereby decades problems elicited the fastest retrieve selections, fives problems the slowest response.

2.5.1 Method

2.5.1.1 Participants

Twenty-four undergraduates from the School of Psychology at Cardiff University were given course credit for their participation. All were native English speakers reporting normal hearing and corrected or normal vision and had not participated in any of the other experiments in this series.

2.5.1.2 Materials & Design

See corresponding section in Experiment 2a, the same set of materials and experimental design was employed in the present experiment.

2.5.1.3 Procedure

The exact same procedure was employed as in Experiment 1b and the same instructions were given to participants as those detailed in Experiment 1a.

2.5.2 Results & Discussion

As in Experiment 1b the only measures taken were the actual *strategy selection* and the *selection latency*. As Figure 2.4 illustrates, similar to the results derived from the dual-phase design, the retrieve strategy was selected more often than the calculate strategy. Furthermore, main effects of sum type were evident upon the percentage of retrieve selections made, $F(2, 22) = 16.81$, $MSE = 156.21$, $p < .001$, thus confirming that selection was influenced by the selection-by-feature effect. As participants had no experience of solving the problems within the experimental design adopted in the current study it was inferred that for a feature to be adopted by this mechanism the act of actually solving the problem is not necessary. A greater number of retrieve selections were evident in the decades condition than the mixed condition (p = .007), decades than fives and mixed than fives conditions (p < .001). To dissociate between selections based upon the selection-by-feature effect and a problem familiarity based mechanism, similar to Experiment 2a, effects of problem familiarity were analysed within sum type levels. A significant positive correlation between problem familiarity and the percentage of retrieve selections in each of the sum type conditions would indicate that selection was influenced by problem familiarity. However, non-significant correlations were reported in each condition (all rs < .52, all ps > .09), indicating that the selection-by-feature account provides the best explanation of performance.

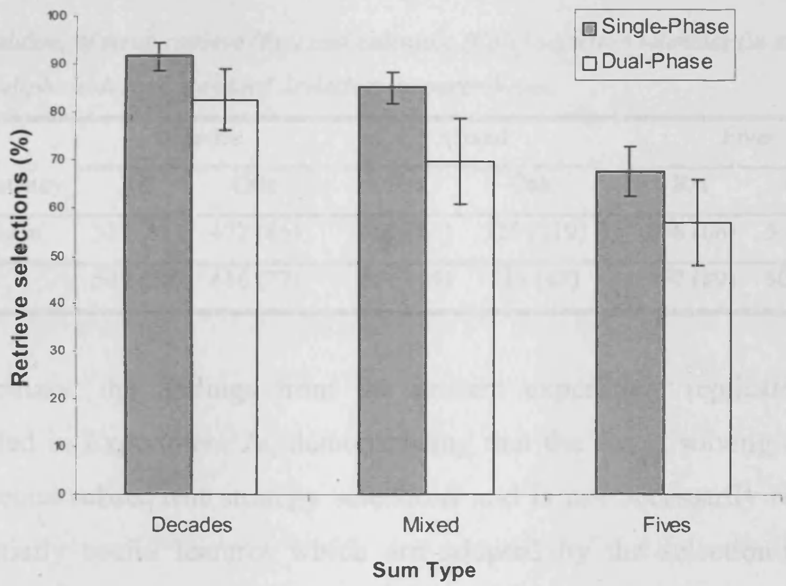

Figure 2.4: The percentage of retrieve strategy selections by condition. Grey bars represent data from the selection-phase design employed in the present experiment, white bars the data from the dual-phase design in Experiment 2a. Error bars represent the standard error of the mean.

In respect to the time taken to select retrieve and calculate selections, the duration of retrieve selections was influenced by sum type, $F(2, 22) = 17.11$, MSE = .002, $p < .001$, whereby retrieve selections were made more rapidly in the decades (M = 527 ms), than the mixed (M = 560 ms) and fives condition (M = 608 ms). When comparing retrieve and calculate selection latencies between experimental designs (selection- and dual-phase), there was no significant difference for retrieve selections, $F(1, 40) = 2.44$, $MSE = .006$, $p = .13$, indicating that the experimental design did not influence the time taken to select either strategy unlike Experiment 1b.

The only other departure from the findings reported in Experiment 2a was that main effects of sum type upon the percentage of late selections were evident, $F(2, 22) = 4.12$, $MSE = .48$, $p = .03$, indicating that a greater percentage of late selections was evident in the fives condition ($M = 7.99\%$) than the mixed ($M = 5.21\%$) or decades ($M = 2.43\%$). This finding in itself is unsurprising and reflects the main effects of sum type upon the duration of retrieve selections, such that faster selections were made for decades problems, hence a lower overall likelihood of producing a late response. Longer latencies were evident in the fives condition, indicating a greater potential for late responses.

Table 2.4.

Summary, by condition, of mean retrieve (Ret) and calculate (Calc) selection latencies (in ms) for selection- and dual-phase design. Standard deviations in parentheses.

	Decades		Mixed		Fives	
Selection Latency	Ret	Calc	Ret	Calc	Ret	Calc
Selection-phase	527 (53)	472 (86)	560 (54)	525 (119)	608 (66)	561 (91)
Dual-phase	509 (52)	416 (77)	556 (45)	515 (48)	567 (89)	506 (62)

In summary, the findings from the present experiment replicate the key findings reported in Experiment 2a, demonstrating that the act of solving a problem does not influence subsequent strategy selections and is not necessarily required to identify potentially useful features which are adopted by the selection-by-feature mechanism. That the same conclusion was drawn in Experiment 1b suggests that this finding can be generalised across different types of problem and strategy selections (i.e., retrieve and calculate). Furthermore, all of the key findings in Experiment 2a were replicated here, establishing some key empirical signatures for the first time in this paradigm. Specifically, that sum type influenced the percentage of retrieve selections supports the case that problem features were not identified through the act of solving prior problems. This confirms that the dual-phase design is an appropriate test of selection for a range of problems. In addition, calculate selections were made more rapidly than retrieve selections, and sum type was shown to influence the duration of retrieve strategy selections such that selections for decades problems were made more rapidly than the other conditions. This confirms that rather than being sensitive to familiarity or feature manipulations *per se*, selection latencies are more sensitive to encoding demands, where decades numbers are encoded more rapidly. The final experiment presented in Chapter 2 is designed to investigate whether the act of selecting a strategy or being required make a strategy selection before solving the problem influences the amount of time taken to actually solve the problem and/or the accuracy of given answers.

2.6 EXPERIMENT 3

Experiment 3, which investigates the solution-phase in isolation (i.e., the *solution-phase design*), is designed as a test of the corresponding measures taken in the dual-phase methodology reported in Experiments 1a and 2a. Findings from Experiments 1b and 2b indicate that the act of solving a problem did not influence subsequent strategy selections made within the experiment. The current experiment investigates the opposite eventuality, that the act of selecting a strategy before being asked to solve the problem influences the manner in which the problem is then solved. The dual-phase design is contingent upon the fact that the order of processing (strategy selection, then strategy execution) reflects the natural order undertaken by problem solvers in real-world problem-solving (Reder & Ritter, 1992; Schunn et al, 1997). Specifically, that the strategy selection made in the selection-phase of the dual-phase design determines which strategy is executed in the solution-phase. However, it may be that the transition between phases within each trial requires a re-initiation of the selection process at the outset of the solution-phase in the dual-phase design. If this were the case it would be predicted that there would be no difference in solution latencies between experimental designs (solution-phase vs. dual-phase), as in both designs the solution-phase of each trial requires selection and then solution processes. Conversely, the possibility remains that participants in the dual-phase design do not need to make another strategy selection when at the outset of the solution-phase. Accordingly, if this were the case solution latencies would be consistently shorter in the dual-phase design than the selection-phase design. The design also affords an opportunity to replicate the effects of the experimental manipulations employed in

Experiment 1a (sum familiarity and answer familiarity) and Experiment 2a (sum type; decades, mixed and fives) upon the solution phase of the design.

To test the two key aims of the current study for these eventualities the stimuli used in Experiments 1a and 2a were employed here. Participants were required to solve problems without making a making a predicted strategy selection in a distinct phase first. To identify whether the order of processing in the dual-phase design mirrors the order of processing in normal arithmetic problem solving, two comparisons are presented. Firstly, the pattern of influence exerted by the experimental variables from Experiment 1a (sum and answer familiarity) and Experiment 2a (sum type). If the interaction between sum and answer familiarity in Experiments 1a and effects of sum type in Experiment 2a are revealed this would indicate that selecting a strategy in a distinct phase in each trial prior to solving a problem does not influence the problem-solving process. Secondly, by comparing solution latencies across designs, it will be apparent whether in the solution-phase of the dual-phase design the selection process is re-engaged to choose a solution strategy.

2.6.1 Method

2.6.1.1 Participants

Twenty-four undergraduates from the School of Psychology at Cardiff University were given course credit for their participation. All were native English speakers reporting normal hearing and corrected or normal vision and had not participated in any of the other experiments in this series.

2.6.1.2 Materials & Design

The stimuli presented in the current experiment were presented in Experiments 1a (sum familiarity/answer familiarity) and 2a (decades, mixed and fives), see corresponding sections in Experiment 1a and 2a for details of the stimuli construction. The problems from each of the conditions were presented in a pseudo-randomised ordering and therefore not blocked by the experiment of origin. The experiment commenced with 5 practise questions, followed by a total of 112 experimental trials.

2.6.1.3 Procedure

The experiment was programmed and run in Visual Basic 6.0. Participants were advised that they would be presented with a series of sums, one-by-one and that they would be required to solve each problem as quickly and accurately as possible. The experiment was self-paced, each trial commencing when participants pressed the enter key to initiate the lead-in to the trial. The lead-in commenced with a fixation mark ("X + X"), positioned in the centre of the screen, which flashed 3 times, each flash interleaved by 850 ms. On what would have been the fourth appearance of the fixation mark the problem appeared in its place. The problem remained visible on the screen until participants pressed the enter key to confirm their response which was entered using the numerical keypad on a standard qwerty keyboard. Both solution latencies and the given sum solution were automatically recorded by the runtime program.

2.6.2 Results & Discussion

Two measures were taken, the *solution latency* and *sum solution.* Solution latencies were recorded from the point at which the problem was initially presented on the screen until enter was pressed to confirm the answer. In the dual-phase studies, solution latencies were tagged to the preceding strategy selection made in the selection phase. As the selection-phase was not tested in this design solution latencies were analysed without the strategy tag. To facilitate a like-for-like comparison between experimental designs, solution latencies from the dual-phase studies were reanalysed, collapsed across the strategy selection tag thus matching the outputs from the current study. The following analyses will present the findings pertaining to the independent variables from Experiments 1a and 2a separately.

2.6.1.1 Sum Familiarity and Answer Familiarity

Only 7.61% of problems were answered incorrectly (compared to 7.4% in the dual-phase design) and 5.66% of responses were not made within 10 s (4.7% in the dual-phase design). Similar to the findings from Experiment 1a, where solution latencies tagged to calculate selections were influenced by the interaction between sum and answer familiarity, the same interaction was significant here, $F(1, 23) = 12.13$, $MSE = .07$, $p = .002$. Simple effects indicated, similar to Experiment 1a, that effects of answer familiarity were only evident in familiar (as opposed to unfamiliar) problems, $F(1, 23) = 49.4$, $p < .001$, and that for these types of problems, the act of selecting a strategy prior to solving the sum (as in the dual-phase design) does not influence performance when solving the problem. A mixed measures ANOVA with sum and answer familiarity as within subjects variables and *design* (solution- vs. dual-phase) as a between subjects variable revealed a non-significant difference between solution latencies across experimental designs, $F(1, 46) = .4$, $MSE = 3.18$, $p = .53$. If solution latencies were significantly shorter in the dual-phase than the single-phase design this would speak to the notion that participants only had to make a selection once, rather than twice in each trial. However, as there was a non-significant difference between the two designs (see Figure 2.5) this suggests that the transition between the selection- and solution-phases in the dual-phase design is characterised by a re-initiation of the selection-phase of the solution process, contrary to the predictions of the SAC model.

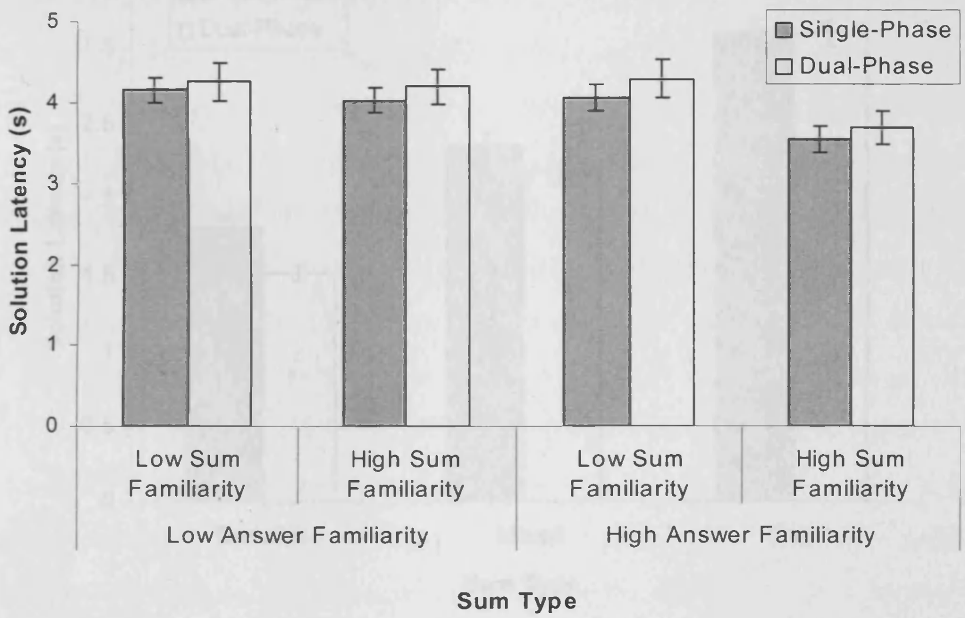

Figure 2.5: Solution latencies in each sum type condition. Grey bars represent data from the solution-phase design employed in the present experiment, white bars the data from the dual-phase design in Experiment 1a. Error bars represent the standard error of the mean.

2.6.2.2 Sum Type

Only 1.62% of sum solutions were incorrectly answered (compare to 4.85% in the dual-phase design) and 7.64% of latencies exceeded 10 s (5.2% in the dual-phase design). Main effects of sum type were evident upon solution latencies in the current study replicating those observed in Experiment 2a, $F(2, 22) = 230.65$, $MSE = .05$, $p < .001$. This confirms that the finding that the act of selecting a strategy before solving a problem does not change the pattern of influence exerted by the experimentally manipulated variable sum type. Similar to the dual-phase design, decades problems were solved more rapidly than mixed or fives problems, and that mixed problems were solved more rapidly than fives problems (all *p*s < .001). When contrasting the mean solution latencies across experimental designs (see Figure 2.6), a mixed measures ANOVA with sum type and design (solution- vs. dual-phase) revealed a non-significant difference in latencies, $F(1, 46) = 3.1$, $MSE = .31$, $p = .08$. Accordingly, as there was a non-significant difference between latencies in the single- and dual-phase designs, similar to the findings from the sum and answer familiarity problem set, it is possible that the selection process is re-initiated at the start of the solution-phase in the dual-phase methodology.

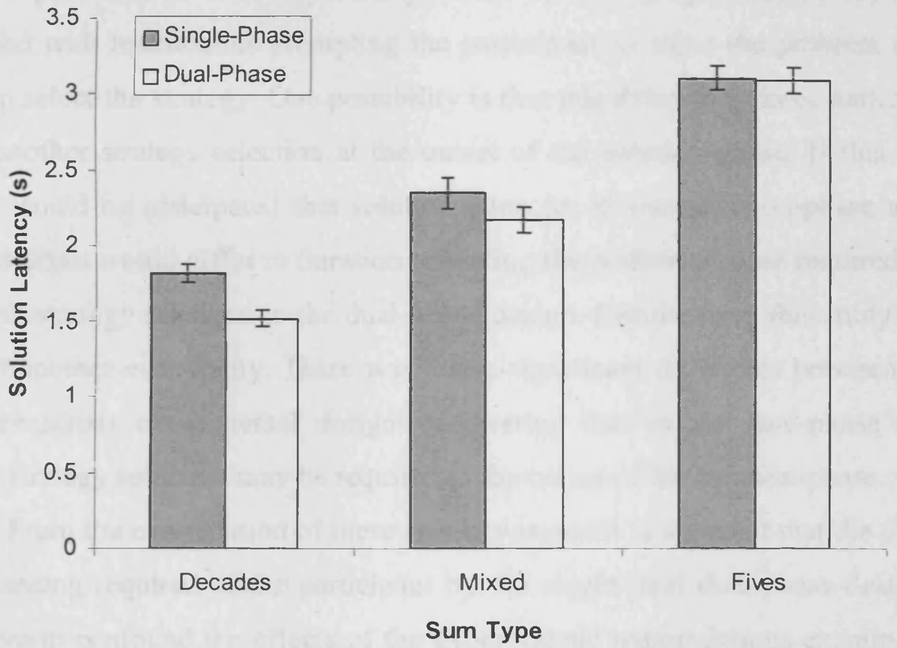

Figure 2.6: Solution latencies in each sum type condition. Grey bars represent data from the solution-phase design employed in the present experiment, white bars the data from the dual-phase design in Experiment 2a. Error bars represent the standard error from the mean.

In summary, the pattern of effects exerted upon solution latencies in the current study replicates those evident in Experiments 1a and 2a. Accordingly, it is apparent that the solution processes employed in the dual-phase design are not distorted by the act of making a predicted strategy selection before solving the problem. This confirms that the same pattern of influence exerted by sum familiarity, answer familiarity and sum type upon solution latencies occurs whether participants are required to make predicted strategy selections in a distinct phase prior to solving the problem or are only required to solve the problem.

The findings from the present experiment also are revealing of the processing undertaken in the transition between the two phases in the dual-phase design. In the dual-phase design it is assumed by Reder and colleagues (Reder & Ritter, 1992; Schunn et al, 1997) that the two phases comprising the problem-solving process (i.e., selection then solution), as identified in previous experiments, are dissociable and follow each other in a set order. Reflecting this demarcation in the dual-phase design the solution-phase immediately follows the selection-phase. Reder and colleagues propose that for this experimental design selections made in the selection-phase determine which strategy is deployed in the solution-phase of the design. However, between phases there is a slight delay in the runtime programme as the screen is refreshed with instructions prompting the participant to solve the problem replacing those to select the strategy. One possibility is that this delay may force participants to make another strategy selection at the outset of the solution-phase. If this were the case it would be anticipated that solution latencies in the selection-phase and dual-phase designs would differ in duration reflecting the additional time required to make a second strategy selection in the dual-phase design. Results from this study however support another eventuality. There was a non-significant difference between solution latencies across experimental designs suggesting that in the dual-phase design a second strategy selection may be required at the outset of the solution-phase.

From the examination of these two key issues it is apparent that the difference in processing required of the participant by the single- and dual-phase designs does not serve to confound the effects of the experimental manipulations examined so far upon solution latencies. In the dual-phase design it may not be the case that retrieve/calculate selections made during the selection-phase determine the strategy selected to actually solve the problem. However, effects of sum type, answer familiarity and problem familiarity upon solution latencies replicate the findings revealed in Experiments 1a and 2a. This confirms that both designs can be used with a large degree of confidence to assess the accuracy, i.e., the relation between predicted strategy selections and solution latencies.

2.8 GENERAL DISCUSSION

The five experiments presented in Chapter 2 have examined the processes by which strategies are selected in mental arithmetic problems. The findings presented in these experiments demonstrate that the dual-phase experimental design is an effective and reliable tool for measuring the performance of the selection mechanism in arithmetic problem solving. Experiments 1a and 2a showed that strategy selections in different types of problems are influenced by mechanisms sensitive to qualitatively distinct problem components (i.e., problem familiarity and specific problem features). Furthermore, that the strategy selection-phase temporally precedes the solution-phase in problem solving. Experiments 1b, 2b and 3 confirm these findings, illustrating that the dual-phase design does not confound the accuracy of responses derived from the selection- and solution-phases of the design, supporting the notion that a distinct strategy selection phase precedes strategy execution. This confirms that problem familiarity and selection-by-feature effects are based upon pre-experimental familiarity and features identified and employed prior to the experiment, presumably in prior real-world problem-solving episodes. Together these experiments support the Adaptive class of selection models, in particular the SAC model (Reder & Ritter, 1992; Schunn et al, 1997). A brief summary of the key findings from the five experiments will follow, with particular emphasis upon how the results contribute to the limited body of experimental research into strategy selection. Specific reference will be made to the type of factors that influence the selection process and the mechanisms required to process these factors.

Results from Experiments 1a, 1b, 2a and 2b provide concrete support for the Adaptive class of strategy selection models (i.e., SAC, CMPL, ASCM, SCADS and SCADS*). Two avenues of investigation lead to this conclusion. I turn first to the rapidity with which strategy selections were made. To recapitulate, on average, selection latencies ranged between 450 ms and 650 ms and were consistently inside the 850 ms window indicating that selections were made well before a solution strategy could be completed. Consequently, it was not possible for participants to solve the problem and use hindsight to identify an appropriate strategy selection, the approach necessarily advocated by the Automaticity models in rapid selection tasks. In the ACT-R (Lebiere & Anderson, 1998) and DOA (Siegler & Shrager, 1984) simulations, following the Obligatory Activation assumption (Logan, 1988) the retrieval strategy is automatically applied upon presentation of a problem. Consequently, accurate predicted strategy selections (as required in selection-phase of the dual-phase design) can only be made when the retrieval production fails to return an answer, at which point it would be apparent to the simulation whether the answer could be retrieved or not. Similarly, in the ITAM model (Logan, 1988; 2002), where retrieval and calculation classes of strategy are run in parallel, racing to produce an answer to the problem, accurate selection in the dual-phase experimental design could only be returned upon completion of one of the strategy pathways.

A further account of rapid strategy selection was also examined in Experiment 1a. Rather than waiting for the retrieval production to complete it may be that an early read of the problem's answer, which reflects the degree of progress made by the retrieve production in finding the correct answer, is used to guide responses in the rapid strategy selection task (see also Reder & Ritter, 1992; Schunn, et al, 1997 for similar argument). This possibility is not specified within any of the Automaticity selection models, but it is a logical extension (see Reder & Ritter, 1992) in tasks where responses are required within a timescale which precludes the full execution of the retrieval strategy. However, the prediction that problems with more familiar answers, and hence answers which should be retrieved more rapidly, should elicit a greater percentage of retrieve selections was not borne out by the data.

The second avenue of investigation centred upon the role played by problem familiarity in determining the strategy selections made in Experiments 1a and 1b. In accordance with the predictions of the SAC model, it was found that familiar problems elicit a greater percentage of retrieve strategy selections (hence fewer calculate selections) than unfamiliar problems. This suggests that the final level of activation (i.e., comprising base level + spreading activation) at the most active problem node in memory elicited by encoding a problem's terms predicted performance in the selection task. It should be noted that based upon the outputs of a regression analyses presented in Appendix B an account of selection in which answer magnitude influences selection cannot be entirely discounted. However, converging evidence presented in Experiment 1a provides supports the problem familiarity account. The Automaticity models which are based upon an obligatory search for a problem's solution (Logan, 1988), were unable to account for the problem familiarity effect, paralleling item familiarity effects evident in a number of other paradigms which require rapid memory searches (Yonelinas, 2002). The CMPL account

(Rickard, 1997; 2004) similarly struggles to account for this finding. Candidate strategies (retrieve and calculate) are activated in memory as a by-product of problem encoding. The competition between the search for the answer by the retrieval strategy and the answer of the first step of the calculate algorithm (as in CMPL calculation algorithms are decomposed into a series of retrieval productions) determines which strategy is selected. Once activation elicited by either strategy in the search for the problem's solution breaches a threshold, the strategy is allowed to execute to completion, inhibiting the action of the losing strategy. Accordingly, as selection is determined by the activation elicited by the search for a problem's answer, rather than the activation elicited by the problem itself, the CMPL is unable to account for the problem familiarity effect.

The findings from Experiment 2a and 2b serve to question the ecological validity of the SAC model, illustrating that problem familiarity does not account for predicted strategy selections on all types of problem. Here, it was apparent that particular problem features influence the strategy selection process. The covariates analysis in Experiment 2a illustrated that a model accounting for both problem features and problem familiarity provided the best fit to the data. That there were no problem familiarity effects within each of the sum types levels indicates that the particular problem feature took precedence over sum familiarity as the key determinant of selection in these problems. The SAC model and other models such as ACT-R have acknowledged that the presence of specific problem features influences selection, however, they are only specified as ad-hoc assumptions in the model. Only the SCADS* model formally specifies a pathway through which the selection-by-feature account could operate, positing that a feature detection mechanism runs in parallel with encoding processes identifying task relevant and irrelevant features (Siegler & Araya, 2005). The mechanism tracks the utility of each feature identified and records performance in the presence and absence of the feature to determine whether subsequent strategy selections should be determined by specific features. However, in essence this mechanism is still beset by the same limitations of the rudimentary ad-hoc assumptions employed by a number of models. The mechanism scans the problem for features that are specified at the outset of the model run and cannot assimilate new features online. Furthermore, as the model is fundamentally based upon the prior success of strategy applications, where success is indexed by the selection of a strategy producing a correct response, the key predictor of selection would be answer familiarity which was shown to have no impact upon the selection process in Experiment 1a.

A further finding, that calculate strategy selections were made more rapidly than retrieve selections, did not fit with any of the existing accounts of strategy selection. This effect was evident in problems which elicited both a high percentage of retrieve (Experiments 2a and 2b) and calculate (Experiments 1 and 1b) selections indicating that it could not be attributed to the type of problem presented, or indeed the mechanism used to determine selections (i.e., problem familiarity or selection-by-feature). One possible account of this effect is that participants' expectations may influence selection latencies. For example, in Experiment 1a most problems elicited a high percentage of calculate selections. Accordingly participants may have become accustomed to choosing calculate, a switch from the 'default' strategy (i.e., calculate) to retrieve may have incurred a time cost during the selection-phase. However, in the fives sum type condition in Experiment 2a, where there was a more equitable selection of the retrieve (M = 56.9%) and calculate (M = 43.4%) strategies, the effect was still present thus ruling out this possibility. Also, the effect was not promoted by experimental design. Calculate selections were made more rapidly than retrieve selections in the two experiments which tested the selection-phase in isolation (Experiments 1b and 2b). Nor was it apparently contingent upon the dominant strategy selection, as calculate selections were made more rapidly when calculate was the dominant strategy (Experiment 1a) and retrieve was the dominant strategy (Experiment 2a). The finding in itself is counter-intuitive and contradicts the positions advocated by the existing selection mechanisms. For example, in the SAC model during the selection-phase the level of activation elicited by the encoding of the problem terms produces a feeling-of-knowing (FoK) which in turn is subject to a single-counter threshold (see Nelson & Narens, 1990 for discussion). If the FoK breaches the criterion, reflecting a relatively high level of activation, the retrieve strategy will be selected, conversely if activation is low, the calculate strategy will be selected. As SAC operates with a single threshold criterion (i.e., the decision whether to apply the retrieve strategy or not), in respect to the temporal dynamics of the selection process, both retrieve and calculate selections should therefore be made simultaneously. However, this is not the case in the experiments reported in the present chapter, nor in the original Game Show studies where an unprincipled relationship was evident between the duration of retrieve and calculate selections (see

Table 1.1 in Chapter 1). Nelson and Narens (1990) also propose that the FoK response may be subjected to separate thresholds, one responsible for information pertaining to items that are known, the other for items that are not known. A detailed consideration of this hypothesis will be made in Chapter 3.

The empirical work in this chapter has served to evaluate the key predictions of the Automaticity and Adaptive class of models, establishing that the SAC model provides the most comprehensive account of how strategies are selected to date. Two key problem-level manipulations (i.e., pre experimentally derived problem familiarity and problem features) were shown to influence selection for the first time in the mental arithmetic literature. Furthermore, the dual-phase methodology was decomposed confirming that there was no bias in either phase attributable to the order of processing it required indicating that the outcomes derived from this design are consistently representative of those used in real-world problem-solving. In the following empirical chapter, sensitivity of the selection mechanism to problem- and task-level manipulations is examined in greater detail. The empirical work serves to extend the key predictions supported by the empirical outcomes detailed in the present chapter whilst building a more coherent overview of the functionality of the selection mechanism in real-world problem-solving episodes.

CHAPTER THREE

3.0 ABSTRACT

Three key issues were examined in Chapter 3 arising from the problem-level effects evident in Chapter 2. When two problem feature manipulations were employed neither feature was found to influence selection (Experiment 4). However, in Experiment 5a, when a problem feature and problem familiarity manipulation was used an interaction between these two cues demonstrated that different cue configurations serve to control the effects of specific cues upon selection. Furthermore, problem familiarity effects upon selection were unaffected when conscious processing during the course of the experiment was precluded and task instructions designed to bias retrieve or calculate selections were employed (Experiment 6a). However, problem feature effects were shown to be reliant on conscious processes in Experiment 5a but not 5b, and selections were biased by task instructions (Experiment 6b). These findings were evaluated in light of the existing models of selection and where necessary candidate mechanisms were identified in light of the limitations of these models.

3.1 INTRODUCTION

The empirical drive of this thesis so far has examined the existing accounts of selection whilst identifying the types of factors that influence retrieve and calculate strategy selections. In Chapter 2 it was shown that the problem-solving process is comprised of two distinct phases. In the first phase strategies are selected upon the basis of problem familiarity or specific problem features (i.e., the selection-by-feature effect). This is followed by a solution-phase, where the chosen strategy is executed in an attempt to solve the problem (see also Reder & Ritter, 1992; Schunn et al., 1997). Using a task in which rapid strategy selections were required, findings from Chapter 2 clearly indicate that existing simulations of the selection process are unable to account in full for the findings detailed. In light of this conclusion the experimental work presented in Chapter 3 was designed to examine three key issues arising from the empirical work in Chapter 2, rather than directly evaluating predictions derived from the Adaptive and Automaticity accounts of selection. Of course, the findings from following experiments can be contrasted to the predictions of these models, but fundamentally the issues explored in the present chapter have received little, if any coverage in the existing models.

The first issue examined in this chapter centres upon the selection-by-feature effect identified in Experiments 2a and 2b. In two experiments reported in this chapter an opportunity was afforded to examine the empirical signature, i.e., the effects of problem features upon strategy selection. The second issue examined centres upon the notion that in real-world problem-solving the selection mechanism must be able to make rapid strategy selections for problems that comprise a number of cues to selection. Accordingly, in Experiments 4 and 5a two different configurations of problem-level factors are manipulated to examine how selections are determined when there are competing cues to strategy selection. The final key issue addressed in this chapter focuses upon how the context in which a problem is solved influences the selection process. Specifically, using two task-level interventions in which the availability of conscious processes (Experiments 5a and 5b) and task instructions (Experiments 6a and 6b) are manipulated it will be evident whether features of the processing context interact with the influence exerted upon selection by problem familiarity and selection-by-feature effects. In the remainder of this section the thrust of these issues will be outlined in greater detail.

