The National University of Ireland

NUI MAYNOOTH

Ollscoil na hÉireann Má Nuad

Mathematics Experience and Format-specific Effects in Numerical Cognition

Thesis submitted to the Department of Psychology, Faculty of Science and Engineering, in fulfilment of the requirements for the degree of Doctor of Philosophy, National University of Ireland Maynooth

Justé Anca Koller B.A. (Hons)

August 2012

Research Supervisor: Dr Fiona Lyddy

Head of Department: Dr Fiona Lyddy

Table of Contents

Acknowledgements .... **i**
Publications and Presentations Based on this Work .... **ii**
Abstract .... **iii**
List of Figures .... **vi**
List of Tables .... **viii**
Glossary of Terms .... **x**

Chapter 1 .... **1**
General Introduction **1**

1.1. The Relationship between Numerical Cognition and Language 3
1.2. The Dissociation between Language and Numerical Abilities 8
1.3. Development of the Number Symbol–Concept Relation 9
1.4. The Influences of Surface Format on the Core Numerical System 12
1.5. Theoretical Accounts of Format-independent Number Representation 15
1.6. Theoretical Accounts of Format-specific Number Representation 16
1.7. Cognitive Interference: Processing Differences between Digits and Words 18
1.8. Cognitive Interference Modelled by the Size Congruity and Symbolic Distance Effects 23
1.9. Automaticity of Processing and Cognitive Interference 26
1.10. Individual Differences in Number Processing 27
1.11. The Current Research 31

Chapter 2
Experiment 1: Cognitive Interference in a Digit–Word Counting Task: The Role of Mathematics Experience in Format-specific Processing **37**

2.1. Introduction 37
2.2. Method 42
2.2.1. Participants 42
2.2.2. Apparatus and Materials 42
2.2.3. Procedure 46
2.2.4. Ethical Considerations 50
2.3. Results 51
2.3.1. Accuracy 52
2.3.2. Reaction Times 53
2.3.3. Interference 55
2.3.4. Facilitation 56
2.4. Discussion 57

Chapter 3
Experiment 2: Size Congruity and Distance Effects in a Digit–Word Number Comparison Task at Different Levels of Mathematics Experience 60

3.1. Introduction 60
3.2. Method......................65
3.2.1. Participants 65
3.2.2. Apparatus and Materials......................65
3.2.3. Procedure......................67
3.3. Results 69
3.3.1. Accuracy......................70
3.3.2. Congruency 73
3.3.3. Distance Effects......................78
3.4. Discussion 85

Chapter 4
Experiment 3: Congruency Effects in a Quantifier Word Counting Task as a Function of Mathematical Experience 89

4.1. Introduction 89
4.2. Method......................93
4.2.1. Participants 93
4.2.2. Apparatus and Materials......................93
4.2.3. Procedure......................95
4.3. Results 97
4.3.1. Accuracy......................97
4.3.2. Reaction Time 98
4.3.3. Interference and Facilitation......................100
4.4. Discussion 101

Chapter 5
Experiment 4: Format-specific Effects in Arithmetic at Different Levels of Mathematics Experience: Evidence from Eye-tracking 107

5.1. Introduction 107
5.2. Method......................116
5.2.1. Participants 116
5.2.2. Apparatus and Materials......................118
5.2.3. Procedure......................121
5.2.4. Ethical Considerations......................123
5.3. Results 123
5.3.1. Part 1: Addition and Subtraction 124
5.3.2. Part 2: Multiplication and Division 134
5.4. Discussion 144

Chapter 6
Experiment 5: The Interaction between Encoding and Answer-retrieval Stages of Arithmetic: Format and Operation Effects at Different Levels of Mathematics Experience 149

6.1. Introduction 149
6.2. Method 154
6.2.1. Participants 154
6.2.2. Materials and Apparatus 155
6.2.3. Electroencephalography (EEG) materials and procedure 156
6.2.4. Procedure 157
6.2.5. Ethical Considerations 159
6.3. Results 160
6.3.1. Behavioural data 160
6.3.1.1. Accuracy 160
6.3.1.2. Reaction Time (RT) 161
6.3.2. Event-related Potentials (ERPs) 162
6.3.2.1. Amplitude 163
6.4. Discussion 175

Chapter 7 181
General Discussion 181

7.1. Overview 181
7.2. Format-specific Encoding 182
7.3. Format-specific Arithmetic 185
7.4. The Influence of Mathematics Experience on Format-specific Numerical Cognition 192
7.5. Implications of the Current Research 198
7.6. Limitations of the Current Research and Outlook for Future Research 200
7.7. Conclusions 202

References 204
Appendices 219

Acknowledgements

This research was funded by the Irish Research Council under the Embark Initiative.

I am very grateful to a number of people who have supported me in the completion of this work. Firstly, I would like to thank my supervisor, Dr Fiona Lyddy, whose guidance and complete support throughout my research has been incredibly helpful. I am also very grateful to Prof. Stephen Buckley from the Department of Mathematics and Statistics for his supervisory role and advice with regards to mathematics.

For his involvement in the preparation and analyses of the EEG experiment, I am very thankful to Dr Richard Roche. I would like to sincerely thank Mr Derek Walsh for providing technical support with the eye-tracking and EEG equipment. Thanks are also due to Siliang Tang for his advice on eye-tracking.

Finally, I would like to thank all the participants who took time out of their schedules to participate in my experiments.

Publications and Presentations Based on this Work

Koller, J. & Lyddy, F. (under review). Self-efficacy and working memory capacity as predictors of numeracy.

Koller, J. & Lyddy, F. (2011). Operation specific effects of number format as a function of math experience: an eye tracking study. Paper presented at the Psychological Society of Ireland's 41st Annual Conference, Galway, November, 2011.

Koller, J. & Lyddy, F. (2010). Interference from reading in a digit–word counting task: the role of mathematical experience in digit processing. In Concannon-Gibney, T., Kelly, A. & Willoughby, K. (Eds.) *Literacy in the 21st century: Perspectives, challenges and transformation.* Dublin: Reading Association of Ireland.

Koller, J. & Lyddy, F. (2010). Congruency effects in a quantifier word counting task as a function of mathematical experience. Paper presented at the Psychological Society of Ireland's 40th Annual Conference, Athlone, November 2010.

Koller, J. & Lyddy, F. (2009). Understanding number: symbols, language and numerical cognition. *Reading News*, 3-7.

Koller, J. & Lyddy, F (2009). Numerical and Physical distance effects in a digit–word interference task. Poster presented at the Psychological Society of Ireland's 39th Annual Conference, Wexford, November 2009.

Koller, J. & Lyddy, F. (2009). Stroop Interference and Facilitation in a numerical word–digit task as a function of mathematical education. Paper presented at the Reading Association of Ireland's Annual Conference, Dublin, September 2009.

Abstract

A persistent issue in numerical cognition research is how the format of numerical information influences numerical processing. The format-independent view postulates that information from various formats (e.g. ‘3’ or ‘three’) is represented in a uniform numerical code and that format should thus have no influence on number manipulation. The format-specific view assumes separate representational pathways for arabic digits and number words, which come into play during number processing as well as manipulation. Five experiments were undertaken with methods ranging from behavioural measures of reaction time to more refined measures of cognitive processes such as eye-tracking and event-related potentials (ERPs). In each experiment, effects of format were investigated at different levels of mathematics experience, in order to examine how the processing of numbers might differ in this regard.

The first three experiments focused on basic number processing and processing differences that can occur for arabic digits, number words and quantifier words. In Experiment 1, a modified counting Stroop task was employed to investigate cognitive interference of arabic digits and number words. Participants took longer to respond on incongruent trials (e.g. **4 4 4**; how many numbers are present? Correct response: ‘3’) relative to neutral (e.g. *** * ***; Correct response: ‘3’) and congruent (e.g. **3 3 3**; Correct response: ‘3’) trials. Individuals with high mathematics experience showed greater interference on digit trials, whereas no effect of mathematics experience was found for word trials (e.g. **three three**; respond ‘2’). This suggests that the influence of format on number processing can be regulated by mathematics experience.

Experiment 2 investigated this effect further by considering numerical (e.g. **5 2**; which number is higher?) and physical size (e.g. **5 2**; which number is physically bigger?) comparisons of digit and word stimuli. For both formats, participants responded faster on trials with a large numerical distance (e.g. **2 7**) compared to trials with a small numerical distance (e.g. **2 3**) suggesting that specific number meanings are accessed spontaneously from digits and number words, however the size congruity effect only occurred for digit stimuli. Individuals with greater mathematics experience showed an overall advantage for numerical comparison, regardless of format.

Based on the findings from Experiments 1 and 2, Experiment 3 modified the counting Stroop task (Experiment 1) to investigate if mathematics experience can be related to the processing of quantifier words (e.g. many, few, each). Stimuli were presented as either specific (e.g. **both both**; correct response '2') or general (e.g. **some some**) quantifier words and participants were required to count the items on-screen. While the effects were minimal in comparison with Experiment 1, any effects related to the congruity of the stimuli only emerged for the highly mathematics experienced participants, suggesting the involvement of number experience in quantifier word processing, and in turn for extracting number meaning from language in general.

As the first three experiments demonstrated format-specific effects in basic number processing, the second part of the thesis investigated these effects for more advanced numerical processing such as arithmetic. The second part of the thesis also employed more refined measures of cognitive processing (eye-tracking and event-related potential [ERP] technology) to investigate effects that might not be evident from behavioural data alone. Experiment 4 employed eye-tracking technology to compare effects of problem size, operation and format at different levels of mathematics experience. Fixation patterns supported the format-specific view of number processing by suggesting that in comparison with digit-format, word-format impeded the use of direct memory retrieval in arithmetic, an effect that seemed to be more pronounced for individuals with low mathematics experience. Eye-tracking data also supported behavioural data as well as self-report data that have been noted in reports on strategy use in arithmetic. From this, inferences were made regarding the degree to which surface format influences subsequent calculation processes and how this might be moderated by mathematics experience.

Experiment 5 investigated the interaction between the encoding and answer-retrieval stages in digit- and word-format arithmetic by separating the presentation of the first operand and the rest of the equation in a true–false verification task (e.g. '**3**' and '**x 4 = 12**'; correct response 'true'). Before each test block, participants were told which operation was to follow (addition or multiplication). ERP findings suggested that operands presented in the same format were encoded in the same way, with effects of operation only emerging during the second part of the equation, after participants had seen the operation sign ('+' or 'x'). Regardless of format, the High Maths group showed greater left anterior potentials for multiplication than addition, suggesting an advantage for arithmetic fact retrieval.

In the final chapter of the thesis the findings are discussed in relation to existing theoretical accounts on the influence of format in numerical cognition, with specific focus on the benefit of considering mathematics experience in this regard.

List of Figures

Figure 2.1. *Mean RTs (± SEM) across congruent, neutral and incongruent stimuli for (a) the Low Maths (n = 20) and (b) the High Maths group (n = 20).* 53

Figure 2.2. *Mean disparity in RTs between congruent and neutral (facilitation, shown above the x axis) and incongruent and neutral (interference, shown below the x axis) conditions in the Low and High Maths groups for digit and word stimuli (± SEM).* 56

Figure 3.1. *Mean RTs (± SEM) in the physical comparison task across congruent, neutral and incongruent trials for (a) the Low Maths (n = 24) and (b) the High Maths group (n = 20).*....................74

Figure 3.2. *Mean RTs (± SEM) in the numerical comparison task across congruent, neutral and incongruent trials for (a) the Low Maths (n = 24) and (b) the High Maths group (n = 20).* 76

Figure 3.3. *Distance effects in the physical comparison task across task-relevant (physical distance) and -irrelevant (numerical distance) dimensions for (a) the Low Maths and (b) the High Maths group (± SEM).*....................79

Figure 3.4. *Distance effects in the numerical comparison task across task-relevant (numerical distance) and -irrelevant (physical distance) dimensions for (a) the Low Maths and (b) the High Maths group (± SEM).*....................83

Figure 4.1. *Mean RTs (± SEM) across congruent, neutral and incongruent trials for One, Two and General Quantifier words for (a) the Low Maths (n = 15) and (b) the High Maths (n = 14) group.*....................99

Figure 5.1. *Presentation examples of digit and word format equations* 120

Figure 5.2. *Mean RTs across format, problem size and maths group (Low, Middle or High) in (a) the addition and (b) the subtraction task (± SEM).* 126

Figure 5.3. *Mean RTs across format, problem size and maths group (Low, Middle or High) in (a) the multiplication and (b) the division task (± SEM).*..........136

Figure 6.1. *Examples of true digit and word format equations.* 159

Figure 6.2. *Mean RTs (± SEM) across format and operation for the Low and High Maths groups.*....................161

Figure 6.3. *The 32-channel ERP montage showing the scalp locations of electrodes F3, F4, P3 and P4.* 163

Figure 6.4. *Grand mean waveforms of the Low Maths group over the left and right anterior and posterior scalp in the time window 0 – 60 ms and 70 – 140 ms post-stimulus for the first operand in digit format equations.* 165

Figure 6.5. *Grand mean waveforms of the Low Maths group over the left and right anterior and posterior scalp in the time window 0 – 60 ms and 70 – 140 ms post-stimulus for the first operand in word format equations.* 166

Figure 6.6. *Grand mean waveforms of the High Maths group over the left and right anterior and posterior scalp in the time window 0 – 60 ms and 70 – 140 ms post-stimulus for the first operand in digit format equations.* 167

Figure 6.7. *Grand mean waveforms of the High Maths group over the left and right anterior and posterior scalp in the time window 0 – 60 ms and 70 – 140 ms post-stimulus for the first operand in word format equations.* 168

Figure 6.8. *Event-related potentials of the Low Maths group over the left and right anterior and posterior scalp in the time windows 100 – 180 ms and 270 – 440 ms post-stimulus for the second part of digit format equations.* 170

Figure 6.9. *Event-related potentials of the Low Maths group over the left and right anterior and posterior scalp in the time windows 100 – 180 ms and 270 – 440 ms post-stimulus for the second part of word format equations.* 171

Figure 6.10. *Event-related potentials of the High Maths group over the left and right anterior and posterior scalp in the time windows 100 – 180 ms and 270 – 440 ms post-stimulus for the second part of digit format equations.* 172

Figure 6.11. *Event-related potentials of the High Maths group over the left and right anterior and posterior scalp in the time windows 100 – 180 ms and 270 – 440 ms post-stimulus for the second part of word format equations.* 173

List of Tables

Table 2.1. *Mean percentages of errors across congruent, neutral and incongruent conditions for the Low and High Maths groups.* 52

Table 2.2. *Means and standard deviations of RTs on congruent, neutral and incongruent conditions for the Low and High Maths groups.* 54

Table 3.1. *Mean percentages of errors across congruent, neutral and incongruent conditions in the physical and numerical comparison tasks.* 71

Table 3.2. *Means and Standard Deviations of RTs on incongruent trials in the Physical Comparison Task for the Low (L) and High (H) Maths groups.* 82

Table 3.3. *Means and Standard Deviations of RTs on incongruent trials in the Numerical Comparison Task for the Low (L) and High (H) Maths groups.* .82

Table 4.1. *Mean percentages of errors for the Low and High Maths groups across congruent, neutral and incongruent quantifier word trials.* 98

Table 5.1. *Mean percentages of errors in the addition and subtraction tasks.* ..125

Table 5.2. *Mean RT disparity (ms) between small and large problems (problem size effect) in the addition and subtraction tasks.* 128

Table 5.3. *Means and standard deviations of number of fixations in the addition and subtraction tasks* 130

Table 5.4. *Mean disparity in number of fixations between small and large problems (fixation count problem size effect) in the addition and subtraction tasks.* 131

Table 5.5. *Means and standard deviations of fixation duration in the addition and subtraction tasks.* 133

Table 5.6. *Mean disparity in fixation duration between small and large problems (problem size effect) in the addition and subtraction tasks.* 134

Table 5.7. *Mean percentages of errors in the multiplication and division tasks.* 135

Table 5.8. *Mean disparity in RT between digit and word format problems in the multiplication and division tasks.* 138

Table 5.9. *Means and standard deviations of number of fixations in the multiplication and division tasks.* 139

Table 5.10. *Mean discrepancy in number of fixations between digit and word format problems in the multiplication and division tasks.* 141

Table 5.11 *Means and standard deviations of fixation durations in the multiplication and division tasks.* 143

Glossary of Terms

Additive Viewpoint of Arithmetic: The view that in arithmetic, the two stages of operand encoding and answer-retrieval operate independently of one another. Once numbers from different formats have been encoded to underlying number meanings, any subsequent calculation processes are thus thought to operate independently of encoding.

Arithmetic Fact Retrieval: The process of retrieving answers to arithmetic equations from memory.

Automaticity of Processing: The degree to which the processing of a certain stimulus occurs automatically, even if it is instructed to be ignored under task demands.

Cognitive Interference: Where two stimulus features are processed simultaneously and the processing of one feature impedes the processing of the other. In the original colour Stroop task (Stroop, 1935), for example, this effect refers to the slowed response on trials where colour and word meaning mismatch (e.g. Blue; respond 'red').

Developmental Dyscalculia: A deficit in numerical processing that is specifically related to severe impairments in learning arithmetic.

Digit: The arabic numeral representation of a number (e.g. '4').

Distance Effect: In numerical comparison, the distance effect demonstrates that the time taken to compare two numbers is a function of the numerical distance between the two numbers. For example, it is easier to compare numbers that are numerically further apart (e.g. **2** vs. **9**) than numbers that are numerically close (e.g. **2** vs. **3**).

Electroencephalography (EEG): A technique used to study the electrical activity of the brain that can be measured by placing electrodes on the scalp.

Encoding: The process of accessing number meanings from symbolic numerical notations (e.g. '**3**' or '**three**').

Event-related Potentials (ERPs): Variations in amplitude that reflects changes in brain activation in response to specific stimuli.

Format-independent Processing: The view that numerical information from various different formats is translated to a uniform amodal number representation and that similar processing takes place for different numerical formats (see also the additive view of arithmetic).

Format-specific Processing: The view that different symbolic numerical notations (e.g. arabic numerals and number words) are processed along separate pathways and not necessarily translated to a uniform amodal representation for numbers from all formats.

Interactive Viewpoint of Arithmetic: The view that the stages of operand encoding and answer-retrieval in arithmetic interact with one another, such that encoding conditions, such as operand format, have a direct influence on answer-retrieval strategies (see also format-specific processing).

Numeracy: Proficiency with basic numerical and probability concepts and the ability to apply these skills to real-world scenarios.

Numerical Distance: The numerical difference between two numbers on a number line. The numbers '1' and '3', for example have a numerical distance of '2'. The numbers '1' and '6' have a numerical distance of '5'.

Numerosity: The number of the objects in a collection.

One-To-One Correspondence: The process of matching the items in one set with the items of a second set so that each item is paired with one other item.

Operation Effect: The differences in performance or brain activation between addition, subtraction, multiplication or division.

Physical Distance: The difference in physical size between two stimuli.

Problem Size Effect (PSE): The increase in errors and response time in arithmetic as the magnitude of the operands in an equation increases.

Processing Bias: More automatic processing that develops for certain stimuli due to extensive practice, memory and exposure.

Size Congruity Effect: The increase in response time when comparing two numbers and number meaning is incongruent with the physical sizes of the numbers (e.g. **2** **5**; which number is numerically higher?).

Stroop Facilitation: The faster response on congruent trials in Stroop tasks. This occurs in the counting Stroop task, for example, when number meaning matches the number of items to be counted (e.g. **3 3 3**; respond '3').

Stroop Interference: The slowed response on incongruent trials in Stroop tasks. This occurs, in the counting Stroop task, for example, when number meaning mismatches the number of items to be counted (e.g. **4 4 4**; respond '3').

Subitizing: The ability to quickly and spontaneously perceive the number of items presented in a small set (up to 3 or 4 items). This process differs from the more effortful process of counting larger sets of objects.

Task-irrelevant Stimulus Features: The features in Stroop tasks that are to-be-ignored under task demands.

Task-relevant Stimulus Features: The features in Stroop tasks that are to-be-attended to under task demands. For example, in the counting Stroop task, the number of items is task-relevant and number meanings are task-irrelevant (to-be-ignored).

Transcoding: The process of reading, writing and understanding numbers from various symbolic formats.

Chapter 1

General Introduction

The study of numerical cognition involves the understanding of numbers and the mental processes involved in number representation, manipulation and calculation. Numerical cognition research aims to understand numerical processing, and also to highlight the features that aid numerical competence and proficiency. Since our concept of numbers form such a central part of every day life, the origins of our 'number sense' is of great interest to theorists (Dehaene, 1997). Leading theorists agree that people seem to possess an innate sense of number (e.g. Butterworth, 1999; Dehaene, 1997) that can be likened to spontaneous cognitive processes such as colour perception. Knowledge of symbolic number (e.g. '3' or 'three') builds on this basic number sense and enables more complex mathematical functions (e.g. mental arithmetic).

The assumption that numerical cognition is closely linked to language has influenced many theories of number processing, and in turn, the hypothesis that numerical abilities emerge from linguistic abilities (Dehaene, 1992). Whereas studies of animal and infant numerical cognition (e.g. Boysen & Capaldi, 1993; Wynn, 1992) suggest a language-independent sensitivity to number, symbolic numerical representation is essential for complex, uniquely human, numerical functions (e.g. Dehane, 1997). Successful numerical cognition thus require reading, writing and understanding numbers in various different formats, a set of skills referred to as transcoding (Dehaene, 1997). In adulthood, the human brain constantly transcodes between numerical formats, reflecting a long learning history of associating certain symbols with certain concepts (e.g. Butterworth, 1999; Deacon, 1997). In time, extensive practice and memory allow number meanings to be accessed automatically from different numerical symbols (e.g. Bush, Whalen, Shin & Rauch, 2006). The building blocks of competent numerical cognition in adulthood are thus basic numerical skill (e.g. Halloway & Ansari, 2009; Kaufman, Handl, & Thoeny, 2003) and knowledge of linking symbolic numerical formats with underlying number concepts (e.g. Dehane, 1997; Gilmore, McCarthy & Spelke, 2007).

A central debate in numerical cognition research is how different numerical symbolic notations (e.g. ‘3’ versus ‘three’) influence the manipulation of numbers. Some theorists such as McCloskey and Macaruso (1995) argue that all numbers are represented in an underlying uniform code regardless of their symbolic input (e.g. ‘3’ or ‘three’), and that the same processes therefore take place for the manipulation of numbers presented in different formats. Others, such as Campbell and Clark (1988; see also Campbell & Alberts, 2009) argue for format-specific number representations and that the surface format directly influences number processing and calculation. Since evidence in support of both views exists (e.g. Campell & Alberts, 2009; Zhou, 2011) there is still debate on where, and under which task demands, surface format is most influential. In favour of the format-specific view, Campbell and Alberts (2009) argued that arithmetic performance reflects experience and practice with the operand format in question, which suggests the potential utility of considering individual differences in this regard. However, the research on adult numerical cognition to date has not considered individual differences related to mathematics experience and how it might regulate the influence of format in numerical cognition. Furthermore, as most of the evidence in support of the format-specific and format-independent views comes from studies of arithmetic, it seems that more basic number processing has been overlooked in such studies. The current thesis investigated symbolic numerical processing and manipulation at different levels of mathematics experience in adulthood. With the aim to provide a clearer view of how numerical information is accessed from different symbolic formats, effects of format were investigated for more basic processes such as counting or number comparison, as well as more advanced processes such as arithmetic. Effects of format can serve to identify the extent to which numerical concepts are processed independently from input format (e.g. Campbell & Alberts, 2009). By including a wide range of tasks, such effects can be informative as regards the cognitive architecture of calculation processes, as well as basic symbolic numerical representation (Bassok, 2001; Campbell & Alberts, 2001; Landy & Goldstone, 2007).

1.1. The Relationship between Numerical Cognition and Language

A long-standing issue in the field of numerical cognition has been whether it is our capacity for language that allows us to manipulate numbers or whether these skills function independently from language. Theoretical accounts differ in this regard, reflecting underlying differences in how the language–concept relationship is viewed. Some adopt a strong Whorfian hypothesis (e.g. Simon, 1997) and others argue that language only facilitates certain aspects of numerical cognition (e.g. Dehaene, 1997). Inferences made from psycholinguistic research adapted for the study of numerical cognition strongly depend on which view is supported. Studies of Amazonian tribes, for example, whose languages lack counting words (Saxe, 1981;

Wassmann & Dasen, 1994), demonstrate that language differences do not necessarily predict conceptual differences. However, processing differences between different numerical formats (e.g. arabic digits or number words; e.g. Roelofs, 2006) illustrate that certain symbols can activate underlying number meanings more readily than others, and the strength of this concept–symbol connection can influence subsequent information processing (e.g. Dehaene, 1997).

While symbolic numerical representation is a uniquely human characteristic, this ability is thought to stem from a core numerical knowledge system common to animals, infants and human adults. Considering evolutionary and developmental evidence, theorists such as Hauser and Spelke (2004) argue that core knowledge systems evolved to form the basis for these advanced knowledge systems that are exclusive to humans. Regarding numerical cognition, these domains involve an exact system for representing small magnitudes and an approximate system for representing large magnitudes (Dehaene, 1997; Hauser & Spelke, 2004). Language, however, does not seem to underpin the number representation that human adults share with pre-verbal infants and non-human primates. Departing from the cause and effect view of the language–thought relationship, symbolic notation is rather thought to organise core knowledge systems into meaningful relationships (Gleitman & Papafragou, 2005), and thus aids the development of formal knowledge (e.g. the knowledge of natural numbers).

On the other hand, some theorists hold that numerical concepts rely exclusively on language and culture (e.g. Simon, 1997) and that the human brain has evolved to process many forms of magnitude and not numerical information specifically. Such views emphasise the importance of counting words in numerical competence and argue that exact number discrimination depends on language. Some animal species and pre-verbal infants do, however, engage in exact number processing, in the absence of language (see Boysen & Capaldi, 1993; Gallistel & Gelman, 1992, Wynn, 1992).

Studies of Amazonian tribes whose language does not possess counting words are often cited by theorists who postulate the central involvement of language in number development (Dehaene, 1997). In the absence of counting words, it is predicted that children in these cultures will not develop a true concept of numerosity. However, Dehaene (1997), for example, noted that to solve calculations, pupils in a New Guinea school often pointed to different parts of their bodies, which represent different numbers. The representation of numbers can thus circumvent number words (Saxe, 1981; Wassmann & Dasen, 1994) so that language differences need not necessarily reflect conceptual differences. As Gelman and Butterworth (2005) point out, cultural differences in such studies, which were unaccounted for, could also have contributed to the differences in performance.

If our concept of number is thought of as an innate perceptual sense, which can be likened to automatic processes such as colour perception or spatial awareness (Spelke & Dehaene, 1999), language should not be necessary for this system to exist. Our sensitivity to numerical quantities, indeed, seems to be an automatic perceptual process represented in processing pathways in the inferior parietal cortex (Dehaene, Molko, Cohen & Wilson, 2004). Evidence from animal and infant numerical discrimination also provides compelling evidence for this innate ability to process number. This effortless process, thought to be the basis of the core analogue numerical stream, allows some animals to perform simple numerical discriminations (Boysen & Capaldi, 1993; Gallistel & Gelman, 1992).

Further support for the existence of a language-independent number sense comes from studies which showed that infants can discriminate between (small) numbers of objects, actions and sounds (e.g. Antell & Keating, 1983; Starkey & Cooper, 1980; Starkey, Spelke & Gelman, 1990; Wynn, 1996). Wynn (1992; 1996) reported a number of experiments, which suggested that infants could correctly anticipate simple addition and subtraction problems, a finding which has been widely replicated (Baillargeon, 1994; Koechlin, Dehaene & Mehler, 1997; Simon, Hespos & Rochat, 1995; Wynn, Bloom & Chiang, 2002). In using a habituation paradigm, Wynn's experiments showed that five-month-olds were sensitive to changes in the number of objects presented visually. The habituation paradigm, a robust measure of infants' expectations in visual perception (Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981), relies on an infant's tendency to look for longer at certain stimuli than at others (e.g. a new or unexpected change in the visual field). In an experiment that utilised a 1+1 operation, for example, the infant firstly saw an object being placed on a platform and then an upward rotating screen hid the object from view. After this, the infant saw a hand placing another identical object behind the screen and an empty hand leaving the stage. When the screen came down, the platform either contained the correct number (two objects) or the incorrect number (one object) of objects (Wynn, 1992). Infants tended to look for significantly longer at incorrect (unexpected) compared to correct (expected) outcomes. The infants' sensitivity to number was also found to be quite specific, suggesting that numerical discrimination took place. In an experiment which used the 1+1 operation, infants looked significantly longer at an outcome of three objects than an outcome of two objects. It thus seems that the infants were not simply expecting the display to contain more objects than it initially did, but rather that they engaged in more precise numerical discrimination (Wynn, 1992).

In favouring the view that numerical concepts are exclusively language and culture-dependent, Simon (1997) argued that infants' apparent numerical ability reflects mere surprise at a change in the visual scene, as opposed to actual numerical discrimination. Koechlin et al. (1997) also posed the question of whether or not the infants were only sensitive to the spatial locations of objects instead of the specific number of objects. Their study, which provided evidence against this argument, involved objects placed on a rotating plate located on the platform behind the screen. The same results as in Wynn's studies were obtained, suggesting that the infants did not merely look longer at the presence of an object in an unexpected location. Instead, the infants seemed to be particularly sensitive to the numerosity of the display, supporting the argument of infant numerical discrimination. Similar observations have also been demonstrated with animals (e.g. Boysen & Capaldi, 1993; Gallistel & Gelman, 1992), which supports the view of an innate language-independent numerical system, common to infants, human adults and some animal species. Importantly, the apparent precision of the infants' numerical discrimination suggests that humans have an innate ability for numerical processing per se, which seems to be independent from spatial and language processing.

1.2. The Dissociation between Language and Numerical Abilities

Since accurate manipulation of specific numbers often seems to occur even when specific words for these numbers are unavailable, the difference between the analogue and specific core numerical systems should not be viewed as entirely language-based (Gelman & Butterworth, 2005). The analogue and specific core systems are thought to activate separate brain mechanisms under certain circumstances, but are not normally mutually exclusive (Stanescu-Cosson, Pinel & Van de Moortlele et al., 2000). When a difficult calculation is performed, for example, the two systems are activated in order to perform the operation. However, language-based representations seem to be essential in order to perform operations beyond the number three (Dehaene, 1997; Gelman & Gallistel, 1978; Saxe, 1981). In a review of experimental and neuroscientific evidence, Gelman and Butterworth (2005) address this issue and support the argument that numerical concepts have neural and developmental roots that are language-independent.

If numerical abilities operate independently of language function, a dissociation might be predicted between language and numerical cognition. This was found by Butterworth, Cappelletti and Kopelman (2001) who described a patient with semantic dementia who had relatively spared numerical ability. Despite impaired semantic memory and reading of non-number words, the patient, I.H., was able to read and write most number words, and could transcode (a property thought to rely on language ability; see Dehaene, 1992) from written or spoken number words to arabic digits and vice-versa. He also had relatively spared calculation abilities despite compromised language function. Furthermore, since I.H. was severely impaired on other high frequency items, such as naming pictures of common objects (Butterworth et al., 2001), his preserved transcoding ability did not seem to reflect the fact that the association of certain numbers with certain words is highly practised. This observed dissociation between language and numerical cognition supports the view that number is represented in semantic memory as a domain-specific category (e.g. Caramazza & Shelton, 1998).

This argument is also supported by findings of other conditions, which shows a double dissociation of language function and numerical cognition. Despite having good language skills, children with William's syndrome perform poorly on relatively simple number tasks (Ansari & Karmiloff-Smith, 2002; Paterson, Girelli, Butterworth & Karmiloff-Smith, 2006; Udwin, Davies, & Howlin, 1996). The reverse effect is also found, with most children with developmental dyscalculia generally not showing language impairments (see for example Lewis, Hitch & Walker, 1994; Ostad, 1998). If numerical ability relied on language, children with such literacy deficits should not be expected to have intact numeracy and vice versa. Although co-morbidity of developmental numeracy and literacy deficits is relatively high, the data show that the majority of those with a literacy deficit have relatively spared numeracy (see Butterworth, 2005).

1.3. Development of the Number Symbol–Concept Relation

Despite having an innate language-independent number sense, people seem to possess the capacity to *spontaneously* link symbols with concepts, a uniquely human characteristic. Although various animal species such as chimpanzees (e.g. Boysen & Capaldi, 1993) and rats (e.g. Church & Meck, 1984) have proved capable of symbolic numerical representation, it is important to note that this was only achieved through extensive training and not spontaneously as is the case with humans. Humans are capable of derived performance and generalising rules and concepts to new situations (e.g. language function), a property which animal species lack (e.g. Deacon, 1997). Considering the concept of natural number, for example, even the most extensively trained chimpanzees fail to fully master this concept. Humans, however, seem to posses an early capacity for linking numbers with symbolic notations, a property that formal mathematics instruction builds on. Gilmore and colleagues (2007), for example, showed that in the absence of arithmetic instruction, children could perform simple symbolic calculations. Children were presented with the following problem, for example: “Sarah has fifteen sweets and she gets nineteen more. John has fifty-one sweets. Who has more?” The children’s answers were relatively accurate and did not seem to rely on guessing strategies. Performance was also as accurate as in research using similar problems in non-symbolic form (e.g. Barth, LaMont, Lipton & Spelke, 2005). This suggests that children are capable of translating between symbolic and non-symbolic numerical concepts, before they are able to represent exact numbers symbolically. Performance dropped once they were asked to provide an exact, as opposed to an approximate answer, suggesting that the children’s performance lies in the use of the non-symbolic number system to solve approximate symbolic problems (Barth et al., 2005).

This transition from the analogue representational system of infants (e.g. Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981; Wynn, 1992) to the explicitly trained language-based system, is needed in order to perform arithmetic on exact numbers that exceed the number three (Dehaene et al., 1999). Failure to make this transition from the analogue to the specific symbolic number representation seems to be associated with mathematical impairments. Rousselle and Noël (2007) found, for example, that children with mathematical learning difficulties only displayed impairments in conditions that employed arabic digits (symbolic number magnitude) compared to conditions that employed collections of items (non-symbolic number magnitude). Also, in a number Stroop task variant that compared the physical sizes of arabic numerals (e.g. **3** **7**; which font size is bigger?), children automatically seemed to access number magnitude, suggesting that the deficit does not lie in accessing number magnitude, but rather in specifically accessing number magnitude from symbols (Rousselle & Noël, 2006). There was also no evidence found for a difference in performance between children with mathematical learning difficulties and children with mathematical learning difficulties co-morbid with reading difficulties. It could thus be argued that the deficit is a more general learning impairment, specifically related to the association of certain meanings with certain symbols, rather than number per se. It is also worth mentioning that the association of words with number concepts can be more difficult for larger numbers. Dehaene (1997) notes that in the history of language development, naming the numerals 1 – 3 was probably as easy as naming perceptual properties such as ‘hot’ or ‘cold’. The fact that the brain processes ‘oneness’, ‘twoness’ or ‘threeness’ as effortlessly as other perceptual properties could thus make it easier to associate symbols with these number meanings. Beyond the number ‘three’, numerical meanings thus take on a less exact mental representation (Dehaene, 1997).