In Chapter 2, the familiarity of a problem's terms and specific problem features was shown to influence strategy selection. Whilst problem familiarity effects have been identified in a number of studies (see Reder & Ritter, 1992; Schunn et al, 1997), selection-by-feature effects have not previously been detailed. Two experiments in Chapter 3 provided an opportunity to examine how problem features influence selection in greater detail, as little is known about how these effects actually are realised in rapid strategy selection tasks. Experiments 2a and 2b demonstrated selection-by-feature effects such that the sum type (i.e., decades, mixed and fives problems) manipulation influenced the percentage of retrieve strategy selections returned. However, there has been no empirical evidence presented to date indicating whether problem feature manipulations influence calculate selections as well as retrieve selections or whether they influence selection in unfamiliar problems, as opposed to the familiar problems (i.e., decades, mixed and fives) employed in Experiments 2a and 2b. Furthermore, the covariates analysis conducted in Experiment 2a revealed that problem familiarity covaried significantly with the percentage of retrieve strategy selections returned. From this it may be inferred that strategy selections attributed to the action of the selection-by-feature mechanism in those studies may have been incorrect. Addressing these outstanding questions and with the intent to develop a clearer understanding of the functional characteristics of the selection-by-feature effect, in Experiments 4 and 5a problem features were manipulated in unfamiliar problems designed to elicit a high percentage of calculate selections.

The analytical approach adopted in the Chapter 2 was based upon the predictions derived from existing models of strategy selection. However, all of these models, barring the SCADS* model assert that selection in the dual-phase experimental paradigm is determined by a single factor, such as problem familiarity (i.e., SAC) or the familiarity of the problem's answer (ACT-R, ITAM, DOA, CMPL). In the SCADS* model (Siegler & Araya, 2005), four sources of historical data, termed as *global, featural, problem-specific* and *novelty* data, are evaluated when making a strategy selection. The relative weighting of each of the data sources is entered into a stepwise regression which computes the predicted strength of each candidate strategy. Unfortunately, due to the complexity of this simulation and its reliance upon modelling large data sets it was not possible to obtain testable predictions from this account of cue combination that could be assimilated into the empirical approach employed within this thesis. As there are no other empirical or theoretical accounts of how multiple sources of information (or cues to selection) might combine to influence selection within the arithmetic strategy-selection literature there is no foundation upon which to ground further consideration of this issue. Excepting Experiments 1a and 1b, where effects of problem familiarity and not answer familiarity were evident, in Chapter 2 problem-level manipulations (i.e., problem familiarity or problem feature) were examined in separate experiments where other factors shown to influence selection were controlled. However, this approach is unlikely to mirror processing in real-world scenarios where the number and type of factors that influence selection are not controlled. Accordingly, to tap this facet of the selection mechanism in full it is necessary to understand how multiple cues, of different types, serve to elicit retrieve and calculate strategy selections. For example, when making a selection, do particular types of problem-level manipulation take precedence over others, or does the influence exerted by both types of manipulation interact? Experiment 4 examines the influence exerted upon selection by two problem features in a single experiment. In extension to this rationale, Experiment 5a examines whether problem feature or problem familiarity manipulations take precedence, or whether their influence upon selection is derived from an interaction between the two factors

When considering how the context in which a problem is solved influences strategy selection, within the problem-solving literature it is proposed that there are a range of candidate factors, aside from problem-level manipulations, that may influence strategy selection. For example, in a complex mental arithmetic problem-solving task Cary and Carlson (1999) demonstrated that the availability of working memory aids dictated which strategy was actually used to solve a problem. If an aid (i.e., pen and paper) was available, the strategy participants selected to solve the problem took advantage of the aid. Alternatively, if there was no aid available, participants used a problem-solving strategy that minimised the demands placed upon working memory resources. From this example, it is apparent that the influence exerted by the context (i.e., the availability of an aid) in which the problem is solved is inseparable from way the problem is actually solved and therefore the strategy chosen to solve the problem. In Chapter 2 of this thesis, the potential influence exerted by the context in which a problem was solved can be appreciated by examining an overview of the experimental designs employed and the findings revealed. To illustrate, when considering the findings revealed in Chapter 2, the results from experiments employing the dual-phase design (Experiments 1a and 2a) were successfully replicated in experiments testing the selection-phase (Experiments 1b and 2b) and solution-phase (Experiment 3) in isolation. From this analysis it was demonstrated that effects of the processing context, i.e., *whether the problem is solved after making a strategy selection or not*, also if *participants are only required to solve the problem rather than making a predicted strategy selection before solving the problem* did not impact problem familiarity or selection-by-feature effects. Accordingly, it was from this overview of the experiments that a task-level (or the context in which strategies are selected, or problems solved) did not interact with the influence exerted by problem familiarity of problem features upon selection

Three experiments in this chapter sought to identify two further components of the problem-solving episode that may influence selection. Experiments 5a and 5b investigated whether conscious processes influence problem familiarity and selection-by-feature effects. In the Chapter 2 effects of both problem-level manipulations were attributed to the action of the selection mechanism. However, a hitherto unexplored possibility is these effects are facilitated by other cognitive mechanisms, in particular the selection-by-feature effect. Drawing upon findings from related paradigms where memorial searches determine performance, such as the Feeling of Knowing literature (Botvinick et al, 2001; Koriat, 1993; Koriat, Ma'ayan & Nussinson, 2006; Nelson & Narens, 1990; Schwartz & Metcalfe, 1992) and recognition memory literature (e.g., Hirschman & Henzler, 1998; Yonelinas, 2001; 2002) it was hypothesised that a conscious appraisal of a problem's terms (Cary & Reder, 2002) may be responsible for the selection-by-feature effect. In Experiments 5a and 5b, a secondary task (articulatory suppression) was employed in order to impair a conscious appraisal of the specific factors inherent in a problem and of the familiarity of the problem during the course of the selection task. By inhibiting this process it will become evident whether strategy selections influenced by problem familiarity or problem features stem from the selection mechanism itself, or a separate process contingent upon conscious processing during the task. In Experiments 6a and 6b a further type of task-level manipulation was employed. Specifically, the type of instructions given to participants was manipulated to examine whether biases in instructions are reflected in selections. Previous mental arithmetic problem-solving studies have shown that task instructions emphasising usage of the retrieve strategy increased the number of self-reported retrieve selections made by participants after they solved problems (Kirk & Ashcraft, 2001, see also Blöte, Van der Burg & Klein, 2001). Similarly, task instructions emphasising usage of the calculate strategy served to elicited calculate selections more often.

The findings from these studies, where possible, are contrasted against predictions derived from the existing models of strategy selection. Due to the limitations inherent in these models in respect to their specificity and ability to account for the findings presented in the thesis to date, the findings from these studies serve to detail the strategy selection mechanism to a greater resolution.

3.2 EXPERIMENT 4

In Chapter 2 evidence has been presented to suggest that strategy selection is influenced by two key factors, problem familiarity and specific problem features. In Experiments 1a and 1b, selection was influenced by a rapid assessment of the familiarity of a problem's terms. The stimuli used in these experiments elicited a high percentage of calculate selections and there were no problem feature manipulations employed which may have influenced selection. In contrast, retrieve and calculate selections in Experiments 2a and 2b selection were sensitive to the selection-by-feature effect such that obvious problem features (i.e., decade, mixed and fives) influenced selection. The problems presented in that experiment were relatively high in familiarity in contrast to those used in Experiments 1a and 1b and elicited a high percentage of retrieve selections. Bridging the manipulations employed in these experiments this study was designed to investigate two hypotheses. Firstly, to identify whether in unfamiliar sums, where problem familiarity is controlled and problem features are evident, if the selection-by-feature effect influences calculate selections. As prior experiments have only examined the selection-by-feature effect in decades, mixed and fives problems the current study will reveal whether the effect is limited to highly familiar problems (i.e., decades, mixed and fives) or whether its influence extends to all problems upon the familiarity dimension. Furthermore, selection-by-feature effects have only been observed in problems which elicit a high percentage of retrieve selections. As the problems presented in this study do not contain any numbers divisible by 5 or 10 into integers it is predicted that the calculate strategy will be selected most often. Problem feature effects would indicate that the selection-by-feature effect can also be generalised to calculate selections.

Secondly, an opportunity is afforded to identify further types of problem feature that influence selection. To date, within the strategy selection literature only two problems features have been shown to influence selection: *operator type (*Reder & Ritter, 1992*),* where problems with different operators are presented within a task. Also the *sum type* manipulation used in Experiments 2a and 2b in this thesis (i.e., decades, mixed and fives problems). In the current study participants were presented with unfamiliar double-digit addition problems, similar to those used in Experiments 1a and 1b, for which it was predicted that a high percentage of calculate selections would be returned. Two specific problem features were manipulated, accordingly, any variation in the percentage of calculate selections made across conditions would denote the influence of the particular problem features upon strategy selection. Problems were constructed from two classes of problem feature, the odd or even status of the addends and the relative magnitude of the addends. In non-decades and non-fives sums such as 21 + 23 or 22 + 24 each addend in a problem may be categorised by its odd or even status. Evidence from arithmetic solution verification tasks (e.g., 26 + 22 = 59, true or false?) which employed problems which did not contain decade or fives addends suggests that adults (Lemaire & Fayol, 1995; Lemaire & Reder, 1999) and children (Lemaire & Siegler, 1995) are sensitive to the odd or even status of the addends when required to verify the accuracy of a problems solution. Accordingly, for the manipulation *addend status* three types of problems were constructed containing addends which were both even numbers, both odd numbers and problems with one odd and one even addend. On a structural level, in this manipulation the three levels were designed to mirror the manipulation of decades, fives and mixed problems presented in Experiments 2a and 2b. It was anticipated that problems comprising two even addends would be solved more rapidly than problems comprising one or two odd addends. This would be indicative of a violation of the problem size effect (Groen & Parkman, 1972) which stipulates that solution latencies increase in a linear trend with the magnitude of the addends. Such a finding would indicate that the selection mechanism utilises the odd or even status of the addends as a cue to strategy selection.

The second factor manipulated here, the *relative magnitude* of the addends, was operationalised for two purposes. Firstly, to identify whether the mechanism responsible for the selection-by-feature effect is also sensitive to this problem feature. If so, it would be predicted that problems with addends of a similar magnitude would elicit a greater percentage of retrieve selections than problems with disparate addends. Secondly, whereas, other manipulations of problem features (i.e., sum type, and addend status) focus upon the correspondence between the problem features inherent in both addends (e.g., *both addends are decades numbers*) this feature would be identified by a more complex assessment of the relationship between the two numbers. It may be that one addend is processed in the context of another, where the smaller number is compared to the larger addend which acts as a starting point, or anchor, for the assessment. Evidence from numerical processing paradigms suggests that if this were the case, a positive correlation between processing time and relative addend magnitude would be expected (Brysbaert, 1995; Brysbaert, 2005; Fayol & Seron, 2005). Faster strategy selections would be predicted for problems with a similar, rather than disparate magnitude, as less time is required to count from the larger number to the smaller number. Alternatively, a second possibility is that the magnitude of one addend is compared directly to another, with neither addend acting as an anchor. In number comparison studies (e.g., Moyer & Landauer, 1967; Ratinck & Brysbaert, 2002), a process commonly associated with arithmetic problem solving (Dehaene, 1992), participants are able to identify 2 as the smaller number in the pair 2-8 more rapidly than in 2-3. Accordingly, if problems comprising addends of a disparate magnitude elicit faster strategy selections than those with a similar magnitude, this would indicate that the addends are being compared with each other in the strategy selection process.

Participants engaged in the same dual-phase methodology as used in Experiments 1a and 2a. Six experimental conditions were constructed based upon a factorial manipulation of two types of problem feature, *addend status* (*odd*, *even*, and *mixed* odd and even addend pairs) and *relative addend magnitude* (*similar* or *disparate*). Effects of addend status upon the percentage of calculate selections made would indicate that problem features influence calculate strategy selections (rather than only retrieve selections) and also problems relatively low in familiarity. Furthermore, effects of the relative addend magnitude manipulation upon time taken to select the calculate strategy will provide insight into whether number comparison or anchoring processes are used by the selection process to analyse problem features.

3.2.1 Method

3.2.1.1 Participants

Twenty-four undergraduates from the School of Psychology at Cardiff University were given course credit for their participation. All were native English speakers reporting normal hearing and corrected or normal vision and had not participated in any of the other experiments in this series.

3.2.1.2 Materials & Design

The addends in each problem summed to less than 100 and were drawn from a sample of numbers ranging from 12 to 87. All of the problems in the stimulus set were double-digit addition problems and presented in an identical fashion to Experiments 1a and 2a using the dual-phase design. Two variables were investigated; addend status (even, odd or mixed) and the relative magnitude of the addends. Even problems comprised two even number addends (e.g., 16 + 18), odd problems comprised two odd number addends (e.g., 17 + 19). Mixed problems comprised one odd and one even number addend (e.g., 16 + 19), the ordering of which was randomised in the problems presented within the stimuli set. The difference in magnitude between the first and second addends in similar relative magnitude problems ranged from 1- 7 (e.g., 23 + 24 or 31 + 38). Addends in the wide magnitude conditions differed by a minimum of 32 and a maximum of 73. The six experimental conditions were contrasted in a repeated measures design. Participants completed 88 trials in total, comprising 12 problems from each experimental condition which were preceded by 16 practice trials (see Appendix E, table E3, for stimuli).

3.2.1.3 Procedure

See Experiment 1a for detailed exposition of the procedure employed in the current study, also for details of the instructions given to participants at the outset of the experiment.

3.2.2 Results

3.2.2.1 Scoring Procedure

As in Experiments 1a and 2a four measures were recorded during the experiment; *strategy selection*, *strategy selection latency, problem solution* and

CARDIFF UNIVERSITY PRIFYSGOL CAERDYDD

solution latency. Full details of how and when the measures were recorded can be found in the corresponding section in Experiment 1a.

3.2.2.2 Strategy Selection

A repeated measures ANOVA indicated that there was no significant effect of either manipulation; addend status, $F(2, 22) = 1.87$, $MSE = 1.22$, $p = .18$ or relative addend magnitude, $F(1, 23) = 1.30$, $MSE = 1.73$, $p = .27$, upon the percentage of late responses. This confirms that the experimental manipulations did not influence participants' propensity to produce late responses.

To identify the influence exerted by the two manipulations of problem feature a 2 (relative magnitude; similar, disparate) x 3 (addend status; odd, even and mixed) repeated measures ANOVA was conducted upon the percentage of calculate selections made by participants. This reveals, as Figure 4.1 illustrates, that selection of this strategy was insensitive to both the relative magnitude of the addends; $F(1, 20) < .001$, $MSE = 23.15$, $p = 1$, and addend status, $F(2, 19) = .09$, $MSE = 51.84$, $p = .92$, demonstrating that neither problem feature influenced the selection process. As neither problem feature manipulation was shown to influence calculate selections an item-level covariates analysis was run to identify the item-level factors that may have influenced selection. The same procedure and four models tested in Experiment 1a were applied to this data set (see also Appendix A). The magnitude model (magnitude of the first and second addends as covariates; $df = 7$, $AIC = 364.02$) and familiarity model (familiarity of first and second addends as covariates; $df = 7$, $AIC = 362.4$) provided comparable fits to the data. As did a model containing both sum familiarity and answer familiarity as covariates ($df = 7$, $AIC = 365.85$), each model sharing the same number of parameters. A model containing only problem familiarity as a covariate ($df = 6$, $AIC = 359.66$) provided a significantly better fit to the data than the other three models ($ps = .01$), suggesting that problem familiarity contributed to strategy selections in the present experiment in the absence of problem feature effects sought here.

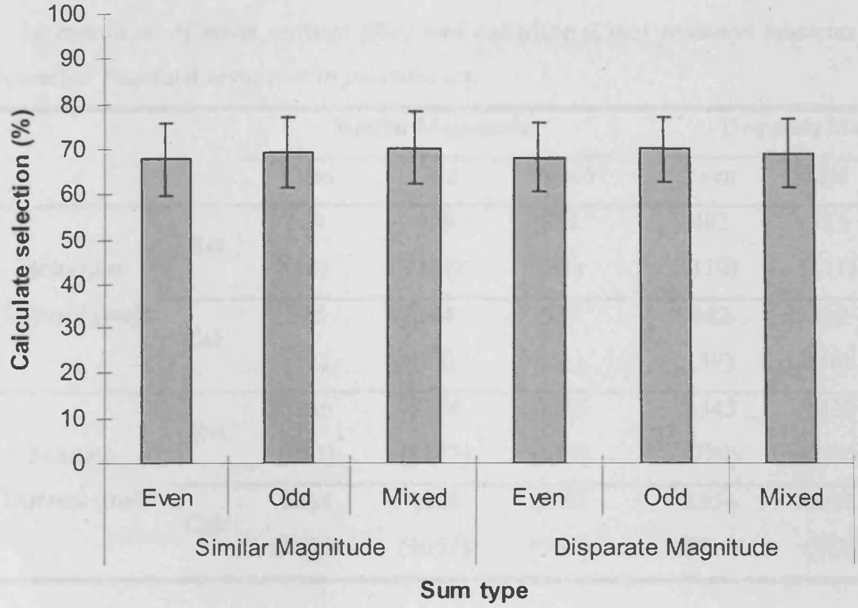

Figure 3.1: The percentage of calculate strategy selections returned in each condition. Error bars represent the standard error of the mean.

3.2.2.3 Selection Latency Analysis

In total, 11.3% of the strategy selections were not returned within the 850 ms timeframe allowed for responses in this phase of the experiment. A 2 (relative magnitude; similar, disparate) x 3 (number status; odd, even, mixed) repeated measures ANOVA run on the mean calculate selection latencies revealed null effects of addend status, $F(2, 18) = 0.6$, $MSE = .004$, $p = .94$, and relative magnitude, $F(1, 19) = 0.33$, $MSE = .003$, $p = .57$. Similarly, no effects of addend status or relative magnitude were evident upon the relatively small percentage of retrieve selection latencies; $F(2, 7) = 0.22$, $MSE = .007$, $p = .81$, and $F(1, 8) = 0.29$, $MSE = .006$, $p = .87$.

Table 3.1.

Summary, by condition, of mean retrieve (Ret) and calculate (Calc) selection latencies (in ms) and solution latencies. Standard deviations in parentheses.

		Similar Magnitude			Disparate Magnitude		
		Even	Odd	Mixed	Even	Odd	Mixed
Selection Latency (ms)	Ret	539 (78)	499 (111)	521 (63)	492 (110)	533 (111)	507 (86)
	Calc	525 (92)	544 (47)	537 (51)	542 (59)	532 (56)	551 (67)
Solution Latency (ms)	Ret	3236 (633)	3708 (1177)	3565 (483)	2543 (789)	3429 (751)	3063 (513)
	Calc	4664 (1536)	4508 (1051)	3752 (589)	3556 (410)	3901 (571)	3921 (575)

3.2.2.4 Solution Latency Analysis

Only 6.28% of solution latencies were excluded for exceeding 10 s and 4.1% of problems were solved incorrectly. As Table 4.1 illustrates, solution latencies tagged to retrieve selections made in the selection-phase were significantly shorter than those tagged to calculate selections $F(1, 27) = 30.28$, $MSE = .49$, $p < .001$. However, the mean solution latencies tagged to retrieve selections exceeded 2.4 s, suggesting that calculation procedures were used to solve these problems. A 2 x 3 repeated measures ANOVA revealed a significant interaction between relative magnitude and addend status for solution latencies tagged to calculate strategy selections, $F(2, 18) = 5.66$, $MSE = .49$, $p = .01$. Post-hoc comparisons demonstrate that there was a simple effect of addend status (i.e., even, odd, mixed) upon solution latencies in problems comprising addends with a relatively similar magnitude, $F(2, 18) = 4.34$, $p = .03$, and also problems with addends of a disparate magnitude, $F(2, 18) = 7.62$, $p = .004$. The interaction also revealed that problems comprising addends of a similar relative magnitude which had mixed addends were solved more rapidly than problems with odd addends ($p = .01$). Furthermore, that for problems with a disparate relative addend magnitude even problems were solved more rapidly than both mixed ($p = .02$) or odd problems ($p = .01$). From this it is tentatively suggested that problems with either one or two even addends in are solved faster than problems with two odd addends, irrespective of the relative magnitude of the addends in a problem. More tellingly however, simple effects of relative addend magnitude were only found in odd, $F(1, 19) = 8.14$, $p = .01$, and even addend problems, $F(1, 19) = 5.93$, $p = .03$, such that problems with disparate addends magnitude were solved more rapidly than problems with a similar relative magnitude.

3.2.3 Discussion

Experiments 2a and 2b demonstrated that particular problem features (i.e., the presence of decades and fives addends) influenced the strategy selection process via a selection-by-feature sensitive mechanism. In the current study the selection mechanism was found to be insensitive to the odd or even status of the addends and the relative magnitude of the problems, indicating that the selection-by-feature effect did not influence selection. The null effect of relative addend magnitude upon selection latencies indicates that the selection mechanism does not set one addend as an anchor against which the other addend is compared. Furthermore, these null effects demonstrate that selection is not contingent upon a comparison between the magnitudes of both addends in a problem. These findings fit with the covariates analysis which revealed that problem familiarity, rather than the magnitude of the first and second addends covaries significantly with selection of the calculate strategy. This suggests that both addends in a problem are viewed as a whole, rather than as two separate addends and that selection of the calculate strategy in the current study was influenced by problem familiarity rather than specific problem features.

When considering the null effects of the two problem features manipulated it may be that participants simply did not appreciate the specific problem features. In the experiments reported in Chapter 2, effects of sum type (decades, fives and mixed) and problem familiarity were evident not only in strategy selections, but also in solution latencies. This suggests that a linkage between the factors that influence the selection process and solution latencies exists and, as highlighted in Chapter 1, stands as a key indicator of accurate performance. It is proposed that the linkage between factors that influence strategy selection and solution latencies are likely to stem from two sources. Firstly, it may be that a conscious evaluation or appreciation of the particular problem features inherent in a problem (or stimulus set) maybe required. Within a problem there are many features that could potentially influence selection which could be appreciated within the 850 ms time frame imposed upon strategy selections in this design. For example, the colour or size of the addends (Siegler & Araya, 2005) or the operator type (Reder & Ritter, 1992). Considering that sum type effects in Experiment 2a and problem familiarity effects in Experiment 1a were reflected in both the percentage of strategy selections and solution latencies it seems that selection-by-feature effects may only occur if the individual appreciates that problems with a particular feature are solved more rapidly than problems without that feature. Hence, a key signature of the selection-by-feature effect may be effects of a particular problem feature upon both strategy selection and solution latencies. However, the interaction between addend status and relative addend magnitude revealed in solution latencies in the present study was not mirrored in the percentage of calculate strategies reported in this study questioning such a notion.

A distinct possibility is that the proposed correspondence between effects of the experimental manipulations upon both strategy selection and solution latencies was not evident here due to the complexity of the experimental design. In the current study the factorial manipulation of addend status (comprising 3 levels) and relative addend magnitude (comprising 2 levels) may have served to reduce the salience of these problem features. To illustrate, the factorial manipulation necessitated that there were 6 different problem conditions within the stimulus set and accordingly the problem features may have not been apparent to participants. To examine this eventuality and the results from the covariates analysis run in Experiment 4, in Experiment 5a only one problem feature (addend status) was employed in tandem with a manipulation of problem familiarity. Furthermore, Experiment 5a examined whether a conscious appreciation of the problem features inherent in a set of problems is necessary for the selection-by-feature effect to influence selection. To do so a secondary task, articulatory suppression, was used to preclude conscious processing within the task.

3.3 EXPERIMENT 5a

This study was designed to address two key issues that arose in Experiment 4. Firstly, to examine why the problem features manipulated in that study failed to influence strategy selections, and secondly, to investigate whether consciously directed processes were responsible for the identification of the problem features employed responsible for selection-by-feature effects. In respect to the first aim, it was shown in Chapter 2 that effects of the experimental manipulations in the selection-phase were largely mirrored by effects of the same manipulation upon solution latencies. For example, in Experiment 1a, a higher percentage of retrieve selections were made for relatively familiar problems and those problems were solved more rapidly than unfamiliar problems. Similarly, in Experiment 2a, a greater percentage of retrieve selections was made for decades than fives problems and solution latencies were significantly shorter in the decades condition than fives condition accordingly. However, in Experiment 4, there were null effects of the two types of problem features manipulated upon both the selection- and solution-phases of the experiment.

Three possible explanations for the null effects evident in the previous experiment are considered. Turning first to the most straightforward explanation, it may be that the participants were simply insensitive to the relative magnitude of the addends paired within a problem. However, it was anticipated that the addend status manipulation would influence selection as this manipulation is analogous to the sum type (decades, mixed and fives) manipulation used in Experiment 2a and 2b. To illustrate, in both manipulations, the addend pairs in a problem could be considered in respect to a categorisation of the types of addends comprising the pair. For example, *both addends are decades numbers,* or *both addends are even numbers.* Alternatively, the null effects evident in Experiment 4 may be attributed to the complexity of the 2 (relative magnitude) x 3 (addend status) design employed in the previous study. In Experiments 2a and 2b, the only problem feature manipulated (i.e., sum type; decades, mixed and fives problems) may have been more easily apparent to participants. In contrast, in Experiment 4, each problem was comprised of one of the three addend status levels and one of the two relative addend magnitude levels. The presence of these features may not have been apparent to participants, as essentially there were 6 different types of problems constructed from the factorial manipulation of the two variables within the stimulus set. To mitigate this eventuality, a more simplistic factorial design was employed to identify whether the selection-by-feature effect influences strategy selection in unfamiliar problems (i.e., non-decade/fives problems) as well as in familiar problems (i.e., decade/fives problems). Accordingly, the addend status manipulation employed was comprised of only two levels (even and mixed problems) and the relative addend magnitude manipulation was jettisoned from the design. It was predicted that selection-by-feature effects would be revealed in the percentage of calculate selections returned such that a greater percentage of calculate selections — and hence fewer retrieve selections — will be made for mixed rather than even problems.

A third consideration arising from Experiment 4 was that the familiarity of a problem's terms may serve to moderate the selection-by-feature effect. It may be that problem features are only used to inform selection when it is not possible to distinguish between strategy selections upon the basis of problem familiarity alone, especially, when problems are high in familiarity. In the previous experiment, all of the problems were relatively low in familiarity in comparison to those presented in Experiments 2a and 2b (i.e., decades, mixed and fives problems). To test this eventuality, the problem familiarity manipulation was re-introduced in order to examine whether the familiarity of a problem moderates the impact of the selection-by-feature effect. If only problems which are relatively high in familiarity (e.g., decades, mixed and fives problems), as opposed to relatively low in familiarity (i.e., those presented in the current study) exhibit selection-by-feature effects then it would be predicted that addend status effects on calculate selections would not be found in this study. However, if selection-by-feature effects influence strategy selections in problems at both ends of the familiarity dimension that addend status effects would be present in the present study.

An alternative interpretation of this hypothesis is that the influence of the selection-by-feature effect may be limited to retrieve strategy selections, rather than just high familiarity problems. If problem feature effects do not influence calculate selections this would be suggestive of the notion that separate selection mechanisms determine retrieve and calculate strategy selections. This notion is contrary to the specifications of the SAC model which proposes a single-counter selection mechanism based upon the FoK response elicited by the relative activation at the problem nodes (Schunn et al, 1997). This threshold account determines whether the retrieve strategy should be selected or not. Accordingly, if selection-by-feature effects are limited to retrieve strategy selections this would indicate that the mechanism which governs retrieve and calculate selections is sensitive to different types of information. As the problems employed in the present experiment were all unfamiliar (in comparison to decades, mixed and fives problems), with a comparable level of familiarity to the problems presented in Experiments 1a and 1b, it was predicted that the calculate strategy would be selected most often. Significant effects of the problem feature addend status upon the percentage of calculate strategy selections reported would support the notion that a single-counter selection mechanism determines both retrieve and calculate selections as per the predictions of the SAC account.

In respect to the second aim of the current study, to identify the cognitive mechanisms responsible for strategy selection, a secondary task, articulatory suppression, was included in the experimental design as a between subjects manipulation. Whereas previous experiments presented in this chapter focus on the problem-level influences (i.e., problem familiarity and problem features) upon strategy selection, termed *intrinsic* factors by Cary and Reder (2002), a secondary task was implemented in this study to examine *extrinsic* (i.e., task-level) features. Types of task-level features that may influence selection include the broader context in which each trial is processed, task instructions, and the prior history of success with a strategy (Cary & Reder, 2002). In their review of a range of tasks proposed to rely upon metacognitive monitoring and control processes Cary and Reder (2002) examined the boundary between conscious and metacognitive contributions (i.e., FoK) to selection. Hitherto, this issue has received limited empirical investigation in the literature. Examining the process through which new strategies are discovered, Siegler and Stern (1998) examined the role played by insight, distinguishing between conscious and unconscious pathways to strategy discovery in their investigation of arithmetic problem solving in children. It should be noted that although the strategy discovery process differs significantly from the feature detection crucial to the selection-by-feature effect, there may be commonalities between the two. One useful approach to insight in problem-solving derived from Siegler and Stern's (1998) research suggests that the discovery of new problem solving techniques implies a "prior conscious insight" (p. 378, Siegler & Stern, 1998, see also Gick & Lockhart, 1995). This is indicative of a potential role for conscious processing in selection tasks, however, in their review chapter, Cary and Reder (2002) conclude that:

> People's actions are influenced by features of the task, features of the environment, and their success at using specific strategies. Remarkably, this adaptation often occurs without any awareness of what procedures or strategies are being used, the base rates of types of stimuli in the environment, or the success of a given strategy...If people in the field maintain the position that metacognition requires conscious awareness, then we would argue that cognitive monitoring and strategy selection often occur without metacognitive intervention. (p.76)

From these two accounts it is unclear whether conscious processes are necessary for the selection of strategies in arithmetic problems. Specifically, whether problem features are detected through conscious insight, or whether (by implication) the effect of problem features upon selection is realised through consciously procedures rather than the implicit mechanisms through which problem familiarity effects are purported to rely upon (Reder & Ritter, 1992; Schunn at al, 1997). Reder and Cary's assertion is based upon a range of tasks with different characteristics. Their examination of problem-level factors is predominantly founded upon the outcome of the Game Show studies (Reder, 1987; Reder & Ritter, 1992; Schunn et al, 1997). In this task the time limit in the strategy selection phase (i.e., 850 ms) serves to favour the contribution of more automatic and unconscious processes. Insufficient time is afforded for conscious processes to direct behaviour and return responses within the time available. Conversely, their identification of task-level factors was derived from a review of more complex tasks than the arithmetic strategy selection task employed in this thesis and in the Game Show studies. These experiments included a building sticks task (Lovett & Anderson, 1996) and an experiment in which participants were required to judge the plausibility of statements based upon the content of a story that had just been read (Reder, 1987). Also, an air traffic control task (Reder & Schunn, 1999) where positive correlations between individual differences in strategy adaptivity, working memory capacity and inductive reasoning abilities were evident, points to the contribution of conscious processing to performance. However, a shortfall in their analysis is evident in the structure of the dissociation they have drawn. Problem-level influences are derived from methodologies where speeded responses (i.e., the Game Show studies) are required; while candidate task-level factors were derived from non-speeded responses in which conscious processing was afforded. To address this limitation, delineating between consciously directed processes and metacognitive processes, such as the FoK mechanism, in the present experiment both factors are investigated under speeded response conditions. Here the time limit imposed upon the selection-phase prevented participants from solving the problem then employing recollective processes to identify the strategy they used to solve the problem, or indeed the strategy they should have used to solve the problem. Accordingly, strategy selections made during this time can be attributed to metacognitive (e.g., the SAC model) or automatic procedures (e.g., ACT-R and ITAM). In this study the dual-phase design was employed and the secondary task, articulatory suppression, was used to examine the contribution of conscious processing to strategy selection.