The importance of this association is also seen in infants' treatment of quantifier words (e.g. "some", "many", "both" etc.), a symbolic format that is not explicitly trained to relate to specific numerosities. In an exploration of singular/plural morphology in children's language acquisition, Barner, Chow and Yang (2009) found a significant correlation between quantifier knowledge and numeral knowledge. However, quantifier knowledge does not seem to facilitate the acquisition of numeral knowledge, supporting the argument that these two concepts develop independently, at least to some degree. Infants distinguish between numerals and other quantifiers early on in development and only assign exact meanings to numerals, using quantifier words to gather general information about the semantic qualities of a noun. Specifically, 3- to 5-year-olds only assigned an exact meaning to the word 'one', whereas the word 'a' took on a more general meaning, not necessarily relating to only one entity. Young children also often took the word 'some' to refer to a whole set of objects as opposed to just a portion, suggesting that children understand the core meanings of these quantifiers, but require extensive training to learn how these words contrast with other words (Barner et al., 2009).

1.4. The Influences of Surface Format on the Core Numerical System

Once the explicitly trained symbolic number system is in place, numerical information can mainly be represented in two formats, namely arabic digits and number words (Fias, Reynvoet & Brysbaert, 2001). According to Cohen, Dehaene and Verstichel (1994) there are no reasons to conclude that number words (e.g. one, two or three) are processed differently from other words. However, in light of the language-independent number sense that humans seem to possess (e.g. Dehaene, 1995), this might not be the case. Hurewitz, Papafragou, Gleitman and Gelman (2006) argue that acquiring number terms, compared to other words, could be especially difficult for young learners as they do not describe any individual properties in the environment, but refer to sets of objects. In other words, number words are more abstract than other words (Butterworth, 1999). Number words are also unique in the sense that they can conform to various different word classes, depending on context. Sometimes the word ‘two’, for example, is used as a noun and sometimes as an adjective (Frege, 1974) and the type of objects that are quantified also differ from situation to situation. People use the recursive property of language extensively in order to generalise from instance to instance that regardless of the nature of the objects, the number is always ‘two’ (Hurewitz et al., 2006).

Whether or not arabic numerals are processed in a similar or different way to number words, however, remains uncertain. Most of the neuropsychological evidence suggests that the two formats are processed along separate pathways. Dehaene and Cohen (1995) showed that different neuronal pathways are involved in reading digits and words. Split-brain studies indicate, for example, that the left hemispheric visual system recognises both formats, whereas the right hemisphere only recognises simple arabic digits. Furthermore, even the left hemispheric pathway is sub-divided into many specialised networks, with the lesion of one of these, for example, resulting in the impairment of visual word recognition, but not arabic digit, object or face recognition (e.g. Anderson, Damasio & Damasio, 1990; Greenblatt, 1973). The rare reverse case has also been reported by Cipolotti, Warrington and

Butterworth (1995) where arabic numeral reading was impaired, but word reading was intact.

A further line of evidence for processing differences between different numerical formats is the issue of how numerical surface format affects arithmetic performance, a persistent issue in cognitive research (e.g. Campbell, Parker & Doetzel, 2004). Both neuropsychological evidence and research with normal populations have yielded mixed reports in this regard, suggesting that if such a processing difference exists, it is still uncertain where it lies. Cohen and Dehaene (1994) and Sokol, McCloskey, Cohen and Aliminosa (1991) presented evidence for format-independent arithmetic performance in some acalculic patients, whereas others presented evidence of format-specific arithmetic skills (e.g. Kashiwagi, Kashiwagi & Hasegawa, 1987; McNeil & Warrington, 1994). Similar mixed reports were found for normally functioning adults, with arguments for both format-independent arithmetic (Noël & Seron, 1992; Rickard, Healy & Bourne, 1994) and format-specific arithmetic (Bernardo, 2001; Blankenberger & Vorberg, 1997; Campbell & Alberts, 2009). Several studies have also suggested that the problem size effect, namely an increase in response latencies and errors accompanying an increase in magnitude of the numbers in an arithmetic problem, is larger with problems written in verbal format (e.g. three + eight) than in problems written in digit format (e.g. 3 + 8; Campbell, 1994; Campbell & Alberts, 2009). This effect has been demonstrated across a number of different languages, including French, Dutch, English and Chinese (Campbell et al., 1999; Noël, Fias & Brysbaert, 1997) and suggests that a difference in processing of the two formats indeed exists and that this difference influences arithmetic performance directly.

While there seems to be clear differences in how number words and arabic digits are represented in the brain, considerable debate still exists on how surface format influences the retrieval of answers to arithmetic equations. Two main lines of argument exist in this area, with disagreement on the extent to which access to underlying magnitude meaning is important for numerical activities. Since numerical surface formats (e.g. digits or number words) symbolically represent magnitude or quantity, it is of great interest to investigate how this symbol–concept relationship functions across formats (Fias, Brysbaert, Geypens & d'YdeWalle, 1996) and how it influences subsequent number manipulation. How surface format influences numerical cognition is a central debate in the literature, with previous research disagreeing on whether different formats are processed along common or separate pathways (Zhang, Si, Zhu & Xu, 2010).

1.5. Theoretical Accounts of Format-independent Number Representation

The first line of argument assumes that regardless of input format, numbers are all represented in a uniform abstract code, which enables similar numerical functions to be performed across different input formats (e.g. Gallistel & Gelman, 1992; McCloskey, 1992; McCloskey & Macaruso, 1995). This view, postulated by McCloskey's Abstract Code Model (McCloskey, 1992; McCloskey & Macaruso, 1995) assumes the necessity of accessing underlying magnitude information before numbers can be compared, manipulated or processed in any way (Fias et al., 1996). Numerical information from various input formats is transcoded into an abstract, amodal code. This code allows access to magnitude information and calculation procedures, from which numerical information is translated to a format-specific output code. In arithmetic, for example, this view postulates that the input format has no influence on any subsequent retrieval or calculation processes, since these all operate from the same uniform magnitude code (e.g. Zhou, 2011). In support of this view, neuropsychological studies have shown deficits in brain damaged patients that varied with arithmetic operation rather than input format (McCloskey, 1992; McCloskey & Macaruso, 1995). Arithmetic performance thus varied depending on whether addition or subtraction took place, but not whether or not the problem was presented in arabic digit or number word format.

1.6. Theoretical Accounts of Format-specific Number Representation

The opposite argument assumes format-specific representation input codes, without the need for an amodal abstract code from which all number information is accessed. The two main models that advocate this view are Dehaene's Triple-Code Model (1992b) and Campbell and Clark's Encoding Complex Model (1988; Campbell, 1994).

In Dehaene's Triple-Code Model (Dehaene, 1992; Dehaene & Cohen, 1995) numerical information is represented in three distinct ways, namely a verbal, arabic or amodal magnitude code, and each of the three codes is specialised for specific numerical functions. The verbal code, for example, is involved in retrieval of arithmetic facts, the magnitude code in quantity comparisons and the arabic code in performing calculations on multi-digit numbers. When numerical information is presented it can be translated between the three codes depending on the function that is required. For example, if an arithmetic fact is not available from memory, the operands can be transcoded to the arabic and amodal codes and arithmetic facts can be retrieved accordingly (e.g. 17 + 5 = 17 + 3 + 2 = 20 + 2; LeFevre, Bisanz & Daley et al., 1996). According to this view access to the underlying amodal magnitude representations is only necessary for some numerical activities, whereas others can function without it. Schmithorst and Brown (2004) provided fMRI evidence for the triple-code model in complex arithmetic by showing that three separate components emerged that corresponded to the hypothesised functions of the three codes. A few commentators also support this view of the co-existence of format-dependent and format-independent processing pathways. Nieder, Diester and Tudusciuc (2006), for example, found that some neurons in the intraparietal sulcus are sensitive to numerical magnitude, but not numerical format, whereas other neurons are specifically sensitive to format.

Campbell and Clark's encoding complex view (Campbell, 1994; Campbell & Clark, 1988; 1992) is a slightly different approach, which assumes purely modality-specific representations and rejects the notion of a uniform abstract code for any numerical processing or manipulation to take place. This view argues that, for example, different surface formats influence calculation procedures not just quantitatively, but qualitatively such that arabic digits and number words directly promote the use of different strategies in arithmetic (e.g. Campbell & Alberts, 2009). Unlike the abstract code and triple code models, this view argues that the encoding and retrieval/calculation conditions of arithmetic problem solving closely interact with one another, in the absence of a central amodal code. Evidence for this view comes from Campbell and colleagues' experiments which showed a format x problem size interaction in arithmetic problems: impeded performance on word format problems was found to be even greater on problems with large operands (e.g. 'nine + fifteen' vs 'one + two'). Although problems in word format are generally more difficult to perform, the fact that this effect was enhanced for large problems led Campbell and Epp (2005) to argue that these word format costs could not be attributed to the encoding of the operands, but that format influences the retrieval of the answer directly.

1.7. Cognitive Interference: Processing Differences between Digits and Words

Most of the support for the above mentioned models comes from studies of adult arithmetic. However, lower level numerical activities, such as number comparison or counting, have not been investigated in this regard, with direct comparison between arabic digits and number words generally not featuring in such experiments. As a consequence, models such as McCloskey's (1986) and Campbell and Clark's (1988; 1992) can explain format-specific effects of number manipulation very well, but do not explain how semantic access is gained from the presentation of a single digit or word, for example (Dehaene, 1992).

In the field of psycholinguistics, studies of cognitive interference have been extensively used to model processing differences between two stimulus features, a method that seems well suited to studying basic processing differences between numerical surface formats. Cognitive interference refers to the phenomenon where two stimulus features are processed simultaneously and the processing of one stimulus feature slows down the processing of the other stimulus feature (e.g. **four**

four four; How many words are present? see Bush et al., 2006). During such tasks, depending on the basis for responding, one stimulus feature is task-relevant and the other stimulus feature is task-irrelevant (to be ignored). It is the conflict between these two dimensions that results in cognitive interference and in turn, a slowed response (e.g. Tang, Critchley & Glaser et al., 2006). Experimental tasks that model cognitive interference thus illustrate the degree to which processing of one stimulus feature activates representations of the other stimulus feature even if this feature is to be ignored under task instructions. Where number words and arabic digits are concerned, cognitive interference tasks can be informative as regards the degree to which processing of the two formats overlaps or differs (e.g. Bush et al., 2006; Tang et al., 2006).

The original colour/word Stroop interference task showed that it took longer for participants to name the colour of the ink that words were written in, when ink-colour and colour-word did not match (e.g. 'BLUE' written in red ink; correct response is red; Stroop, 1935), known as the colour Stroop effect. The Stroop task has also subsequently been specifically adapted to study numerical dimensions. Windes (1968) introduced an enumeration Stroop task and found that performance was slower when stimuli to be counted were arabic numerals that were incompatible with the number of items presented (e.g. **4 4 4**, correct response is 'three'). This effect has been widely replicated with robust results (e.g. Flowers, Warner & Polansky, 1979; Pavese & Umiltà, 1998; Shor, 1971). More recently, Bush, et al. (1998; 2006) developed a similar counting Stroop task with number words for use in fMRI settings (e.g. **four four four**, correct response is 'three'). During this task the highly automatic process of reading is placed in competition with subitizing, the automatic enumeration process that takes place for a small number or items (e.g. Bush et al., 2006).

The counting Stroop task has also been used to illustrate that number knowledge is extracted from language that does not explicitly refer to specific numbers. In a modified counting Stroop task, for example, participants took longer to indicate that there was only one word presented when the word was plural (e.g. **CATS**, correct response is 'one'), than it did when the word was singular (e.g. **CAT**, correct response is 'one'; Berent, Pinker & Tzelgov et al., 2005). In some cases it also took longer to indicate that two words were on the screen when the words were singular (e.g. **CAT CAT**, correct response is 'two'). Overall, the numerical Stroop task seems to be particularly sensitive to underlying number meanings that are accessed from words and arabic digits.

A small number of numerical Stroop studies have also employed different numerical formats for comparison in the same task. Roelofs (2006), for example, reported Stroop-like interference in a study that examined the naming of dice, digits and number words. In a series of experiments, arabic digits were presented alongside incongruent dot patterns (e.g. **3 ● ●**) or number words (e.g. **3 two**) and participants were asked to either name the number represented by the digit, dot pattern or number word, while ignoring the other task-irrelevant incongruent digit, dot pattern or number word. Dot patterns did not affect word or digit naming latencies to a significant extent, however, words affected digit naming latencies and digits affected word naming latencies to the same (significant) extent. These results suggested that digit naming was achieved in a manner similar to word naming, as opposed to dice naming. Such findings might be expected as dot patterns on dice would be relatively unfamiliar numerical representations of number. Since arabic digits and number words are both symbolic numerical representations, their processing can be expected to be more similar compared to dot patterns, which represent number analogically.

However, other Stroop interference findings which did not include dot patterns found that digits produced similar interference to pictures, rather than words (Fias et al., 2001; Reynvoet, Brysbaert & Fias, 2002). During this task, digit and word pairs were presented together (e.g. **3 four**) and the left–right positions altered randomly across trials. Similar to the experiment of Roelofs's (2006), participants had to respond to either digit or word meaning while ignoring the task-irrelevant digit/word. The pattern of performance mirrored previous findings of word/picture interference, which led Fias et al. (2001) to argue that arabic digits and pictures are processed similarly. The basis for this conclusion was that digit naming was disrupted with the presence of an incongruent number word, whereas number word naming was not found to be disrupted with the presence of an incongruent digit. Similarly, in picture–word Stroop tasks, picture naming is generally disrupted by the presence of an incongruent word, whereas word naming is not disrupted by the presence of an incongruent picture to a great extent (e.g. Alario, Segui & Ferrand, 2000; Starreveld & La Hej, 1996). The authors argue that the similarity in processing of arabic digits and pictures stems from the fact that both digit and picture naming occur through a semantic route, whereas word naming can occur without gaining access to underlying number meaning.

CHAPTER 1

The process of word reading involves mapping letters to sounds, whereas with digits there is no such explicit letter-sound mapping opportunity. In the case of words, access to word meaning can follow letter-sound mapping or can occur independently if the words are highly practised (e.g. Plaut, McClelland, Seidenberg & Patterson, 1996). Semantic mediation is more direct for digits, since it operates in the absence of letter-sound mapping. Activation of digit meaning is therefore unavoidable, whereas with words only letter-sound mapping need to be present for accurate naming. Given the different findings of Roelofs et al. (2006) and Reynvoet et al. (2002), whether or not spontaneous semantic activation takes place with number words seem to depend on specific task instructions and situations. Neuropsychological studies show, for example, that in situations where letter-sound knowledge is lost, as is seen with patient I.H. (Butterworth et al., 2001), accurate reading and spelling can be achieved by a meaning-mediated process alone.

Overall, these findings suggest that compared to dot patterns, arabic digits and number words are read similarly. In comparison with each other, however, an advantage seems to exist for digit compared to number word processing (e.g. Campbell et al., 1999; Noël et al., 1997). This is likely to be due to time consuming letter-sound mapping and phonological activation that occur during number word reading. In addition, people encounter countless combinations of letters every day, which can form words to refer to a countless number of concepts, whereas arabic digits are usually only used in the context of number. This association between underlying number concepts and arabic digits is thus more practised than this association with number words, which could underpin the observed format-specific processing differences.

1.8. Cognitive Interference Modelled by the Size Congruity and Symbolic Distance Effects

Cognitive interference tasks have also been used to study other dimensions that relate to number concepts, such as numerical magnitude and physical size. Besner and Coltheart (1979) originally described this task that placed these two dimensions into competition with one another. These tasks typically involve two arabic digits, with varying physical sizes and numerical magnitudes presented together. The participant has to indicate which number (left or right) is either physically or numerically larger, depending on task requirements. On congruent trials the physically larger numeral is also the numerically larger numeral, e.g. **5** **2** (correct response 'left' in both physical and numerical comparison tasks), whereas on incongruent trials the physically larger numeral is the numerically lower numeral, e.g. **5** **2** (correct response 'right' in physical comparison task and 'left' in numerical comparison task). Cognitive interference is typically measured as an increase in response latencies and errors on incongruent trials relative to congruent trials. The size congruity effect demonstrates that when an arabic digit is presented, underlying access is gained to number meaning, which is closely related to physical size, underlying the interference on incongruent trials. The size congruity effect has also been demonstrated to a lesser extent with number words (Cohen-Kadosh, Henik &

Rubinstein, 2007; 2008) suggesting that a similar process takes place with number word processing.

Cognitive interference in the size congruity effect seems to arise due to the automatic spontaneous processing of certain stimulus features, regardless of task-relevance. The original magnitude Stroop task has been widely adapted to study this phenomenon (e.g. Banks, 1977; Dehaene, 1989; Parkman, 1971) and to illustrate that number magnitude seems to be automatically accessed when a numeral is presented despite it being the unattended task dimension (e.g attend to numerical magnitude and ignore physical size; Girelli, Lucangeli & Butterworth, 2000). Moyer and Landauer (1967) suggested that arabic numerals are converted to an analogue representation, which enables a physical comparison between the two numbers (Besner & Coltheart, 1967). The specific semantic number that the numeral refers to is therefore converted to a more general analogue item, with perceptual properties, which allows a physical comparison to take place.

Related to size congruity is the symbolic distance effect, which Moyer and Landauer (1967) originally modelled to show that "the time to make the judgement is a function of the numerical distance (difference) between the numbers" (p.105). The symbolic distance effect demonstrates that it is generally more difficult to discriminate between stimuli that are similar than between stimuli that are dissimilar. In number comparison tasks, the time taken to make a judgment is thus inversely related to the numerical distance between the two numbers. In the number Stroop task, for example, participants displayed faster response latencies when the two digits for comparison were numerically further apart (e.g. **2 9**) compared to two digits that were numerically closer together (e.g. **7 9**). The presence or absence of the symbolic distance effect is viewed as evidence for the degree to which the two stimulus features (numerical magnitude and physical size) are processed autonomously, and can in turn provide insight into processing differences between different numerical surface formats. Importantly, numerical and physical distance can be varied parametrically (e.g. small, medium and large physical/numerical distances) beyond a mere smaller/ larger classification, which allows a more in depth investigation into how physical size and magnitude representations overlap (Tang et al., 2006).

Findings from the symbolic distance effect strongly support the notion of a pre-verbal, spontaneous capacity for comparing items hierarchically. A symbolic distance effect has been found, for example, for abstract linguistic dimensions where the two concepts do not relate to quantification or magnitude. Friedman (1978), for example, found a symbolic distance effect when participants were instructed to choose the better or worse of low imagery word pairs such as “hate versus peace” or “hate versus pressure”. Thus, when two numbers are compared, underlying magnitude representations need not necessarily be activated in order to make a decision. The items could merely be related to each other by a verbal code that organises objects in a hierarchical set (see Cohen-Kadosh et al., 2007). We might thus predict differences in size congruity and physical/numerical distance effects between arabic digits and number words, based on the difference with which the formats activate underlying magnitude meanings.

1.9. Automaticity of Processing and Cognitive Interference

Kadosh and colleagues (2007) used event-related potentials (ERP) technology in a Stroop task variant to show that the conflict between the two stimulus features (physical size and numerical magnitude) is not completely resolved until the response is initiated (Kadosh, Kadosh & Linden et al., 2007). Regarding the degree to which processing of these two features are shared across cognitive systems, this finding supports the argument that distinct mechanisms for physical size and numerical magnitude exist, which enables a comparison of the processing of the two features. Most notably, magnitude processing seems to be modulated by both shared and distinct neural substrates, depending on task requirements.

It could be argued that the physical dimension (relating to perceptual properties; e.g. Berent et al., 2005) is more related to the analogue numerical system and that the numerical dimension (adhering to language dimensions; e.g. Berent et al., 2005) is more related to the specific numerical system. Therefore, in a physical comparison task (e.g. **2 4**, correct response 'left') a size congruity effect is not observed for young children as they have not been exposed to arabic numerals to such an extent that this symbol–concept relation can be accessed automatically. In adulthood, however, sufficient experience with arabic numerals has taken place, which results in Stroop interference during incongruent trials (the size congruity effect). On the other hand, the analogue perceptual number representation (e.g. physical size) is already in place in infancy (e.g. Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981), which results in Stroop interference on incongruent trials when physical size is the "to be ignored" dimension.

In light of the differences in size congruity effects between adults and young children, similar differences might even be observed in adulthood at different levels of mathematics experience. The automaticity of a cognitive process is not a phenomenon that is either present or absent, but rather one that exists on a continuum and develops gradually across time depending on practice and experience (Logan, 1985; MacLeod & Dunbar, 1988; Schiffrin, 1989). Arguably, if experience with numerical information in arabic digit format increases the automaticity with which underlying number meanings are accessed from arabic digits, it would result in greater cognitive interference in numerical Stroop tasks. The degree of interference could thus be related to the degree of experience with the format in question. Individual differences in numerical processing might therefore be useful in studying the influences of surface-format in numerical cognition.

1.10. Individual Differences in Number Processing

While the studies mentioned above have focused on how numerical information is represented and manipulated, individual differences in number processing have generally not been considered in this regard. If increased exposure to certain stimuli can produce information processing biases, as has been shown with emotional Stroop task paradigms (e.g. Edwards, Burt & Lipp, 2006), individual differences relating to mathematics should also influence numerical information processing. Patterns observed in emotional Stroop task paradigms demonstrate, for example, that anxious individuals show an involuntary attentional bias for anxiety related stimuli (e.g. Edwards et al., 2006). This bias is thought to result from increased focused attention and memory for anxiety related stimuli above other stimuli. Thus if, as according to Ashcraft (2006), automaticity of processing for certain stimuli (e.g. anxiety related words) develops as a result of rehearsal and memory, increased mathematics experience could lead to a similar processing “bias” for numerical information.

Although numerical competence is a complex skill, relying on various abilities (e.g. Mazzocco, 2008), the argument that practice and memory lead to increased proficiency with numbers is held by most leading theorists in the area of numerical cognition. Dehaene (1997), for example, argues that it is unlikely that some individuals are biologically predisposed to be mathematics proficient and emphasises the role of memory and practice. For ‘prodigies’, for example, numbers are so practised that the presentation of nearly every number activates learned facts stored in memory about that number. In such cases, Dehaene (1997) argues, it is the extensive exposure and practice with numbers that result in their superior abilities, rather than a predisposed numerical aptitude. Similarly, Butterworth (1999) is of the view that there is no evidence relating mathematics achievement to innate intellectual advantages. Instead, the best predictor of mathematics achievement is practice and training. Furthermore, Stevenson and Stigler (1992) noted that the emphasis placed on innate numerical ability varies cross-culturally. In Japan, for example, effort and learning is emphasised in school performance, whereas American parents often emphasise innate talents and limitations. These cultural differences seem to profoundly influence mathematics achievement, with the Japanese showing an advantage in numerical achievement compared to the American; which further strengthens the case for practice and memory.

CHAPTER 1

Acquiring advanced numerical concepts seem to be particularly difficult, in comparison with language acquisition, for example, which emphasises the need for extensive practice and rehearsal in order to master numerical concepts. Learning to count is easy for children as they are already competent in the necessary activities that they need to engage in to achieve this, such as searching, verbal labelling and one-to-one correspondence. However, equations beyond simple addition require skills that humans are ill-equipped for such as memorisation of large numbers and remembering various different facts that are easily confused with one another (e.g. multiplication tables; Dehaene, 1997). In comparison with literacy development, which mostly involves adding new words to existing concepts of word classes and grammar, mathematical abilities often require developing completely new skills that add on to previously acquired skills, but are conceptually distinct (LeFevre, 2000). Number representation and calculation, for example, require the abilities to read, write and transcode between different symbolic numerical notations (Deheaene, 1992). It seems that at this point in development, mathematics education and cultural variables would greatly influence numeracy. Formal numerical manipulation thus requires an "increasingly sophisticated understanding of numerosity" (Butterworth, 2005, p. 15).

Studies of individual differences in numerical cognition have mostly come from a developmental perspective. However, studying adult samples can be informative of how the experiences encountered earlier in life can influence later numerical information processing. For example, great variability in the processing of basic probability and numerical concepts (numeracy) exists among even highly educated adult populations (e.g. Jukes & Gilchrist, 2006; Lipkus, Samsa & Rimer, 2001; Peters, Västfjäll & Slovic et al., 2006). This suggests that while children can acquire the necessary skills for performing formal mathematics, they may still not be able to apply these skills to novel situations in adulthood (Dehaene, 1997).

Overall, there seems to be a lack of consideration for mathematics experience in adult numerical cognition studies of both lower level number processing (e.g. number comparison), as well as more advanced number manipulation (e.g. calculation). Mathematics experience, however, seems important to consider, as differences in exposure to numerical information should influence the automaticity with which underlying number meanings are accessed from symbolic formats. For example, if individuals with greater mathematics experience are more proficient at accessing number meaning from a variety of different numerical formats (e.g. arabic digits, number words, quantifier words etc.) it would lend more support to models which assume an underlying analogue code for all numbers (e.g. McCloskey's Abstract Code model, 1992). If all numbers are translated to an internal amodal code an advantage with numbers should not discriminate between formats. On the other hand, if processing differences between numerical formats (e.g. arabic digits and number words) differ at different levels of mathematics experience, it would be more in line with the accounts which postulate that different numerical formats assume separate representational codes, without the need for a uniform analogue code (Campbell & Clark's Encoding Complex Model, 1995; Dehaene's Triple Code model, 1992). Practice with a particular format would thus strengthen its processing (e.g. arabic digits), while not necessarily influencing the processing of another (e.g. number words).

1.11. The Current Research

The current study investigated the influences of format and mathematics experience across a wide range of numerical functions. By taking into account individual differences related to numeracy and mathematics education, the study explored the possibility that format effects in adult number processing and manipulation could be regulated by mathematics experience.

Support for models of symbolic number representation such as the abstract code model (McCloskey, 1992) and the encoding complex model (e.g. Campbell & Clark, 1992) have mostly come from studies of arithmetic. More research is thus needed to relate processing differences between formats to early numerical processing such as magnitude comparison or subitizing. While a small number of studies have compared the reading of arabic digits and number words, Stroop tasks investigating number–size comparisons and subitizing have not compared different formats directly, with experiments mostly focusing on either arabic digits or number words. The first three experiments in the current thesis examined such basic numerical processing by adapting Stroop tasks to investigate the processing differences that might emerge for arabic digits, number words and quantifier words in the English language.

In chapter 2 (Experiment 1) the original counting Stroop task was adapted to include arabic digits for comparison with number words. The increase in RT on incongruent (e.g **four four four**, respond ‘three’) relative to neutral (e.g. **cat cat cat**, respond ‘three’) trials was investigated for each format and at high and low levels of mathematics experience, based on participants’ Irish Leaving Certificate performance and results on a numeracy test. The view was explored that greater experience with mathematics could result in an advantage for processing numerical information, and that this might further vary between arabic digits and number words. Since the Stroop task has been widely used in other individual differences domains (see Chapter 2) it seemed well-suited to the study of individual differences in numerical information processing.

Chapter 3 (Experiment 2) addressed a similar question by considering format-specific effects in terms of size congruity and symbolic distance at different levels of mathematics experience. By modifying the task developed by Tang et al. (2006; see Chapter 3), arabic digits as well as number words were investigated in physical (e.g. **two** **five**, which number is physically bigger?) and numerical comparison tasks (e.g. **two** **five**, which number is numerically bigger?). This experiment addressed the question of whether or not the dimensions of physical size and numerical magnitude are processed similarly and the role that stimulus format and mathematics experience can play in this regard.

Based on the results from Experiments 1 and 2, Experiment 3 (Chapter 4) investigated whether or not greater mathematics experience can result in an advantage in extracting numerical information from language more generally. The counting Stroop task used in Experiment 1 was adapted for studying quantifier words with specific (e.g. **both**) and general (e.g. **some**) number meanings. Since quantifier words do not express number meanings as explicitly as number words or arabic digits, it is not certain whether or not quantifier word processing follows a more numerical or linguistic processing route. The role of number knowledge in quantifier word processing has not been explored to a great extent. However, in development, number knowledge seems to be central to quantifier word knowledge. Differences in quantifier word processing related to adults' mathematics experience were thus explored.

Overall, Experiments 1 to 3 considered basic number encoding and how mathematics experience can influence this process. Subsequent to encoding, various other functions take place, such as calculation and arithmetic fact retrieval. To investigate these processes, Experiments 4 and 5 considered the role of operand format and mathematics experience in performing mental arithmetic. In addition to this, eye-tracking and event-related potential (ERP) technology were used in Experiments 4 and 5 respectively, as these measures have been shown to be sensitive to effects that might not be evident from behavioural measures alone (e.g. Merkley & Ansari, 2010). Different stimuli can be processed along separate routes, but can still yield similar behavioural patterns (e.g. Zhang et al., 2010; Zhou, 2011). More sensitive measures such as eye-tracking and ERP technology were therefore employed in the second part of the thesis alongside behavioural measures of accuracy and reaction time.

As mentioned above, two opposing viewpoints exist on how the encoding and answer-retrieval stages of arithmetic relate to one another. Recent studies favour both the additive viewpoint (e.g. MCloskey's abstract code; Zhou, 2011) of format-independent answer retrieval as well as the interactive viewpoint of format-specific answer retrieval (Campbell & Alberts, 2009). The role of mathematics experience has, however, not been explored in these studies, and it seemed an important variable to consider in the study of adult arithmetic. In Chapter 5 (Experiment 4), a study of Campbell and Alberts (2009) was replicated in order to examine the influence of operand format on the calculation strategies used in arithmetic. Campbell and Alberts (2009) investigated whether the format of the operands directly influences the strategies that participants reported using (e.g. direct memory retrieval or calculation), or if relatively similar calculation processes take place for arabic digits and number words, with their results supporting the former view. Since shortcomings have been noted with self-reports (Kirk & Ashcraft, 2001), Experiment 4 employed eye-tracking measures to investigate if the findings of Campbell and Alberts (2009) could be supported. Specifically, the experiment tested whether or not measures of fixation count and fixation duration reflect similar interactions of format, operation and problem size as was noted in the self-reports of Campbell and Alberts's (2009) participants.

While overall relatively little research has been conducted using eye-tracking in the study of numerical cognition, it has been a useful tool in studying information processing in reading (e.g. Inhoff, 1984, 1985; Rayner & Pollatsek, 1987) and thus seems well suited for the study of numerical processes. A recent interest in using eye-tracking to study numerical cognition specifically has also emerged (e.g. Merkley & Ansari, 2010; Moeller, Neuburger & Kaufman, 2009), as eye-tracking can provide a more extensive measure of information processing than reaction time and accuracy (Desroches, Joanisse & Robertson, 2006).

CHAPTER 1

While addressing the question regarding the relationship between the different stages in arithmetic, a more in depth analysis of the interaction between the encoding and answer-retrieval stages was conducted in the final experiment. To do this, Experiment 5 replicated an event-related potentials (ERPs) study of Zhou (2011), which aimed to separate the presentation of the encoding and retrieval phases of arithmetic equations in a true/false verification task that presented addition and multiplication equations in separate blocks. In this study, the equation '3 + 2 = 5', for example, was presented as '3' and '+ 2 = 5' on separate presentation-screens (or 'three' and '+ two = five'). This allowed the effects of operation, format and mathematics experience at the encoding and answer-retrieval stages to be investigated separately. Zhou (2011) noted a dissociation in how addition and multiplication is mentally represented even during the encoding phase where participants only saw a single digit operand on-screen. In support of the additive view of arithmetic (e.g. McCloskey's abstract code model), multiplication and addition operands presented in the same format are encoded differently, which allows the relevant arithmetic facts to be retrieved. If the interactive view of arithmetic (e.g. Campbell & Clark's encoding complex model) were supported, the dissociation between arithmetic operations should only emerge subsequently to the encoding phase, since addition and multiplication operands presented in the same format should be encoded similarly (Zhou, 2011). The final experiment (Experiment 5) investigated the event-related potentials at the encoding and retrieval phases separately, while controlling for mathematics experience and presenting equations in digit as well as word format, unlike Zhou's (2011) study which only involved arabic digit operands. By including two formats, the specific effects of operation and mathematics experience that emerge for each format and at each level of arithmetic could be compared.

The overall objective of the current research was to investigate how numerical information is accessed from symbolic formats, and how this might differ at different levels of mathematics experience. By investigating these effects for various numerical functions and utilising a wide range of measures, the aim was to gain a more comprehensive view of the mental representation of numbers. The early experiments (Experiments 1 to 3) investigated basic number encoding, which formed the basis for investigating format effects in more complex numerical cognition, such as calculation (Experiments 4 and 5). The following chapter (Experiment 1) set out to explore the role of mathematics experience in format-specific processing in a simple counting task. By investigating cognitive interference of arabic digits and number words, the automaticity of processing of the two formats could be directly compared.

Chapter 2

Experiment 1: Cognitive Interference in a Digit–Word Counting Task: The Role of Mathematics Experience in Format-specific Processing

2.1. Introduction

Widespread evidence supports the argument for format-specific influences on performance in tasks where participants are engaged in numerical processing. As mentioned in Chapter 1, format-specific effects have been found in arithmetic studies (e.g. Campbell, 1994; Campbell & Alberts, 2009) highlighting differences in performance between digits and number words. Little evidence exists, however, to link format-specific processing with early numerical processing such as subitizing: the rapid enumeration process that takes place for a small number of items (1 to 4; Dehaene, 1997). While numerous studies have investigated such basic numerical processing, these have mostly focused on one or the other format, but have not compared the processing of digits and number words directly.

According to Ganor-Stern and Tzelgov (2008) the most effective way to study mental representations is to investigate if their processing is automatic even when participants are instructed to ignore them as part of a task. Stroop interference tasks, for example, make use of the observation that when words are read, access to underlying word meaning is generally unavoidable. Windes (1968) originally introduced an enumeration Stroop task and found that participants were slower to count stimuli when they were incompatible arabic digits (e.g. **3 3**; respond '2'). Other studies have subsequently replicated this effect (Flowers et al., 1979; Pavese &

Umiltà, 1998; Shor, 1971) and have also demonstrated the Stroop effect for number words (Bush et al., 1998; 2006). With the aim to draw on the known success of the original colour–word Stroop task (Stroop, 1935), Bush et al. (1998) designed the counting Stroop task to study the neural basis of informational conflict in functional magnetic resonance imaging (fMRI) settings. Since speaking requires head movements that are not tolerated by fMRI, the original colour–word Stroop task was not suitable. Arbitrarily labelling response-buttons with colour names was also not ideal as this would have added undesired cognitive complexity to the task. The counting Stroop was thus created in response to these shortcomings as it allows button-press responses (within the subitizing range) that do not require speech. The task requires participants to count the number of (identical) words on-screen while ignoring the number meanings of the words (Bush et al., 1998, 2006). Trials where number and word meaning are incongruent (e.g. **four four four**; respond ‘3’) result in a slowed response compared to neutral (e.g. **cat cat cat**; respond ‘3’) and congruent (e.g. **four four four four**; respond ‘4’) trials. While the two formats have not been compared directly, this effect has also been shown with arabic digits (e.g. Muroi & McLeod, 2004) suggesting that when presented with a digit or a number word, access to underlying word meaning is an unavoidable, automatic process.