Under articulatory suppression the participant is required to continuously repeat a vocalisation (e.g., '*xyz*') whilst simultaneously performing the primary task. In a number of other experimental paradigms including serial recall (Jones, Hughes & Macken, 2006; Macken & Jones, 1995), recognition memory tasks (see Yonelinas, 2002 for discussion) and mental arithmetic problem solving (Lee & Kang, 2002; Logie, Gilhooly & Wynn, 1996; Seitz & Schumann-Hengsteler, 2000, 2002) suppression has been shown to partial out the contribution of consciously directed operations or the "inner voice" (see p. 436, Macken & Jones, 1995) to processing on a primary task. It has been suggested in previous experiments in this thesis that the rapidity with which strategy selections can be made (i.e., within 850 ms) indicates that the selection task is completed by metacognitive or automatic processes. Therefore, it was not expected that participants' ability to complete the selection task will be impaired by the suppression task. However, it will prevent participants from any deeper cognitive processing of the stimuli during the selection task, especially during the short intervals between problems (i.e., during the lead in before problem presentations). In particular, discovering new problem features through conscious insight (e.g., Siegler & Stern, 1998).

In related paradigms and following the predictions of the SCADS* model (Siegler & Araya, 2005), for a particular problem feature to influence selection it must first be identified by the individual. The suppression task was employed to prevent participants from identifying the problem feature during the experiment. Accordingly, if problem features in this selection task are identified by a conscious insight or evaluation of the stimuli received during a task, the strategy selections made by participants in the suppression condition will be insensitive to the problem feature addend status. However, if the selection-by-feature effect is reflected in the percentage of calculate selections returned under suppression, this will speak to the notion that conscious processes are not required for the identification of new and useful problem features. Based upon this hypothesis it was decided that participants in both suppression conditions (suppression vs. no suppression) should not have had any prior experience with the dual-phase selection task. Furthermore, it was ensured that the practise questions presented at the outset of the experiment did not resemble those presented in the test phase itself. This approach ensured that participants were not primed with an appreciation that addend status may be manipulated during the task.

Participants engaged in the dual-phase design similar to that employed in Experiment 4 of this chapter but with one notable change. In the classic dual-phase design each trial comprised of a selection-phase immediately followed by a solution-phase. However, the introduction of the suppression task necessitated a revision to this design. The selection- and solution-phases were blocked such that participants made predicted strategy selections for all of the problems in the stimulus set, maintaining continuous suppression throughout the course of each block. Then, once all of the selection-phase blocks were complete, the same set of problems were represented (but in a pseudo-randomised order) and participants were required to solve the problems (i.e., the solution-phase). As in the selection-phase, in the solution-phase problems were blocked. This arrangement meant that participants maintained suppression both whilst making selections and between trials in the selection-phase ensuring that conscious processes could not intrude at any point in the trial blocks.

Two variables, addend status (even and mixed problems) and problem familiarity (low and high) were manipulated as within subjects manipulations, each comprising two levels. Articulatory suppression was manipulated as a between subjects variable. It was predicted that effects of addend status in the selection-phase in the absence of suppression would reveal that the selection-by-feature effect is not limited to highly familiar problems (e.g., decades and fives) but also influences selection in problems which are unfamiliar. This would be indicative of the ubiquity of the effect and that null effects of problem features could attributed to the complexity of the feature manipulation employed in Experiment 4. Furthermore, this finding would demonstrate that the selection-by-feature effect is not limited solely to retrieve strategy selections but also calculate selections. Effects of suppression upon performance in the selection-phase would be revealing of the cognitive mechanisms that contribute to the selection-by-feature effect. It was predicted that there may be an interaction between the suppression condition and the two manipulations, problem familiarity and addend status. Specifically, suppression may attenuate effects of addend status, reflecting the contribution of the conscious evaluative processes required to identify a problem feature before it influences selection.

3.3.1 Method

3.3.1.1 Participants

Twenty-four participants from the School of Psychology at Cardiff University were given course credit or payments in return for their participation. All were native English speakers reporting normal hearing and correct or normal vision.

3.3.1.2 Materials & Design

All the problems presented were double-digit addition sums comprising two addends drawn from a sample ranging from 12 to 49. None of the addends were divisible into integers by 5 or 10, there were no tied addends (i.e., 23 + 23) and each addend pair was from the same decade class (e.g., 23 +29). Please refer to Appendix E, table E4, for a list of the stimuli employed in this experiment.

Two manipulations of stimulus type were employed, a sum familiarity manipulation and a problem feature manipulation. As in Experiments 1a and 1b twenty-four problems were presented in each sum familiarity condition (low vs. high). A one-way ANOVA was used to ensure that there was a significant difference in mean sum familiarity ratings between the low and high conditions, $F(1, 46) = 12.852$, $p < .001$. In respect to the problem feature manipulation, in each level of sum familiarity there was an equal number of problems with even addends and one even (i.e., *even*) and one odd addend (i.e., *mixed*). To ensure that any reported effects of addend status were not masquerading as an effect of an underlying factor (i.e., answer familiarity, the familiarity of the first or second addend) separate repeated measures ANOVAs were run to confirm that within both levels of sum familiarity non-significant differences between even and mixed status conditions were evident (all Fs < 3.65, all ps $> .07$).

For the articulatory suppression task participants were required to continuously vocalise a repeated loop of three consonants *x y z,* pronounced '*zee*' rather than '*zed*'. Rather than vocalising the consonants aloud, participants were asked to whisper the consonants. This way it is easier for participants to maintain a consistent and rapid speed of the repetition (Macken & Jones, 1995; Jones, Hughes & Macken, 2006). For counterbalancing purposes participants were randomly allocated to one of four suppression conditions in which they were required to do suppression during both the selection- and solution-phase blocks, during only one of the phase blocks, or not at all. Accordingly, in each phase 12 participants were required to suppress and 12 participants did not suppress.

3.3.1.3 Procedure

A modified version of the classic dual-phase design employed in Experiments 1a and 2a was used in the present experiment. Whereas in the classic design, on each trial participants are presented with a problem, made a predicted strategy selection, then solved the problem, here the two distinct phases were completed separately. In the first half of the experiment, the selection-phase, participants made strategy selections for each of the problems in the stimulus set which were presented one after the other. In the second half, the solution-phase, participants were presented the same problems again and were required to solve them. Results from experiments in Chapter 2 indicated that the same effects were evident in the selection-phase and dual-phase designs. Accordingly, testing phase blocks independently (i.e., the selection-phase or solution-phase) should not compromise the influence exerted by the experimental variables, nor contaminate a comparison of the effects derived from the current experiment and those in the dual-phase studies in Chapter 2.

The selection-phase of the experiment operated in a similar fashion to that described in Experiments 1b and 2b. Briefly, participants were presented with a problem after a short lead-in and were then required to indicate whether they would retrieve or calculate the problem's answer if they were required to solve it. The strategy selection (retrieve or calculate) was recorded as was the time taken to select the strategy. If a selection was not made within the time limit of 850 ms the lead-in to the next trial started and the response was marked as *late*. After making a strategy selection, or a *late* response, the lead-in to the next trial immediately commenced. The experiment opened with 30 practise questions designed to accustom participants to the selection task. A further 10 practise questions followed, those participants not required to use suppression during the experiment completed them as normal. But participants required to suppress completed the problems in tandem with articulatory suppression.

During the experimental trials, each block of 10 problems, which were organised in a novel pseudo-randomised order for each participant, was punctuated by a pause in the program. In the no suppression condition, a message prompting participants to take a brief pause appeared, and to click on the start button when they were ready to recommence. For those required to suppress, a prompt indicating that the suppression should cease appeared. To recommence the experiment participants clicked on a start button when they were ready to recommence the suppression and the experiment. The longest time participants were required to perform continuous suppression during this phase of the study was approximately 45 s. This feature was designed to reduce fatigue from the repeated articulations thus allowing the participants to maintain a rapid and continuous rate of suppression during the task. The rate and accuracy of suppression was monitored by the researcher who was present throughout the course of the study. Participants deviating below the required rate were requested to speed up the suppression at an appropriate pause in the primary task.

Upon completion of the selection-phase of the experiment, the solution-phase opened with 10 practise trials, the first 5 of which all participants completed without suppression, for the second 5 trials suppression was required dependant upon the participants' allocated condition. Each of the problems presented in first half of the study were presented again in the solution phase in a pseudo-randomised order. Participants were requested to solve the problems as quickly and accurately as possible. After typing the answer in using the keyboard and pressing the enter key to confirm the entry a lead-in to the next problem started. As in the selection-phase, after each block of 10 trials, the runtime programme paused, allowing all participants a brief rest. The experiment recommenced when participants pressed enter, or clicked the start button to continue.

3.3.2 Results

3.3.2.1 Scoring Procedure

Similar to the dual-phase design, four measures were taken during the experiment. In the selection phase, two measures were automatically recorded by the runtime program; the strategy selection (retrieve or calculate) and the time taken to select a strategy, the selection latency. In the solution-phase the participants' given answer was recorded as was the time taken to enter the answer, the solution latency.

3.3.2.2 Strategy Selection

As Figure 4.2 illustrates, participants chose the calculate strategy more often than the retrieve strategy. This finding is as predicted given the familiarity of the problems presented in the current study when compared to those presented in Experiments 2a and 2b. A 2 (sum familiarity; low vs. high) x 2 (addend status; even vs. mixed) x 2 (suppression in selection-phase; yes vs. no) mixed measures ANOVA run on the percentage of calculate strategy selections yielded two significant interactions between the experimental variables. There was a significant interaction between sum familiarity and addend status, $F(1, 20) = 9.76$, $MSE = 62.36$, $p = .005$. When examining this interaction further, simple effects reveal that addend status — and by implication the selection-by-feature mechanism — only influenced selection in the problems relatively high in familiarity within the stimulus set, $F(1, 20) = 6.17$, $p = .022$. A greater percentage of retrieve selections (and hence fewer calculate selections) were made for even than mixed problems for these higher familiarity problems. Furthermore, simple effects indicated that problem familiarity effects were present in even problems, $F(1, 20) = 10.16$, $p = .005$, but not mixed problems, $F(1, 20) = .02$, $p = .9$, such that in even/high familiarity problems fewer calculate selections were made than in even/low familiarity problems. The complexity of this interaction suggests that participants did not solely rely upon one cue to selection at the exclusion of another but that cue which actually influences selection is specific to particular problems.

Crucially, this interaction revealed the influence of the selection-by-feature mechanism in problems, which in comparison to the decade, mixed and fives problems presented in Experiments 2a and 2b, were relatively low in familiarity. Furthermore, the results show that the influence exerted by the selection-by-feature effect upon strategy selection is not limited to retrieve strategy selections (as evident in Experiment 2a and 2b), but can also influence selection of the calculate strategy. This finding supports the SAC account which proposes that there is a common selection mechanism responsible for both retrieve and calculate selections, rather than separate mechanisms, one responsible for retrieve selections and one responsible for calculate selections (e.g., Nelson & Narens, 1990).

Turning to the effects of the suppression task upon strategy selection, a significant interaction between sum familiarity and suppression condition, $F(1, 20) = 4.6$, $MSE = 62.35$, $p = .045$, upon the percentage of calculate strategy selections reported was revealed. Fewer calculate strategy selections (hence more retrieve selections) were made for relatively familiar, rather than unfamiliar problems, when participants were required to suppress, $F(1, 20) = 7.12$, $MSE = 224.39$, $p = .02$. This demonstrates that under suppression the problem familiarity manipulation influenced strategy selection. Conversely, without suppression, as in Experiments 1a and 1b, there were significant effects of sum familiarity, $F(1, 20) = 8.59$, $MSE = 132.38$, $p = .008$, upon calculate selections. This suggests that the presence of suppression served to inhibit the selection-by-feature effect as there were no effects of addend status upon selection in the suppression condition.

When the implications of the two interactions revealed in the current study are considered in tandem, an interesting picture of the relationship between problem familiarity and problem features emerged. The interaction between addend status and problem familiarity first and foremost denoted the influence of the selection-by-feature mechanism upon selection, albeit its influence was localised to relatively familiar problems within the stimulus set. On the other hand, the interaction between sum familiarity and suppression condition indicates that sum familiarity effects were only present in the suppression condition. Taken together in the problems presented in the current study it was evident that without suppression both the problem feature, addend status, and problem familiarity influenced selection. However, under suppression the influence of the problem feature was attenuated and in its absence problem familiarity influenced selections.

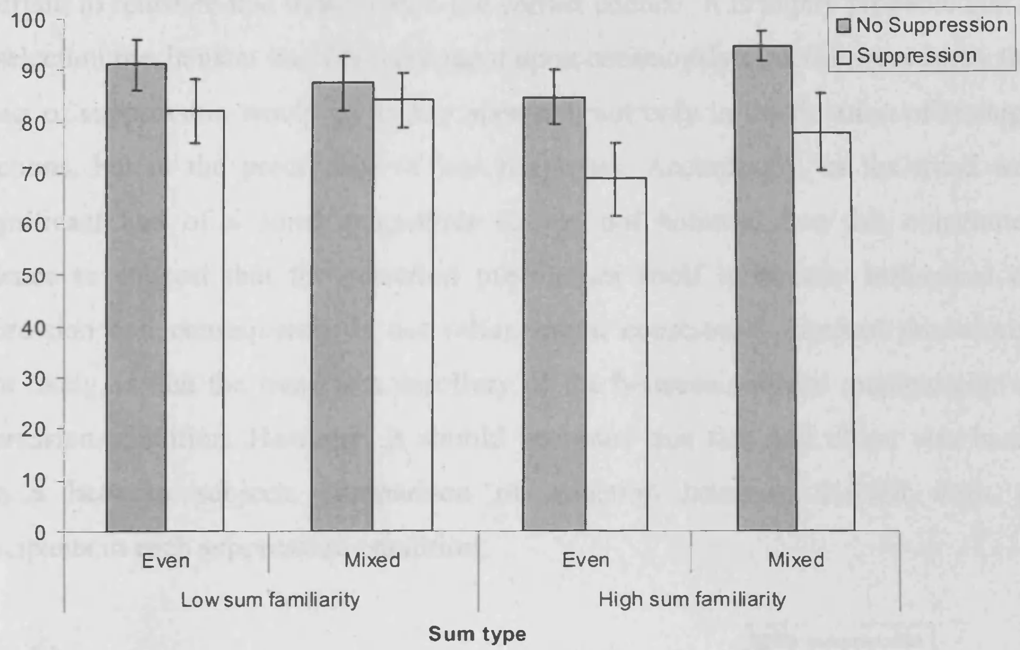

Figure 3.2: The percentage of calculate strategy selections returned either with or without suppression (error bars represent the standard error of the mean).

3.3.2.3 Selection Latency Analysis

Participants required to suppress produced a greater percentage of late responses than those not required to suppress, $F(1, 22) = 5.7$, $MSE = 1.87$, $p = .03$. However, to put this finding in perspective, the percentage of late responses in the experiment amounted to only 5.6% of all responses. When compared to other experiments presented in this thesis this figure is relatively low. As in previous experiments where comparable stimuli were employed (i.e., Experiment 1a and 1b) the mean duration of calculate strategy selections was insensitive to sum familiarity, $F(1, 22) = .02$, $MSE = .001$, $p = .9$. However, dissimilar to Experiments 2a and 2b, where calculate selection latencies were sensitive to the problem feature manipulation of sum type, here latencies were insensitive to the manipulation of addend status, $F(1, 22) = .003$, $MSE = 002$, $p = .96$. Furthermore, suppression did not impact latencies significantly, $F(1, 22) = 2.84$, $MSE = .03$, $p = .11$. This supports the notion that the selection mechanism itself is immune to disruption from the suppression task. However, as Figure 4.3 demonstrates, in general, under suppression participants took marginally longer to make calculate selections (approximately 50 ms). This may be attributable to the effect of the manipulation upon the selection mechanism but it is important to reinstate this trend within the correct context. It is highly probable that if the selection mechanism itself is contingent upon consciously directed procedures, the impact of suppression would be highly apparent, not only in the duration of strategy selections, but in the percentage of late responses. Accordingly, as the trend was insignificant and of a small magnitude it was not believed that this constituted evidence to suggest that the selection mechanism itself is directly influenced by suppression and consequently is not reliant upon consciously directed procedures. More likely, is that the trend is a corollary of the between subjects manipulation of suppression condition. However, it should be noted that this null effect was based upon a between subjects comparison of selection latencies derived from 12 participants in each suppression condition.

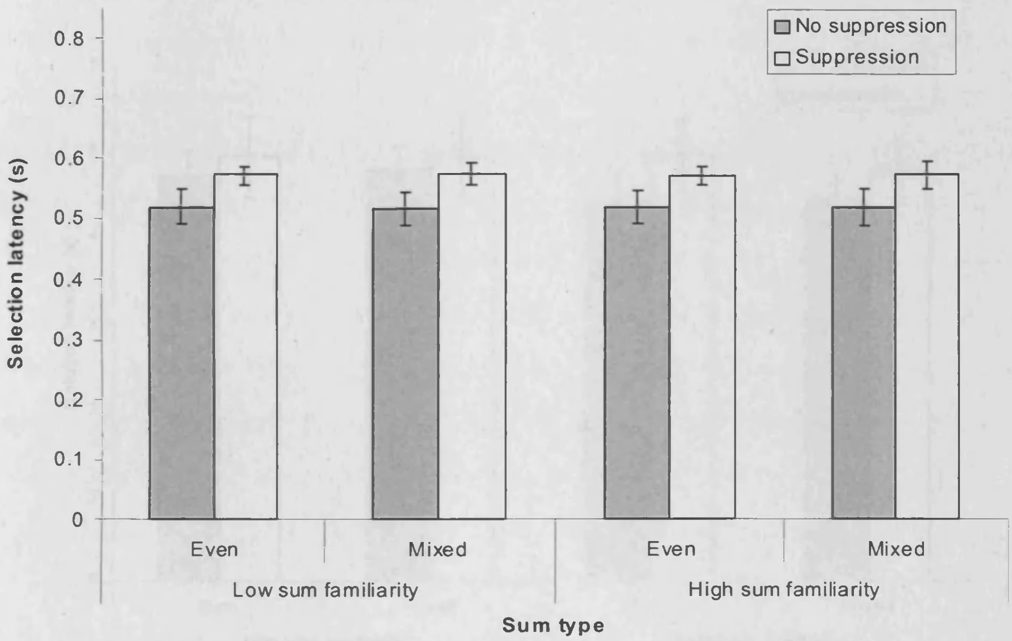

Figure 3.3: Calculate strategy selection latencies either with or without suppression (error bars represent the standard error of the mean).

3.3.2.4 Solution Latency Analysis

Only 2.8% of sum solutions were not made within the time limit of 10 s and 18.1% of given answers were incorrect. In the no suppression condition 8.4% of responses were incorrect, while for those participants required to suppress 8.9% of sum solutions were incorrect. To analyse solution latencies tagged to calculate selections made in the selection-phase a 2 (sum familiarity; low vs. high) x 2 (addend status; even vs. mixed) x 2 (suppression in solution phase; yes vs. no) repeated measures ANOVA was conducted. Effects of sum familiarity were revealed, $F(1, 22) = 5.53$, $MSE = .27$, $p = .03$, indicating that familiar problems were solved more rapidly than unfamiliar problems thus replicating the effects evident in Experiment 1a. Null effects of addend status, $F(1, 22) = .08$, $MSE = .35$, $p = .79$, and suppression condition, $F(1, 22) = .35$, $MSE = 5.58$, $p = .56$, were also found. In Figure 4.4 it is apparent that in general under suppression participants took longer to solve problems than participants not required to suppress. It is likely that the null effects of the suppression condition can be attributed to individual differences. A number of studies in arithmetic problem-solving have demonstrated that performance in mental arithmetic tasks may vary considerably across individuals (see also Campbell & Xue, 2001; Hecht, 1999; Jackson & Coney, 2007; Lefevre & Kulak, 1994; LeFevre, Kulak & Bisanz, 1991; Little & Widaman, 1995).

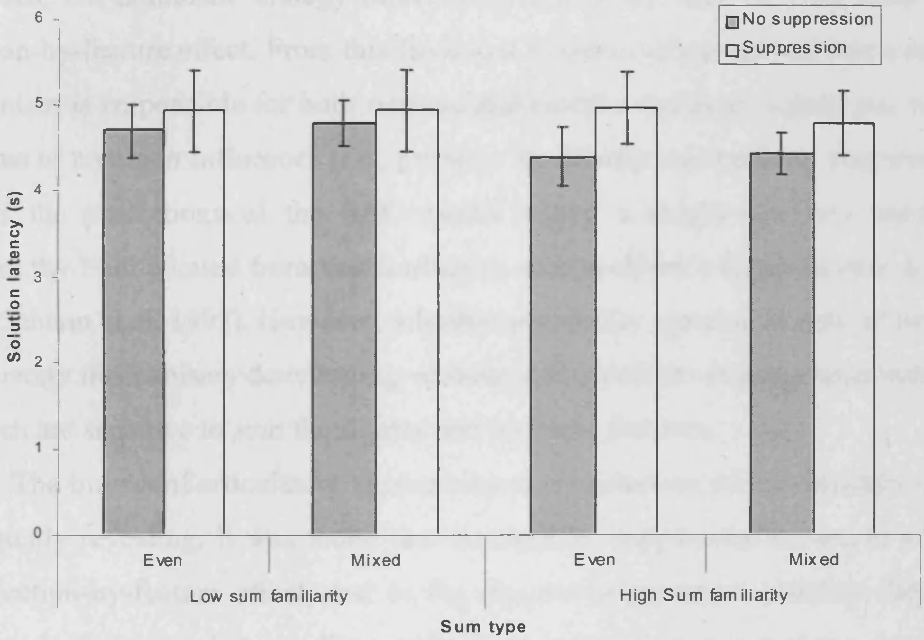

Figure 3.4: The time taken to solve problems (s) in trials tagged to a calculate strategy selections made in the selection-phase of the design. Error bars represent the standard error of the mean.

3.3.3 Discussion

The results from the present study address the candidate issues identified as responsible for the null effect of addend status evident in Experiment 4. To recapitulate, three explanations for this null effect were considered within the current experiment. Firstly, that the null effects were a consequence of the complex factorial manipulation addend status and relative addend magnitude. Specifically, that due to the range of problem types in the stimulus set participants were unable to identify either of the factors, or the levels operationalised within each problem feature manipulation. Secondly, that the selection-by-feature effect is limited to familiar problems, such as the decade, mixed and fives problems employed in Experiment 2a and 2b, rather than the unfamiliar problems employed in the present experiment and also Experiment 4. Thirdly, that the selection-by-feature effect is limited to retrieve strategy selections and does not extend to calculate selections. By reducing the complexity of the problem feature manipulation and introducing a problem familiarity manipulation, an interaction between sum familiarity and addend status was revealed. This indicates for the first time the presence of the selection-by-feature effect in unfamiliar problems presented in this study. Also this demonstrates that calculate selections, the dominant strategy selection in this study, can be influenced by the selection-by-feature effect. From this finding it is tentatively proposed that a common mechanism is responsible for both retrieve and calculate strategy selections, which is sensitive to common influences (i.e., problem familiarity and problem features). This follows the predictions of the SAC model where a single selection mechanism assesses the FoK elicited from the familiarity of a problem's terms (Reder & Ritter, 1992; Schunn et al, 1997). However, a further possibility remains: it may be that there are separate mechanisms determining retrieve and calculate strategy selections, both of which are sensitive to sum familiarity and problem features.

The impact of articulatory suppression upon selection of the calculate strategy was equally revealing. It was found that articulatory suppression served to attenuate the selection-by-feature effect, and in the absence suppression problem familiarity influenced strategy selection. From this it is inferred that conscious processing, occurring between trials was responsible for the identification of that problem feature. Furthermore, that suppression did not influence the time taken to select a strategy, nor the percentage of late responses indicates that the selection mechanism is immune to disruption from this secondary task[5]. Building on this premise, the findings from the current study are revealing of the process by which problem features used by the selection mechanism are identified. In Schunn et al's (1997) Experiment 2 they demonstrated that when presenting participants with problems without any pre-

[5] It should be noted that this between subjects comparison is built upon a sample of 12 participants in each suppression condition. However, this same null effect was replicated in Experiment 5b

experimental familiarity (i.e., problems with multiplication or *sharp* operators), by repeatedly solving these problems, familiarity with the problem components was increased. The linkage between solving a problem, updating the familiarity of the problem and subsequent strategy selections based upon the familiarity of the problem was reflected in the greater number of retrieve strategy selections made in contrast to problems which were only solved infrequently. Previous research has indicated that problem familiarity is derived from the act of solving a problem (Reder & Ritter, 1992; Schunn et al, 1997). Accordingly, the relationship between problem familiarity effects in both the selection- and solution-phases of the design stands as a key component of performance in the dual-phase experimental design. However, it is not clear whether the selection-by-feature effect is contingent upon the individuals' experience of solving the problems or a pre-experimental awareness of a host of potentially useful problem features.

In the present study participants completed both phases of the experiment in isolation, making predicted strategy selections in the selection-phase before moving on to solve the problems in the solution-phase. Similarly, in Experiment 2b which used the single phase experimental design (i.e., section-phase only), selection-by-feature effects were evident. In both designs, when making the predicted strategy selection, participants had no prior experience of solving those or similar problems. Accordingly, it seems reasonable to conclude that the identification of problem features useful to the selection mechanism is not necessarily derived from the act of solving a problem as is apparent in respect to problem familiarity. Hence it is proposed that selection-by-feature effects upon both strategy selections and solution latencies may not be a key signature of performance in this experimental paradigm as is the case with problem familiarity effects (see Schunn et al, 1997).

To test this notion further, and clarify the relationship between the two factors of sum familiarity and problem features in strategy selection, the following experiment also employs articulatory suppression to inhibit the selection-by-feature effect but in a set of problems rich in problem features and high in pre-experimental familiarity, namely the decades, mixed and fives problems used previously in Experiments 2a and 2b. From this experiment it should be evident whether the selection-by-feature effect is reliant upon feedback derived during the course of the experiment or on pre-experimental experience.

3.4 EXPERIMENT 5b

The present experiment was designed to examine the process by which problem features are identified. In Experiment 5a it was shown that articulatory suppression mitigated the selection-by-feature effect by attenuating effects of the problem feature manipulation addend status. This finding was tentatively attributed to the notion that the influence of addend status upon selection was contingent upon an inter-trial conscious evaluation of the problem features inherent in the stimulus set. The alternative hypothesis, that the suppression task impacted the selection mechanism itself was refuted by the finding that selection latencies were not affected by the suppression task.

This was explored by testing familiar problems rich in problem features in the same experimental design as employed in the Experiment 5a. The sum type manipulation was used to examine whether suppression attenuated selection-by-feature effects in decades, mixed and fives problems. As stated in Experiment 2a and 2b, the sum type manipulation presents a set of problem features which are highly apparent to individuals, easily identifiable and rich in pre-experimental familiarity. In Experiment 2b the influence of sum type upon retrieve strategy selections successfully replicated the effects evident in Experiment 2a accordingly sum type is seen as a robust effect. In Experiment 5a it was shown that suppression inhibited effects of addend status upon selection. From this it was inferred that this problem feature was identified by participants during the course of the experiment under conditions in which a conscious appraisal used to identify particular problem features was not impeded by suppression. However, it may also be the case, similar to problem familiarity, where prior experience with solving specific problems serves to increase their familiarity, that problem features are not only identified during the course of an experiment, but are already known from previous processing episodes. This scenario is particularly likely with decades, mixed and fives stimuli as these problems are encountered frequently and as such it is probable that individuals are highly accustomed to using the features in these problems to facilitate problem-solving performance. Accordingly, it was predicted that if the identification of the problem features that contribute to selection-by-feature effects are solely reliant upon conscious processes operating during the course of the experiment then suppression will attenuate the selection-by-feature effect. Alternatively, null effects of suppression upon the sum type manipulation would indicate that the problem features influencing selection were already known by the participant such that identification of the feature was not required during the course of the experiment.

As in Experiment 5a, participants first made predicted strategy selections for each of the problems in the stimulus set, then solved all of the problems. In addition, the current study also provided an opportunity to replicate a key finding from Experiment 5a. In that experiment it was shown that selection latencies were largely immune to interference from the suppression task. From this it was inferred that the influence of the suppression task could not be localised to the processes responsible for actually making the strategy selection, but the consciously directed mechanism which operates during the course of the experiment, between trials, and is potentially responsible for identifying problem features.

3.4.1 Method

3.4.1.1 Participants

Twenty-four participants from the School of Psychology at Cardiff University were given course credit or payments in return for their participation. All were native English speakers reporting normal hearing and correct or normal vision.

3.4.1.2 Materials & Design

The same type of stimuli detailed in Experiments 2a and 2b were employed in this study. To recapitulate, the addends in all of the problems were drawn from a sample of decades and fives numbers from which three conditions were constructed, *decades* (e.g., 20 + 50), *fives* (e.g., 25 + 50) and *mixed* (i.e., 25 + 55). All problems summed to less than 100 and did not comprise of any tied addends (e.g., 25 + 25). Twenty-four problems were prepared for each sum type condition comprising 12 novel problems and 12 further problems created by switching the order of the addends in the novel problems (see Appendix E, table E5, for stimuli).

As in the previous experiment, participants were randomly allocated to one of four suppression conditions in which they were required to carry out suppression in both the selection- and solution-phases of the experiment, in only one phase, or in neither phase. The selection-phase of the experiment commenced with 30 practise questions, followed by a further 10 practise questions which were completed with or without articulatory suppression dependant upon which condition participants were allocated to. None of the practise questions comprised addends divisible by 5 or 10, ensuring the experimental manipulation of sum type manipulation was first encountered by participants during the test phase.

The solution-phase of the experiment started with 10 practise questions. None of the participants were required to practise suppression for the first 5 problems, while in the last 5, those required to, completed these trials under suppression. After each block of 10 trials in the selection- and solution-phase of the experiment the program paused, providing participants with an opportunity to take a break from the suppression, and recommence the experiment when ready.

3.4.1.3 Procedure

The exact same runtime procedure and instructions were issued to participants as detailed in the previous experiment.

3.4.2 Results

3.4.2.1 Scoring Procedure

See corresponding section in Experiment 5a for further detail.

3.4.2.2 Strategy Selection

As in Experiments 2a and 2b where sum type was also manipulated the retrieve strategy was selected more often than the calculate strategy (see Figure 4.5). Furthermore, a 3 (sum type; decades, mixed and fives) x 2 (suppression condition; no suppression vs. suppression) mixed measures ANOVA indicated that the percentage of late selections recorded was immune to both sum type and suppression condition, $F(2, 21) = 1.69$, $MSE = .73$, $p = .21$, and $F(1, 22) = .16$, $MSE = 3.09$, $p = .69$, respectively. Using the same ANOVA a significant interaction between sum type and suppression condition upon the percentage of retrieve selections made was found, $F(2, 21) = 4.52$, $MSE = 265.4$, $p = .02$. Simple effects demonstrated that sum type influenced selection irrespective of whether participants were required to suppress, $F(2, 21) = 34.46$, $p < .001$, or not, $F(2, 21) = 9.24$, $p = .001$. These two findings not only replicate the effect of sum type on selection evident in Experiments 2a and 2b, but contrast the effects of suppression evident in Experiment 5a, where suppression attenuated the effects of the problem feature manipulation addend status.