Cognitive interference occurs on incongruent trials in the Stroop task when two stimulus features are processed simultaneously and the processing of one impedes the processing of the other, reflecting the extent to which the processing of the two features overlaps (e.g. Bush et al., 1998, 2006). In the counting Stroop task the two highly automatic processes of subitizing and reading are placed in competition, such that number words that are incongruent in meaning interfere with the counting process. The increased response latency on incongruent trials relative to neutral trials (e.g. **dog dog dog**; respond '3') acts as a measure of cognitive interference. Another effect, deemed Stroop facilitation is also often noted where faster responses occur on congruent (e.g. **three three three**, respond 'three') relative to neutral trials (e.g. Bush et al., 2006). The counting Stroop effect thus demonstrate that incongruent numerical stimuli slow down the counting process, congruent numerical stimuli speed up the counting process and number-neutral stimuli do not influence the counting process.

The focus of counting Stroop tasks has mostly been the study of informational conflict in general rather than numerical cognition per se. As such, format-specific processing has generally not been of interest and mathematics experience has not been considered in these tasks. The current experiment considered these factors: if arabic digits and number words are processed differently, differences in cognitive interference in the counting Stroop task might be predicted between the two formats. Research has mostly included only one format without comparing formats directly. The original counting Stroop task (Bush et al., 1998; 2006), for example, which was designed to measure cognitive interference in fMRI settings only focused on number words. The few Stroop studies that have included different numerical formats (as described in Chapter 1, p. 19 – 21) have suggested that different processes take place for the naming of digits and number words. Generally, number words seem to spontaneously gain access to phonological codes upon which access to semantic information can follow (Damian, 2004). Arabic digits seem to gain more direct access to semantic codes, and only subsequently gain access to lexical information. However, these distinctions are not always as clear-cut even if these are the typical routes of processing for these formats. Dual-route models of word reading, for example, argue that access to underlying word meanings can either follow on from letter-sound mapping or occur directly (e.g. Coltheart, 2005; Plaut et al., 1996), which seems to mirror some of the accounts of digit naming (e.g. Butterworth, 1999; Dehaene, 1992, 1997). Word frequency should also be expected to play a role in the extent to which spontaneous access to underlying meaning is achieved. Similar to arabic digits, high frequency nouns (such as small number words) are likely to be read through a conceptually driven route, reflecting a strong symbol–concept relation established through extensive exposure. It is this automatic conceptual processing that gives rise to cognitive interference in the Stroop task.

The processing that takes place when an arabic digit or number word is read can thus be classified as 'automatic' when its meaning interferes with the task at hand even when it is to be ignored under task instructions (e.g. Ganor-Stern & Tzelgov, 2008). However, since automaticity of processing is not an "all–or–nothing" process, but rather exists on a continuum (MacLeod & Dunbar, 1988), the degree to which a format (e.g. digit or number word) spontaneously activates underlying magnitude representations could be related to the individual's experience with the format in question. The Stroop task has been widely utilised in individual differences research to highlight any processing biases that might occur as a result of rehearsal and memory. The emotional Stroop task (Edwards et al., 2006; Williams, Mathews &

MacLeod, 1996), for example, demonstrated that anxious individuals tend to show an involuntary attentional bias towards anxiety related stimuli. In this adaptation of the original colour–word Stroop task (Stroop, 1935), anxious individuals took longer to name the colour that threat-related words (e.g. 'panic' or 'coffin') were printed in compared to 'neutral' words (e.g. 'plate' or 'button'). Similar findings have also been noted for addiction (see Cox, Fadardi & Pothos, 2006, for review). With regards to numerical cognition, individuals with greater mathematics experience could display similar selective processing for certain numerical stimuli. It could thus be hypothesised that individuals with differing mathematical histories, reflecting different levels of practice, memory efficiency and education, could display differences in the automaticity with which number meanings are accessed from symbolic formats. Although the Stroop task has been widely used to reflect individual differences in selective processing, it has not been applied to mathematics experience in this regard. The current study tested this hypothesis by measuring differences in cognitive interference in a digit–word counting task. Participants were divided into 'high' and 'low' mathematics experience groups, based on self-reported performance in the Irish Leaving Certificate mathematics examination and participants' performance on a numeracy test was also assessed. Greater cognitive interference was predicted for individuals with greater experience with numbers, based on the assumption that these individuals could show selective processing for numerical stimuli. This effect was also expected to differ between arabic digits and number words, reflecting the relative automaticity with which underlying number meaning is accessed from the two formats.

2.2. Method

2.2.1. Participants

Forty participants took part in the experiment (age 18 – 29; *M* = 21.22; *SD* = 3.34). Participants were divided into a High Maths and a Low Maths group based on Irish Leaving Certificate performance. Those who reported an obtained grade higher than a C3 for higher level mathematics were assigned to the High Maths group (*n* = 20; 13 men and 7 women). The rest of the participants were assigned to the Low Maths group (*n* = 20; 9 men and 11 women). Participants who had studied foundation level mathematics or who reported reading difficulties were excluded from the study. Most of the participants in the High Maths group reported grades in the A/B range (*n* = 15). All of the participants in the Low Maths group had studied ordinary level Leaving Certificate mathematics and most reported grades in the B/C range (*n* = 14). All participants spoke English as their first language and had normal or corrected-to-normal vision.

2.2.2. Apparatus and Materials

Participants completed the counting Stroop task as well as a numeracy test, a measure of numerical self-efficacy (the Subjective Numeracy Scale) and a number of working memory span tasks.

Counting Stroop task. The stimuli for the computerised counting Stroop task were presented on a 15-inch LCD monitor linked to a computer. Each stimulus was positioned centrally on the screen and subtended between approximately 1 to 1.9 degrees of visual angle. Programming for the task was done in *Superlab®,* which recorded all participant input and reported reaction times in milliseconds (ms) as well as accuracy. The stimuli consisted of four stimulus types, namely number words and digits (the number stimuli), and animal names and symbols (the neutral stimuli). The four number words employed were: ‘one’, ‘two’, ‘three’ and ‘four’, the digits were: ‘1’, ‘2’, ‘3’ and ‘4’, the animal names were: ‘dog’, ‘cat’, ‘mouse’ and ‘bird’ and the symbols were: ‘?’, ‘*’, ‘@’ and ‘#’. The word stimuli were selected on the basis of being common words within a single semantic category, balanced for word length and part of speech (all are nouns), as was employed by Bush et al. (2006). The symbol stimuli used in the neutral condition were selected on the basis that they did not resemble digits or evoke numbers in some way. The stimuli and instructions were presented in black print against a white background.

Some trials involved the meaning of the word or digit matching the number of items presented on-screen (congruent trials), for example: ‘**2 2**’ (correct response: 2). Other trials involved the meaning of the word or digit not matching the number of items presented on-screen (incongruent trials), for example: ‘**three three three three**’ (correct response 4). Some trials contained non-numerical words or symbols (animal names and neutral symbols), which were presented one to four times on any given trial (neutral trials).

The stimulus sets consisted of 12 different combinations of the 6 different stimuli categories, which were neutral digit, congruent digit, incongruent digit, neutral word, congruent word and incongruent word. Each set was presented twice resulting in each participant being presented with a total of 144 trials. Trials were presented in a quasi-random order.

Numeracy test. Participants completed a 17-item numeracy test adapted from Lipkus et al. (2001). The test consisted of 17 problems, ranging from easy to difficult, to measure participants' proficiency with probabilities, proportions and percentages (see Appendix 2). Ten of the equations were taken from Lipkus et al.'s (2001) study (e.g. which of the following represents the biggest risk of getting a disease? 1%, 10% or 5%). A further 7 problems were devised to include more difficult calculations (e.g. If the bill came to € 42 and I gave the waiter € 50 as payment, after deducting a 10% tip, how much change will I get?). Participants were given eight minutes to complete as many of the problems as possible. A blank sheet of paper was provided to work out the answers. The experimenter also had a sheet of paper for each participant to record demographic information regarding the participant's age, gender and self-reported Irish Leaving Certificate mathematics performance. While self-reported Leaving Certificate mathematics performance was used to assign participants to different groups of mathematics experience, the numeracy test was used to assess differences in numerical ability between the two groups.

Digit forward span task. Sequences of different digits were presented aurally by means of a voice recording and after each sequence there was a pause during which the participant repeated the digits out loud (Aleman & Van't Wout, 2007; Oberauer, Süβ & Schulze, 2000). The digits forward task consisted of 16 trials of different list length: 3 digits (3 trials); 5 digits (5 trials); 7 digits (5 trials); 9 digits (3 trials). The number of correct digits recalled in the correct order was recorded for each item and the percentage items correctly recalled acted as a score for the scale. A similar scoring method was used by Oberauer et al. (2000) as it has the advantage over traditional scoring methods of gaining scores for individual items.

Digit backward span task. During this task participants repeated each heard sequence in reversed order: 3 digits (3 trials); 5 digits (5 trials); 7 digits (5 trials); 8 digits (3 trials). Scoring was similar to the Digit Forward span task.

Sentence span task. The computerised sentence span task was programmed in *Superlab*® to utilise a dual-task paradigm where the participant made true or false judgments on simple sentences while remembering the last word of each sentence (see Oberauer et al., 2000). Following the procedure of Oberauer et al. (2000), short trivially true or false sentences were used, where the last word of each was a familiar noun of no more than 3 syllables (e.g. Cats chase mice). Each sentence was presented on the screen for 3 seconds, followed by a 1-second interval. The participant's task was to indicate via button press, during the 4 seconds, whether the sentence was true or false and also to remember the last word of each sentence. After a few sentences the computer instructed the participant to write down the last word of each sentence, in sequence, on an answer sheet. The participant then pressed the space bar to continue the task. The task included 2 practice sentences and 5 test trials (25 sentences overall ranging from word list lengths 3 – 7), presented in ascending order of list length. The number of correct words written down in the correct order acted as a score for each trial. The overall score for the scale was computed as the percentage total words correctly recalled across the five trials.

Subjective numeracy scale (SNS: Fagerlin, Zikmund-Fisher & Ubel et al., 2007). This eight-item self-report questionnaire measured self-perceived efficacy to perform a variety of mathematical tasks (four items of the SNS ability subscale e.g. "How good are you at working with fractions?") as well as preference for information presented in prose versus numerical form (four items of the SNS preference subscale e.g. "When reading the newspaper, how helpful do you find tables and graphs that are parts of a story?"). A six-point Likert scale (e.g.*1 = not at all helpful; 6 = extremely helpful*) was used, with total scores ranging from 4 – 24 on each of the subscales (see Appendix 3).

2.2.3. Procedure

The experiment took place individually for each participant in a quiet room with a PC and two chairs. Upon arrival each participant was told that the study would investigate the processing differences between different numerical formats. The experimenter explained that as part of the experiment the participant was also required to complete some calculations and memory tasks. The participant then received an informed consent form (see Appendix 1). Once the consent form was signed, the experimenter handed the participant the 17-item numeracy test (Appendix 2) and a blank sheet of paper and a pen. The blank sheet of paper could be used to work out the answers and it was made clear that it would be discarded at the end of the experiment. Participants were told that they would only be given a limited amount of time and that they should aim to answer as many questions as possible. The experimenter did not tell participants how much time they had (8 minutes). Participants were also told to start at the beginning and to continue on from there, but to skip a question if it could not be answered. The experimenter told the participant to commence the test once the experimenter had left and that they would be called when the time was finished. A timer was used to keep track of the time. After eight minutes the experimenter announced that the time was finished and collected the answer sheet. The participant then completed the Subjective Numeracy Scale (SNS). After completion of this, the experimenter noted the participant's age, gender, obtained grade in Leaving Certificate mathematics (e.g. A, B, C etc.) and the level of Leaving Certificate mathematics studied (Higher or Lower level).

The experimenter then explained that the next part of the experiment would be a computerised task and asked the participant to sit facing the computer. Participants were instructed to read the on-screen task instructions, but not to commence the task until the experimenter had instructed them to do so. The following message appeared on the screen:

A NUMBER OF ITEMS WILL APPEAR ON THE COMPUTER SCREEN.
AT ANY ONE TIME THERE WILL BE ONE, TWO, THREE OR FOUR ITEMS.

YOUR TASK IS TO COUNT THE NUMBER OF ITEMS ON THE SCREEN AND PRESS THE APPROPRIATE DIGIT RESPONSE KEY BY PRESSING 1, 2, 3 OR 4 ON THE KEYBOARD

SOMETIMES THE MEANING OF THE ITEMS WILL CLASH WITH THE NUMBER OF ITEMS THAT ARE PRESENTED. YOU SHOULD TRY TO IGNORE THE MEANING OF THE STIMULUS AND JUST COUNT THE NUMBER OF ITEMS THAT ARE PRESENT.

BOTH SPEED AND ACCURACY ARE IMPORTANT.

PRESS THE SPACE BAR WHEN YOU ARE READY TO SEE SOME PRACTICE TRIALS.

IF YOU ARE READY TO BEGIN PRESS THE SPACE BAR

After the participant had read these instructions the experimenter emphasised again that both speed and accuracy were important and showed the participant which keys on the keyboard to use (keys 1 – 4 on the left of the keyboard). The participants were also instructed to use the index and middle fingers of each hand to respond. The experimenter remained in the room as the participant completed two practice trials. Once it was clear that the task instructions were understood, the experimenter left the room and the participant commenced the task by pressing the space bar. A total of 144 trials were presented in a quasi-random order. Each stimulus remained on-screen until the participant responded by pressing a key on the keyboard. An inter-stimulus interval of 1000 ms (blank white screen) was used.

Once the participant had completed this task, the experimenter returned and explained that the next part would involve some memory tasks. The experimenter then gave verbal instructions to the digit span task. Participants were told that they would hear sequences of digits and that it was their task to repeat the heard sequence out-loud after each sequence. It was emphasised that the order of the digits were important and that participants should only repeat the digits that they could remember. The participant was asked to sit with their back to the experimenter so as not to be distracted by the experimenter’s movement in recording the participant’s responses. The experimenter remained in the room and recorded the number of correct digits recalled in the correct positions on a form with the correct digit sequences. After this, the experimenter explained that the next task would involve the same procedure except that this time the participant should report the heard digits in reversed order. The same procedure was followed as for the forward digit span task.

After the completion of the digit span tasks, the participant was asked to turn around again to face the computer screen, on which instructions to the sentence span task were presented. Participants were told that they would see short sentences on the screen and that it was their task to indicate as quickly as possible, after each sentence, whether it is true or false. To indicate that the sentence was true, participants were instructed to press the 'd' key on the left of the keyboard. To indicate that the sentence was false, participants were instructed to press the 'k' key on the right of the keyboard. In addition to this, participants were also told to try and remember the last word of each sentence. It was explained that after a few sentences a computer screen would appear with the words: "Now write down the last word of each sentence. Press the space bar to continue". Participants were provided with a pen and an answer sheet indicating each test block, with slots provided to write down the words. Once participants had completed the practice trials successfully and it was evident that the task instructions were understood, the experimenter left the room and the participant commenced the task by pressing the space bar.

Statistical analyses focused on changes in accuracy and reaction times in terms of the congruency of the stimuli. Interactions between congruency, format and maths group were investigated and were thought to be reflective of the degree to which automaticity of processing differed between digits and number words and between the two maths groups.

2.2.4. Ethical Considerations

The current research was granted ethical approval from the University Ethics Committee. Before the experiment began, the experimenter made it clear that participation was completely voluntary and that the participant could decide to withdraw from the study at any stage during the experiment. The experimenter also provided an e-mail address that the participant could contact if they wished to withdraw their data from the study and made it clear that this could be done up until the results were published. None of the participants chose to withdraw from the study.

As described in the informed consent form (Appendix 1), the participant was assured that all data and information provided during the experiment would be kept confidential. The experimenter made it clear that during the experiment, all data would immediately be coded so that participants could only be identified by a participant code number. Before the study commenced, the experimenter also checked if participants had any visual, auditory or reading difficulties that might interfere with the tasks. Participants were made aware that the study does not involve any medical treatment, counselling or diagnosis; but that it aimed to investigate processing differences between different numerical formats and the role that other variables related to mathematics can play in this.

After completion of all the tasks, the experimenter thanked each participant for their participation. The experimenter made it clear that only group data were of interest and that no individual scores would be considered in the analyses. These ethical considerations also applied to the subsequent experiments.

2.3. Results

Reaction times (RTs) in the counting Stroop task were recorded as time taken (ms) to press the 1, 2, 3 or 4 key on the keyboard after the stimulus appeared. The mean RTs were calculated for each stimulus category for the High and Low Maths groups. Errors were also recorded and were excluded from the RTs analysis (4.27 % of the overall data).

An independent samples t-test indicated that the High Maths group (M = 11.8, SD = 3.59) outperformed the Low Maths group (M = 9.05, SD = 3.38) on the numeracy test, $t(38) = 2.49$, $p = .017$. Overall, men (M = 11.73, SD = 3.82) also outperformed women (M = 8.83, SD = 2.96) on the numeracy test, $t(38) = 2.63$, $p = .012$.