Pairwise comparisons indicate that a significant effect of suppression upon retrieve selections was only evident for fives problems (p = .02). A greater percentage of retrieve selections (and hence a lower percentage of calculate selections) was made under suppression than without suppression (see Figure 4.5). It could be that suppression effects can be localised to fives problems. However, it is not possible to corroborate this finding conclusively from the data derived from this study. A further possibility is that this finding, rather than reflecting the influence of articulatory suppression upon selection, may simply reflect the abnormally low frequency of retrieve selections made by participants who were not required to suppress. To illustrate, in Experiment 2a (where participants were not required to suppress) retrieve was selected in 56.9% of fives problems. Whereas in the present study participants not required to suppress only selected the retrieve strategy in 30.9% of fives problems. Furthermore, the lower rates of retrieve selection in this study could not be localised to the performance of a small minority of participants as 66% of participants chose retrieve in 41.6% of trials or less.

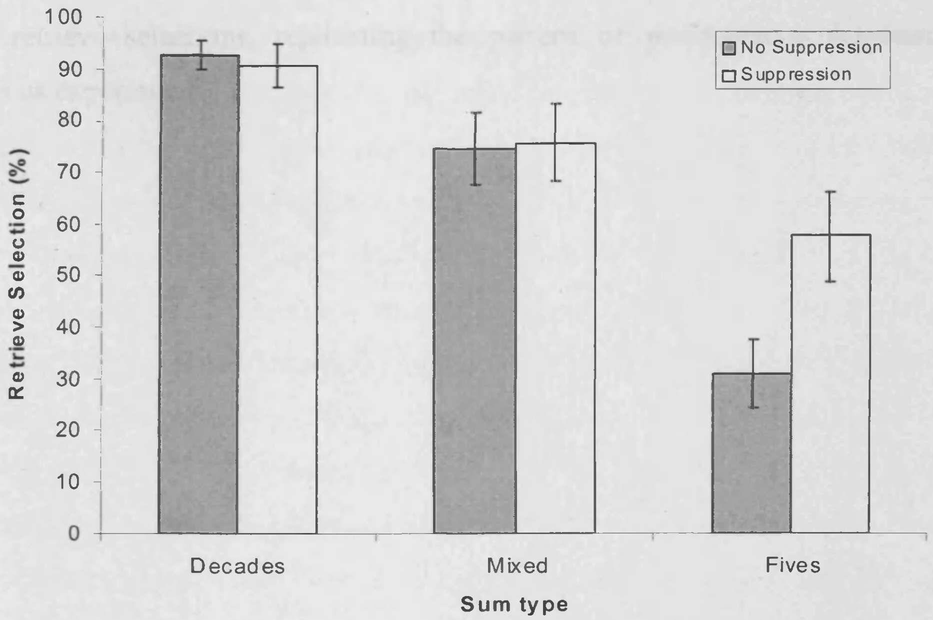

Figure 3.5: The percentage of retrieve strategy selections returned either with or without suppression (error bars represent the standard error of the mean).

3.4.2.3 Selection Latency Analysis

In total, 3.93% of responses were not made within the time limit, indicating that participants were easily able to make strategy selections with and without suppression, well within 850 ms. Breaking down this figure, 2.1% of responses were not made within the time limit in the no suppression condition while 1.8% of responses were late in the suppression condition. A mixed measures ANOVA run with sum type as a within subjects factor and suppression condition as a between subjects factor revealed that sum type influenced the duration of retrieve strategy selections, $F(2, 20) = 8.26$, $MSE = .001$, $p = .002$. Similar to Experiments 2a and 2b, post hoc comparisons indicated that the retrieve strategy was selected more rapidly in the decades than mixed or fives condition, and more rapidly in the mixed than fives condition (all *ps* $< .05$). Between subjects analysis replicated the findings from Experiment 5a indicating that there were null effects of suppression condition, $F(1, 21) = .35$, $MSE = .01$, $p = .56$, demonstrating that the time taken to select the retrieve strategy was not influenced by the suppression manipulation. Similar to the previous experiment, from this finding it is inferred that the suppression task did not exert a direct influence upon the selection mechanism. Due to the low frequency of calculate strategy selections the mean duration of calculate selections were not analysed. However, as Table 4.2 illustrates calculate selection latencies were consistently longer than retrieve selections, replicating the pattern of performance demonstrated in previous experiments.

Table 3.2.

Summary by condition of mean calculate (Calc) and retrieve (Ret) strategy selections (in %), selection latencies (in ms) and solution latencies (in ms) split by sum type condition and articulatory suppression condition. Standard deviations are presented in parentheses.

		Decades		Mixed		Fives	
		No Supp	Supp	No Supp	Supp	No Supp	Supp
Selection Latency	Calc	523	567	587	582	565	571
		(52)	(67)	(82)	(95)	(67)	(76)
	Ret	497	515	534	533	569	523
		(52)	(73)	(61)	(84)	(57)	(83)
Solution Latency	Calc	1758	1808	2227	2221	2632	3438
		(381)	(778)	(609)	(464)	(903)	(649)
	Ret	1663	1673	2163	2219	2575	2970
		(390)	(287)	(685)	(447)	(1140)	(529)

3.4.2.4 Solution Latency Analysis

Only 1 response in the whole dataset was not made within the time limit of 10 s set for responses in this phase of the experiment. In total only 7.35% of sum solutions were incorrect, comprised of 5.2% of responses in the no suppression condition, 2.1% in the suppression condition. The difference between the two conditions can be largely accounted for by the performance of one participant in the no suppression condition whose errors alone accounted for 3.1% of the total errors made in this phase in the whole experiment. A mixed measures ANOVA revealed significant effects of the sum type manipulation upon solution latencies which were tagged to both retrieve and calculate selections made during the selection-phase, $F(2, 20) = 35.47$, $MSE = .14$, $p < .001$, and $F(2, 8) = 19.08$, $MSE = .123$, $p = .001$, respectively. In both, problems were solved more rapidly in the decades than mixed or fives, and in the mixed than fives (all *ps* < .002). Articulatory suppression did not influence the time taken to solve problems tagged to retrieve, $F(1, 21) = .61$, $MSE = .95$, $p = .44$ or calculate selections, $F(1, 21) = .03$, $MSE = .89$, $p = .88$, made in the selection-phase of the experiment. Similar to the previous experiment, despite the presence of a trend, individual differences are posited as responsible for the null effects of suppression upon solution latencies tagged to retrieve selections.

3.4.3 Discussion

The results detailed in the previous section support the notion that generally speaking the problem features responsible for selection-by-feature effects are not necessarily identified by a conscious appraisal of the problems during the course of the experiment. Although there was a significant interaction between sum type and suppression condition only retrieve selections in the fives condition were influenced by suppression. However, the abnormally low level of retrieve selections, in comparison to those reported in the same sum type condition in Experiments 2a and 2b, serves to question whether suppression actually influenced selection in these problems. Null effects of suppression upon selection in decades and mixed problems supports the notion that problem features used for these problems were already known by participants prior to the experiment. Furthermore, that they were employed by the selection mechanism without the contribution of consciously directed procedures. This finding contrasts the conclusions drawn in Experiment 5a where the significant interaction between addend status and suppression condition revealed that participants identified problem features during the course of the experiment. However, replicating the effects of suppression upon selection latencies in the previous study, it was apparent that the selection mechanism itself was immune to influence from the suppression task.

One caveat with the notion that selection is influenced by pre-experimentally derived problem features is that the possibility still remains that selection may be influenced by unconscious processes. Returning to Siegler and Stern's (1998) analysis of insight in strategy discovery, it may be that problem features are identified during the course of the experiment but at an "implicit, unconscious, metaprocedural level" (p. 378, Siegler & Stern, 1998, see also Karmiloff-Smith, 1992) immune to interference from the suppression task. Accordingly, it is possible that conscious evaluative processes are not responsible for the identification of problem features which result in selection-by-feature effects. The findings from the present experiment, taken in tandem with those from the previous study, appear to support the position forwarded by Cary and Reder (2002), that there is no clear distinction between conscious and metacognitive processing in strategy selection. The most useful description of the interplay between these two processes is that their deployment was opportunistic. In Experiment 5a it was concluded that articulatory suppression inhibited the action of consciously directed processes operating over the course of the experiment. These conscious processes were proposed to be responsible for identifying features inherent in the problems presented in the stimulus set that were subsequently employed by the selection mechanism to direct retrieve/calculate selections. While in this study articulatory suppression did not impact the discovery of new problem features as the selection mechanism was already using known problem features derived from prior problem solving episodes.

In an experiment detailed in Appendix D, the possibility that unconscious mechanisms are responsible for the identification of problem features during the course of the experiment was examined. This notion has already been highlighted in the specifications of the SCADS* mechanism, where to briefly recapitulate, perceptual and encoding processes invoked as responsible for the identification of problem features (Siegler & Araya, 2005). A secondary task designed to inhibit a comparison between the features of each addend inherent in a problem (e.g., *the first and second addends in a problem are both even numbers*) was employed. Selection- and solution-phases were blocked (as in the current study) and the same set of problems tested in Experiments 5a was examined. The secondary task was derived from the Vigilance literature where participants are commonly required to monitor a continuous presentation of letters, words or other stimuli, reporting the occurrence of a critical signal (Ballard, 1996, see also Auburn, Chapman & Jones, 1987). Critical to performance in symbolic vigilance tasks is the ability to encode the signal, detecting changes in the signal and identifying 'hits' (Ballard, 1996). To appreciate a hit the individual is required to retrieve the higher order semantic category of each signal. For example, if the signal was the word 'carrot', the higher order semantic category would be 'vegetable'. They must then maintain that category name in short-term memory while comparing the semantic category of the current signal to the prior signal. In accordance with the rapidity with which signals were presented (i.e., 1 per second) participants would have been unable to use conscious processes to complete this task. During blocks of trials a continuous presentation of words drawn from a number of categories were presented auditorily. A critical signal was defined as the occurrence of three items from the same category in a row. Participants made retrieve or calculate strategy selections and at the end of each trial block indicated whether a signal was present. When detecting a problem feature (such as those used to influence selection in the addend status manipulation) it is assumed that the individual will extract a host of features from each of the addends in the problem. For example, *the first addend is an even number*, *the second addend is an even number*. Then search for a correspondence between the features identified in each addend, *both addends are even numbers*, by comparing the activated feature. It was predicted that if feature detection in problems is contingent upon perceptual and encoding mechanisms operating during the course of the experiment that this secondary task would inhibit the influence exerted by the problem feature addend status. However, there was a non significant interaction between the vigilance task condition and the problem feature addend status. Accordingly, it may be that perceptual and encoding mechanisms are not solely responsible for the identification of problem features during the course of the experiment. However, further investigation of this hypothesis is required to isolate the action of mechanisms with the potential to identify and compare the types of features that can be extracted from a problem. A more detailed exposition of this experiment is presented in Appendix D.

In summary, the effects of articulatory suppression upon strategy selection revealed in Experiments 5a and 5b suggest that the selection-by-feature effect may be contingent upon conscious processes identifying problem features during the course of the study. The results from Experiment 5a demonstrate that effects of addend status were mitigated when the contribution of conscious processes to performance was precluded by the articulatory suppression secondary task. Conversely, those participants not required to suppress in that study demonstrated effects of addend status. However, as Experiment 5b demonstrated, it was found that inhibiting conscious procedures does not necessarily mitigate problem feature effects. Suppression had no effect upon the influence exerted by the sum type (i.e., decades, mixed and fives problems) manipulation on retrieve strategy selections. Following the predictions of the SCADS* model, in an experiment presented in Appendix D of this thesis, the possibility that encoding and perceptual processes were responsible for feature detection, rather than conscious mechanisms, was examined. This experiment was conducted to investigate whether the null effects of suppression evident in Experiment 5b were suggestive of the notion that unconscious processes (i.e., encoding and perceptual mechanisms) were actually responsible for feature detection, rather than conscious processes as suggested by the results from Experiment 5a. However, as there was no evidence of an interaction between the vigilance task designed to inhibit feature detection and addend status. This suggests that as in

Experiment 5a the null effects of suppression could be attributed to the proposition that participants were already aware of the problem features applicable to decade, mixed and fives problems and hence did not need to identify any problem features during the course of that experiment. In the following experiment a further type of task-level manipulation is employed to examine the susceptibility of selection to task instructions.

3.5 EXPERIMENT 6a

Building upon the distinction drawn between problem-level and task-level manipulations of strategy selection, the present experiment was designed to examine a further type of task-level manipulation. Findings from a range of studies have shown that strategy selection in problem-solving tasks is highly susceptible to task instructions. For example, in a series of maze navigation puzzles, Gardner and Rogoff (1990) used instructions emphasising speed over accuracy and vice-versa. They found that children used less advanced planning strategies when speed was emphasised and more advanced planning strategies when accuracy was emphasised. Returning to mental arithmetic problem-solving, Kirk and Ashcraft (2001) reported that a between subjects manipulation of instruction type served to bias self-reported indications of strategy selection. Testing single-digit addition and multiplication problems, participants were required to solve the problem then identify which strategy they actually used to solve the problem. In a between subjects design three sets of task instruction were examined, one group received instructions emphasising the use of direct retrieval to solve problems, another emphasising the use of calculation procedures and a further set of instructions derived from LeFevre, Sadesky and Bisanz (1996) which did not explicitly favour either strategy. They found that by manipulating the emphasis of the instructions self-report responses followed suit. For example, participants receiving instructions emphasising use of the direct retrieval strategy reported using calculation procedures on only 9% of addition problems and 4% in multiplication trials. Conversely, those who received instructions emphasising calculate procedures reported using calculate strategies on 62% of their addition trials and 38% of multiplication trials when presented with the same problems. A number of authors have expressed doubts with the reliability of self-report measures (e.g., Ericsson & Simon, 1993; Kirk & Ashcraft, 2001 and Payne, 1994). This concern stems from the notion that the mental processes responsible for selecting a strategy and deriving a solution are not necessarily accessible to conscious processes (see Cary & Reder, 2002 for discussion of this issue). Accordingly, the accuracy with which retrospective strategy selections are identified may be questionable. From Kirk and Ashcraft's (2001) findings it is unclear whether the task instruction effect (see also Blöte, Van der Burg & Klein, 2001) occurs as a function of the conscious and evaluative processes used to reconstruct the prior problem-solving episode as part of the self-report response, or alternatively, whether unconscious and implicit processes are responsible for the effect.

In the present study, similar to Experiments 1b and 2b, the selection-phase design was employed but with a between subjects manipulation of task instructions to examine whether they influence the selection mechanism. From Kirk and Ashcraft's (2001) study it seems likely that the influence of task instructions upon self-reported strategy selections was contingent upon the action of consciously directed procedures used retrospectively to identify strategy selection in an earlier processing episode. However, Cary and Reder (2002) propose that task instruction manipulations operate at an unconscious level. In this study a modification to the methodology used previously in this thesis (where participants were required to choose either retrieve or calculate) was made such that participants were allocated to one of two *task instruction* conditions. Depending upon which condition they were randomly allocated to, participants indicated whether when solving the problem they would *retrieve the answer, yes or no*, or *calculate the answer yes or no*. Analogous to the manipulation used by Kirk and Ashcraft (2001), it was predicted that the two types of instruction would elicit a bias towards the strategy identified in the instruction. Specifically, it was predicted that participants responding to the retrieve yes/no instruction would favour selection of the retrieve strategy over those responding to the calculate yes/no instruction. The time limit imposed on the selection-phase of 850 ms in this design served to prevent participants from solving the problem and using conscious recollective processes to identify the strategy used to solve the problem. Accordingly, significant effects of the task instruction manipulation upon selection would be indicative of the action of the selection mechanism, rather than the consciously directed procedures used to identify strategy selections retrospectively.

The same stimulus manipulations as detailed in Experiment 1a and 1b (i.e., sum familiarity and answer familiarity) were used in this study. In those experiments a robust effect of the sum familiarity manipulation was evident such that a greater percentage of calculate selections was made for relatively unfamiliar problems than familiar problems. Null effects of the answer familiarity manipulation indicated that the familiarity of a problem's solution did not influence the selection process. The current experiment also provided an opportunity to examine the mechanics of the FoK response and how the activations elicited by a problem result in retrieve and calculate strategy selections. This issue has been touched upon briefly in the discussion of results originating in previous experiments (e.g., Experiment 5a) but has not received direct empirical investigation in this thesis. However, due to the centrality of the FoK mechanism to the approach to selection adopted in Chapter 3 it is worthy of further examination. In the SAC model (Reder & Ritter, 1992; Schunn et al, 1997), it is proposed that the relative level of activation at the most active node in memory determines the strength of the FoK. Accordingly, for activations elicited by familiar problems there is a greater disparity between activation at the most active node and the second most active node than evident in unfamiliar problems (Schunn et al, 1997). The FoK is subject to a threshold mechanism against which the strength of the FoK is compared to a preset threshold. If the FoK strength exceeds the threshold the retrieve strategy is selected. Alternatively, if the FoK fails to breach the threshold level the retrieve strategy is not selected (i.e., the calculate strategy is chosen). This mechanism is analogous to the *single-counter* FoK hypothesis proposed by Nelson and Narens (1990) where a FoK is based upon the availability of information in memory. The single-counter hypothesis is contingent upon the information accumulated about an item known in memory, which in this experimental paradigm, rather than the recognition paradigm used by Nelson and Narens (1990) to examine the FoK mechanism, would elicit a retrieve selection when activation levels exceed the threshold. Of course, by default, the mechanism can also elicit calculate selections but only as a consequence of insufficient information accumulating over time to exceed the threshold.

An alternate conception of FoK is proposed by Nelson and Narens (1990), derived from their examination of an apparent paradox in research conducted by Hart (1965). Hart (1965) illustrated that participants were unable to recall the solutions to general knowledge questions could indicate very rapidly whether they knew the answer or not. Furthermore, that these evaluations were positively correlated to performance in later recognition tests of the unrecalled items. This suggested that participants knew about the absence or presence of an item in memory without having to complete a memory search for the item. From this Nelson and Narens (1990) proposed a *dual-counter* hypothesis, where one FoK component taps information in memory and the other tapping information pertaining to what is not present in memory. Relating this conception back to strategy selections made in the experimental paradigm used in this study, calculate selections do not occur as a default but as a consequence of accumulating evidence indicating that an item is not resident in memory. While, as before, retrieve selections are based upon the activation elicited by information indicating that the item is stored in memory. Separate FoK counters index these activations whose return is subject to a *difference threshold* which discerns the difference in the two values. Where a positive value returned (i.e., the item is known in memory) in this experimental paradigm a retrieve strategy selection would be made. Conversely, a negative return would elicit a calculate selection. Accordingly, it was reasoned that if the dual-counter hypothesis is responsible for FoK responses in this strategy selection paradigm, it may be the case that the type of information which can be extracted from a problem contributes to either knowledge of facts stored in memory, or not stored in memory. Therefore, it was predicted that if sum familiarity effects were evident in only one of the response type conditions (i.e., calculate yes/no and not retrieve yes/no), this would indicate that separate selection mechanisms influence retrieve and calculate strategy selections.

To address a potential flaw in the design of Experiments 1a and 1b the answer familiarity manipulation was reintroduced. A number of models of the strategy selection process detailed previously suggest that the familiarity of a problem's answer influences strategy selection (i.e., ACT-R, ITAM and DOA). In these accounts the act of encoding a problem automatically initiates a search for the problem's solution (i.e., the Obligatory Activation Assumption, see Logan, 1988). Previously, the null effects of the answer familiarity manipulation upon selection found in Experiment 1a and 1b indicate that the memorial search for the answer does not influence selection. However, effects of this manipulation may be revealed by the task instruction manipulation employed in this experimental design if separate FoK counters determine retrieve and calculate selections. Specifically, as the time limit in the selection-phase prevents the participant from solving the problem before making a selection, answer familiarity has been conceived as a predictor of the progress made in finding a solution before a strategy selection had to be made (Reder & Ritter, 1992). Accordingly, in problems with familiar answers (as opposed to unfamiliar answers), greater progress in retrieving a solution will be made up to the point at which this process is truncated and a selection returned based upon this level of progress. It was thus predicted that effects of the answer familiarity manipulation would be more likely in the retrieve yes/no instruction condition as answer familiarity stands as an indicator of progress made in retrieving the answer to a problem in accordance with the Obligatory Activation Assumption (Logan, 1988).

A further between subjects manipulation, termed *key position,* was used to ensure that there was no bias in selection attributable to the relative positioning of the *yes* and *no* keys on the keyboard (i.e., on the left or right side of a standard qwerty keyboard). In summary, it was predicted that effects of task instruction would indicate whether the unconscious mechanisms responsible for strategy selections are solely determined by problem-level manipulations (i.e., problem or answer familiarity), but also features of a task (i.e., task instructions). Null effects of key position would confirm that any task instruction effects were not attributable to a bias in selections arising from the positioning of the yes and no keys on the keyboard. From the two within subjects manipulations, sum familiarity and answer familiarity, effects of these variables would be revealing of the operation of the selection mechanism. If the percentage of calculate selections elicited from the two task instruction conditions both exhibited effects of sum familiarity this would support the SAC model's account of selection. Alternatively, if sum familiarity effects were only evident in one of the task instruction conditions this would support the notion that separate mechanisms are responsible for retrieve and calculate selections. Furthermore, effects of answer familiarity may interact with task instruction if separate FoK components determine retrieve and calculate selections.

3.5.1 Method

3.5.1.1 Participants

Twenty-four participants from the School of Psychology at Cardiff University were given course credit or payment in return for their participation. All were native English speakers reporting normal hearing and corrected or normal vision and had not participated in any of the other thesis experiments.

3.5.1.2 Materials & Design

The same type of stimulus manipulations were used as in Experiments 1a and 1b (sum familiarity and answer familiarity). All sums were double-digit addition problems comprising two addends drawn from a sample ranging from 12 to 49 (see Appendix E, table E6, for stimuli). None of the addends were divisible by 5 or 10, there were no tied addends (i.e., 23 + 23) and each addend pair was from the same decade class (e.g., 23 +29). The two variables, sum and answer familiarity, were contrasted in a repeated measures design, in which participants completed 16 practise trials followed by 64 experimental trials, 16 trials in each condition, the order of which was pseudo-randomised for each participant. Separate repeated measures ANOVAs ensured that there was a significant difference between the low and high levels of problem familiarity, $F(1, 15) = 105.34$, $MSE = 4142.89$, $p < .001$ and answer familiarity; $F(1, 15) = 412.63$, $MSE = 427.50$, $p < .001$.

Task instructions were manipulated as a between subjects variable, 12 participants were asked to indicate whether they would retrieve the answer, yes or no, the other 12 were asked whether they would calculate the answer, yes or no. Furthermore, the positioning of the 'yes' key was manipulated as a between subjects variable. For half of the participants (n = 12) the yes key was positioned on the left side of the keyboard (on the z key) and no on the m key. For the second half the positions were reversed.

3.5.1.3 Procedure

The same selection-phase procedure was used as in Experiments 1b and 2b. To recapitulate, after a short lead-in participants were presented with a problem and asked to make a predicted strategy selection within 850 ms. However, the design employed in the current experiment included one notable change. Whereas previously participants were required to indicate what strategy they would use to solve a problem by selecting either retrieve or calculate, here participants were either asked whether they would retrieve the answer and responded by choosing yes or no. Participants in the other task instruction condition were asked if they would calculate the answer and responded with yes or no. Once the strategy selection had been made participants were prompted to initiate the lead-in to the next trial.

3.5.2 Results

3.5.2.1 Scoring Procedure

Two measures were derived; the *strategy selection* which, dependant upon the response type condition they were allocated to (i.e., would you retrieve the answer, or would you calculate the answer), participants responded with a yes or no selection. It should be highlighted at this juncture that *retrieve yes* and *calculate no* selections both equated to the same underlying strategy selection (i.e., retrieve), *calculate yes* and *retrieve no* selections to the selection of the calculate strategy. To minimise confusion, in the analysis, reported strategy selections will be identified by the underlying strategy selection they represent (i.e., retrieve or calculate) and will be split by task instruction condition (calculate y/n, or retrieve y/n). The *strategy selection latency* was also recorded; this measured the time taken to make the strategy selection. Unless detailed to the contrary, all statistical analyses will employ mixed measure ANOVAs using a 2 (selection question; retrieve y/n vs. calculate y/n) x 2 (yes key position; left vs. right) x 2 (sum familiarity; low vs. high) x 2 (answer familiarity; low vs. high) design. In each of the between subjects conditions there were 12 participants and where insignificant, interactions are not reported.

3.5.2.2 Strategy Selection

As in Experiments 1a and 1b a higher percentage of calculate strategy selections were returned than retrieve selections (see Figure 4.5). To test the relationship between calculate and retrieve selections a mixed measures ANOVA was conducted to identify whether any of the variables manipulated made participants return a greater number of late responses. Null effects of the problem-level manipulations sum and answer familiarity and the task-level manipulations, key position and task instruction, (all Fs < .72, all ps > .39) revealed that the manipulated variables failed to compromise the dependant relationship between the two underlying strategies retrieve and calculate. This replicates findings from all of the previous experiments which have demonstrated that the dependant relationship between retrieve and calculate selections is unaffected by the experimental manipulations.

Similar to previous experiments presented in this thesis which employed the sum familiarity manipulation, the familiarity of the problem was shown to influence calculate selections, $F(1, 20) = 20.83$, $MSE = 4.61$, $p < .001$. Pairwise comparisons demonstrate that the calculate strategy was selected more often in unfamiliar than familiar problems (p < .001). Furthermore, null effects of answer familiarity upon the percentage of calculate strategy selections made, $F(1, 20) = .55$, $MSE = 3.7$, $p = .47$, replicate the findings from Experiments 1a and 1b. Both of these findings serve to replicate the pattern of performance exhibited in Experiments 1a and 1b which employed identical stimuli. When considering the impact of task instructions on the task, the interaction between sum familiarity and task instruction, $F(1, 20) = .33$, $MSE = 4.61$, $p = .58$, was insignificant. This suggests that sum familiarity had comparable effect upon the strategy selection irrespective of the task instruction participants received. There was a non-significant interaction between task instruction and answer familiarity, $F(1, 20) = .01$, $MSE = 3.7$, $p = .92$. From this there is no evidence to suggest that separate mechanisms are responsible for retrieve and calculate strategy selections. Also, the task instruction manipulation failed influence calculate strategy selections, $F(1, 20) = .17$, $MSE = 35.4$, $p = .69$, and null effects of key position reveal that the spatial positioning of the yes and no keys on the keyboard did not serve to bias participants responses in favour of one response key over the other, $F(1, 20) = .005$, $MSE = 35.4$, $p = .95$.

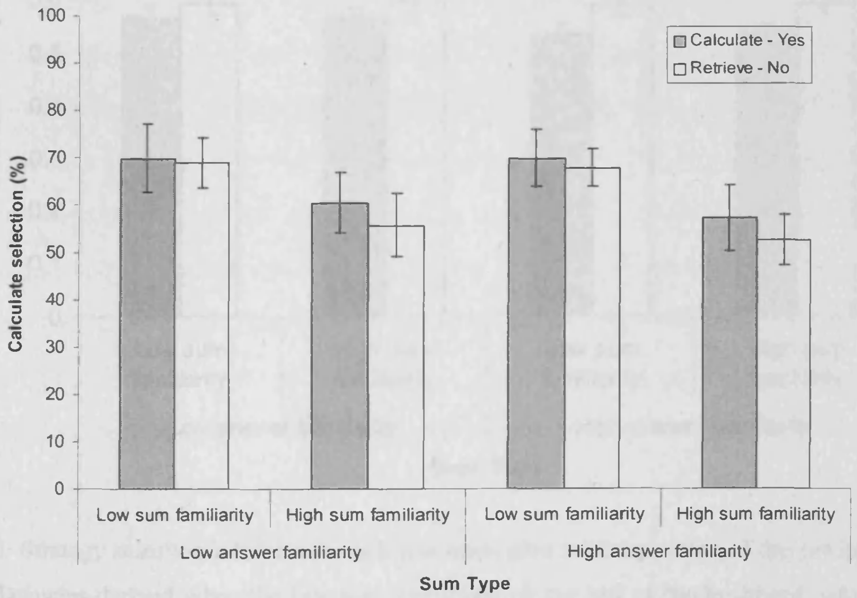

Figure 3.5: The percentage of retrieve strategy selections made in each of the sum type conditions for participants making calculate strategy selections (i.e., selecting 'calculate yes' vs. 'retrieve no'). Grey bars represent responses to the selection question 'calculate y/n' where the participants chose 'calculate no' (i.e., to use the retrieve strategy). White bars represent responses when participants were asked 'retrieve y/n' and selected 'retrieve yes'. Error bars represent the standard error of the mean.

3.5.2.3 Strategy Selection Latency

Only 10.6% of responses were not made within the 850 ms window allocated for strategy selections in the current experiment. It was found that sum familiarity did not influence the amount of the time taken to select the calculate strategy, $F(1, 20) = .01$, $MSE < .001$, $p = .91$. However, marginally significant effects of the answer familiarity manipulation were evident, $F(1, 20) = 4.55$, $MSE = .01$, $p = .05$, such that calculate strategy selections were made more rapidly in problems with relatively familiar answer than problems with unfamiliar answers. Although this level of significance is marginal non-significant effects of answer familiarity were evident in Experiments 1a and 1b accordingly, further consideration of this result is not developed in this thesis. Non-significant effects of task instruction and yes key position upon the time taken to select the calculate strategy were found, $F(1, 20) = .54$, $MSE = .02$, $p = .47$ and $F(1, 20) = 2.27$, $MSE = .02$, $p = .15$, respectively.

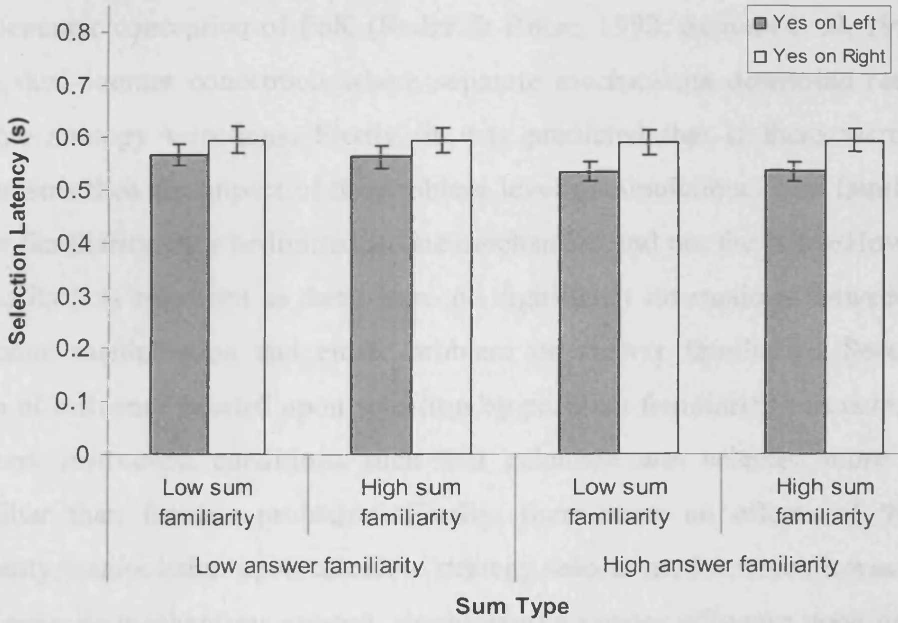

Figure 3.6: Strategy selection latencies in each condition split by the position of the yes key. Grey bars represent latencies derived when the key was positioned on the left of the keyboard, white bars when the yes key was positioned on the right. Error bars represent the standard error of the mean.