The High Maths group also showed higher self-perceived numeracy ability (M = 4.64, SD = 1.17), than the Low Maths group (M = 3.75, SD = 1.19), $t(38) = 2.38$, $p = .023$, suggesting that participants' assessments of their own numerical ability was relatively accurate. Regarding working memory, the High Maths group performed better on sentence span (M = 80.6, SD = 15.04), $t(38) = 2.35$, $p = .024$, and backward digit span (M = 72.02, SD = 14.42), $t(38) = 2.41$, $p = .021$, than the Low Maths group (*Sentence Span M* = 67.6, SD = 19.59; *Backward Digit Span M* = 62.27, SD = 10.92). No significant advantage for the High Maths group was found for forward digit span (*High M* = 83.57, SD = 8 and *Low M* = 79.22, SD = 12.3). The High Maths group thus showed an advantage for the storage and transformation functions of working memory, whereas the two groups showed similar short term memory function.

2.3.1. Accuracy

Errors were classified as trials where a participant indicated the wrong number of items presented on-screen (i.e. pressing the wrong key). Table 2.1 presents the mean percentage of errors made across the different stimulus categories for the Low and High Maths groups.

Table 2.1. *Mean percentages of errors across congruent, neutral and incongruent conditions for the Low and High Maths groups.*

Maths Group	Congruent Digit	Neutral Digit	Incongruent Digit	Congruent Word	Neutral Word	Incongruent Word
Low	2.5	3.54	8.33	1.46	2.92	7.71
High	0.62	1.67	12.71	0.62	1.46	7.71
Average	1.56	2.61	10.52	1.04	2.19	7.71

Overall, most errors were made on incongruent trials and least errors were made on congruent trials. Participants also made more errors on incongruent digit than incongruent word trials, an effect that was more evident for the High Maths group. A 2 x 3 x 2 mixed (between–within) ANOVA was conducted to analyse the differences in error rates between the different stimulus types (digits and words), congruency levels (congruent, neutral and incongruent) and Maths groups (Low and High). A main effect was found for congruency, $F(2, 76) = 45.03$, $p < .001$, with a medium associated effect size (partial eta squared = 0.54), that is, overall more errors occurred on incongruent conditions. A main effect was also found for format at the $p = .05$ level, $F(1, 38) = 4.09$, with a small effect size (partial eta squared = 0.097) indicating that slightly more errors were made for digit than word stimuli overall. No other effects were significant.

Paired samples t-tests with Bonferroni adjustments showed that the increase in errors on incongruent (relative to neutral) trials was significant for the High Maths group for both digits, $t(19) = -4.91$, $p < .001$, and words, $t(19) = -3.42$, $p < .001$. Similarly, for the Low Maths group, errors on incongruent digit, $t(19) = -3.52$, $p < .001$, and word trials, $t(19) = -4.52$, $p < .001$, were more frequent than errors on neutral trials. No significant difference in error rates was found between congruent and neutral trials.

2.3.2. Reaction Times

Congruency. Figure 2.1 and Table 2.2 present the mean correct RTs across the stimulus categories for the Low and High Maths groups.

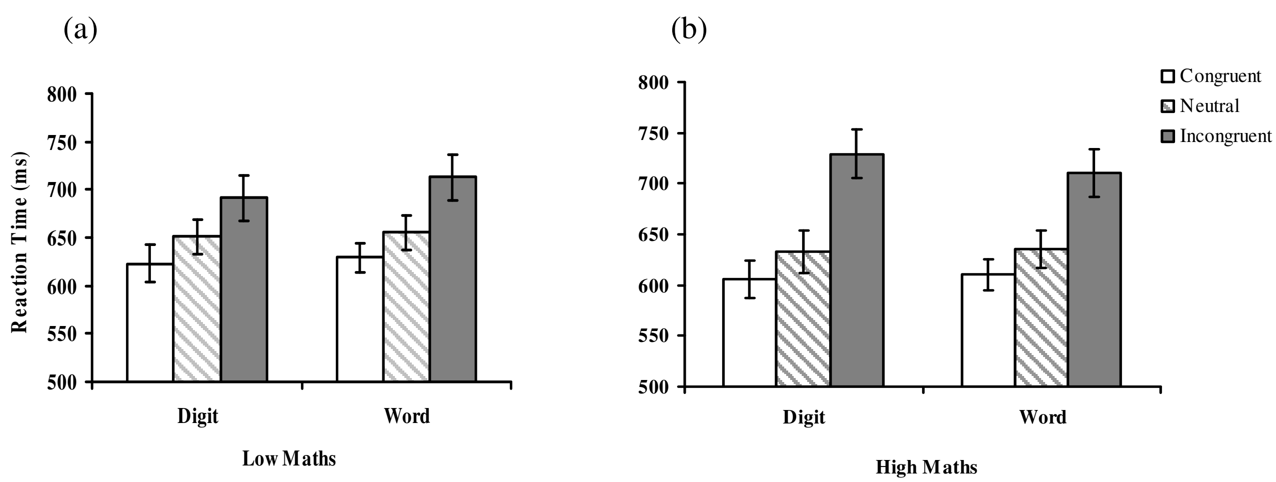

Figure 2.1. *Mean RTs (± SEM) across congruent, neutral and incongruent stimuli for (a) the Low Maths (n = 20) and (b) the High Maths group (n = 20).*

The overall patterns reflected an increase in RT from congruent to neutral to incongruent trials and this pattern was relatively similar for digit and word stimuli as presented in Figure 2.1 and Table 2.2.

Table 2.2. *Means and standard deviations of RTs on congruent, neutral and incongruent conditions for the Low and High Maths groups.*

Maths Group	Congruent Digit	Neutral Digit	Incongruent Digit	Congruent Word	Neutral Word	Incongruent Word
Low Maths	623.11	651.04	691.35	629.57	655.16	712.83
	(88.15)	(79.96)	(93.74)	(80.98)	(86.31)	(94.07)
High Maths	605.53	632.94	729.1	610.24	635.81	710.43
	(84.02)	(95.01)	(105.8)	(68.8)	(82.53)	(105.13)
Average	614.32	642	710.22	619.91	645.49	711.63
	(85.47)	(87.16)	(100.5)	(74.8)	(83.93)	(98.47)

A 3 x 2 x 2 mixed between-within groups ANOVA was conducted to analyse the effects of congruency (congruent, neutral and incongruent trials), format (digits and words) and Maths group (Low and High Maths) on RTs. A main effect was found for congruency, $F(2, 76) = 90.06$, $p < .001$, with a large associated effect size (partial eta squared = 0.72). No main effect was found for Maths group suggesting that the overall response latencies did not differ significantly between the High ($M = 654.01$) and Low ($M = 660.51$) Maths groups. However, a significant congruency x Maths group interaction effect was found, $F(2, 76) = 4.53$, $p = .014$ (partial eta squared = 0.11), suggesting that the High Maths group was more affected by the congruency of the stimuli. No further main or interaction effects were found.

For digit stimuli, Bonferroni corrected dependent t-tests showed that the RT increase on incongruent relative to neutral trials was significant for High Maths, $t(19) = -9.19$, $p < .001$, but not Low Maths participants. For word stimuli RTs on incongruent trials were significantly slower than RTs on neutral trials for both High, $t(19) = -6.25$, $p < .001$, and Low Maths participants, $t(19) = -4.2$, $p < .001$. RTs on neutral conditions did not differ with format or Maths group. No significant difference was found in RTs between congruent and neutral trials.

To summarise, the RT data showed that participants took overall longer to count the items on-screen when digit/word-meaning mismatched the number of items than on trials where they did match or where digit/word-meaning was neutral. While the overall patterns of performance were quite similar for the High and Low Maths groups, a congruency x Maths group interaction suggested that the RT increase on incongruent trials was greater for the High Maths group. While both groups showed this effect for number word stimuli, only the High Maths group showed this effect for word as well as digit stimuli. For digit stimuli the High Maths group showed approximately double the RT difference between incongruent and neutral trials compared to the Low Maths group (96.16 ms difference vs. 40.31 ms difference). To investigate the congruency x maths group interaction further, difference scores were calculated as the discrepancy in RT between incongruent and neutral trials and are presented in the next section.

2.3.3. Interference

Stroop Interference scores were calculated for each participant by subtracting the mean RTs on neutral conditions from the mean RTs on incongruent conditions (presented in Figure 2.2). A 2 x 2 mixed (between–within) ANOVA was conducted to analyse the differences in interference between the two formats and the two groups. Overall, no significant difference in interference was found between digit (M = 68.24, SD = 65.34) and word (M = 66.15, SD = 57.44) stimuli. A main effect was found for Maths group, $F(1, 38) = 84.77$, $p = 0.017$ (partial eta squared = 0.141) indicating that the High Maths group showed overall greater interference. No further significant effects were found. However, independent t-tests (Bonferroni corrected) showed that this group difference in interference was due to the digit stimuli, with significantly greater digit interference found for the High (*M* = 96.164, *SD* = 46.79) than the Low Maths group (*M* = 40.31, *SD* = 70.23), $t(38) = 2.96$, $p = .005$. The difference in interference for word stimuli (*High M* = 74.62 and *Low M* = 57.67) was not significant ($p = 0.357$).

2.3.4. Facilitation

Facilitation scores were similarly calculated as the disparity between RTs on congruent and neutral trials. The facilitation effect, however, was not significant since the RT data showed no significant difference in RT between congruent and neutral trials for word stimuli. A 2 x 2 ANOVA with format and Maths group as factors were conducted on the facilitation data. No significant main or interaction effects were found. Figure 2.2 presents the mean facilitation scores along with the mean interference scores.

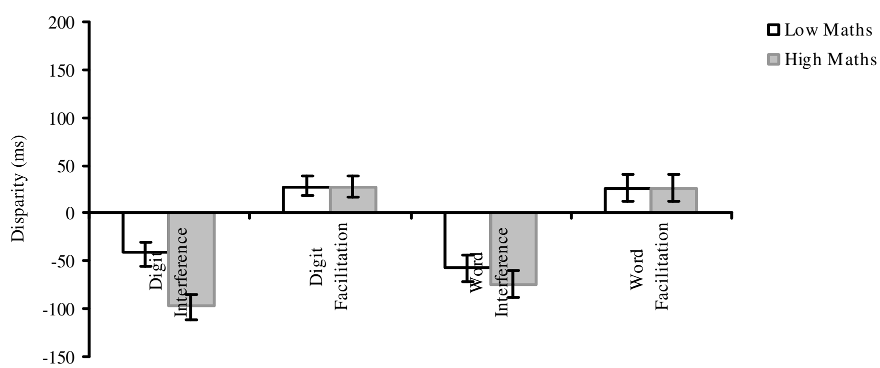

Figure 2.2. *Mean disparity in RTs between congruent and neutral (facilitation, shown above the x axis) and incongruent and neutral (interference, shown below the x axis) conditions in the Low and High Maths groups for digit and word stimuli (± SEM).*

In summary, the interference data show that interference from digit stimuli differed between individuals in the High and Low Maths groups. For individuals with greater mathematics experience (High Maths), incongruent digit meanings interfered with the counting process (e.g. 4 4 4; respond '3') producing slower RTs on incongruent digit trials. While both groups showed this interference effect for incongruent number word stimuli, only those in the High Maths group also showed this effect for arabic digit stimuli. For word stimuli, the pattern of interference was relatively similar for those in High and Low Maths groups.

2.4. Discussion

Previous research has been rather inconclusive as regards processing differences between digits and words. In accordance with the counting Stroop literature, the Stroop interference effect was found for number words (Bush et al., 1998; 2006) in the current study, demonstrating that number meaning is readily accessed from number words and that it interferes with the counting process on incongruent trials. With regards to arabic digit stimuli, only individuals with greater mathematics experience showed an interference effect. While previous studies have noted interference effects of incongruent arabic digit stimuli (e.g. Flowers et al., 1979; Pavese & Umiltà, 1998; Shor, 1971), individual differences relating to mathematics experience have previously not been considered. The current study suggests that format-specific processing differences can emerge when participants' mathematics experience is considered. These findings might be interpreted as a heightened appreciation for numerical information in digit format that results from extensive exposure to arabic digits through learning. As automaticity of processing exists on a continuum (MacLeod & Dunbar, 1988), the more experience gained with a symbolic format, the stronger the symbol–concept relation becomes and the more automatically its underlying number meaning is activated. It is this automatic access to underlying number meanings that interferes with the counting process. Only when incongruent numerical stimuli are to be counted does interference occur and not when the items to be counted are number neutral (e.g. 4 4 4 vs. * * *). Any observed between-groups processing differences thus reflect the relative strengths of these relations between symbols and concepts.

As previous Stroop task findings have yielded mixed results regarding processing differences between digits and words (Fias et al., 2001; Reynvoet, Brysbaert & Fias, 2002; Roelofs, 2006), the strongest evidence for format-dependent processing comes from studies of arithmetic and brain damage studies that yielded double dissociations of processing. Such studies (e.g. Campbell, 1999; Campbell & Epp, 2005) support the notion that surface format affects ‘later’ numerical processing such as calculation (arguably because digits place fewer demands on working memory resources than words), but do not provide much evidence to suggest that such effects are present in ‘early’ numerical cognition (e.g. subitizing). The current findings suggest that format-specific effects can be present even in early number processing, if an individual’s experience with the symbolic format is considered. Overall, arabic digits and number words slowed down the subitizing process to relatively the same extent, suggesting similar processing of the two formats. However, when individual differences related to mathematics experience were considered, the findings suggest a heightened appreciation of arabic digit format that accompanies High Maths experience.

As the current measure of assigning participants to the two groups was primarily based on self-reported Leaving Certificate mathematics performance, a more robust measure could yield even greater group differences. It could be the case that highly mathematics competent and experienced individuals were included in the Low Maths group and vice versa. However, High Maths participants did, on average, outperform Low Maths participants on the numeracy test, storage and transformation working memory and numerical self-efficacy, supporting concrete differences between the two groups. The two groups were slightly unbalanced in terms of sex, with most of the High Maths group consisting of men (13 men and 7 women) and most of the Low Maths group consisting of women (9 men and 11 women), which could have contributed to the overall advantage for men on the numeracy test. However, gender differences are generally not known to affect Stroop performance (e.g. see MacLeod, 1991).

In summary, the study showed that in the counting Stroop task individuals with greater mathematics experience found it more difficult to ignore task-irrelevant digit stimuli, whereas both groups found it equally difficult to ignore task-irrelevant word stimuli. Such individual differences, reflecting experience with numbers, thus seem to play a role in format-specific processing differences that emerge during basic numerical functions such as subitizing. In Chapter 3 this effect was investigated further to see if it would also hold for other basic numerical functions such as magnitude and size comparisons.

Chapter 3

Experiment 2: Size Congruity and Distance Effects in a Digit–Word Number Comparison Task at Different Levels of Mathematics Experience

3.1. Introduction

Many studies of number processing have investigated the size congruity (Henik & Tzelgov, 1982) and symbolic distance (Moyer & Landauer, 1967) effects. The size congruity effect is a Stroop-like phenomenon where the processing of one stimulus feature impedes the processing of another, resulting in cognitive interference. Besner and Coltheart (1979) originally described a task that placed the two dimensions of numerical magnitude and physical size in competition with one another. Two arabic digits, with varying physical sizes and numerical magnitudes are presented together for comparison and the participant makes a judgement based on either physical size or numerical magnitude. On congruent trials the physically larger numeral is also the numerically higher numeral, e.g. '**5** **2**'; whereas on incongruent trials the physically larger numeral is the numerically lower numeral, e.g. '**2** **5**'. Cognitive interference is typically measured as an increase in response latencies and errors on incongruent trials relative to congruent trials or neutral (e.g. **2** **5**) trials. The size congruity effect reflects the automaticity of processing of the task-irrelevant dimension, namely physical size in numerical comparison and numerical magnitude in physical size comparison (e.g. Besner & Coltheart, 1979; Dehaene, 1992; Henik & Tzelgov, 1982; Schwarz & Heinze, 1998). The degree of interference thus indicates the degree to which the processing of physical size and numerical magnitude overlaps.

The symbolic distance effect is often investigated alongside the size congruity effect to illustrate that in-depth processing takes place, beyond a mere small–large comparison of number magnitude (e.g. Tang et al., 2006). In a landmark study, Moyer and Landauer (1967) modelled the symbolic distance effect, which states that when two numbers are compared "the time to make the judgement is a function of the numerical distance (difference) between the numbers" (p.105). For example, it is easier to compare two numbers that are numerically further apart than two numbers that are numerically close (e.g. '**2 7**' vs **'2 3'**). Numerical Stroop tasks, unlike the original colour–word Stroop task (Stroop, 1935) allow the parametric variations of the two conflicting stimulus features (physical size and number magnitude). Distance effects can thus reflect a level of processing that surpasses the classification of numbers as either small or large, and shows that the magnitudes of each number are distinctly encoded (Tzelgov, Meyer & Henik, 1992). In short, the symbolic distance effect demonstrates that it is generally more difficult to discriminate between two stimuli that are similar than between two stimuli that are dissimilar (Tang et al., 2006).

Moyer and Landauer (1967) suggested that when a symbolic numeral is read it is converted to an analogue representation that allows a physical comparison between the two numerals to take place, similar to other spontaneous perceptual processes (e.g. comparing the sizes of two objects). This effect seems to be the case for both symbolically presented numbers such as number words (e.g. Foltz, Poltrock & Potts, 1984) and analogically presented numbers such as dot patterns (Buckley & Gillman, 1974).