The same pattern of results emerged in respect to the time taken to select the retrieve strategy. There were null effects of sum familiarity $F(1, 13) = 3.16$, $MSE = .001$, $p = .1$, and marginally significant effects of answer familiarity, $F(1, 13) = 4.93$, *MSE* = .002, p = .05. Similar to the duration of calculate strategy selections, faster selections were made for problems with more familiar answers than problems with unfamiliar answers (see Figure 4.6).

3.5.3 Discussion

The findings reported in this study fail to provide any evidence to suggest that task instructions influence rapid retrieve or calculate strategy selections. This supports the notion that task instruction effects are realised through consciously directed processes following the conclusions drawn from Kirk and Ashcraft's (2001) verbal protocol study. Furthermore, the positioning of the yes and no keys on the keyboard did not serve to bias selection responses either. In respect to the second issue examined in this study, three findings converge upon the notion that the single-counter hypothesis determines strategy selections in this experimental paradigm. This serves to support the predictions of the SAC account of selection which favours a single-counter conception of FoK (Reder & Ritter, 1992; Schunn et al, 1997) rather than a dual-counter conception where separate mechanisms determine retrieve and calculate strategy selections. Firstly, it was predicted that if there were separate mechanisms then the impact of the problem level manipulations, sum familiarity and answer familiarity, may be limited to one mechanism and not the other. However, this eventuality was ruled out as there were no significant interactions between the task instruction manipulation and either problem or answer familiarity. Secondly, the pattern of influence exerted upon selection by problem familiarity was comparable in both task instruction conditions such that calculate was selected more often for unfamiliar than familiar problems. Finally, there were no effects of the answer familiarity manipulation upon calculate strategy selections, for which it was predicted that if separate mechanisms existed, would exert a greater influence upon participants in the retrieve yes/no condition than those in the calculate yes/no condition.

Before considering the theoretical ramifications of the two key outcomes revealed in the present experiment in greater detail, in the following experiment the same task instruction manipulation was used. To examine whether the susceptibility of the selection mechanism to task instructions manipulation is moderated by the types of problems presented (i.e., problems in which problem familiarity effects are evident, or problems in which problem feature effects are evident) the sum type manipulation (i.e., decades, mixed and fives problems) was employed in place of the sum familiarity and answer familiarity manipulations. Furthermore, manipulation of the task instruction variable facilitated a further investigation of the single- versus dual-counter hypothesis of FoK.

3.6 EXPERIMENT 6b

In the previous experiment it was shown that task instructions emphasising selection of either the retrieve or calculate strategy had no impact upon predicted strategy selections. Furthermore, there was no evidence to suggest that a dual-counter account of the FoK mechanism provided a better fit to the pattern of results than the single-counter account. In the present experiment the same two issues detailed in the previous experiment were examined. To briefly recapitulate, the task instruction variable (i.e., retrieve yes/no vs. calculate yes/no) was employed to investigate what types of task-level manipulation influence the strategy selection mechanism. Findings from a mental arithmetic study (e.g., Blöte, Van der Burg & Klein, 2001; Kirk & Ashcraft, 2001) revealed that self-reported measures of strategy use were susceptible to the influence of task instructions (see also Gardner & Rogoff, 1990, for similar result but in a maze navigation study). Specifically, task instructions emphasising usage of the retrieval strategy elicited a high percentage of self-reported retrieval selections made after solving the problem. Similarly, instructions emphasising the use of calculate strategies produced a bias in responses to selection of the calculate strategy. However, null effects of task instructions upon selection in Experiment 6a serve to question whether the selection mechanism is indeed responsive to manipulations of the task (i.e., task instructions), rather than the problem itself (i.e., problem familiarity or problem features). One possible explanation for this null effect resides in the difference between the measures of strategy selections obtained in this experimental paradigm and those produced in the Kirk and Ashcraft (2001) study. In the selection-phase experimental design, responses are produced rapidly and there is no opportunity afforded to use conscious processes to identify what strategy was actually used to solve the problem. As shown by the articulatory suppression manipulation employed in Experiments 5a and 5b, the selection latencies and the percentage of late responses illustrate that the selection mechanism itself is apparently immune to interference from secondary tasks. Conversely, the self-report responses collected by Kirk and Ashcraft (2001) are contingent upon a conscious evaluation of a prior processing episode. Here the participant must identify what strategy they actually used to solve the problem and in the eventuality that information is not accessible, guess the most likely candidate (Kirk & Ashcraft, 2001). Accordingly, it may be the case that the influence of the task instruction manipulation is realised through the contribution of conscious processes used to reconstruct a prior processing episode, rather than the unconscious processes shown to drive strategy selections in previous experiments presented in this thesis.

In respect to the second facet of selection examined in Experiment 6a, the issue of whether a single-counter or dual-counter account of the FoK mechanism provides the best description of the findings, the same rationale was applied in this study. If selection-by-feature effects, which have been shown to influence selection in decades, mixed and fives problems (see also Experiments 2a, 2b and 5b), are only evident in one of the task instruction conditions (i.e., retrieve yes/no or calculate yes/no) this would be indicative of the operation of separate counters. One counter determining retrieve selections based upon information in memory pertaining to the accumulated knowledge of things that are known. A second counter monitoring knowledge pertaining to things that are not known and resulting in calculate strategy selections.

To examine these two issues, the same selection-phase design was employed here as in Experiment 6a including a between subjects manipulation of task instruction (i.e., retrieve yes/no vs. calculate yes/no) and key position (yes key on the left vs. right side of the keyboard). However, a further hypothesis, that only certain types of problem are sensitive to the task instruction was also examined. Whereas in the previous experiment problem familiarity effects were evident and task instructions did not influence selection, in this study the sum type manipulation (i.e., decades, mixed and fives problems) was employed to identify whether strategy selections influenced by the selection-by-feature effect (rather than problem familiarity) are sensitive to task instructions.

3.6.1 Method

3.6.1.1 Participants

Twenty-four participants from the School of Psychology at Cardiff University were given course credit or payment in return for their participation. All were native English speakers reporting normal hearing and corrected or normal vision and had not participated in any of the other thesis experiments.

3.6.1.2 Materials, Design & Procedure

The same type of stimuli were used as in Experiments 2a and 2b (decades, mixed, fives problems). To recapitulate; both double-digits addends in each problem were either divisible by 5 or 10. Problems in the decades condition contained two addends which were divisible by 10 into integers, problems in the fives condition comprised two addends divisible into integers by 5. Problems in the mixed condition contained one decade and one fives addend. No tied problems were included and all problems summed to less than 100 (see Appendix E, table E7, for stimuli). Participants completed 16 practise questions followed by 12 trials in each sum type condition, the order of which was pseudo-randomised for each participant.

As in the previous study, two further manipulations were run as between subjects variables; task instruction (retrieve y/n vs. calculate y/n) and yes button position (left vs. right of keyboard). Full exposition of the methodology and procedure is provided in the corresponding section presented in the previous experiment.

3.6.2 Results

3.6.2.1 Scoring Procedure

All statistical analyses, unless detailed to the contrary, were conducted using a 2 (task instruction; retrieve y/n vs. calculate y/n) x 2 (yes key position; left vs. right) x 3 (sum type; decades, mixed and fives) mixed measures ANOVAs. Within this design sum type was manipulated within subjects, selection task instruction and yes key position between subjects. To briefly recapitulate, retrieve yes and calculate no selections both equated to the same underlying strategy selection (i.e., retrieve), calculate yes and retrieve no selection to the calculate strategy. To minimise confusion, in the analysis, selection responses will be identified by the underlying strategy selection (retrieve or calculate) and will be split by the task instruction condition (calculate y/n, or retrieve y/n).

3.6.2.2 Strategy Selection

In accordance with findings detailed in prior studies employing the sum type manipulation, the underlying strategy selected most often was the retrieve strategy (see Figure 4.7). A mixed measures ANOVA was conducted, analysing the percentage of late responses in each of the conditions. Null effects of sum type, task instruction and yes key position (all Fs $< .7$, all ps $> .14$) upon the percentage of late responses revealed that the dependant relationship between retrieve and calculate selections was not compromised by the experimental variables.

It was found that there were no significant interactions between any of the manipulations (all Fs < 1.1, all ps > 3.7). However, there was a significant effect of sum type on the percentage of retrieve selections recorded (i.e., retrieve yes and calculate no selections), $F(2, 19) = 14.58$, $MSE = 4.76$, $p < .001$. As Figure 4.7 illustrates and pairwise comparisons confirm, the retrieve strategy was selected more often in the decades than mixed ($p = .013$), decades than fives ($p < .001$) and mixed than fives ($p < .001$) conditions. This replicates the pattern of influence exerted upon selection by the sum type manipulation shown in Experiments 2a, 2b and 5b. Turning to the effects exerted on selection by the task instruction manipulation, significant main effects were evident, $F(1, 20) = 5.4$, $MSE = 12.6$, $p = .03$. Participants indicating whether they would retrieve the answer from memory (yes vs. no) made a greater number of retrieve selections than those who were asked whether they would calculate the answer (p = .03). Null effects of the key position manipulation serve to indicate that selection of the yes or no response was not contingent upon the spatial location of the yes key on the keyboard, $F(1, 20) = .04$, $MSE = 2.62$, $p = .84$.

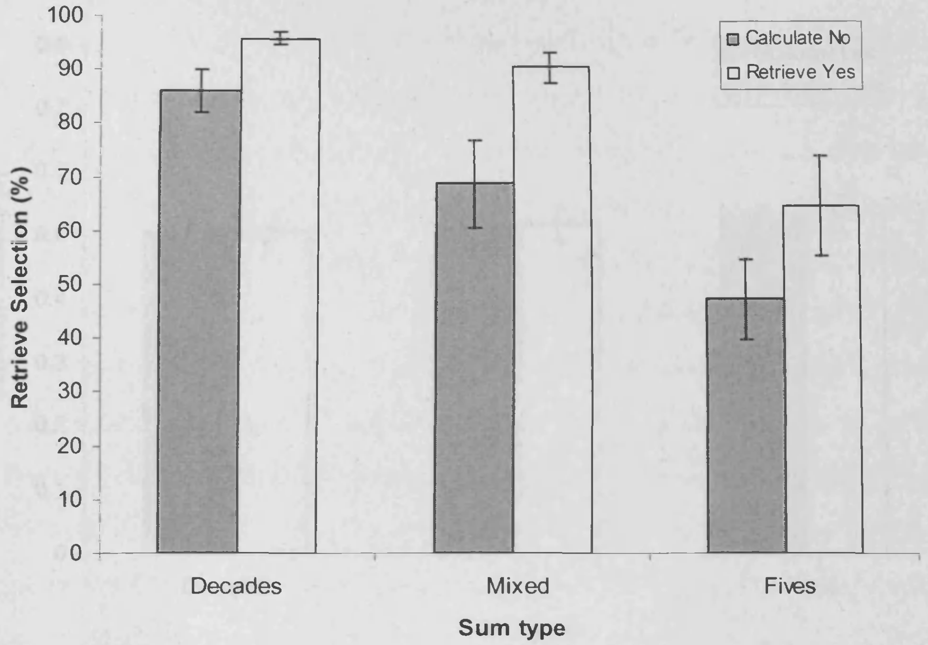

Figure 3.7: The percentage of retrieve strategy selections made in each of the sum type conditions. Grey bars represent responses to the selection question 'calculate y/n' where the participants chose 'calculate no' (i.e., to use the retrieve strategy). White bars represent responses when participants were asked 'retrieve y/n' and selected 'retrieve yes'. Error bars represent the standard error of the mean.

3.6.2.3 Strategy Selection Latencies

In respect to the rapidity with which strategy selections were made, overall, a total of only 6.02% of selections were not made within the 850 ms time limit. Marginally significant effects of sum type were evident upon the time taken to select the retrieve strategy, $F(2, 18) = 3.33$, $MSE = .002$, $p = .06$. This narrowly fails to replicate findings from previous experiments presented in this thesis which employed identical stimuli in which significant effects of sum type upon retrieve selection latencies were reported. Despite the significant effects of task instructions upon retrieve selections there was no effect of this manipulation upon the time taken to select retrieve strategies, $F(1, 19) = .18$, $MSE = 12.59$, $p = .68$. Nor was if found that the spatial location of the yes key position influenced mean retrieve selection latencies, $F(1, 19) = .09$, $MSE = 12.59$, $p = .77$, see also Figure 4.8.

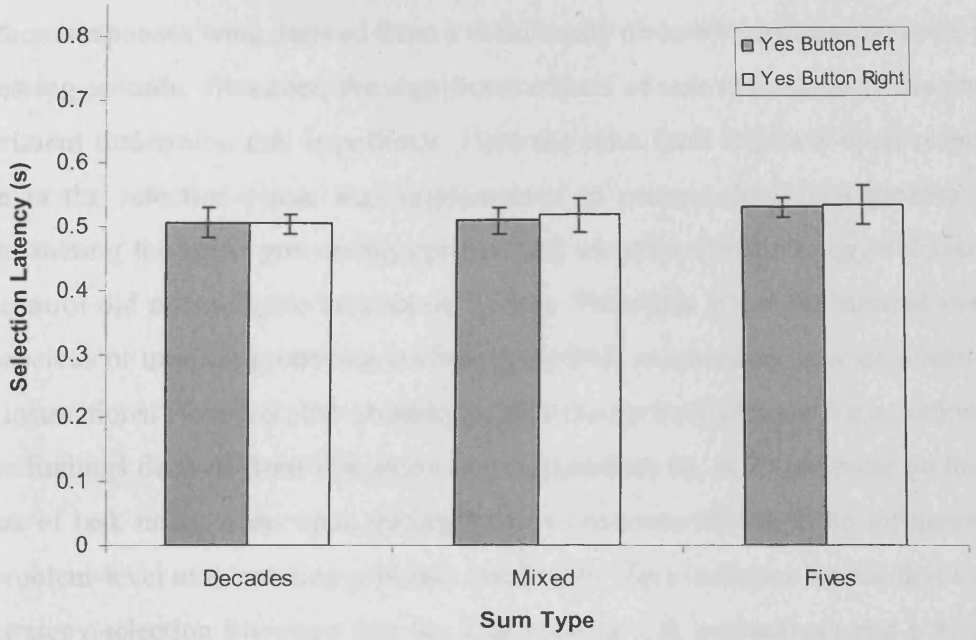

Figure 3.8: Strategy selection latencies in each sum type condition split by the position of the yes key. Grey bars represent latencies derived when the key was positioned on the left of the keyboard, white bars when the yes key was positioned on the right. Error bars represent the standard error of the mean.

3.6.3 Discussion

In summary, it was found that task instructions, dissimilar to Experiment 6a, influenced the process that determines retrieve strategy selections. Participants instructed to predict whether they would retrieve the answer (yes or no) selected the retrieve strategy on a greater percentage of trials than those participants asked to indicate whether they would calculate the problem's solution. In respect to the second issue examined in the present experiment, there was no evidence to support the predictions of the dual-counter account of FoK, where separate counters measure the accumulation of information pertaining to what is known in memory and what is not known (see Nelson & Narens, 1990). Turning first to the effect task instructions upon strategy selection, the findings reported in the current study support Cary and Reder's (2002) conception of strategy selection in that selection was not only influenced by problem-level manipulations (i.e., problem familiarity and/or problem features) but also task-level (i.e., task instructions) manipulations. In the previous study the null effects of task instruction were attributed to the notion that conscious processes were responsible for their effect. As previously identified, instruction effects were reported in the self-report selections made by participants in Kirk and Ashcraft's (2001) study and these responses were derived from a consciously directed reconstruction of a prior processing episode. However, the significant effects of task instruction in the present experiment undermine this hypothesis. Here the time limit imposed upon responses made in the selection-phase was implemented to prevent conscious process from reconstructing the prior processing episode and identifying the strategy selected but this control did not mitigate instruction effects. From this it can be inferred that the unconscious or implicit processes, including the FoK mechanism, may be sensitive to task instructions. However, the ubiquity of this theory is challenged by a comparison of the findings derived from this study and Experiment 6a. In Experiment 6a the null effects of task instructions upon strategy selections were shown to be influenced by the problem-level manipulation problem familiarity. This indicates for the first time in the strategy selection literature that the influence of task instructions and potentially other task-level manipulations upon selection is dependant upon the type of problem.

To date, selection models, in particular the SAC account, have focused upon modelling problem-level effects, rather than task-level manipulations. Accordingly, there is very little specification of how task-level manipulations actually influence selection. Further discussion of this issue and the implications of this finding are presented in further detail in Chapter 4.

3.7 GENERAL DISCUSSION

Whereas in Chapter 2 the empirical work focussed primarily upon two objectives, establishing key empirical phenomena within the paradigm and evaluating the robustness of the dual-phase design, in Chapter 3 a more detailed examination of the selection mechanism is presented. To do so, five experiments examined three key issues which have received little, or no, attention within the arithmetic strategy-selection literature to date. In Experiments 4, 5a and 5b it was shown for the first time in a controlled empirical setting that the influence exerted upon selection by problem-level manipulations is contingent upon the type and number of manipulations operationalised. Furthermore, that underpinning the retrieve and calculate selections made in these experiments a common selection mechanism was in operation. However, it was also revealed that the influence of selection-by-feature effects upon retrieve and calculate selections may be inconsistent. Although in Experiments 6a and

6b some evidence is presented to suggest that problem feature effects are contingent upon not only the type of problem feature/s manipulated, but also the task-level conditions imposed upon the processing episode. To examine whether this outcome could be attributed to action of the selection mechanism itself or to other cognitive functions, the wider context in which problems were solved was also examined in Experiments 5a, 5b, 6a and 6b. From these experiments it was found that although selections were made rapidly, they were still sensitive to other influences, including the availability of consciously directed processes and task instructions. The findings from these experiments emphasise limitations inherent in the familiarity- (problem or answer) based selection models detailed in Chapter 1, as the influence of problem- and task-level factors cannot be dissociated from the selection process. A brief outline of the key findings from each of the experiments presented in this chapter will follow. Particular emphasis will be made to the key empirical phenomena identified in Chapter 2 and the limitations of existing models of strategy-selection.

Results from Experiments 5b and 6b in this chapter serve to support the proposed selection-by-feature effect originally identified in Chapter 2. In Experiment 2a it was concluded that particular features inherent in problems, for example, *both addends in a sum are divisible by 10,* influence the strategy selections mechanism. As evidence of this effect stemmed from one set of stimuli (i.e., decades, mixed and fives problems), although replicated in the selection-phase design, it was necessary to identify the boundaries of this effect. As decades, mixed and fives problems elicited a high percentage of retrieve strategy selections it was considered whether the influence of the selection-by-feature effect is limited to retrieve strategy selections. Furthermore, decades, mixed and fives problems are all relatively familiar problems in comparison to those presented in Experiments 1a and 1b where sum and answer familiarity was manipulated. It may also be the case that the selection-by-feature effect is limited to familiar as opposed to unfamiliar problems. Findings from Experiments 5a demonstrate that selection-by-feature effects can influence the percentage of calculate selections made in unfamiliar problems and are thus not limited to either retrieve selections, or highly familiar problems. From this it is inferred that problem features may influence strategy selection in a range of problem types.

As well as being informative of the ubiquity of the selection-by-feature effect, findings from Experiments 5a and 5b also confirmed that this effect is based upon problem features rather than problem familiarity. In Experiments 2a and 2b, effects of the problem feature manipulation sum type were found in the percentage of retrieve selections reported. However, a covariates analysis demonstrated that problem familiarity covaried significantly with retrieve selections. The import of this finding was undermined by a further analysis run in Experiment 2a which revealed that problem familiarity effects were not present within each level of the sum type manipulation. Confirming this account of the selection-by-feature effect, Experiment 5a employed a design with the power to dissociate the influence exerted upon selection by problem features and problem familiarity. A significant interaction between these two manipulations, such that an effect of the problem feature addend status was only evident in familiar (as opposed to unfamiliar) problems confirmed that the selection-by-features effect is distinct from that of problem familiarity.

As well as establishing the scope of the selection-by-feature effect the experiments in this chapter also provided an opportunity to examine how problem-level manipulations combine to influence selection. As detailed previously, the majority of strategy-selection models stipulate that retrieve and calculate selections are determined by a single factor. However, in reality, problems may be comprised of a number of factors with the potential to influence selection, including problem familiarity and problem features. In Experiment 4 two feature manipulations were employed, addend status (i.e., odd, even and mixed problems) and relative addend magnitude (i.e., similar and disparate). As neither feature was found to influence selection it could be inferred that the null effects of both feature manipulations upon selection could be attributed to the action of one manipulation nullifying the effect of the other. A further eventuality considered was that the complexity of the 2 (relative addend magnitude) x 3 (addend status) factorial design used failed to render either feature apparent to the participant. However, this finding needs to be considered with some caution. The relative addend magnitude manipulation, designed to investigate whether number comparison or anchoring processes were used when identifying features in a problem, was not examined in any of the other experiments presented in this thesis. Nor had this type of manipulation been employed in any other selection tasks using a similar experimental paradigm. Accordingly, it may not be a feature that actually influences selection. In contrast, when comparing the influence exerted by two problem-level manipulations, problem familiarity and addend status, in Experiment 5a a significant interaction was found. This suggests that it is not necessarily the number of problem-level factors inherent in a problem that limits their impact upon selection, but the type of factor. To illustrate, in Experiment 5a the interaction between addend status and problem familiarity revealed that selection in relatively familiar problems was influenced by addend status, but not in relatively unfamiliar problems. More importantly, from the interaction between these two factors, it is proposed that selection is not influenced by one factor at the exclusion of another, but that the influence exerted by one factor is moderated by the other.

Turning to the third issue examined within this chapter, that of the role played by the context in which a strategy-selection is made, the particular focus of this issue was in identifying how context contributes to the selection-by-feature effect. When examining how problem familiarity and selection-by-feature effects occur it is necessary to identify the processes responsible for generating differentials in problem familiarity and responsible for identifying problem features relevant to selection. Turning first to problem familiarity effects, the SAC model provides an apparently sound foundation for appreciating how problem familiarity influences selection. As detailed previously, briefly stated, problem familiarity is determined by the frequency of exposure to problems, or specific components of problems (Reder & Ritter, 1992; Schunn et al, 1997), whereby greater exposure confers greater familiarity. Supporting this account, in Experiment 1a calculate selections were shown to be sensitive to problem familiarity. However, in Experiment 1b, it was evident that problem familiarity effects upon selection were determined by problem familiarity ratings derived prior to the experiment. Problem familiarity effects upon calculate strategy selections were shown to be immune to two task-level manipulations in Experiments 5a and 6a. Articulatory suppression, a secondary task designed to inhibit conscious processes not only during but between strategy selection trials, failed to attenuate problem familiarity effects. Furthermore, in Experiment 6a, the manipulation of task instructions failed to impair problem familiarity effects. Perhaps more importantly, there was no interaction between problem familiarity and the task instruction given to participants. From this it is apparent that effects of problem familiarity are not only robust, but immune to interference from conscious processes and attempts to bias strategy selections using task instruction manipulations.

Conversely, from the experiments presented in the present chapter, selection-by-feature effects were shown to be sensitive to both problem- and task-level manipulations. Findings detailed in Experiment 2a, demonstrate that retrieve selections were influenced by the type of problem presented, decades, mixed or fives. While Experiment 2b demonstrated that the problem features specific to this set of problems are not necessarily identified by the act of solving the problems. Based upon this foundation, in Experiment 5b it was apparent that selection in decades, mixed and fives problems was immune to interference from articulatory suppression. It is concluded that the problem features responsible for the selection-by-feature effect are derived from prior problem-solving episodes and are thus immune to suppression effects, similar to problem familiarity. However, when examining a further type of problem feature, addend status, it was found in Experiment 5a that Articulatory suppression did attenuate the selection-by-feature effect. Accordingly, it is suggested that selection-by-feature effects can be determined by problem features either derived from prior problem-solving episodes, or from features identified during the course of an experiment. Generally stated, it the influence exerted upon selection by problem features is much more sensitive to task-level manipulations than that of problem familiarity.

CHAPTER FOUR

4.1 AIMS OF THE THESIS

Although strategy selection is a key component of the problem-solving process, the level of attention paid to the mechanism responsible for rapid and accurate performance fails to reflect the importance of this cognitive function. To redress this neglect I have endeavoured to demonstrate that examining strategy selection in arithmetic problem-solving is a useful way to observe the operation of not only the selection mechanism but the processes that monitor and control knowledge in general. In this final chapter I review the findings derived from the key themes underpinning the two empirical series presented in this thesis. Broadly speaking, the principle aims of the thesis are threefold. Reflecting the paucity of empirical research in the domain, the opening remit of the thesis was to establish a robust and flexible methodology which could be used to examine the selection mechanism. Addressing the limitations of existing simulations of strategy selection the key tenets of these models were contrasted from a largely exploratory standpoint. Then the findings from these experiments were used to direct further empirical investigation designed to extend and refine key aspects of the selection mechanism itself. By identifying and elucidating critical facets of the selection process it is hoped that a greater understanding of the process by which problems are solved is achieved.

4.2 SUMMARY OF FINDINGS: ESTABLISHING THE KEY EMPIRICAL PHENOMENA

Despite the number of existing strategy selection models the lack of empirical research designed to establish key effects within the paradigm is compelling. Exploiting the dual-phase experimental design developed by Reder and colleagues (Reder, 1987; Reder & Ritter, 1992; Schunn et al, 1997) the empirical series presented in this thesis took a more holistic approach to selection than is evident in existing accounts. By examining the type of factors that influence selections in real-world problem-solving, taken together, the experiments reported in this thesis indicate that selection is more flexible than previously conceived. Specifically, a range of problem-level (i.e., problem familiarity and problem features) and task-level manipulations (i.e., task instructions and articulatory suppression) were shown to influence retrieve and calculate selections. As many of the issues examined in this thesis have not been covered within the arithmetic strategy-selection literature previously, in the following sub-sections I shall detail each of the key issues examined and their implications for the strategy selection process.

4.2.1 Testing the Dual-Phase design

At the outset it was necessary to establish a robust methodology with the power to examine the predictions derived from existing models of selection. Reflecting limitations in the scope of prior empirical research, the findings supporting existing models of selection were derived from a range of arithmetic tasks. These included pseudo-and alphabetic-arithmetic tasks (Compton & Logan, 1991; Rickard, 2004; Rickard & Bourne, 1996; Rickard, Healy & Bourne, 1994), arithmetic inversion problems (Siegler & Araya, 2005) and in the case of the ACT-R, ASCM and DOA, a conventional arithmetic problem-solving experiment. In a series of experiments Reder and colleagues (Reder, 1987; Reder & Ritter, 1992; Schunn et al, 1997) demonstrated that the dual-phase Game Show design provided a suitable methodology with which to investigate the operation of the selection mechanism. Furthermore, within this design there was the potential to integrate a range of problem- and task-level manipulations.

The design itself is contingent upon a clear distinction between selection- and solution-phases and that this demarcation mirrors the natural order of processing in problem-solving. To test this assertion it was necessary to ensure that the outcomes from the design demonstrated that strategy selection could be made rapidly and with a degree of accuracy. Furthermore, that the modifications made to the Reder and colleagues' (Reder & Ritter, 1992; Schunn et al, 1997) dual-phase design did not compromise the rapidity and accuracy of selections. In the selection-phase of the design a time limit of 850 ms was imposed upon selection to force rapid strategy selections. In a simple mental arithmetic problem-solving study Staszewski (1988) had demonstrated that at least 850 ms is required for a direct retrieval of a problem's answer to be completed. Results from the Game Show studies had already illustrated that selections could be made rapidly (i.e., 580 – 760 ms). Based upon the modifications made to the dual-phase design used by Reder and colleagues (Reder & Ritter, 1992; Schunn et al, 1997) it was necessary to replicate these effects to ensure the validity of the design. Overall, in the experiments reported in this thesis it was found that mean selection latencies ranged between 416 – 638 ms and the mean percentage of late responses between 3.9% - 12.2%, indicating that participants were able to select the retrieve or calculate strategy well within that time limit.

Not only could selections be made rapidly but it was also evident that predicted strategy selections were largely accurate. In related experimental paradigms, where predictions of future performance are based upon memory monitoring mechanisms, the accuracy of these mechanisms is deemed to be questionable. However, it should be noted that in the metacognition literature, apparent inaccuracies in Judgements-of Learning and Ease-of-Learning (see Koriat & Ma'ayan, 2005; Koriat, Ma'ayan & Nussinson, 2006; Nelson & Narens, 1996; Nelson & Dunlosky, 1991) are often attributed to a lack of understanding of how these mechanisms function (Koriat, 2006; Koriat, Ma'ayan, Nussinson, 2006; Smith, Shields & Washburn, 2003) rather than inaccuracy *per se*. To assess the accuracy of predicted strategy selections, Reder and Ritter (1992) using *d'* (see Macmillan & Creelman, 2004) and the Goodman-Kruskal gamma correlation (see Nelson, 1986), compared the predicted strategy selection in a trial to the time taken subsequently to solve that problem. By subjecting the strategy selections and solution latencies to the incentive scheme used in those studies hits, misses, false alarms and correct rejections were identified. To illustrate, a hit was identified when the participant chose the retrieve strategy and correctly answered the problem within the time limit allocated to retrievals. False alarms were allocated to trials in which retrieve was chosen but the accompanying solution latency exceeded the time limit within the scheme for retrievals. As the focal remit of this thesis was to identify the factors that influence selection, rather than establishing the accuracy of these selections, based upon the good levels of accuracy established in Reder and colleagues' (Reder & Ritter, 1992; Schunn et al, 1997) studies it was deemed appropriate to take a more simplistic approach to accuracy. Furthermore, by jettisoning the incentive scheme used by Reder and colleagues to stimulate performance for both theoretical and empirical ends (see section 4.2.2), this type of analysis was not possible. Accordingly, the analysis of accuracy was simplified such that retrieve selections should be accompanied by significantly shorter selection latencies than calculate selections. In general it was found that solution latencies were representative of the time that should be taken by either the retrieve or calculate strategy to solve that problem. Again, it was deemed unnecessary to quantify these time periods, as individual differences dictate that solution latencies vary considerably between individuals (Campbell & Xue, 2001; Hecht, 1999; Jackson & Coney, 2007; Lefevre & Kulak, 1994; LeFevre, Kulak & Bisanz, 1991; Little & Widaman, 1995) and that there was no one-fits-all time limit that could be used effectively as a cut-off between retrieve and calculate selections.

In respect to the modification made to Reder and colleagues' (Reder & Ritter, 1992; Schunn et al, 1997) dual-phase design, one notable change was made, demonstrating that performance attributed to the selection- and solution-phases by Reder and colleagues was indeed dissociable. In Experiments 1b, 2b, 3 either the selection-phase or solution-phase were performed in isolation. Findings from Experiments 1b and 2b replicated the results from Experiments 1a and 2a which used the dual-phase design. Also Experiment 3 replicated the pattern of influence exerted by the problem-level manipulations upon solution latencies evident in Experiments 2a and 2b. Therefore it was concluded that the experimental design had the suitable flexibility with which to examine a range of problem- and task-level influences. Also, by replicating the effects found in the dual-phase design, the single-phase designs not only testified to the robustness of the methodology but supported the Reder and colleagues' assumption that performance arising from the selection- and solution-phases of problem-solving is dissociable (Reder & Ritter, 1992; Schunn et al, 1997).

4.2.2 Problem Familiarity as a cue to strategy selection

The problem familiarity effect stands as one of the most controversial, but important findings within the paradigm to date. Considering the lack of empirical research and reliance on modelling techniques within the paradigm, the effect is relatively well established at an empirical level. Reder (1987) illustrated that by priming the terms of a general knowledge question, and hence increasing the familiarity of these primed terms, that individuals are more likely to select the retrieve strategy than calculate. Within the Game Show studies, using a similar technique, the same finding was revealed (Reder & Ritter, 1992; Schunn et al, 1997) in mental arithmetic problems. To recapitulate, in all but one of these studies (Reder & Ritter, 1992; Experiment 1) retrieve and calculate selections were examined in double-digit multiplication or sharp problems. When first presented with these problems, due to their difficulty (e.g., 23 x 46), participants did not make predicted retrieve selections nor were they able to solve them using direct retrieval procedures. With repeated exposure to the problems and repeated solving of these problems, participants became familiarised with the problems and returned a greater frequency of predicted retrieve selections. However, there was a crucial limitation in Reder and colleagues' studies (see Reder & Ritter, 1992; Schunn et al, 1997). The priming methodology and incentive scheme used to reinforce learning served to confound selections for double-digit addition problems in particular. Rather than basing selections upon the actual familiarity of the problem, participants were attempting to beat the incentive scheme, selecting retrieve almost exclusively to maximise their score (Reder & Ritter, 1992; Experiment 1). A further consideration examined was that the use of an incentive scheme and priming technique may be unnecessary as individuals are already familiarised to a certain extent with a vast range of problems. In particular, it may be that this artificial method of inducing familiarity may not mirror the same process used in real-life problem-solving. Accordingly, a different approach was adopted in the experiments presented in this thesis. Based upon number familiarity ratings, problem familiarity ratings were derived and manipulated in a number of experiments. For the first time in the literature, the experiments presented in this thesis converge upon the notion that pre-experimental familiarity, rather than experimentally induced familiarity, also influences retrieve and calculate strategy selections. This supports an account of real-world problem-solving such that prior exposure to problems serves to increment problem familiarity and also influences selection itself. In Experiments 1a and 1b, the retrieve strategy was selected more often in relatively familiar problems than unfamiliar problems. Regression analyses conducted in Experiment 1a (see Appendix B) suggest that instead of problem familiarity the magnitude of a problem's answer may influence selection. However lines of converging evidence provide greater support to the problem familiarity account of selection. Further experimentation, directly manipulating answer magnitude, would be required to rule out the potential contribution of this measure to selection conclusively. Problem familiarity effects were also evident in Experiments 5a and 6a; these will be discussed in greater detail in sections 4.2.4 and 4.2.5. Taken in tandem with the findings from the Game Show studies the problem familiarity effects evident in this thesis demonstrate the ubiquity of this effect. Perhaps more importantly, they illustrate that selection is not only influenced by experimentally induced problem familiarity but extant problem familiarity, derived from a lifetime of problem-solving.

4.2.3 The Selection-by-feature effect

One issue that has received very limited coverage in the existing models of strategy selection is the type of problem-levels factors that influence the selection mechanism. Aside from problem familiarity, only one model, SCADS* (Siegler & Araya, 2005) has sought to accommodate a range of influences upon selection. When examining double-digit addition and multiplications problems, using the dual-phase design, Reder and Ritter (1992) identified that strategy selections were not only influenced by problem familiarity but by a particular problem feature, operator type (i.e., + or x). Whereas the accuracy of selections based purely upon problem familiarity was shown to be high in other studies (see Reder & Ritter, 1992; Schunn et al, 1997) the accuracy of retrieve selections in addition problems was relatively low in comparison to multiplication problems. Participants often selected retrieve but were unable to solve the problems within the time limit set by the incentive scheme. Low levels of accuracy can in part be attributed to the influence of the incentive scheme. Participants realised that the greatest pay-off opportunity was to make a predicted retrieve selection and attempt to solve the problem as rapidly as possible. But the authors also acknowledge that selection was influenced by the type of operator rather than the familiarity of the problem.

Such problem feature effects have been accommodated by ad-hoc assumptions bolted on to account for their influence (i.e., ACT-R and SAC). However, the possibility that a range of features influences selection and that their effect is contingent upon a common mechanism was considered in the present thesis. Accordingly, three different types of problem feature were examined in the experiments reported in this thesis to examine the scope of these effects. In Experiments 2a and 2b, retrieve strategy selection were influenced by the problem feature sum type in that participants utilised features such as *both addends are decades numbers*, or *both addends are fives numbers*. Although problem familiarity was not manipulated in the current study, a covariates analysis revealed that the familiarity of the problem covaried significantly with retrieve selections. Further analyses indicated that within each level of sum type there were no effects of problem familiarity but the possibility remained that problem familiarity effects were misattributed to the problem feature. In Experiment 4, the two problem features manipulated, addend status and relative addend magnitude, had no impact upon selection. Effects of addend status were expected as the manipulation was analogous to the sum type manipulation. However, the relative addend manipulation required a magnitude comparison between both addends in a problem, rather than the identification of a particular feature, such as *both addends are even numbers*. Subsequent effects of addend status in Experiment 5a, albeit with 2 rather than 3 levels and paired with a manipulation of problem familiarity rather than another problem feature manipulation, demonstrated that operator type and sum type are not the only problem feature manipulations to influence selection.

From these studies it is possible to build upon a basic understanding of the types of feature that influence selection. Effects of sum type and addend status demonstrate that participant's utilised semantic features of a problem, such as *both addends are even numbers,* or *both addends are multiples of five.* Null effects of the relative addend magnitude manipulation indicate that magnitude comparisons between the two addends did not influence selection. To confirm this notion it would be necessary to examine the influence of the relative addend magnitude upon selection in isolation, or alternatively paired with a manipulation of problem familiarity (see 4.2.4). Furthermore, whereas sum type effects were realised in highly familiar problems which elicit a high percentage of retrieve selections (i.e., decades, mixed and fives), effects of addend status in relatively unfamiliar problems which predominantly elicit calculate selections indicate the generality of the selection-by-feature effect to a range of problems, problems varying in familiarity and strategy selections.

4.2.4 Cue combination: When problem familiarity and problem features collide

In most mental arithmetic problems there will be a number of cues with the potential to influence selection. This is especially likely in more complex problems such as general knowledge questions (see Reder, 1987). Irrespective of the number of cues, or the complexity of the problem, individuals still must be able to make rapid and accurate strategy selections. For the first time in the paradigm, in a number of experiments, multiple problem-level manipulations were employed to examine whether one manipulation takes precedence over the other, or whether selections are influenced by an interaction between the factors. To do so, three configurations of problem-level manipulation were examined. Experiments 1a, 1b and 6a demonstrated that when two familiarity manipulations were employed (i.e., problem and answer familiarity) problem familiarity took precedence. However, this needs to be treated with some caution as a number of studies have demonstrated that answer familiarity does not influence rapid strategy selections (Reder & Ritter, Schunn et al, 1997, cf. the ACT-R, ITAM and DOA models). To confirm whether problem familiarity takes precedence over answer familiarity it would be necessary to examine answer familiarity effects when manipulated in isolation. Two problem features (addend status and relative addend magnitude) were manipulated in Experiment 4. The null effects of these manipulations were primarily attributed to the complexity of the factorial manipulation used in that experiment where 6 different configurations of problems were required. Accordingly, participants may have been unable to identify either or both of the manipulations, alternatively selection may have been confounded by the influence of both manipulations, indicating that they are unable to interact. Supporting this notion, in Experiment 5a it was evident that selections are influenced by both problem familiarity and problem features (i.e., addend status). More importantly, a significant interaction between these manipulations suggests that the influence of each manipulation is conditional. To illustrate, addend status effects were only evident in relatively familiar problems.

From these three configurations of problem-level manipulations it is apparent that problem familiarity is a particularly robust effect in the presence of other familiarity or feature based manipulations. Conversely, in the presence of other feature or familiarity manipulations problem feature effects appear to be more sensitive as illustrated by the complexity effect.

4.2.5 Selection in context

While the focus of the review presented in this chapter so far, and indeed the empirical work undertaken within the paradigm, has centred upon how problem-level manipulations influence selection, four experiments examined how context influences selection. In the distinction drawn between intrinsic and extrinsic influences upon strategy adaptivity, Cary and Reder (2002) proposed that selection is sensitive to not only problem-level manipulations but task-level manipulations. Following these authors' assertion that conscious awareness is not a pre-requisite for strategy adaptation, in Experiments 5a and 5b the contribution of conscious processes to selection was examined using a secondary task. In Experiment 5a articulatory suppression, used to inhibit consciously directed action during the task, interacted with problem familiarity. Generally, selection in the present experiment was characterised by an interaction between addend status and problem familiarity such that addend status effects were only apparent in familiar (as opposed to unfamiliar problems). However, this interaction between problem familiarity and suppression was particularly revealing as under suppression, effects of problem familiarity were re-instated demonstrating that addend status effects were attenuated by the suppression task. From this it is evident that consciously directed processes are at least in part responsible for addend status effects. It was predicted and shown that the action of the selection mechanism itself was immune to suppression effects following the notion that selection is based upon implicit or unconscious processes such as the FoK mechanism (Reder & Ritter, 1992; Schunn et al, 1997). Accordingly, it is suggested that the locus of suppression effects upon addend status can be localised to the consciously directed processes operating during the course of the experiment. These were presumably responsible for the identification of the problem feature, rather than inhibiting the action of the selection mechanism directly. In Experiment 5b sum type was manipulated to replicate the action of suppression on the selection-by-feature effect but using a different problem feature, sum type. However, strategy selections made for decades and mixed problems were immune to suppression effects. Extending the earlier rationale, it was inferred that consciously directed processes were not required for the identification of these problem features. However, selection in the fives condition was influenced by suppression, but this effect was attributed to the uncommonly low percentage of retrieve selections made in this condition, for which there was no logical explanation.

That the sum type manipulation was largely insensitive to suppression demonstrates that the influence of conscious processes upon selection is contingent upon the type of problem features inherent in a problem. While effects of the addend status feature were reliant upon consciously directed processes occurring during the course of the experiment, sum type effects were not reliant upon the same process.

This is based upon the notion that participants were already aware of the problem feature from problem-solving episodes previous to the experiment. Further research would be beneficial to refine this hypothesis. Specifically, it would be useful to examine whether making participants aware of the addend status manipulation before the experiment would serve to render suppression ineffective.

To examine a further class of context effects proposed by Cary and Reder (2002) a second task-level manipulation was employed to examine the sensitivity of selection to biases in the task instructions. In arithmetic problem-solving and related paradigms a number of studies have illustrated that instructions emphasising the use of one strategy over the other can bias selections (Kirk & Ashcraft, 2001; Blöte, Van der Burg & Klein, 2001; see also Gardner & Rogoff, 1990). However, these effects have not been examined in speeded response conditions therefore the measures taken in these experiments may be contingent upon conscious processes used to reconstruct prior processing episodes (e.g., self-reports). In Experiment 6a, where participants were instructed to indicate whether they would retrieve the answer (yes or no), or calculate the answer (yes or no), problem familiarity effects were immune to the task instruction. In Experiment 6b, there was no significant interaction between task instructions and sum type. However, participants given instructions emphasising the use of the retrieve strategy selected the retrieve strategy more often than participants given instructions biased towards calculate selections. That instructional bias only influenced the effect of sum type and not that of problem familiarity is particularly revealing. From this it was inferred that the selection-by-feature effect is susceptible to task instruction manipulations while selection influenced by problem familiarity is immune to such manipulations.

4.2.5 Summary of key findings

In summary, the key findings revealed from the two empirical series reported in this thesis combine to produce a much more complex overview of selection than previously conceived. It is apparent that problem-solving, on an experimental level, can be successfully lesioned into distinct selection- and solution-phases. Furthermore, the dual-phase experimental design employed in the thesis experiments provides a powerful and flexible methodology with which to investigate the selection process. Using this design, key findings from the Game Show studies were replicated, namely, that strategy selections are made rapidly, with a degree of accuracy and are influenced by problem familiarity. More importantly, familiarity derived by individuals from prior problem-solving episodes influences selection rather than experimentally induced during the course of an experiment. Evidence was also presented demonstrating that feature effects are contingent upon the identification and application of particular semantic characteristics inherent in the problem during selection.

Examining a more realistic problem-solving scenario, where multiple cues are present in a problem, it is tentatively proposed that the competition between two problem features may serve to attenuate the effects of one or both feature. However, selection may be influenced by both problem familiarity and problem features in tandem. Aside from the problem-level manipulations strategy selection, even though made very rapidly, varies in sensitivity to the wider processing context. Problem familiarity effects were largely immune to context manipulations as neither the availability of conscious processes, nor task instruction manipulations influenced selection. Furthermore, the selection mechanism itself was apparently un-reliant upon conscious processes. However, it was proposed that conscious processes are responsible for the identification of problem features during the course of an experiment, but problem features can also be derived from prior processing episodes, and in this instance are unaffected by the availability of conscious processes. In the following section these key phenomena will be used as a benchmark to examine the predictions of the existing accounts of selection.

4.3 THEORETICAL IMPLICATIONS

Despite the relatively large number of strategy selection models, the empirical research undertaken within the paradigm is limited in scope. The empirical series presented in this thesis has sought to broach this remiss through an empirical investigation of the models' predictions. In this section, I return back to the models comprising the Automaticity and Adaptive accounts of selection. The predictions of these accounts will be evaluated in light of the key phenomena detailed in the previous section, and as none of the models are able to account for these phenomena in full, I specify a framework in which candidate mechanisms with the potential account for these phenomena are outlined.

4.3.1 The Automaticity account of strategy selection

A central tenet of the Automaticity account is that individuals automatically attempt to solve problems following the Obligatory Activation Assumption (Logan, 1988). Accordingly, within this account strategies are not chosen *per se* as this assumption circumvents any requirement to select between candidate strategies. In respect to the experimental paradigm used within the experiments reported in this thesis predictions from the three Automaticity models (ACT-R, ITAM and DOA) stipulate that retrieve and calculate selections are determined by a measure derived from the attempted retrieval. For example, in ACT-R (Lebiere & Anderson, 1998) when presented with a problem, the retrieval production searches memory for chunks matching the problem terms. If a chunk in memory, complete with a problem's answer, has sufficient activation, the problem chunk will be overwritten and the answer available for output. A similar process is responsible for strategy selections in the DOA model (Siegler & Shrager, 1984). Here knowledge is structured as a series of associations in memory between the problem and answer. When problems are presented, the first step taken is an attempt to retrieve the answer. The relative concentration of associative activation between a problem and candidate answers activated determines whether the answer will be returned. Highly peaked distributions denote a greater confidence in the accuracy of an answer, whereas relatively even distributions of activation denote uncertainty. Contrasting these two accounts, in the ITAM model (Logan, 1988; 2002) it is proposed that an attempt to access all of the instances in memory of the problem, complete with the answer (i.e., a direct retrieval) is made. Simultaneously, ITAM searches for all stored instances of calculation algorithms applicable to the problem in memory.

Based upon these fundamentals, three streams of investigation failed to provide any support for the Automaticity accounts. All of the models were underpinned by the notion that the most rapid return from a problem-solving episode would be a completed direct retrieval of the problem's solution. Accordingly, following Reder and Ritter's (1992) critique of these accounts, none of the models were able to accommodate the rapidity and degree of accuracy of selections. In all of the thesis experiments it was evident that strategy selections were made consistently within the 850 ms designed to preclude participants from solving problems during this phase. Furthermore, precluding them from making predicted strategy selections upon the basis of a completed direct retrieval participants and using hindsight to identify the actual strategy they used to solve the problem.

A further possibility considered within the Game Show studies was that an early read of the problem's answer could be used to inform selection. It was argued that the attempted answer retrieval could be truncated allowing the return of a predicted strategy selection within the time limit. At that point in time the progress made in finding a solution, potentially indexed by the concentration of activation elicited by a problem, could act as a reliable indicator of the likelihood that the problem could be solved using the retrieve strategy. Accordingly, in problems with relatively familiar answers, it would be predicted that the retrieve strategy would be selected more often. However, Reder and Ritter (1992) found that selection was immune to the technique used to experimentally prime the association between problem and answer. Confirming this effect, findings from Experiments 1a, 1b and 6a also revealed that the familiarity of a problem's answer did not influence selection.

These three findings serve to highlight a fundamental flaw in the specifications of the Automaticity models. This is not to say that the Obligatory Activation assumption (Logan, 1988) is incorrect as it is acknowledged within wider cognition as being a particularly parsimonious way of accounting for a range of memory phenomena. However, it appears to be likely that this process is not responsible for the strategy selections made within this experimental paradigm, and potentially strategy selection in problem-solving in general.

4.3.2 The Adaptive account of strategy selection

Underpinning strategy selection in the Adaptive accounts is the notion that strategies are chosen in a distinct phase before an attempt to solve a problem is engaged. To recapitulate, in this class of models it is appreciated that problems can be and are often solved by a range of different strategies. Furthermore, given the same problem twice in succession individuals may employ different solution strategies (Reder, 1987; 1988; Shrager & Shipley, 1998; Siegler, 1999; Siegler & Araya, 2005; Siegler & Shipley, 1995). In the following sub-sections the key empirical phenomena revealed in this thesis will be applied to the predictions of the CMPL (Rickard, 1997; 2004), SCADS* (Siegler & Araya, 2005) and SAC (Reder & Ritter, 1992; Schunn et al, 1997) models.

4.3.2.1 The Component Power Laws theory (CMPL; Rickard, 1997; 2004)

In many respects the CMPL model bridges the divide distinguishing the Automaticity and Adaptive approaches. In line with the Adaptive account, when presented with a problem, in CMPL either the retrieve or calculate strategy is selected to the exclusion of the other. Fundamental to selection in this account is the approach taken to calculate strategies. Calculate strategies in CMPL are viewed as a series of direct retrievals. When applying this account to the data presented in this thesis, as selection is determined by the action of two competing retrievals, attempting to either solve the problem or the first step of the calculation procedure, it is apparent that the familiarity of a problem's solution — rather than the problem itself — would influence selection. Unlike the Automaticity accounts, it could be argued the model has the potential to return rapid and accurate predicted selections as a function of the competition for activation between retrieve and calculate selections. Furthermore, these responses could be returned with a degree of accuracy. Responses could be based upon activation accumulated for each strategy and hence accuracy would not be contingent upon participants solving problems and then using hindsight to identify the strategy they used to solve the problem. However, in Experiments 1a, 1b, 6a selection was shown to be insensitive to the familiarity of the problems' solution, and thus failing to provide unequivocal support for this account.

4.3.2.2 The Strategy Choice and Discovery Simulation (Siegler & Araya, 2005)*

Before attempting to evaluate the specifications of this model it should be noted that the SCADS* model will be focused upon in this section, rather than the 1st or 2nd generation accounts (i.e., ASCM and SCADS) preceding it. In light of the complexity of the SCADS* simulation I will briefly recapitulate the key components of the model. This family of models (i.e., ASCM, SCADS and SCADS*), similar to the SAC account, stipulate that individuals do not always attempt to retrieve the solution to a problem before applying other strategies (Shrager & Siegler, 1998; Siegler, 1999; Siegler & Araya, 2005; Siegler & Shipley, 1995). Rather than being based upon a single factor, like most of the strategy selection models detailed in this thesis, strategy selections in SCADS* is determined by four different types of information; global, featural, problem-specific, and novelty data. Information pertaining to each of these four types of data is combined for each candidate strategy. The model compares the relative strength of each strategy against its competitors and the candidate strategy with the greatest strength is chosen. Due to the complexity of this approach it is difficult to extract robust empirical predictions of performance from this model. This is particularly problematic because the contribution of each different type of data, to each candidate strategy, within each strategy-selection episode, is difficult to resolve in an empirical setting such as that employed in the experiments reported in this thesis.

To evaluate the SCADS* account, predictions from the model will be pitted against four empirical phenomena revealed in this thesis. Firstly, as strategies are selected in a distinct phase, where either the retrieve or calculate classes of strategy can be selected, this specification can apparently account for the rapidity with which strategy selections can be returned. Furthermore, as selections are made before an attempt to solve the problem is engaged it would also theoretically be possible for the model to return accurate predicted strategy selections. However, one caveat with this conclusion is that the computations that determine selection in SCADS* are relatively complex when compared to the more parsimonious accounts offered by other models. To illustrate, when selecting a strategy, activation elicited by the problem accumulates in parallel for each of the four different sources of data and for each of the candidate strategies in memory. In reality, rather than just being a straight competition between two strategies (i.e., retrieve and calculate), the retrieve strategy will compete for selection against a host of calculation strategies, especially in adult problem-solvers who have a large battery of different calculation (Hecht, 1999; 2002). Such a process may be computationally tractable in the simulations run by Siegler and colleagues. However, it is questionable whether the complexity of the computations required will afford rapid strategy selections in human problem-solvers. To date, there has been no specification of the time the model takes to select strategies; furthermore, existing simulations have been largely based upon data from children, who will have had smaller set of candidate strategies competing for selection than adults. Accordingly, it is unclear, at a practical level, whether using these procedures human problem solvers would be able to return rapid strategy selections within 850 ms. Particularly adult problem solvers who have a more comprehensive battery of candidate strategies than children to delineate between.

Turning to the problem-level manipulations employed in this thesis, although problematic, I shall attempt to extract some predictions from the SCADS* model. In Experiments 1a, 1b, 6a of this thesis it is likely that the contribution of novelty and featural data to selection was minimal. Briefly stated, the problems presented in these studies were relatively unfamiliar in comparison to decades, mixed and fives problems. Furthermore, in these experiments problem features were not manipulated so it is probable that feature data will not influence selection. Following this rationale it is hypothesised that global data, which is based upon the prior history of success of a particular strategy when applied to all problems and problem-specific data which indexes the prior success of a particular strategy on a particular problem would determine selection in these experiments. Generally speaking, as the action of these two data sources is reliant upon the historical success and accuracy of prior strategy applications, strategy selections determined by these factors would be contingent upon the associative link between the problem and its answer (Shrager & Siegler, 1998; Siegler & Shipley, 1995; Siegler, 1999; Siegler & Araya, 2005). As detailed previously, in Experiments 1a, 1b and 6a, selection was not influenced by the familiarity of the problems' solution. It may be that this provides evidence for a substantive case against the SCADS* model, however, this assertion needs to be appreciated with a degree of caution as it is not possible to account for the contribution of novelty or feature data in these types of problem.

A particularly notable facet of the SCADS* model is its ability to account for selection-by-feature effects. Other approaches (i.e., ACT-R and SAC) accommodate such effects by invoking ad-hoc assumptions when required. Such an approach compromises the ecological validity of these simulations as human intervention is required during the modelling procedure to specify particular problem features and to quantify the influence exerted upon selections. For example, in the SAC model a conditional parameter, 'does participant decide to never retrieve for one of the operators? True/False' (Schunn et al, 1997, p.12) was added to the model to account for performance in studies where problems with two different operators were presented. However, this feature was only used to account for performance in a limited number of participants within the data set. In contrast, the feature detection mechanism, running in parallel with the selection mechanism in SCADS* affords a great deal of online flexibility. Siegler and Araya (2005) indicate that the feature detection mechanism may identify features such as the identity of each number in a problem and whether both numbers are equal. Also the mechanism can detect features irrelevant to the problem such as the relative magnitude of the numbers, the colour and size of the numbers in a problem. Accordingly, it is likely that such a mechanism could account for the selection-by-feature effects evident in Experiments 2a, 2b and 5b where problems were comprised of addends that were multiples of five or ten. It should be noted that in this conception the presence of a particular feature does automatically mean that selection is determined by that feature. Using two counters, the model tracks the presence of each feature identified and the success of the strategy applied to solve each problem. It also tracks the success of strategies in the absence of particular features. When there is a difference of sufficient magnitude between the returns in each counter the presence or absence of a particular feature is used to compute the particular strength of a candidate strategy. Accordingly, in Experiment 4, where two feature manipulations were employed, the number of perceptual features, inherent not only in each problem but in the stimulus set as a whole, may have rendered feature detection of no benefit. However, in Experiment 5a, where only one problem feature was manipulated, addend status (with two levels), the feature, *both addends are even numbers,* may have been detected by this mechanism and subsequently influenced selection.

Aside from the complexity effect, a key component of the feature detection mechanism in SCADS* pertains to the consistency of feature effects. When a new problem feature is identified this feature only influences subsequent selections if the problem with that feature was solved correctly and at least 50% faster than usual (Siegler & Araya, 2005). Within SCADS*, perceptual and encoding mechanisms are posited as responsible for feature identification, in Experiments 5a and 5b mixed support for account was revealed. When conscious processes were precluded in Experiment 5a, addend status effects were impaired, while the same secondary task manipulation failed to impact selection-by-feature effects in Experiment 5b in which decade, mixed and fives problems were presented. This finding was attributed to the notion that consciously directed processes were responsible for the identification of problem features during the course of the experiment. This also suggests that some types of problem feature (e.g., *both addends are multiples of ten*) are already known and hence do not require identification during the course of the experiments. These findings cannot be easily accommodated with the SCADS* account as problem features are detected by perceptual mechanisms, presumably operating at a unconscious level during encoding and hence immune to interference from the articulatory suppression secondary task designed to preclude conscious processing.

In summary, it is apparent that the SCADS* account manages to accommodate the key empirical phenomena revealed in this thesis with mixed success. Fundamentally, the model has the potential to return rapid and accurate strategy selections as it is proposed that selections are determined prior to an attempt to solve the problem. However, its apparent reliance upon the associative linkage between a problem and the problem's solution as a predictor of performance in selection tasks serves to undermine the account. Although, in line with the complexity of the account it may be that this evaluative criterion employed within this section is too simplistic as it does not account for the influence of novelty and feature data. Unlike any of the other selection models, SCADS* has the ability to account for selection-by-feature effects. Ideally, further empirical investigation of the key predictions of the model is required to assess claims of this model, however due to the complexity of the account it may be that a fine-grained empirical approach is not possible.

4.3.2.3 The Source of Activation Confusion account (SAC; Reder & Ritter, 1992; Schunn et al, 1997)

As the most detailed account of selection and the focal point for the empirical investigation presented in Chapter 1 of this thesis, the predictions of this model have been examined more rigorously than any of the other models. In all of the thesis experiments it was shown that strategy selections could be made rapidly and with a degree of accuracy, supporting the position advocated by the SAC account that selection occurs in a distinct phase prior to strategy execution. Similar to the findings reported by Reder and colleagues (Reder & Ritter, 1992; Schunn et al, 1997) it was evident that problem familiarity influences selection (see Experiments 1a, 1b and 6a). However, in extension of this rationale, for the first time it was shown that not only problem familiarity induced experimentally, but problem familiarity derived from a long-term history of previous processing episodes influences selection. This suggests that the problem familiarity-based mechanism detailed in the SAC model has the potential to account for problem familiarity effects in real-world problem-solving. However, despite the successes of the model in accounting for these key phenomena, it was unable to accommodate selection-by-feature effects.

4.4 SUMMARY AND CONCLUSIONS

Current specifications of the strategy selection process are derived from computational simulations of a limited range of datasets. As this thesis has demonstrated, it is crucial that such imbalances are redressed by an empirical investigation addressing the scope, veridicality and ecological validity of the models' predictions. The present thesis highlights a number of limitations within the paradigm in respect to the empirical phenomena the models attempt to accommodate and the theoretical approach they adopt to the strategy selection process. Generally speaking the Adaptive class of selection models provided the most comprehensive account of the empirical phenomena presented in this thesis. However, the most notable conclusion that can be drawn from this exercise is that none of the existing models of strategy selection could account in full for these findings. The effects of problem familiarity fitted only the predictions of the SAC model and in particular the FoK mechanism which underpins this model. Problem-level effects also support this account while the selection-by-feature effect potentially could be accounted for by the feature handling component of the SCADS* model. In respect to the latter it was apparent that two processes were responsible for the selection-by-feature effect, a feature detection mechanism and a selection mechanism which can employ problem features during selection to determine retrieve and calculate strategy selections. From the SCADS* simulation it is proposed that automatic perceptual and encoding mechanisms can accommodate feature effects. However from the thesis' experiments it was apparent that when conscious processes were impaired during the course of the study the identification of particular problem features was also impaired. Further investigation of the mechanism through which problem features are integrated into the selection process is required. A computationally complex mechanism is provided within SCADS*, however it may be more beneficial to conceptualise a 'selection-by-feature mechanism' within the SAC based upon the relative parsimony of the account and its ability to model problem familiarity effects. One consideration is that the presence of distinct features in a problem set may influence the positioning of the FoK threshold which determines whether retrieve or calculate selections are returned. For example, problems with addends that are multiples of 5 or 10 may catalyse a shift towards leniency, such that retrieve selections are made more often in the presence of that feature. To examine such a possibility, employing a rationale derived from signal detection theory (Macmillan & Creelman, 2005) I have conducted a couple of pilot studies using a recognition memory paradigm to identify whether criterion shifts can be induced under speeded response conditions (Brown & Steyvers, 2005; Hirschmann & Henzler, 1998). From these studies it will become apparent whether conscious or automatic processes are responsible for selection-by-feature effects. In conclusion, it is apparent that an empirically, rather than theoretically motivated approach to selection is required within the paradigm. Selection was shown to be far more adaptive than previously conceived and in future examination of the selection mechanism it is hoped that subsequent investigation will embrace this notion.

REFERENCES

Anderson, J. R. (1993). *Rules of the Mind.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (2005) Human symbol manipulation within an integrated cognitive architecture. *Cognitive Science, 29*(3), 313-341.

Anderson, J. R., Reder, L. M. & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, 30(3), 221-256.

Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396-408.

Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review, 3,* 231-235.

Ashcraft, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. C. Bisanz, Brainerd, C. J. & Kail, R. (Eds.), Formal Methods in developmental psychology: Progress in cognitive developmental research (pp. 302-338). New York: Springer-Verlag.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44,* 75-106.

Auburn, T.C., Chapman, A.J., & Jones, D.M. (1987) Arousal and the Bakan vigilance task: the effects of noise intensity and the presence of others. *Current Psychological Research & Reviews, 6*(3), 196-206.

Ballard, J.C. (1996). Computerized assessment of sustained attention: a review of factors affecting vigilance performance. *Journal of Clinical Experimental Neuropsychology, 18*(6), 843–863.

Belavkin, R. V., & Ritter, F. E. (2004). *OPTIMIST: A New Conflict Resolution Algorithm for ACT-R.* Paper presented at the Sixth International Conference on Cognitive Modelling, Pittsburgh.

Blake, M. (1973). Prediction of recognition when recall fails: Exploring the feeling-of-knowing phenomenon. *Journal of Verbal Learning and Verbal Behaviour, 5,* 311-319.

Blöte, A. W., Van der Burg, E., & Klein, A. S. (2001). Students' Flexibility in Solving Two-Digit Addition and Subtraction Problems: Instruction Effects. *Journal of Educational Psychology, 93*(3), 627-638.

Botvinick, M., Braver, T., Barch, D. Carter, C. & Cohen, J. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108 (3)*, 624-652

Bourne, L.E., Healy, A.F., Parker, J.T., & Rickard, T.C. (1999) The strategic basis of performance in binary classification tasks: Strategy choices and strategy transitions. *Journal of Memory and Language, 41*(2), 223-252.