Distance effects are typically observed in task-relevant dimensions, namely a physical size distance effect in the physical comparison task and a numerical distance effect in the numerical comparison task (e.g. Fias, Lammertyn & Reynvoet et al., 2003; Pinel, Piazza, LeBihan & Dehaene, 2004; Tang et al., 2006). In the task-irrelevant (to be ignored) dimensions, distance effects can either disappear (Rubinsten, Henik, Berger & Shahar-Shlev, 2002) or reverse (Girelli et al., 2000; Tang et al., 2006). The reversed distance effect operates as follows: during trials with a great difference in the task-irrelevant dimension, a judgement is made quicker and this quick automatic response interferes with the generation of a response to the task-relevant dimension, which is required. Thus, as the symbolic distance in the task-irrelevant dimension increases, time taken to make a judgement based on the task-relevant dimension increases. Processing of physical size and numerical magnitude thus seem to operate in a similar way, but independently of one another (Tang et al., 2006).

Not surprisingly, physical comparison is mostly faster than numerical comparison (e.g. Girelli et al., 2000). Physical size comparison is a spontaneous perceptual process whereas numerical comparison involves transcoding symbols to underlying magnitude meanings before a comparison can take place, which can be a time-consuming process. If physical size and numerical magnitude processing overlap/diverge depending on presentation format, differences in size congruity and distance effects could emerge between digits and number words. Furthermore, if experience with numbers aids the transcoding process from symbols to magnitudes, such effects could be further modulated by mathematics experience.

The present study employed a digit–word number Stroop variant with numerical magnitude and physical size as the two competing dimensions. While the symbolic distance effect has been noted with number words as well as digits, less research has addressed the size-congruity effect for number words. Besner and Coltheart (1979) originally reported the size congruity effect for digit stimuli, but found no effect for word stimuli. The words in this study were presented vertically, however, which confounded results according to Cohen-Kadosh et al. (2007) who found a size congruity effect for number words presented horizontally. However, this study did not include arabic digits for comparison and also only assessed numerical, and not physical, comparison of number words. In a subsequent study that included both digits and verbal numbers, size congruity effects were found for both formats, but to a lesser extent for verbal numbers (Cohen-Kadosh et al., 2008).

The stimuli used in the current task were based on the digit stimuli used by Tang et al. (2006), who highlighted a number of shortcomings with previous numerical magnitude Stroop tasks. Importantly, previous studies did not balance the task-relevant and -irrelevant dimensions (e.g. Girelli et al., 2000; Henik & Tzelgov, 1982). In the study of Girelli et al. (2000), for example, nine single arabic numerals were employed to create two levels of numerical distance, namely numerically ‘close’ and numerically ‘distant’ pairs. These were digit pairs with numerical distances of either 1 (e.g. **4 5**) or 5 (e.g. **4 9**). In the physical dimension, three different font sizes were employed to create physically large, small and neutral stimuli, which resulted in only one level of physical distance on congruent and incongruent trials, where one ‘small’ and one ‘large’ size was presented in each case. Physical and numerical distances were thus unbalanced in terms of level, which made it difficult to judge how much information was available in the task-relevant and -irrelevant dimensions on each trial. For this reason, Tang et al. (2006) parametrically varied both numerical and physical distance: the stimuli comprised of nine different numbers (arabic digits 1 – 9) as well as nine different physical sizes. From this, four numerical distances and four physical distances were created.

The current study used a similar method to create three physical and three numerical distances, but included arabic digits and number words, unlike Tang et al.'s (2006) study, which only considered arabic digits. In line with the number Stroop literature (e.g. Girelli et al., 2000; Tzelgov et al., 1992), Tang et al. (2006) found that participants automatically processed physical size and numerical magnitude of digit stimuli even when it is to be ignored under task instructions. Distance effects were also found suggesting that participants found it easier to respond to trials where the physical size or numerical distance between the numerals were great than when they were small. This suggests that the magnitude of each numeral was encoded distinctly, beyond a mere large–small comparison (Tang et al., 2006).

Tang et al. (2006) also found greater parietal activation for numerical magnitude comparison than physical size comparison and argued that numerical magnitude involves deeper processing than physical size. This was also found by Pinel et al. (2004) suggesting that physical size and numerical magnitude are processed differently. If numerical magnitude requires higher processing than physical size, we can expect an advantage for more mathematics experienced individuals on numerical comparison, but not necessarily on physical comparison. Such group differences would be informative of the role that experience with numbers plays in the automatic transcoding of numerical symbols to magnitudes. It is hypothesised that both digit and word stimuli would show task-relevant distance effects, demonstrating that distinct number magnitudes are accessed from both formats. The presence of size congruity effects could be informative of the degree to which task-irrelevant information interferes with the task, and in turn how the processing of the two formats differ. Chapter 2 showed that with high mathematics experience, an advantage is gained for processing digit stimuli, but not necessarily number word stimuli. Similar format-specific interference patterns were therefore predicted in the current tasks.

3.2. Method

3.2.1. Participants

Forty-five participants took part in the experiment (age 18 – 29; M = 23.22; SD = 3.26). All participants had normal or corrected-to-normal vision and spoke English as their first language. Participants took part in both tasks (physical and numerical comparison) and task presentation was counterbalanced across participants. The same method of assigning participants to High and Low Maths groups as in Chapter 2 (p. 42) was used. The High Maths group consisted of 10 men and 10 women (N = 20) and the Low Maths consisted of 9 men and 15 women (N = 24). The same exclusion criteria were used as in Chapter 2. One participant indicated that they had completed their Leaving Certificate in Irish and their data was therefore excluded from the analysis.

3.2.2. Apparatus and Materials

Stroop Task. Stimuli were presented on a 15-inch LCD monitor linked to a computer. The stimuli were the numbers 2 – 9 in arabic numeral form and written word form. The stimuli and task instructions were presented in black ink against a white background. Eight different numerical magnitudes (the numbers 2 – 9) and 8 different physical sizes were used. Each stimulus was positioned centrally on the screen and subtended between approximately 3.8 to 9.7 degrees of visual angle, with a mean ratio of 20 squared pixels between adjacent physical sizes (smallest was 50 squared pixels; greatest was 190 squared pixels). Word stimuli were in lower case and corresponded to these sizes vertically, with a mean ratio of 20 squared pixels between adjacent sizes (letter sizes were approximately the same as digit sizes). Programming for the task was done in *Superlab®,* which recorded participant input and reported reaction times in milliseconds as well as errors made.

From the 8 numbers (2 – 9) and pixel sizes (50 – 190 squared pixels), 3 numerical distances (ND) and 3 physical distances (PD) were created. Number pairs had a numerical distance of 1 (e.g. **2 3**), 3 (e.g. **2 5**) or 5 (e.g. **2 7**). Physical distance was also manipulated to create PDs of 1, 3 or 5 based on the 8 different physical sizes that were created. There were three trial types employed, namely a) Congruent, where the physically larger digit/word was also numerically larger (e.g**. 4 1**); b) Incongruent, where the physically larger digit/word was numerically smaller (e.g. **four one**); and c) Neutral, where the two digits or words were presented either in the same size (e.g. **2 4**; in the numerical comparison task) or where two identical digits or words were presented in different sizes (e.g. **three three**; in the physical comparison task).

Nine stimulus categories were used overall with 6 stimulus pairs in each. This resulted in 54 congruent word stimuli, 54 congruent digit stimuli, the same number of incongruent stimuli for each stimulus type, 18 neutral word stimuli and 18 neutral digit stimuli. Overall, each task (physical and numerical comparison) contained 252 trials. Participants also completed the same numeracy test, subjective numeracy scale (SNS) and working memory span tasks as described in Chapter 2 (p. 44 – 46).

3.2.3. Procedure

The experiment took place in a small windowless cubicle. Each participant completed both the physical and numerical comparison tasks individually and task presentation was counterbalanced across participants. Care was taken to ensure than no auditory or visual distractions would interfere with participants' task performance. Participants were provided with verbal instructions as to what the study would entail and were then asked to sign an informed consent form (Appendix 1).

Participants were told that two number words or two digits will appear on screen and that they might differ in physical size and/or numerical magnitude. In the physical task, participants were told to indicate which of the two numbers (left or right) were physically larger, while ignoring the meaning of the word or number. The experimenter emphasised that both speed and accuracy were important. Participants were told to use the index finger of each hand to either press the 'd' key (left of keyboard) to indicate that the word/digit on the left is the largest or the 'k' key (right of keyboard) to indicate that the word/digit on the right is the largest. Markers were placed on the keyboard to clearly highlight these response keys. Task instructions were then presented on-screen and participants were given an opportunity to ask questions if the task was not clearly understood. Similar instructions appeared for the numerical comparison task, but numerical comparison was required and physical comparison to be ignored. Participants were given two practice trials to further ensure that they clearly understood what was expected. The practice trials involved one stimulus from the incongruent digit condition and one stimulus from the incongruent word condition. Practice trials were also employed to prevent differences in keyboard familiarity from playing a role, but were limited to two due to the practice effect noted by Bush et al. (1998) in the counting Stroop task. The following instructions appeared for numerical comparison:

**YOU ARE ABOUT TO SEE SOME NUMBER WORDS AND DIGITS ON THE SCREEN.
YOU WILL SEE EITHER TWO WORDS OR TWO DIGITS AT A TIME.
EACH TIME YOU HAVE TO INDICATE WHICH OF THE TWO NUMBERS IS THE HIGHEST WHILE IGNORING THEIR PHYSICAL SIZES.**

**IF THE ONE ON THE LEFT IS THE HIGHER NUMBER PRESS THE 'D' KEY. IF THE ONE ON THE RIGHT IS THE HIGHER NUMBER, PRESS THE 'K' KEY.
SOMETIMES THE PHYSICALLY LARGER NUMBER WILL ALSO BE THE NUMERICALLY HIGHER NUMBER, HOWEVER SOMETIMES THE PHYSICALLY LARGER NUMBER WILL BE THE NUMERICALLY LOWER NUMBER.**

REMEMBER YOU HAVE TO INDICATE WHICH NUMBER IS NUMERICALLY HIGHER WHILE IGNORING THE PHYSICAL SIZE OF THE NUMBERS.

TRY TO GO AS FAST, BUT AS ACCURATELY AS YOU CAN.

PLEASE PRESS THE SPACE BAR TO DO SOME PRACTICE TRIALS.

Once it was clear that the task requirements were understood, the experimenter told the participant to commence the task by pressing the space bar when ready once the experimenter had left. Each stimulus remained on-screen until the participant responded by pressing either the 'd' or the 'k' key. An inter-stimulus interval of a 1000 milliseconds blank white screen was used. Stimuli were presented in a pseudo-random order with digits and words presented in the same test block.

After the participant had completed the first task, the experimenter explained that as part of the experiment the participant was required to complete some calculations. The experimenter then handed the participant the 17-item numeracy test and a blank sheet of paper and a pen. The same instructions were given to complete the numeracy test as is described in Chapter 2. The experimenter also noted each participant's age, gender, obtained grade in Leaving Certificate mathematics (e.g. A, B, C etc.) and the level of Leaving Certificate mathematics studied (Higher or Lower).

After the completion of the numeracy test, participants completed the SNS and the working memory span tasks, which followed the same procedure as described in Chapter 2. Participants then completed either the physical or numerical comparison task depending on which task had already been completed.

Participants were then thanked for their time and participation. The experimenter explained that the study investigated processing differences between arabic digits and number words. The experimenter addressed any remaining questions and emphasised that group, as opposed to individual, data were of interest.

3.3. Results

Reaction times (RTs) in the numerical and physical comparison tasks were recorded as time taken (ms) to press the 'd' or 'k' key on the keyboard after each stimulus appeared. Errors were also recorded and were excluded from the RTs analysis. Overall, 1.28 % of the data was excluded from the physical comparison task and 2.47 % from the numerical comparison task due to errors made. The mean RTs were calculated for each stimulus category for the High and Low Maths groups.

An independent samples t-test indicated that the High Maths group (*M* = 12.85; *SD* = 3.45) outperformed the Low Maths group (*M* = 10.38; *SD* = 2.28) on the numeracy test, $t(42) = 2.85$ $p = .007$. There was no significant gender difference in numeracy test scores.

Similar group differences were found for self-perceived numeracy and working memory as was found in Experiment 1. The High Maths group showed higher self-perceived numeracy ability (*M* = 4.54, *SD* = 0.89), than the Low Maths group (*M* = 3.53, *SD* = 1.14), $t(42) = 3.22$, $p = .002$, suggesting that participants' assessments of their own numerical ability was relatively accurate. Regarding the storage and transformation function of working memory, the High Maths group showed an advantage for backward digit span (*High M* = 76.08, *SD* = 12.51 and *Low M* = 63.67, *SD* = 13.73), $t(42) = 3.12$, $p = .003$. The difference between the groups on sentence span (*High M* = 80.2, *SD* = 16.44 and *Low M* = 71.5, *SD* = 13.52) was approaching significance ($p = .061$). No significant difference between the two groups were found for short-term memory ($p = .22$), measured as forward digit span (*High M* = 85.82, *SD* = 8.79 and *Low M* = 82.28, *SD* = 9.72).

3.3.1. Accuracy

Errors were classified as incorrect responses (i.e. pressing the wrong key). Table 3.1 presents the mean error percentages across the different conditions in the physical and numerical comparison tasks. Errors were minimal overall and were excluded from any subsequent analyses.

Table 3.1. *Mean percentages of errors across congruent, neutral and incongruent conditions in the physical and numerical comparison tasks.*

Maths Group	Congruent Digit	Neutral Digit	Incongruent Digit	Congruent Word	Neutral Word	Incongruent Word
			Physical Comparison			
Low	0.23	0.23	0.62	2.47	0.46	3.09
High	0.19	0.56	1.39	2.50	0.83	0.65
			Numerical Comparison			
Low	0.23	0.46	3.16	1.39	2.08	5.09
High	0.46	0.93	3.80	2.78	2.32	4.074

Physical Comparison Accuracy. On congruent trials performance was less accurate for word stimuli than for digit stimuli, whereas performance on neutral trials was relatively similar for the two formats. Differences between the two groups seemed to emerge on incongruent trials: the Low Maths group made more errors on incongruent word than incongruent digit trials, whereas the reverse pattern seemed to occur for the High Maths group with greater accuracy on incongruent word than incongruent digit trials. A 2 x 3 x 2 mixed between–within groups ANOVA was conducted to analyse the influences of format (arabic digits vs. number words), congruency (congruent, neutral and incongruent trials) and maths group (High and Low) on error rates in the physical comparison task. Overall, performance was more accurate on digit than on word trials, $F(1, 42) = 18.7$, $p < .001$ (partial eta squared = 0.31), and less accurate on incongruent trials, $F(2, 84) = 6.78$, $p = .002$ (partial eta squared = 0.14). However, significant format x congruency, $F(2, 84) = 8.25$, $p = .001$, and format x congruency x maths group interaction effects, $F(2, 84) = 6.85$, $p = .002$, were also found (partial eta squared = 0.16 and 0.14 respectively). No further main or interaction effects were found in the physical comparison task.

Dependent t-tests with Bonferroni corrections showed that for both groups the accuracy on neutral trials did not differ between digit and word stimuli. Congruency did not influence accuracy on digit trials, with relatively similar error rates across congruent, neutral and incongruent trials for both groups. The Low Maths group's performance was more accurate on neutral word trials (e.g. **three three**) than on incongruent (e.g. **two three**), $t(23) = -3.15$, $p = .005$, word trials and the decrease in accuracy on congruent, relative to neutral, word trials approached significance $t(23) = -2.96$, $p = .056$. This is likely to reflect the fact that when two identical words appear the physical size difference can be more obvious, which makes performance less error prone on neutral trials. In line with this, the High Maths group seemed to find it easier to compare neutral over congruent word stimuli, $t(19) = -3.76$, $p = .001$, however, no significant difference in accuracy was found between neutral and incongruent trials. As the task requires rapid responding, it could be the case that High Maths participants are more alert to incongruency, which enables a more cautious and accurate response decision to be made on incongruent trials.

Numerical Comparison Accuracy. In the numerical comparison task, where participants compared the magnitude of the numerals (e.g. **2 5**; which number is numerically higher?) error rates seemed to follow the expected congruency pattern with a linear increase in errors across congruent, neutral and incongruent trials. Overall, participants made more errors on word than on digit trials, but the pattern of performance for the Low and High Maths groups were relatively similar. The same analyses were conducted for error rates in the numerical comparison task as was done for the physical comparison task. Significantly more errors were made on word than on digit trials, $F(1, 42) =$

21.22, $p < .001$ (partial eta squared = 0.34), and overall more errors were made on incongruent trials, $F(2, 84) = 26.42$, $p < .001$ (partial eta squared = 0.39), with no further significant effects. In the High Maths group, the increase in errors on incongruent relative to neutral trials was significant for digit, $t(19) = -3.68$, $p = .002$, but not word trials, with no significant differences in error rates between congruent and neutral trials. For the Low Maths group error rates did not differ as a function of congruency.

Accuracy: Digits vs. Words. In the two tasks overall, more errors were made for word than digit stimuli. Congruency did not affect error rates for digit stimuli in the physical comparison task (error rates were relatively low). For word stimuli, however, the Low Maths group made significantly more errors on incongruent than on neutral trials, whereas the High Maths group showed no difference in this regard. With regards to congruent trials, the High Maths group made significantly more errors than on neutral trials, suggesting that neutral trials were easier to respond to than congruent trials. In the numerical comparison task, congruency only influenced accuracy for the High Maths group: more errors were made on incongruent relative to neutral digit trials.