Bröder, A., Schiffer, S., 2003.Take-the-Best versus simultaneous feature matching: probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General, 132*, 277–293.

Brown, R. M., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behaviour, 12,* 325-337.

Brown, S. & Steyvers, M. (2005) The Dynamics of Experimentally Induced Criterion Shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition. 31*(4), 587-599.

Brysbaert, M. (1995). Arabic number reading: on the nature of the numerical scale and the origin of phonological recoding. *Journal Experimental Psychology: General, 124,* 434-452.

Brysbaert, M. (2005). Number recognition in different formats. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 23-42). Hove: Psychology Press.

Bundesen, C. (1993). The relationship between independent race models and Luce Choice Axiom. *Journal of Mathematical Psychology, 37*(3), 446-471.

Campbell, J. I. D. (1994). Architectures for numerical cognition. *Cognition, 53,* 1-44.

Campbell, J. I. D. (2005). Handbook of Mathematical Cognition. (J. I. Campbell, Ed.). New York: Psychology Press.

Campbell, J. I. D., & Austin, A. (2002). Effects of response time deadlines on adults' strategy choices for simple addition, *Memory & Cognition, 30,* 988-994.

Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Experimental Psychology, 39,* 338-366.

Campbell, J. I. D., & Oliphant, M. (1992). Representation and retrieval of arithmetic facts: A network-interference model and simulation. In J. I. D. Campbell

(Ed.), *The nature and origins of mathematical skills* (pp. 331-364). Amsterdam: Elsevier.

Campbell, J. I. D. (1994). Architectures for numerical cognition. *Cognition, 53,* 1-44.

Campbell, J. I. D., & Timm, J. C. (2000). Adults' strategy choices for simple addition: Effects of retrieval interference. *Psychonomic Bulletin & Review, 7*(4), 692-699.

Campbell, J. I. D., & Fugelsang, J. (2001). Strategy choice for arithmetic verification: effects of numerical surface form. *Cognition, 80,* B21-B30.

Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130,* 299-315.

Campbell, J. I. D., & Gunter, R. (2002). Calculation, culture and the repeated operand effect. *Cognition, 86,* 71-96.

Cary, M., & Carlson, R. A. (1999). External support and the development of problem-solving routines. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*(4), 1053-1070.

Cary, M. & Reder, L. M. (2002). Metacognition in strategy selection: Giving consciousness too much credit. In P. Chambres, M. Izaute, & P.J. Marescaux (Eds.), *Metacognition: Process, Function, and Use*. New York, NY: Kluwer, 63-78.

Cary, M. & Reder, L.M. (2003). A dual-process account of the list-length and strength-based mirror effects in recognition. *Journal of Memory and Language, 49*(2), 231-248.

Church, K., & Hanks, P. (1991). Word Association Norms, Mutual Information and Lexicography, *Computational Linguistics, 16*(1), 22-29.

Clark, J. M., & Campbell, J. I. D. (1991). Integrated versus modular theories of number skills and acalculia. *Brain and Cognition, 17*, 204-239.

Compton, B. J. & Logan, G.D. (1991). The transition from algorithm to retrieval in memory based theories of automaticity. *Memory & Cognition*, *19*, 151-158.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44(*1), 1-42.

Diana, R. A. & Reder, L. M. (2006). The low frequency encoding disadvantage: Word frequency affects processing demands.

Diana, R., Reder, L. M., Arndt. J., & Park, H. (2006). Models of recognition: A review of arguments in favor of a dual process account. *Psychonomic Bulletin & Review, 13*, 1-21.

Erickson, M. A., & Reder, L. M. (1998). The influence of repeated presentations and intervening trials on negative priming. In M. A. Gernsbacher, & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 327-332). Mahwah, N.J.: Erlbaum.

Ericsson, K. A., & Simon, H. A. (1993). Protocol analysis; Verbal reports as data (revised edition). Cambridge, MA: Bradford books/MIT Press.

Fayol, M., & Seron, X. (2005). About Numerical Representations: Insights From Neuropsychological, Experimental, and Developmental Studies. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 3-23). New York: Psychology Press.

Field, A.P. (2005). *Discovering statistics using SPSS.* (2nd ed.). London: Sage.

Gardner, W., & Rogoff, B. (1990). The development of children's improvisational and advance planning skills. *Developmental Psychology, 26,* 480-487.

Geary, D. C., & Wiley, J. G. (1991). Cognitive Addition: Strategy Choice and Speed-of-Processing Differences in Young and Elderly Adults. *Psychology and Aging, 6*(3), 474-483.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal Experimental Child Psychology, 88,* 121-151.

Gick, M. L. & Lockhart, R. S. (1995). Cognitive and Affective Components of Insight. In R.J. Sternberg & J. E. Davidson (Eds.), *The Nature of Insight.* (pp 197 – 228). Cambridge: MIT Press.

Gielen, I., Brysbaert, M., & Dhondt, A. (1991). The syllable-length effect in number processing is task dependant. *Perception & Psychophysics, 50*(5), 449-458.

Gigerenzer, G., Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650-669.

Green, D. R., Cerella, J., & Hoyer, W. J. (2000). *Do strategy probes affect the probability of item retrieval?* Poster presented at the 41st annual meeting of the Psychonomics Society, New Orleans, LA.

Gruneberg, M. M., Monks, J., & Sykes, R. N. (1977). Some methodological problems with feeling of knowing studies. *Acta Psychologica, 41,* 365-371.

Groen, G. J., & Parkman, J. M. (1972). A Chronometric Analysis of Simple Addition. *Psychological Review, 79*(4), 329-343.

Hart, J. T. (1967). Memory and the Memory-Monitoring Process. *Journal of Verbal Learning and Verbal Behaviour, 6,* 685-691.

Hecht, S. A. (1999). Individual solution processes while solving addition and multiplication math facts in adults. *Memory and Cognition, 27*(6), 1097-1107.

Hecht, S. A. (2002). Counting on working memory in simple arithmetic when counting is used for problem solving. *Memory & Cognition, 30*(3), 447-455.

Hirschman, E., & Henzler, A. (1998). The role of decision processes in conscious recollection. *Psychological Science, 9*(1).

Jones, D. M., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory & Language, 54,* 265-281.

Jackson, N., & Coney, J. (2007) Simple arithmetic processing: Surface form effects in a priming task. *Acta Psychologica, 125(1)*, 1-19.

Jones, R., & Van Lehn, K. (1991). A computational model of acquisition for children's addition strategies. In L. Birnbaum & G. Collins (Eds.), *Machine Learning: Proceedings of the Eighth International Workshop* (pp. 65-69). San Mateo, CA: Morgan Kaufmann.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology - Monograph, 74*, 657-690.

Karmiloff-Smith, A., (1992). Beyond modularity: A developmental perspective on cognitive science. Cambridge, MA: MIT Press.

Kirk, E. P., & Ashcraft, M. H. (2001). Telling Stories: The perils and promise of using verbal reports to study math strategies. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27,* 216-230.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*, 609-639.

Koriat, A. (2006). Are we frightened because we run away? Some evidence from metacognitive feelings. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), *Memory and emotion: Interdisciplinary perspectives* (pp. 83-103). Malden, MA: Blackwell.

Koriat, A., & Lieblich, I. (1977). A study of memory pointers. *Acta Psychologica, 41,* 151-164.

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language, 52*, 478–492.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behaviour. *Journal of Experimental Psychology: General, 135(1),* 36-69.

Krueger, L. E. (1986). Why 2 x 2 = 5 looks so wrong: On the odd-even rule in sum verification. *Memory & Cognition, 14,* 141-149.

Lachman, J. L., & Lachman, R. (1980). Age and the actualization of world knowledge. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging* (pp. 313-343). Hillsdale, N.J.: Erlbaum.

Lebiere, C. (1999). The dynamics of cognition: An ACT-R model of cognitive arithmetic. *Kognitionswissenschaft.*, 8 (1), pp. 5-19.

Lebiere, C., & Anderson, J. R. (1998). Cognitive Arithmetic. In J. R. Anderson, & C. Lebiere (Eds.), *The Atomic Components of Thought* (pp. 297-342). New Jersey: LEA.

Lee, K.-M., & Kang, S.-Y. (2002). Arithmetic operation and working memory: Differential suppression in dual tasks. *Cognition, 83,* B63-B68.

LeFevre, J., & Kulak, A. G. (1994). Individual differences in the obligatory activation of addition facts. *Memory & Cognition, 22*, 188 - 200.

LeFevre, J., Kulak, A. G., & Bisanz, J. (1991). Individual differences and developmental change in the associative relations among numbers. *Journal of Experimental Child Psychology, 52,* 256-274.

LeFevre, J.-A., Sadesky, G. S., & Bisanz, J. (1996). Selection Procedures in Mental Addition: Reassessing the Problem Size Effect in Adults. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*(1), 2126-2230.

Lemaire, P., & Fayol, M. (1995). When plausibility judgements supersede fact retrieval: The example of odd-even effect on product verification. *Memory & Cognition, 1,* 134-148.

Lemaire, P., & Reder, L. M. (1999). What affects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory and Cognition, 27*(2), 364-382.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General, 124*, 83-97.

Little, T. D., & Widaman, K . F. (1995). A production task evaluation of individual differences in mental addition skill development: Internal and external validation of chronometric models. *Journal of Experimental Child Psychology, 60,* 361-392.

Logan, G. D. (1988). Toward an Instance Theory of Automatization. *Psychological Review, 95*(4), 492-527.

Logan, G. D. (2002). An Instance Theory of Attention and Memory. *Psychological Review, 109*(2), 376-400.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition, 22 (4)*, 395-410.

Lovett, M. C. & Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology*, 31(2), 168-217.

Luce, R. D. (1959). Individual Choice Behaviour. New York: Wiley.

Luce, R. D. (1963). Detection and recognition. In R.D. Luce, R. R. Bush, & E. Galanter. (Eds.), *Handbook of Mathematical Psychology* (pp. 89-103). New York: Wiley.

Luchins, A. S. (1942). Mechanization in problem solving. Psychological Monographs, 54.

Macken, W. J., & Jones, D. M. (1995). Functional Characteristics of the Inner Voice and the Inner Ear: Single or Double Agency? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(2), 436-448.

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition, 44,* 107-157.

McCloskey, M., & Macaruso, P. (1994). Architecture of Cognitive Numerical Processing Mechanisms – Contrasting Perspectives on Theory Development and Evaluation. *Cahiers de Psychologie - Current Psychology of Cognition, 13(3),* 275-295.

McCloskey, M., & Macaruso, P. (1995). Representing and Using Numerical Information. *American Psychologist, 50*(5), 351-363.

Metcalfe, J. (1986). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition, 12,* 288-294.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition, 15,* 238-246.

Miner, A. C., & Reder, L. M. (1994). A new look at feeling of knowing: Its metacognitive role in regulating question answering. In J. Metcalfe, & Shimamura, A. (Eds.), *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.

Moyer, R., & Landauer, T. (1967). Time required for judgements of numerical inequality. *Nature, 215,* 1519-1520.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*(2), 102-116.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgements of learning (JOLs) are extremely accurate at predicting subsequent recall: The "Delayed-JOL Effect". *Psychological Science, 2,* 267-270.

Nelson, T. O., Leonesio, R. J., Shimamura, A. P., Landwehr, R. F., & Narens, L. (1982). Overlearning and the Feeling of Knowing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8*(4), 279-288.

Nelson, T., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in research and Theory* (Vol. 26, pp. 125-169): Academic Press.

Nino, R. S., & Rickard, T. C. (2003). Practice effects on two memory retrievals from a single cue. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 29*(3), 373-388.

Palmeri, J. T. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 23, 324-54.

Penner-Wilger, M., Leth-Steensen, C. & LeFevre, J-A. (2002). Decomposing the problem-size effect: A comparison of response time distribution across cultures. *Memory & Cognition, 30*(7), 1160-1167.

Park, H., Arndt, J.D., & Reder, L.M. (2006). A contextual interference account of distinctiveness effects in recognition. *Memory & Cognition, 34*(4), 743-751.

Payne, J. W., Bettman, J. R., and Johnson, E. J., (1988). Adaptive Strategy Selection in Decision Making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534-552.

Ratinckx, E., & Brysbaert, M. (2002). Interhemispheric Stroop-like interference in number comparison: Evidence for strong interhemispheric integration of semantic number information. *Neuropsychology, 16,* 217-229.

Reder, L. M. (1982). Plausibility judgements vs. fact retrieval: Alternative strategies for sentence verification. *Psychological Review, 89,* 250-280.

Reder, L. M. (1987). Strategy Selection in Question Answering. *Cognitive Psychology, 19,* 90-138.

Reder, L. M. (1988). Strategic control of retrieval strategies. In G. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 227-259). New York: Academic Press.

Reder, L. M., & Ritter, F. E. (1992). What Determines Initial Feeling of Knowing? Familiarity With Question Terms, Not With the Answer. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(3), 435-451.

Reder, L. M., & Schunn, C. D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. M. Reder (Ed.), *Implicit Memory and Metacognition* (pp. 45-77). Mahwah, N.J.: Erlbaum.

Reder, L. M., & Wible, C. (1984). Strategy use in question-answering: Memory strength and task constraints on fan effects. *Memory and Cognition, 12*, 411-419.

Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General, 126*(3), 288-311.

Rickard, T. C. (2004). Strategy execution in cognitive skill learning: an item-level test of candidate models. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*(1), 65-82.

Rickard, T. C., & Bourne, L. E. J. (1996). Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22*(5), 1281-1295.

Rickard, T. C., Healy, A. F., & Bourne, L. E. (1994). On the cognitive structure of basic arithmetic skills: operation, order, and symbol transfer effects. *Journal* *of Experimental Psychology: Learning, Memory, & Cognition, 20*(5), 1139-1153.

Schunn, C.D. & Reder, L.M. (1998). Strategy adaptivity and individual differences. In D. L. Medin (Ed.) *The Psychology of Learning and Motivation,* Academic Press, 115-154.

Schunn, C. D., Reder, L. M., Nhouyvanisvong, A., Richards, D. R., & Stroffolino, P. J. (1997). To Calculate or Not to Calculate: A Source Activation Confusion Model of Problem Familiarity's Role in Strategy Selection. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*(1), 3-29.

Schneider, W., & Pressley, M. (1989). *Memory development between 2 and 20.* New York: Springer.

Schwartz, B. L., & Metcalfe, J. (1992). Cue Familiarity but not Target Retrievability Enhances Feeling-of-Knowing Judgements. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(5), 1074-1083.

Seitz, K., & Schumann-Hengsteler, R. (2000). Mental multiplication and working memory. *European Journal of Cognitive Psychology, 12(4)*, 552-570.

Seitz, K., & Schumann-Hengsteler, R. (2002). Phonological loop and central executive processes in mental addition and multiplication. *Psychologische Beitrage, 44*, 275-302.

Seyler, D. J., Kirk, E. P., & Ashcraft, M. H. (2003). Elementary subtraction. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29,* 1339-1352.

Shepard, R. N. (1957). Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. *Psychometrika, 22,* 325-345.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116*(3), 250-264.

Siegler, R. S., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. In R. V. Kail (Ed.), *Advances in child development and behavior, Vol. 33* (pp. 1-42). Oxford, UK: Elsevier.

Siegler, R. S., & Jenkins, E. (1989). How children discover new strategies. Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Lemaire, P. (1997). Older and Younger Adults' Strategy Choices in Multiplication: Testing Predictions of ASCM Using the Choice/No-Choice Method. *Journal of Experimental Psychology: General, 126*(1), 71-92.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *The origins of cognitive skills* (pp. 229-293). Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. Halford (Eds.), *Developing cognitive competence: New approaches to process modelling* (pp. 31-76). Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Stern, E. (1998). Conscious and Unconscious Strategy Discoveries: A Microgenetic Analysis. *Journal of Experimental Psychology: General, 127*(4), 377-397.

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). A comparative approach to metacognition and uncertainty monitoring. *The Behavioral and Brain Sciences, 26,* 317-339.

Staszewski, J. J. (1988). Skilled Memory and Expert Mental Calculation. In M. T. H. Chi, Glaser, R. & Farr, M. J. (Eds.), *The Nature of Expertise* (pp. 71-128). New Jersey: LEA Publications.

Sternberg, R. J. (1985). Beyond IQ: A Triarchic theory of human intelligence. New York: Cambridge University Press.

Thorndike, E. L. (1922). *The Psychology of Arithmetic.* New York: The Macmillan Company.

Van Overschelde, J. P., Rawson, K., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50*, 289-335.

Vernon, D., & Usher, M. (2003). Dynamics of Metacognitive Judgements: Pre- and Postretrieval Mechanisms. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*(3), 339-346.

Wheeler, L. R. (1939). A comparative study of the difficulty of the 100 additional combinations. *Journal of Genetic Psychology, 54,* 295-312.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of memory and language 46*, 441-517.

Yonelinas, A. P. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General, 130* (3), 361-379.

Zbrodoff, N. J. & Logan, G. D. (1990). On the relation between production and verification tasks in the psychology of simple arithmetic. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 83-97.

Zbrodoff, N. J. & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. Journal of Experimental Psychology: General, 2, 118-130.

Zbrodoff, N. J., & Logan, G. D. (2005). What everybody finds: The problem size effect. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 331-345). New York: Psychology Press.

APPENDIX A

The following is a covariates analysis to accompany the findings presented in Experiment 1a.

Covariate Analysis

One possibility is that performance was determined by factors that were not directly tested in Experiment 1a, specifically other memorial and or problem features contained in a problem's terms or answer. To rule out this eventuality, using separate Linear Mixed models and specifying the potential confounds as covariates, two models were tested, both of which included sum familiarity and answer familiarity as repeated measures factors. The first model, termed the *magnitude model* sought to identify whether calculate selections were sensitive to the magnitude of the first and/or second addend. The problem size effect (Ashcraft, 1992; Groen & Parkman, 1972; LeFevre, Sadesky & Bisanz, 1996; Penner-Wilger Leth-Steensen & LeFevre, 2002) predicts that solution latencies increase with the magnitude of a problem's answer reflecting a shift in strategy selections from retrieval to calculation procedures. Accordingly, the strategy selection process may be influenced by the magnitude of the first or second addend, however none of the covariates reached significant levels (both Fs $< .34$, both ps $> .57$) demonstrating that the observed percentage of calculate selections was insensitive to the magnitude of the first and second addend ($df = 6$, $AIC = 322.66$).

Another possible confound is the level of familiarity with individual terms within the problem (Campbell & Graham, 1985; Koriat & Lieblich, 1977; Metcalfe, 1986; Metcalfe & Weibe, 1987; Reder, 1987; Schunn et al., 1997). By separately priming the question terms, if was found that participants can be biased towards selecting retrieval rather than calculation (Reder, 1987; Schunn et al., 1997). Familiarity ratings of the first and second addend were entered into the second model (the *familiarity model*). Again, neither measure covaried significantly with the percentage of calculate selections returned (both Fs $< .56$, both ps $> .46$) indicating that strategy selection was insensitive to the familiarity of the first or second addend ($df = 6$, $AIC = 332.29$). Separate chi square tests indicate that a model comprised solely of the experimental factors, sum familiarity and answer familiarity ($df = 4$, $AIC = 314.85$) provided a better fit to the data than the magnitude ($p = .02$) and familiarity models ($p = .02$) confirming that sum familiarity was the best predictor of performance.

APPENDIX B

In Gielen, Brysbaert and Dhondt (1991) a significant correlation between number frequency and number magnitude was evident. Based upon this premise it may be possible that addend magnitude, rather than problem familiarity was responsible for the effects upon selection attributed to problem familiarity in Experiment 1a. To test this eventuality a series of regressions were conducted in which the calculate strategy selection data reported in Experiment 1a was re-analysed with the intention of identifying the relative contribution of problem familiarity and other predictor variables to the strategy selection process.

Four potential predictors of the variance in calculate strategy selections were entered into linear regressions. Problem familiarity (the summed familiarity ratings of both addends) and the familiarity rating of the problem's answer were chosen as they were directly manipulated in Experiment 1a. In light of the correlation between number frequency (or *familiarity*) and magnitude reported in Gielen, Brysbaert and Dhondt (1991) the magnitude of the problems answer was also selected. A further predictor, that of whether the problem necessitated a carry (i.e., 19 + 18) or not (i.e., 13 + 14) when being solved was also chosen as it may have been a feature of the problem particularly apparent to individuals during the selections process. To identify any instances of multicolinearity between these predictors Pearson correlation coefficients were calculated. A significant negative correlation between answer magnitude and problem familiarity, -.877, (n = 64, $p < .001$) was found. To accommodate this finding and preventing the erroneous inflation of standard errors (Field, 2005) in the regression analysis these factors were analysed in separate models. Theoretically, this approach matches that adopted in the computational models of the strategy selection process where problem familiarity, answer familiarity and answer magnitude are analysed as sole determinants of strategy selections. None of the selection models to date provide a rationale to suggest that these factors combine to influence the strategy selection process (see Lebiere & Anderson, 1998; Logan, 2004; Reder & Ritter, 1992; Schunn et al, 1997; Siegler & Araya, 2005).

Separate linear regressions examined each of the predictors in turn. Three of the factors, problem familiarity, answer magnitude and the presence (or absence) of a carry accounted for a significant proportion of the variance in the calculate strategy selections reported in Experiment 1a (see Table B1). Answer magnitude and problem familiarity accounted for a similar proportion of the variance, 46% and 44% respectively, while the carry predictor only accounted for 33%. This suggests that the effects attributed to problem familiarity in Experiment 1a could equally be attributable to the magnitude of a problem answer.

Table B1.
Linear regressions conducted upon calculate strategy selections reported in Experiment 1a.

	B	*SE B*	*ß*	R^2	*Adjusted* R^2
Model 1				.2	.18
Problem familiarity	-.1	.003	-.44***		
Model 2				.21	.2
Answer magnitude	0.07	0.02	.46***		
Model 3				.004	-.01
Answer familiarity	-0.003	0.006	-.06		
Model 4				.11	.09
Carry	-2.48	0.91	-.33**		

* $p < .05$, ** $p < .01$, *** $p < .001$

When examining whether answer magnitude or problem familiarity influenced selection in Experiment 1a two issues need to be considered. Firstly, the problem familiarity effects described in Experiments 1a and 1b replicate those found in the Game Show studies by Reder and colleagues (Reder & Ritter, 1992; Schunn et al, 1997). Unlike the methodology used in this thesis' experiments they employed a priming methodology which allowed them to demonstrate a clear dissociation between problem familiarity and other problem features such as answer magnitude.

To briefly recapitulate, by manipulating problem familiarity in a priming phase Reder and colleagues demonstrated that problems encountered frequently in the priming phase elicited a greater frequency of retrieve strategy selections than those encountered infrequently in the priming phase, irrespective of the magnitude of a problem's answer. This methodology clearly dissociated problem familiarity and answer magnitude, a comparison that is not afforded by the methodology employed with the experiments in this thesis. Here the experiments were designed to examine existing problem familiarity rather than experimentally induced problem familiarity as used in Reder's methodology.

Secondly, on a procedural level the use of answer magnitude as a cue to selection may be problematic. In the dual-phase design employed in this thesis' experiments a time limit of 850 ms was imposed upon the selection phase. This prevented individuals from retrieving a problem's solution and using that as a guide to strategy selection. As Reder and colleagues stated, the accuracy of strategy selections derived from an early read of a problems solution is likely to be significantly compromised (Reder & Ritter, 1992). Furthermore, in the early read approach, selection is determined by the strength of the associative link between a problem and its solution as the act of encoding a problem automatically initiates the search for the answer to that problem (Logan, 1988). Within models of selection, familiarity and associative strength are largely inseparable, accordingly, it would be expected that the familiarity of the answer will determine strategy selection (e.g., CMPL, ITAM, ACT-R) which was not found to be the case in the regression analysis or repeated measures ANOVA analysis.

Although these two sources of evidence provide support for the notion that problem familiarity rather than answer familiarity influences selection the question remains as to why answer magnitude and problem familiarity accounted for a comparable proportion of the variance in the regressions. One possibility may lie in an argument proposed by Siegler and Araya (2005). They examined the frequency with which arithmetic problems are presented to individuals over the course of a lifetime examining, among other sources, educational learning materials. They reported that over an individual's lifetime and when generalised to the population, a negative linear correlation between the frequency of problem presentation (i.e., exposure) and problem size (i.e., answer magnitude) was evident. Smaller problems (i.e., those with a small answer magnitude) were encountered more frequently than larger problems (i.e., large answer magnitude problems). When this is taken in tandem with Reder and colleagues' finding that exposure to a problem serves to increase the familiarity of that problem (Reder & Ritter, 1992; Schunn et al, 1997), a highly significant negative correlation between answer magnitude and problem familiarity would therefore be expected in Experiment 1a. Similarly, both answer magnitude and problem familiarity would be expected to share a comparable amount of variance in calculate selections. From this it is inferred that while answer magnitude and problem familiarity are distinct and dissociable items in memory (see Reder & Ritter, 1992; Schunn et al, 1997) problem familiarity is largely determined by answer magnitude in an opportunistic fashion. Exposure to a problem increments problem familiarity and the frequency of exposure to a problem is predicted by the size of the problem, specifically the size of its answer.

The evidence presented here fails to rule out the possibility that answer magnitude, rather than problem familiarity, influences selection. However, based upon converging lines of evidence derived from previous research by Reder and colleagues, the explanation for the relationship between problem familiarity and answer magnitude provided by Siegler and Araya (2005) and the limitations of the early read account of selection it seems more likely that the problem familiarity account is the most likely of the two alternatives. Further experimentation using a priming methodology similar to that employed by Reder and colleagues may be useful to replicate the findings from the experiments in which problem familiarity has been manipulated in this thesis. This would allow a clear dissociation between the effects of problem familiarity and answer familiarity. Also, direct experimental manipulation of answer magnitude, rather than post hoc analysis, would be beneficial to rule out the contribution of answer magnitude to strategy selection in the dual-phase design used in this thesis.

APPENDIX C

Similar to the approach adopted in Appendix B the strategy selection data recorded in Experiment 2a was re-analysed to identify whether strategy selection should be attributed to the action of the selection-by-feature effect or other problem features (i.e., problem familiarity, answer magnitude or the presence or absence of a carry). A further predictor, termed *sum type* was used to examine the predictions of the selection-by-feature effect identified in Experiments 2a and 2b. This effect is contingent upon the notion that participants use particular features of a problem to determine strategy selections (i.e., the presence of one or two decades or fives addends in a problem).

As illustrated in Table C1 the sum type predictor (model 5) accounted for the greatest proportion of the variance (58%), 15% greater than the carry predictor (model 4; 43%). These findings support the conclusions drawn in Experiments 2a suggesting that selection-by-feature effects influenced strategy selection in problems comprising decades, mixed and fives addends.

Table C1.
Linear regressions conducted upon retrieve strategy selections reported in Experiment 2a.

	B	*SE B*	*ß*	R^2	*Adjusted* R^2
Model 1				.01	-.02
Problem familiarity	-0.006	0.01	-.11		
Model 2				.0	-.03
Answer magnitude	0.0	0.04	.0		
Model 3				.01	-.02
Answer familiarity	-0.005	0.01	-.09		
Model 4				.44	.43
Carry	-4.63	0.89	-.67***		
Model 5				.59	.58
Sum Type	-3.08	.44	-.77***		

*** $p < .001$

APPENDIX D

Vigilance experiment

The following experiment was originally outlined in Experiment 5b of Chapter 3. Experiments 5a and 5b presented in Chapter 3 of this thesis investigated the contribution of the selection-by-feature effect to strategy selection. In the current study, a secondary task, derived from the Vigilance literature was employed to examine whether feature detection is actually determined by a unconscious mechanism, potentially encoding and perceptual mechanisms (Siegler & Araya, 2005). By impeding the action of consciously directed processes in Experiment 5a it was revealed that suppression impaired addend status effects (i.e., even and mixed odd and even problems), while in Experiment 5b, sum type effects (i.e., decades, fives and mixed decades and fives problems) were immune to interference from suppression. The secondary task employed in the present study was designed to impair the unconscious processes responsible for identifying and comparing the problem features that can be derived from a problem.

An equal number of participants were either presented with the same problems as employed in Experiment 5a (problem familiarity and problem feature) or those presented in Experiment 5b (sum type) and the same blocked dual-phase design as employed in those studies was used here. It was predicted that a significant interaction between either of the problem features employed (addend status or sum type) and the secondary task condition would demonstrate that unconscious perceptual and encoding mechanisms may be responsible for feature identification (Siegler & Araya, 2005) rather than consciously procedures as evident in Experiment 5a.

Method

Participants

Twenty-four participants from the School of Psychology at Cardiff University were given course credit or payments in return for their participation. All were native

English speakers reporting normal hearing and correct or normal vision and had not participated in any of the other thesis experiments.

Materials & Design

Refer to the corresponding sections in Experiments 5a and 5b for detail of the problem-level manipulations employed in the present study. All of the participants were required to complete the background auditory monitoring task throughout the duration of the strategy selection task. The auditory items used in the monitoring task were drawn from eight different categories (musical instruments, drinks, music types, weather, fish, birds, vegetables and animals) each of which contributed 12 different words (from Van Overschelde, Rawson and Dunlosky, 2004). These auditory items were recorded using *Sound Forge* in a female voice, each item measuring 500 ms in length and across all items pitch and intensity were controlled. When constructing the sound files each auditory item was separated by 500 ms of silence to ensure the words were easily identifiable.

Within each block of 6 problems in the strategy selection phase participants were exposed to between 38 s (i.e., 16 items) and 41 s (i.e., 20 items) of background sound depending upon how long was taken to make the strategy selections over the course of the trial block. For each participant a pseudo-randomised ordering of auditory items was constructed where targets (i.e., 3 consecutive items from the same category) were always from the same two categories, vegetables and animals. The precise constituents of each target differed for each participant as did the position of the targets, pseudo-randomly ranging from the first to the twelfth trial block. None of the targets were located in the opening 6 auditory items in a list, or the final 6 items.

Procedure

The same experimental design was employed as in Experiments 5a and 5b where the participants made strategy selections for all of the problems, then solved all of the problems in the stimulus set. Accordingly, the same instructions were relayed to participants with few notable departures from Experiment 5a. The experiment started with a practice phase designed to accustom participants with the identification of auditory targets in the monitoring task. Six auditory practice lists of were presented in isolation (i.e., as a primary task) and participants were required to indicate after each list whether they heard a target or not. Targets were present in 3 of the 6 lists over a pair of headphones. At the end of the presentation of each list immediate feedback was provided as to whether the target was correctly identified (i.e., a hit), the absence of a target was correctly identified (i.e., a correct rejection), a target was identified even though the list did not contain one (i.e., false alarm) or a target was missed (i.e., a miss).

Upon completion of this task the selection-phase of the experiment commenced with 16 practice questions. For the opening 8 questions there was no background sound to monitor, the final 8 were accompanied by the background sounds to monitor. During the experimental trials participants were advised to focus on the primary task (i.e., the strategy selection task) but also monitor the background sounds for three items in a row from the same semantic category. After every 6 problems the runtime program paused and asked participants to identify whether they heard a target in the background sound. Participants responded 'yes' or 'no' and unlike the practice trials for the monitoring task did not receive feedback as to the accuracy of the response. To resume the experiment participants clicked on the start button to when ready.

When the selection task was completed for all of the problems in the stimulus set the solution-phase of the experiment opened with 8 practice questions. Each of the problems presented in the selection-phase were represented in a pseudo-randomised order. Participants were requested to solve the problems as quickly and accurately as possible. After typing the answer in and pressing the enter key to confirm the answer the lead-in to the next problem started. In this phase there were no background sounds to monitor.

Results

Scoring Procedure

To identify the effects of the monitoring task employed in the strategy selection-phase the findings from the present experiment were contrasted to those reported in the conditions without articulatory suppression in Experiments 5a and 5b which acted as a control condition for the analyses conducted here. The same measures as in previous experiments were recorded; the strategy selection (retrieve or calculate), the selection latency, the sum solution and solution latency. In addition a measure of performance on the monitoring task was also taken using a basic sensitivity measure employed in signal detection theory (see Macmillan & Creelman, 2005) where responses are recorded as hits, misses, false alarms, correct rejections.

Effects of problem familiarity and addend status upon selection

As Figure D1 indicates, in accordance with previous studies presented in this thesis which have employed the same stimuli, the calculate strategy was selected in a greater percentage of trials than the retrieve strategy. A 2 (sum familiarity; low vs. high) x 2 (addend status; even vs. mixed) x 2 (monitoring condition; monitoring vs. no monitoring) mixed measures ANOVA, similar to Experiment 5a revealed a significant interaction between the two variables sum familiarity and addend status, $F(1, 22) = 17.75$, $MSE = 82.53$, $p < .001$. Simple effects indicate that addend status only influenced relatively familiar problems, $F(1, 22) = 19.17$, $p < .001$, such that a greater percentage of retrieve selections were made on even than mixed status problems. Effects of sum familiarity were evident in both levels of addend status, even and mixed respectively, $F(1, 22) = 7.83$, $p = .01$, and, $F(1, 22) = 6.76$, $p = .016$. More familiar problems elicited a greater percentage of retrieve (hence less calculate) strategy selections.

Main effects of monitoring condition revealed that the monitoring task served to reduce the overall percentage with which the calculate strategy was selected, $F(1, 22) = 6.33$, $MSE = 1166.29$, $p = .02$. There was no significant interaction between either addend status and sum familiarity and monitoring condition, (both ps $> .16$). This suggests that the monitoring task predisposed participants to select the retrieve strategy on a greater percentage of trials than those participants who were not required to monitor. However, as there is no obvious explanation of that effect evidence of an interaction between either of the manipulations and monitoring condition would be more revealing of the impact the monitoring task levied upon strategy selection.

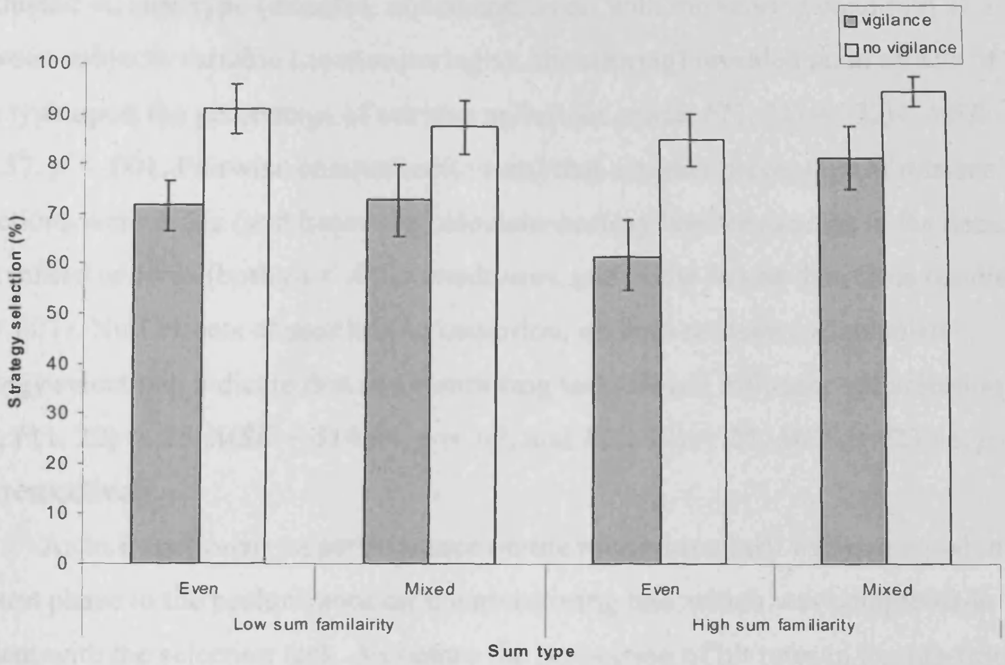

Figure D1: Mean percentage of calculate strategy selections. Error bars represent the standard error of the mean.

To assess the level of performance in the monitoring task when it was performed in isolation (i.e., in the pre-test phase) and in tandem with the selection task the frequency of hit and correct rejections were analysed. When performed in isolation, during the pre-test phase there were 5 trials in which 3 lists comprised targets, whereas when performed in tandem with the selection task, there was a total of 12 lists, 2 of which included targets. The evident probability of correctly identifying a target as a hit in the pre-test phase was .58, whilst during the selection task the probability was .46. The probability of making a correct rejection was also equitable, .91 in the pre-test phase and .89 during the selection-phase indicating that performance levels in the monitoring task did not decrease notably in the dual-task scenario. It was anticipated that the proportion of targets identified would have been approaching 1 as easily discernable targets (i.e., vegetables and fruits) were employed.

Effects of the sum type manipulation upon selection

Similar to previous experiments which have employed this set of stimuli, the retrieve strategy was selected most often (see Figure D2). A mixed measures ANOVA conducted on sum type (decades, mixed and fives) with monitoring condition as a between subjects variable (no monitoring vs. monitoring) revealed main effects of sum type upon the percentage of retrieve selections made, $F(2, 21) = 93.04$, $MSE = 202.57$, $p < .001$. Pairwise comparisons reveal that a higher percentage of retrieve selections were made (and hence the calculate strategy less frequently) in the decades, than mixed or fives (both *ps* < .001) conditions, and in the mixed than fives condition ($p < .001$). Null effects of monitoring condition, on both retrieve and calculate strategy selections indicate that the monitoring task did not influence the selection task, $F(1, 22) = .25$, $MSE = 514.94$, $p = .62$, and $F(1, 12) = .25$, $MSE = 423.03$, $p = .62$, respectively.

As in Experiment 6a performance on the monitoring task was compared in the pre-test phase to the performance on the monitoring task which was completed in tandem with the selection task. As before the proportion of hit rates in the pre-test phase (.49) and selection-phase (.5) were comparable as were the proportions of correct rejections, .88 in the pre-test phase and .85 in the selection phase.

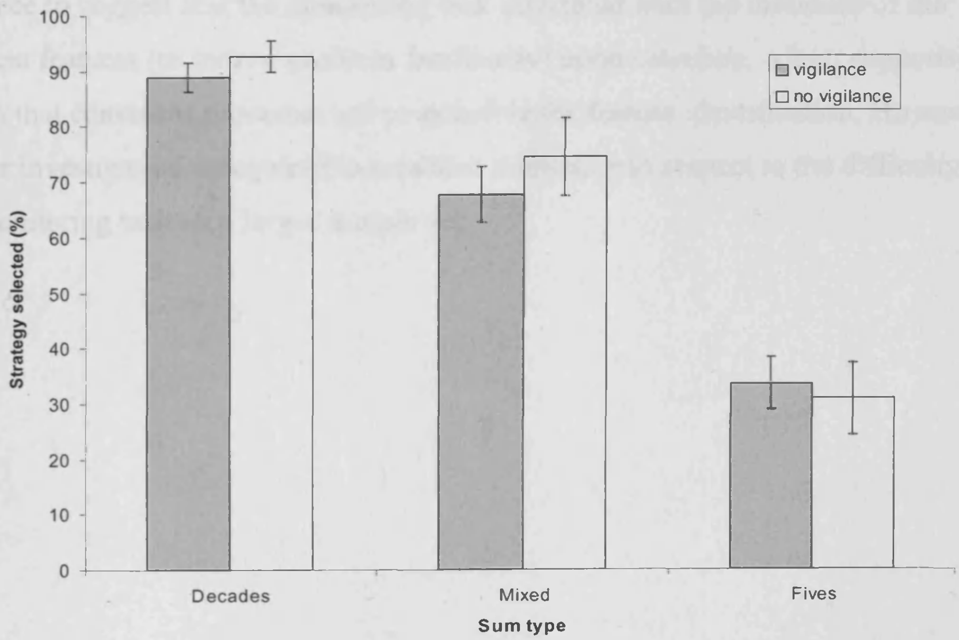

Figure D2. Mean percentage of retrieve strategy selections. Error bars represent the standard error of the mean.

Discussion

In summary, the failure to identify a significant interaction between the either addend status or sum type and the monitoring task condition is suggestive of one of two outcomes. Firstly, that the selection-by-feature effect is immune to interference from the monitoring task, specifically the requirement to identify and compare features in the addends in a problem. Alternatively, it may have been the case that participants were ignoring the monitoring task and focused on completing the selection task. This conclusion is supported by the relatively low hit rates revealed in both studies (approximately .5), although in both studies correct rejection rates were relatively (approximately .9). As the hit and correct rejection rates were similar to those revealed in the pre-test phase where the monitoring task was performed in isolation it may be the case that the monitoring task was too difficult for participants, as performance did not noticeably depreciate when participants completed the selection task in tandem with the monitoring task. Accordingly, at this juncture any conclusions drawn from this study have to be proposed tentatively. There was no evidence to suggest that the monitoring task interfered with the influence of the problem features (or indeed problem familiarity) upon selection, which supports the notion that conscious processes are responsible for feature identification. However, further investigation is required to establish a baseline in respect to the difficulty of the monitoring task on a larger sample set.

APPENDIX E

In the following tables the stimuli used in the thesis experiments are presented. Please note that in Experiment 3 the stimuli used in Experiments 1a and 2a were used.

Table E1.
Stimuli presented in Experiments 1a and 1b accompanied by familiarity ratings of Addend A (i.e., the first addend in the problem), Addend B (i.e., the second addend in the problem), the Problem (i.e., the familiarity of addend A plus Addend B) and the Answer for each problem.

Sum	Answer	Sum Familiarity group	Answer familiarity group	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
26 + 23	49	Low	Low	245	200	445	175
28 + 23	51	Low	Low	225	200	425	160
23 + 29	52	Low	Low	200	200	400	140
27 + 29	56	Low	Low	240	200	440	185
28 + 29	57	Low	Low	225	200	425	165
34 + 33	67	Low	Low	205	215	420	195
34 + 39	73	Low	Low	205	190	395	145
37 + 39	76	Low	Low	175	190	365	190
38 + 39	77	Low	Low	195	190	385	200
44 + 49	83	Low	Low	165	175	340	145
42 + 43	85	Low	Low	220	130	350	200
42 + 44	86	Low	Low	220	165	385	150
44 + 43	87	Low	Low	165	130	295	145
46 + 48	94	Low	Low	220	175	395	165
47 + 49	96	Low	Low	230	175	405	200
48 + 49	97	Low	Low	175	175	350	160
23 + 27	50	Low	High	200	240	440	330
26 + 29	55	Low	High	245	200	445	235
28 + 27	55	Low	High	225	240	465	235
32 + 37	69	Low	High	260	175	435	325
33 + 37	70	Low	High	215	175	390	245
34 + 37	71	Low	High	205	175	380	240
38 + 33	71	Low	High	195	215	410	240
38 + 37	75	Low	High	195	175	370	360
42 + 48	90	Low	High	220	175	395	360
43 + 47	90	Low	High	130	230	360	360
44 + 46	90	Low	High	165	220	385	360
42 + 49	91	Low	High	220	175	395	235
44 + 47	91	Low	High	165	230	395	235
44 + 48	92	Low	High	165	175	340	245
46 + 49	95	Low	High	220	175	395	245
48 + 47	95	Low	High	175	230	405	245
14 + 17	31	High	Low	320	290	610	140
18 + 13	31	High	Low	365	350	715	140
14 + 19	33	High	Low	320	275	595	215
16 + 17	33	High	Low	325	290	615	215
16 + 18	34	High	Low	325	365	690	205
18 + 19	37	High	Low	365	275	640	175

Table E1 (continued)

Sum	Answer	Sum Familiarity group	Answer familiarity group	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
22 + 26	48	High	Low	260	245	505	175
22 + 27	49	High	Low	260	240	500	175
22 + 29	51	High	Low	260	200	460	160
24 + 27	51	High	Low	265	240	505	160
24 + 28	52	High	Low	265	225	490	140
24 + 29	53	High	Low	265	200	465	125
26 + 27	53	High	Low	245	240	485	125
26 + 28	54	High	Low	245	225	470	195
32 + 33	65	High	Low	260	215	475	200
32 + 34	66	High	Low	260	205	465	215
14 + 13	27	High	High	320	350	670	240
12 + 17	29	High	High	315	290	605	200
16 + 13	29	High	High	325	350	675	200
13 + 17	30	High	High	350	290	640	320
14 + 16	30	High	High	320	325	645	320
13 + 19	32	High	High	350	275	625	260
14 + 18	32	High	High	320	365	685	260
16 + 19	35	High	High	325	275	600	240
18 + 17	35	High	High	365	290	655	240
17 + 19	36	High	High	290	275	565	260
22 + 23	45	High	High	260	200	460	270
22 + 28	50	High	High	260	225	485	330
36 + 33	69	High	High	260	215	475	325
32 + 38	70	High	High	260	195	455	245
34 + 36	70	High	High	205	260	465	245
32 + 39	71	High	High	260	190	450	240

Table E2.
Stimuli presented in Experiments 2a and 2b accompanied by familiarity ratings of Addend A (i.e., the first addend in the problem), Addend B (i.e., the second addend in the problem), the Problem (i.e., the familiarity of addend A plus Addend B) and the Answer for each problem.

Sum	Answer	Sum type	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
10+50	60	Decade	445	330	775	290
10+60	70	Decade	445	290	735	245
10+70	80	Decade	445	245	690	260
10+80	90	Decade	445	260	705	360
20+40	60	Decade	370	300	670	290
20+50	70	Decade	370	330	700	245
20+60	80	Decade	370	290	660	260
20+70	90	Decade	370	245	615	360
30+40	70	Decade	320	300	620	245
30+50	80	Decade	320	330	650	260
30+60	90	Decade	320	290	610	360
40+50	90	Decade	300	330	630	360
15+25	40	Fives	345	370	715	300
15+35	50	Fives	345	340	685	330
15+45	60	Fives	345	270	615	290
15+55	70	Fives	345	235	580	245
15+65	80	Fives	345	200	545	260
15+75	90	Fives	345	360	705	360
25+35	60	Fives	370	340	710	290
25+45	70	Fives	370	270	640	245
25+55	80	Fives	370	235	605	260
25+65	90	Fives	370	200	570	360
35+45	80	Fives	340	270	610	260
35+55	90	Fives	340	235	575	360
10+55	65	Mixed	445	235	680	200
10+75	85	Mixed	445	360	805	200
20+35	55	Mixed	370	340	710	235
20+55	75	Mixed	370	235	605	360
20+75	95	Mixed	370	360	730	245
30+25	55	Mixed	320	370	690	235
30+45	75	Mixed	320	270	590	360
40+25	65	Mixed	300	370	670	200
40+45	85	Mixed	300	270	570	200
50+25	75	Mixed	330	370	700	360
50+45	95	Mixed	330	270	600	245
60+25	85	Mixed	290	370	660	200

Table E3.

Stimuli presented in Experiment 4 accompanied by familiarity ratings of Addend A (i.e., the first addend in the problem), Addend B (i.e., the second addend in the problem), the Problem (i.e., the familiarity of addend A plus Addend B) and the Answer for each problem.

Sum	Answer	Odd / Even / Mixed group	Similar / Disparate Magnitude group	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
13+15	28	Odd	Similar	350	345	695	225
13+19	32	Odd	Similar	350	275	625	260
15+19	34	Odd	Similar	345	275	620	205
23+25	48	Odd	Similar	200	370	570	175
23+29	52	Odd	Similar	200	200	400	140
25+29	54	Odd	Similar	370	200	570	195
33+35	68	Odd	Similar	215	240	455	235
33+39	72	Odd	Similar	215	190	405	220
35+39	74	Odd	Similar	240	190	430	200
43+45	88	Odd	Similar	130	270	400	205
43+49	92	Odd	Similar	130	175	305	245
45+49	94	Odd	Similar	270	175	445	165
12+46	58	Odd	Disparate	315	220	535	170
13+45	58	Odd	Disparate	350	270	620	170
13+53	66	Odd	Disparate	350	125	475	215
13+61	74	Odd	Disparate	350	165	515	200
13+69	82	Odd	Disparate	350	325	675	170
13+77	90	Odd	Disparate	350	200	550	360
13+85	98	Odd	Disparate	350	200	550	205
15+43	58	Odd	Disparate	345	130	475	170
15+51	66	Odd	Disparate	345	160	505	215
15+59	74	Odd	Disparate	345	150	495	200
15+67	82	Odd	Disparate	345	195	540	170
15+83	98	Odd	Disparate	345	145	490	205
12+17	29	Mixed	Similar	315	290	605	200
14+15	29	Mixed	Similar	320	345	665	200
16+13	29	Mixed	Similar	325	350	675	200
22+29	51	Mixed	Similar	260	200	460	160
24+27	51	Mixed	Similar	265	240	505	160
28+23	51	Mixed	Similar	225	200	425	160
32+37	69	Mixed	Similar	260	175	435	325
34+33	67	Mixed	Similar	205	215	420	195
36+39	75	Mixed	Similar	260	190	450	360
42+47	89	Mixed	Similar	220	230	450	235
44+43	87	Mixed	Similar	165	130	295	145
48+45	93	Mixed	Similar	175	270	445	150
12+55	67	Mixed	Disparate	315	235	550	195
12+63	75	Mixed	Disparate	315	160	475	360
12+71	83	Mixed	Disparate	315	240	555	145
12+79	91	Mixed	Disparate	315	155	470	235
12+87	99	Mixed	Disparate	315	145	460	280
13+44	57	Mixed	Disparate	350	165	515	165
13+52	65	Mixed	Disparate	350	140	490	200
13+62	75	Mixed	Disparate	350	165	515	360
13+70	83	Mixed	Disparate	350	245	595	145
13+78	91	Mixed	Disparate	350	140	490	235

Table E3 (continued)

Sum	Answer	Odd / Even / Mixed group	Similar / Disparate Magnitude group	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
13+86	99	Mixed	Disparate	350	150	500	280
12+16	28	Even	Similar	315	325	640	225
14+16	30	Even	Similar	320	325	645	320
16+18	34	Even	Similar	325	365	690	205
22+26	48	Even	Similar	260	245	505	175
24+26	50	Even	Similar	265	245	510	330
26+28	54	Even	Similar	245	225	470	195
32+36	68	Even	Similar	260	260	520	235
34+36	70	Even	Similar	205	260	465	245
36+38	74	Even	Similar	260	195	455	200
42+46	88	Even	Similar	220	220	440	205
44+46	90	Even	Similar	165	220	385	360
46+48	94	Even	Similar	220	175	395	165
12+47	59	Even	Disparate	315	230	545	150
12+54	66	Even	Disparate	315	195	510	215
12+62	74	Even	Disparate	315	165	480	200
12+70	82	Even	Disparate	315	245	560	170
12+78	90	Even	Disparate	315	140	455	360
12+86	98	Even	Disparate	315	150	465	205
14+44	58	Even	Disparate	320	165	485	170
14+52	66	Even	Disparate	320	140	460	215
14+60	74	Even	Disparate	320	290	610	200
14+68	82	Even	Disparate	320	235	555	170
14+76	90	Even	Disparate	320	190	510	360
14+84	98	Even	Disparate	320	215	535	205

Table E4.
Stimuli presented in Experiment 5a accompanied by familiarity ratings of Addend A (i.e., the first addend in the problem), Addend B (i.e., the second addend in the problem), the Problem (i.e., the familiarity of addend A plus Addend B) and the Answer for each problem.

Sum	Answer	Sum Familiarity group	Even / Mixed group	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
19+42	61	Low	Mixed	275	220	495	165
28+37	65	Low	Mixed	225	175	400	200
23+48	71	Low	Mixed	200	175	375	240
27+44	71	Low	Mixed	240	165	405	240
59+18	77	Low	Mixed	150	365	515	200
18+63	81	Low	Mixed	365	160	525	205
37+44	81	Low	Mixed	175	165	340	205
43+38	81	Low	Mixed	130	195	325	205
58+23	81	Low	Mixed	170	200	370	205
57+26	83	Low	Mixed	165	245	410	145
37+48	85	Low	Mixed	175	175	350	200
29+64	93	Low	Mixed	200	195	395	150
44+28	72	Low	Even	165	225	390	220
48+26	74	Low	Even	175	245	420	200
34+48	82	Low	Even	205	175	380	170
38+44	82	Low	Even	195	165	360	170
44+38	82	Low	Even	165	195	360	170
48+34	82	Low	Even	175	205	380	170
26+58	84	Low	Even	245	170	415	215
28+56	84	Low	Even	225	185	410	215
46+38	84	Low	Even	220	195	415	215
34+58	92	Low	Even	205	170	375	245
54+38	92	Low	Even	195	195	390	245
38+56	94	Low	Even	195	185	380	165
19+36	55	High	Mixed	275	260	535	235
18+43	61	High	Mixed	365	130	495	165
32+29	61	High	Mixed	260	200	460	165
27+36	63	High	Mixed	240	260	500	160
46+17	63	High	Mixed	220	290	510	160
47+24	71	High	Mixed	230	265	495	240
47+36	83	High	Mixed	230	260	490	145
64+19	83	High	Mixed	195	275	470	145
18+73	91	High	Mixed	365	145	510	235
67+24	91	High	Mixed	195	265	460	235
68+24	92	High	Mixed	235	265	500	245
17+76	93	High	Mixed	290	190	480	150
34+18	52	High	Even	205	365	570	140
36+18	54	High	Even	260	365	625	195
22+34	56	High	Even	260	205	465	185
34+22	56	High	Even	205	260	465	185
36+26	62	High	Even	260	245	505	165
28+36	64	High	Even	225	260	485	195
26+46	72	High	Even	245	220	465	220
16+58	74	High	Even	325	170	495	200
36+46	82	High	Even	260	220	480	170
24+68	92	High	Even	265	235	500	245
66+26	92	High	Even	215	245	460	245
74+18	92	High	Even	200	365	565	245

Table E5.
Stimuli presented in Experiment 5b accompanied by familiarity ratings of Addend A (i.e., the first addend in the problem), Addend B (i.e., the second addend in the problem), the Problem (i.e., the familiarity of addend A plus Addend B) and the Answer for each problem.

Sum	Answer	Sum Type	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
10+30	40	Decade	445	320	765	300
20+30	50	Decade	370	320	690	330
30+20	50	Decade	320	370	690	330
20+40	60	Decade	370	300	670	290
40+20	60	Decade	300	370	670	290
50+10	60	Decade	330	445	775	290
20+50	70	Decade	370	330	700	245
30+40	70	Decade	320	300	620	245
40+30	70	Decade	300	320	620	245
50+20	70	Decade	330	370	700	245
60+10	70	Decade	290	445	735	245
10+70	80	Decade	445	245	690	260
20+60	80	Decade	370	290	660	260
30+50	80	Decade	320	330	650	260
50+30	80	Decade	330	320	650	260
60+20	80	Decade	290	370	660	260
70+10	80	Decade	245	445	690	260
10+80	90	Decade	445	260	705	360
20+70	90	Decade	370	245	615	360
30+60	90	Decade	320	290	610	360
40+50	90	Decade	300	330	630	360
50+40	90	Decade	330	300	630	360
60+30	90	Decade	290	320	610	360
70+20	90	Decade	245	370	615	360
15+25	40	Fives	345	370	715	300
25+15	40	Fives	370	345	715	300
15+35	50	Fives	345	340	685	330
35+15	50	Fives	340	345	685	330
15+45	60	Fives	345	270	615	290
25+35	60	Fives	370	340	710	290
35+25	60	Fives	340	370	710	290
45+15	60	Fives	270	345	615	290
15+55	70	Fives	345	235	580	245
25+45	70	Fives	370	270	640	245
45+25	70	Fives	270	370	640	245
55+15	70	Fives	235	345	580	245
15+65	80	Fives	345	200	545	260
25+55	80	Fives	370	235	605	260
35+45	80	Fives	340	270	610	260
45+35	80	Fives	270	340	610	260
55+25	80	Fives	235	370	605	260
65+15	80	Fives	200	345	545	260
15+75	90	Fives	345	360	705	360
25+65	90	Fives	370	200	570	360
35+55	90	Fives	340	235	575	360
55+35	90	Fives	235	340	575	360
65+25	90	Fives	200	370	570	360
75+15	90	Fives	360	345	705	360

Table E5 (continued)

Sum	Answer	Sum Type	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
15+20	35	Mixed	345	370	715	340
20+15	35	Mixed	370	345	715	340
15+30	45	Mixed	345	320	665	270
20+25	45	Mixed	370	370	740	270
30+15	45	Mixed	320	345	665	270
15+40	55	Mixed	345	300	645	235
20+35	55	Mixed	370	340	710	235
25+30	55	Mixed	370	320	690	235
35+20	55	Mixed	340	370	710	235
15+50	65	Mixed	345	330	675	200
25+40	65	Mixed	370	300	670	200
40+25	65	Mixed	300	370	670	200
45+20	65	Mixed	270	370	640	200
20+55	75	Mixed	370	360	730	360
25+50	75	Mixed	370	330	700	360
30+45	75	Mixed	320	270	590	360
35+40	75	Mixed	340	300	640	360
50+25	75	Mixed	330	370	700	360
60+15	75	Mixed	290	345	635	360
45+40	85	Mixed	270	300	570	200
70+15	85	Mixed	245	345	590	200
40+55	95	Mixed	300	235	535	245
50+45	95	Mixed	330	270	600	245
65+30	95	Mixed	200	320	520	245

Table E6.
Stimuli presented in Experiment 6a accompanied by familiarity ratings of Addend A (i.e., the first addend in the problem), Addend B (i.e., the second addend in the problem), the Problem (i.e., the familiarity of addend A plus Addend B) and the Answer for each problem.

Sum	Answer	Sum Familiarity group	Answer Familiarity group	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
23+29	52	Low	Low	200	200	400	140
26+23	49	Low	Low	245	200	445	175
27+29	56	Low	Low	240	200	440	185
28+23	51	Low	Low	225	200	425	160
28+29	57	Low	Low	225	200	425	165
34+33	67	Low	Low	205	215	420	195
34+39	73	Low	Low	205	190	395	145
37+39	76	Low	Low	175	190	365	190
38+39	77	Low	Low	195	190	385	200
42+43	85	Low	Low	220	130	350	200
42+44	86	Low	Low	220	165	385	150
44+43	87	Low	Low	165	130	295	145
44+49	83	Low	Low	165	175	340	145
46+48	94	Low	Low	220	175	395	165
47+49	96	Low	Low	230	175	405	200
48+49	97	Low	Low	175	175	350	160
23+27	50	Low	High	200	240	440	330
26+29	55	Low	High	245	200	445	235
28+27	55	Low	High	225	240	465	235
32+37	69	Low	High	260	175	435	325
33+37	70	Low	High	215	175	390	245
34+37	71	Low	High	205	175	380	240
38+33	71	Low	High	195	215	410	240
38+37	75	Low	High	195	175	370	360
42+48	90	Low	High	220	175	395	360
42+49	91	Low	High	220	175	395	235
43+47	90	Low	High	130	230	360	360
44+46	90	Low	High	165	220	385	360
44+47	91	Low	High	165	230	395	235
44+48	92	Low	High	165	175	340	245
46+49	95	Low	High	220	175	395	245
48+47	95	Low	High	175	230	405	245
14+17	31	High	Low	320	290	610	140
14+19	33	High	Low	320	275	595	215
16+17	33	High	Low	325	290	615	215
16+18	34	High	Low	325	365	690	205
18+13	31	High	Low	365	350	715	140
18+19	37	High	Low	365	275	640	175
22+26	48	High	Low	260	245	505	175
22+27	49	High	Low	260	240	500	175
22+29	51	High	Low	260	200	460	160
24+27	51	High	Low	265	240	505	160
24+28	52	High	Low	265	225	490	140
24+29	53	High	Low	265	200	465	125
26+27	53	High	Low	245	240	485	125
26+28	54	High	Low	245	225	470	195
32+33	65	High	Low	260	215	475	200

Table E6 (continued)

Sum	Answer	Sum Familiarity group	Answer Familiarity group	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
32+34	66	High	Low	260	205	465	215
12+17	29	High	High	315	290	605	200
13+17	30	High	High	350	290	640	320
13+19	32	High	High	350	275	625	260
14+13	27	High	High	320	350	670	240
14+16	30	High	High	320	325	645	320
14+18	32	High	High	320	365	685	260
16+13	29	High	High	325	350	675	200
16+19	35	High	High	325	275	600	240
17+19	36	High	High	290	275	565	260
18+17	35	High	High	365	290	655	240
22+23	45	High	High	260	200	460	270
22+28	50	High	High	260	225	485	330
32+38	70	High	High	260	195	455	245
32+39	71	High	High	260	190	450	240
34+36	70	High	High	205	260	465	245
36+33	69	High	High	260	215	475	325

Table E7.

Stimuli presented in Experiment 6b accompanied by familiarity ratings of Addend A (i.e., the first addend in the problem), Addend B (i.e., the second addend in the problem), the Problem (i.e., the familiarity of addend A plus Addend B) and the Answer for each problem.

Sum	Answer	Sum Type	Familiarity Addend A	Familiarity Addend B	Problem Familiarity	Answer Familiarity
10+50	60	Decade	445	330	775	290
10+60	70	Decade	445	290	735	245
10+70	80	Decade	445	245	690	260
10+80	90	Decade	445	260	705	360
20+40	60	Decade	370	300	670	290
20+50	70	Decade	370	330	700	245
20+60	80	Decade	370	290	660	260
20+70	90	Decade	370	245	615	360
30+40	70	Decade	320	300	620	245
30+50	80	Decade	320	330	650	260
30+60	90	Decade	320	290	610	360
40+50	90	Decade	300	330	630	360
15+25	40	Fives	345	370	715	300
15+35	50	Fives	345	340	685	330
15+45	60	Fives	345	270	615	290
15+55	70	Fives	345	235	580	245
15+65	80	Fives	345	200	545	260
15+75	90	Fives	345	360	705	360
25+35	60	Fives	370	340	710	290
25+45	70	Fives	370	270	640	245
25+55	80	Fives	370	235	605	260
25+65	90	Fives	370	200	570	360
35+45	80	Fives	340	270	610	260
35+55	90	Fives	340	235	575	360
10+55	65	Mixed	445	235	680	200
10+75	85	Mixed	445	360	805	200
20+35	55	Mixed	370	340	710	235
20+55	75	Mixed	370	235	605	360
20+75	95	Mixed	370	360	730	245
30+25	55	Mixed	320	370	690	235
30+45	75	Mixed	320	270	590	360
40+25	65	Mixed	300	370	670	200
40+45	85	Mixed	300	270	570	200
50+25	75	Mixed	330	370	700	360
50+45	95	Mixed	330	270	600	245
60+25	85	Mixed	290	370	660	200

CARDIFF UNIVERSITY ★ PRIFYSGOL CAERDYDD ★