3.3.2. Congruency

Physical Comparison Task. Figure 3.1 presents the mean correct reaction times (RTs) across congruent, neutral and incongruent trials in the physical comparison task. The RTs for digit stimuli showed an increase across congruent, neutral and incongruent trials. For word stimuli, RTs seemed relatively slow across all three congruency levels and RTs on both congruent and incongruent trials were slightly slower than RTs on neutral trials.

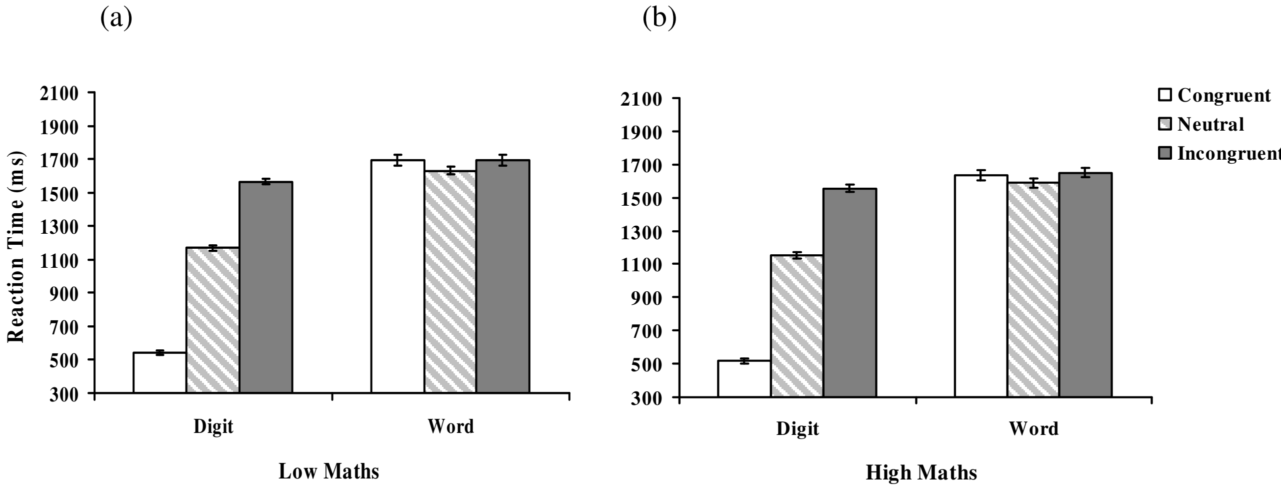

Figure 3.1. *Mean RTs (± SEM) in the physical comparison task across congruent, neutral and incongruent trials for (a) the Low Maths (n = 24) and (b) the High Maths group (n = 20).*

A 2 x 2 x 3 mixed between–within groups ANOVA was conducted on the mean RTs in the physical comparison task. The factors were Maths group, format and congruency. Main effects were found for congruency, $F(2, 84) = 2420.93$, $p < .001$, and format, $F(1, 42) = 2144.38$, $p < .001$, with large associated effect sizes (partial eta squared = 0.98 and 0.98 respectively). A significant format x congruency interaction effect was also found, $F(2, 84) = 2978.87$, $p < .001$, with a large associated effect size (partial eta squared = 0.99) indicating that congruency influenced RT on digit trials more than on word trials. No further main or interaction effects were found. Overall, the two groups showed similar response patterns suggesting that the High Maths group had no significant advantage for physical size comparison.

Dependent t-tests (Bonferroni corrected) showed that for digits, RTs on congruent stimuli (*High M* = 518.62, *SD* = 66.86; *Low M* = 541.38, *SD* = 70.44) were significantly faster than RTs on neutral stimuli (*High M* = 1152.62, *SD* =

74.6; *Low M* = 1169.78, *SD* = 78.9). This was found for both the High, $t(19) = 63.22$, $p < .001$, and Low maths groups, $t(23) = 112.286$, $p < .001$.

RTs on incongruent digit stimuli (*High M* = 1556.02, *SD* = 100.46; *Low M* = 1566.04, *SD* = 90.82) were significantly slower than RTs on neutral digit stimuli for both High, $t(19) = -37.97$, $p < .001$, and Low maths groups, $t(23) = -40.24$, $p < .001$. RTs on digit stimuli thus reflected the Stroop effect, namely a faster response to congruent stimuli (e.g. **2 5**) and a slower response to incongruent stimuli (e.g. **2 5**) relative to neutral stimuli (e.g. **2 2**).

Congruency also affected RTs on word stimuli, however RTs were relatively slow across all three congruency levels. In the Low Maths group, RTs on neutral trials (*M* = 1631.93, *SD* = 127.18) were slightly faster than RTs on both congruent (*M* = 1694.09, *SD* = 172.71), $t(23) = -3.89$, $p = .001$, and incongruent (*M* = 1692.76, *SD* = 161.96) word trials, $t(23) = -3.48$, $p = .002$. This followed the pattern suggested by the accuracy data in the Low Maths group, namely fewer errors on neutral stimuli overall. For the High Maths group, congruency did not seem to influence RTs on word trials: RTs were relatively similar on neutral (*M* = 1591.15, *SD* = 127.49) congruent (*M* = 1633.79, *SD* = 144.36), and incongruent (*M* = 1650.03, *SD* = 127.22) trials.

To summarise, in the physical comparison task, performance on digit stimuli reflected the Stroop effect, namely faster responses on congruent trials and slower responses on incongruent trials relative to neutral trials. Congruency effects were much less evident for word stimuli with relatively slow RTs across congruent, neutral and incongruent trials. The Low Maths group responded faster on neutral word trials than on congruent and incongruent word trials, whereas no effect of congruency was found for the High Maths group in this regard.

Numerical Comparison Task. Figure 3.2 presents the mean correct reaction times (RTs) across congruent, neutral and incongruent trials in the numerical comparison task.

(a) (b)

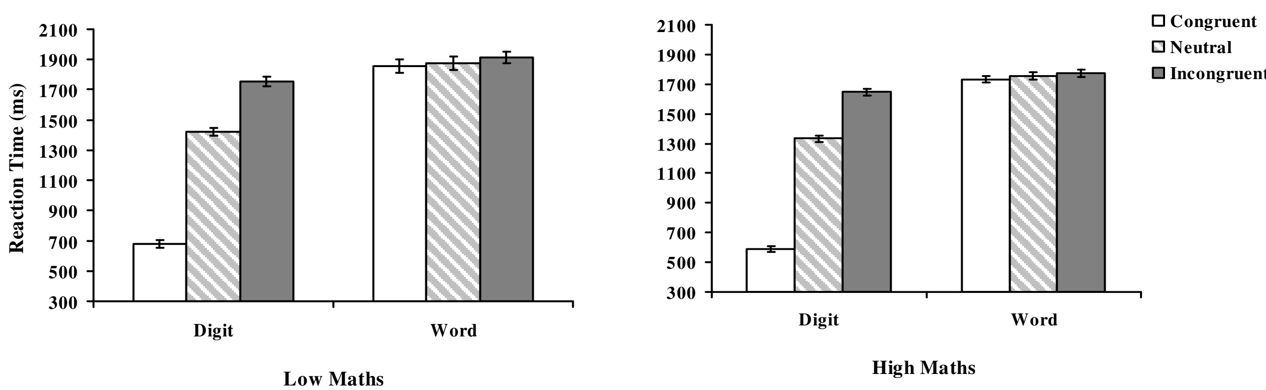

Figure 3.2. *Mean RTs (± SEM) in the numerical comparison task across congruent, neutral and incongruent trials for (a) the Low Maths (n = 24) and (b) the High Maths group (n = 20).*

A similar pattern was observed for digit stimuli in numerical comparison as was seen in physical comparison, namely an increase in RT across congruent, neutral and incongruent trials, whereas RTs on word trials were relatively slow overall. An analysis of variance was conducted for RTs in the numerical comparison task. Main effects were found for congruency, $F(2, 84) = 3034.36$, $p < .001$, and format, $F(1, 42) = 2472.42$, $p < .001$ (partial eta squared = 0.99 and 0.99). A significant congruency x format interaction effect was also found, $F(2, 84) = 2440.16$, $p < .001$, with a large associated effect size (partial eta squared = 0.98) showing that congruency influenced RTs on digit trials more than on word trials. A significant main effect was also found for Maths group, $F(1, 42) = 7.15$, $p = .011$ (partial eta squared = 0.15) indicating that the High Maths group showed faster performance overall.

Taking digit stimuli, dependent t-tests (with Bonferroni corrections) showed that the Stroop effect was significant. In the Low Maths group RTs on congruent digit stimuli (M = 676.55, SD = 120.59), were significantly faster, $t(23) = 55.414$, $p < .001$, and RTs on incongruent digit stimuli (M = 1754.56, SD = 159.35), were significantly slower, $t(23) = 55.41$, $p < .001$, than RTs on neutral digit stimuli (M = 1418.97, SD = 123.43). Similarly, in the High Maths group, RTs on congruent digit stimuli (M = 591.64, SD = 84.92) were significantly faster than RTs on neutral digit stimuli (M = 1332.88, SD = 104.07), $t(19) = 70.409$, $p < .001$, and RTs on incongruent digit stimuli (M = 1646.8, SD = 105.36) were significantly slower than RTs on neutral digit stimuli, $t(19) = -29.63$, $p < .001$. RTs on word stimuli were relatively slow overall and did not differ significantly across congruent, neutral and incongruent trials.

Congruency: Digits vs. Words. To summarise, congruency effects were overall considerably more prominent for digit than word stimuli in both physical and numerical comparison, with relatively slow RTs across all three levels of congruency for word stimuli. Taking the physical comparison task, the RTs on digit stimuli showed that incongruent digit meanings slowed down the size comparison and congruent digit meanings facilitated the size comparison. A different pattern emerged for number words in the physical comparison task with slower RTs overall. When comparing the physical sizes of two number words, Low Maths participants also seemed to find it easier when the two words were identical than when the two words were different (e.g. **three three** vs. **two three**) regardless of congruency.

While the congruency patterns were relatively similar for the two maths groups, the High Maths group made faster numerical comparisons, whereas no group differences were found for physical comparison. The possibility that the High Maths group were merely faster responders overall was thus ruled out (in which case faster responses would also have been expected in the physical task). In the numerical comparison task, the Stroop effect occurred for digit stimuli, whereas for word stimuli RTs did not differ significantly with congruency.

3.3.3. Distance Effects

Further analyses were conducted on the RTs on incongruent trials to test for distance effects. While the size congruity effect demonstrates the degree to which participants are able to ignore task-irrelevant numerical/physical stimuli, the distance effect provides a more refined measure of the automaticity of numerical processing. In number comparison tasks, the presence of distance effects demonstrates that the magnitude of each number has been processed distinctly (Tang et al., 2006; Tzelgov et al., 1992). Thus if each number is encoded beyond a mere small–large classification, the distance effect shows that it is generally easier to compare numbers that are numerically further apart (e.g. **2 7**) than numbers that are numerically closer together (e.g. **2 3**). In the current task three numerical distances (ND) and three physical distances (PD) were used. Number pairs had a numerical distance of 1 (e.g. **2 3**), 3 (e.g. **2 5**) or 5 (e.g. **2 7**). Physical distance was also manipulated to create PDs of 1, 3 or 5 based on pixel size. To investigate distance effects, incongruent trials were selected for analysis as they involve the simultaneous processing of two competing stimulus features (see Tang et al. 2006). Tables 3.2 and 3.3 present the mean correct RTs on incongruent trials at each level of physical and numerical distance in the two tasks. Figures 3.3 and 3.4 present the overall physical distance (PD) and numerical distance (ND) effects in the two tasks. Distance effects were classified as task-relevant (PD in the physical comparison task and ND in the numerical comparison task) or -irrelevant (ND in the physical comparison task and PD in the numerical comparison task).

Physical Comparison Task. In the physical comparison task, participants seemed to respond faster on trials where the physical distance was great than on trials where the physical distance was small (e.g. **4 6** vs. **4 6**) and this effect occurred for both digit and word stimuli. Numerical distance (the task-irrelevant dimension) did not seem to influence RT with relatively similar responses when numerical distance was great (e.g. **2 7**) and when it was small (e.g**. 2 3**). Figure 3.3 presents the distance effects in the physical comparison task for a) the Low Maths group and b) the High Maths group.

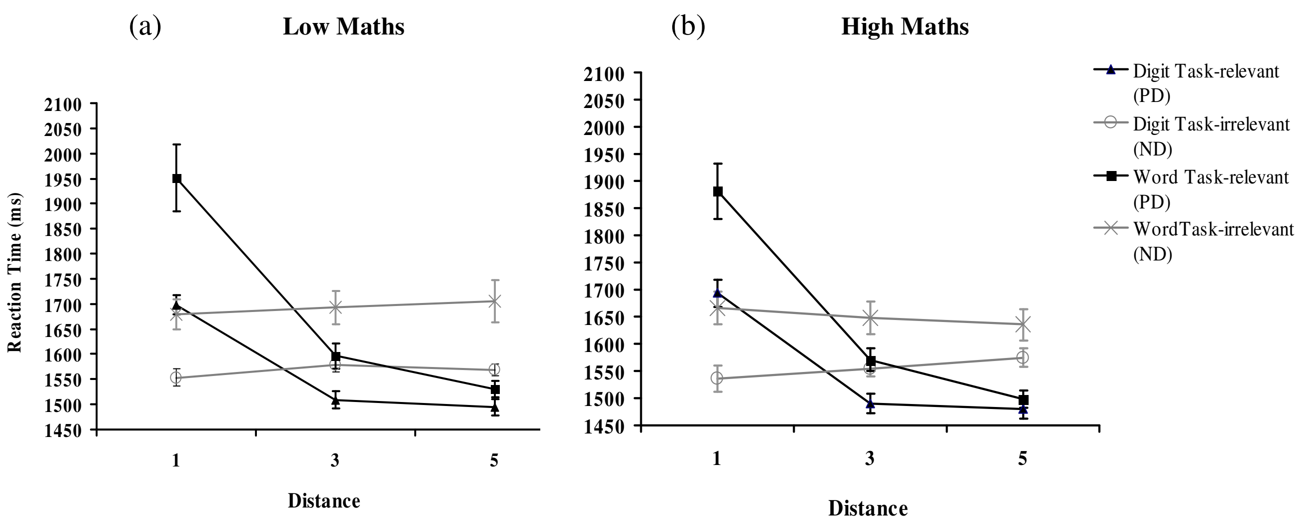

Figure 3.3. *Distance effects in the physical comparison task across task-relevant (physical distance) and -irrelevant (numerical distance) dimensions for (a) the Low Maths and (b) the High Maths group (± SEM).*

A 2 x 2 x 3 x 3 mixed between–within groups ANOVA was conducted with the factors format (digits and words), Maths group (high and low), physical distance (1, 3 and 5) and numerical distance (1, 3 and 5). A main effect was found for format, $F(1, 42) = 61.63$, $p < .001$, indicating overall faster response times on digit stimuli (partial eta squared = 0.59). RTs decreased significantly as physical distance became larger, $F(2, 84) = 172.59$, $p < .001$, (partial eta squared = 0.8). A significant format x physical distance interaction effect was also found, $F(2, 84) = 32.64$, $p < .001$, (partial eta squared = 0.44) suggesting that physical distance influenced the two formats differently. No effects were found for numerical distance or for Maths group.

Paired samples t-tests with Bonferroni corrections were conducted to investigate the significance of the distance effects. At task-relevant level, RTs on digit stimuli decreased as physical distance increased from PD 1 (*High M* = 1693.79, *SD* = 111.13; *Low M* = 1698.02, *SD* = 93.09) to PD 3 (*High M* = 1490.45, *SD* = 80.2; *Low M* = 1508.7, *SD* = 84.52), but not significantly from PD3 to PD 5 (*High M* = 1480.39, *SD* = 78.13; *Low M* = 1494.26, *SD* = 80.4). This decrease was significant from PD 1 to PD 3 for both High, $t(19) = 7.36$, $p < .001$, and Low Maths groups, $t(23) = 7.1$, $p < .001$, with overall similar responses for both PD3 and PD5.

For word stimuli RTs decreased as physical distance increased from PD 1 (*High M* = 1881.19, *SD* = 230.34; *Low M* = 1951.02, *SD* = 323.75) to PD 3 (*High M* = 1570.55, *SD* = 93.63; *Low M* = 1596.58, *SD* = 121.93) to PD 5 (*High M* = 1497.22, *SD* = 71.65; *Low M* = 1530.68, *SD* = 81.02). This decrease in RTs was significant from PD 1 to PD 3 for both High, $t(19) = 8.904$, $p < .001$, and Low, $t(23) = 7.42$, $p < .001$, maths groups and also from PD 3 to PD 5 for both High, $t(19) = 7.65$, $p < .001$, and Low, $t(23) = 3.76$, $p = .001$, maths groups. Overall, a relatively similar pattern of performance was found for the two groups with no significant advantage gained by the High Maths group on physical comparison.

To summarise, physical distance effects were found for digits and words: as the size-difference between the two stimuli became greater, the easier it became for participants to make a size comparison. However, different distance effects were observed for the two formats. For word stimuli, RTs decreased linearly as physical distance became greater (i.e. from PD 1 to 3 to 5), whereas RTs on digit stimuli was only significantly impeded when the size difference between the two numerals was small (PD 1). On subsequent physical distance levels (PD 3 and PD 5), with a greater size difference, responses were relatively faster overall and did not differ significantly. While physical comparisons of number words were slower overall, this difference in RT between digits and words became smaller as physical distance increased. Thus on trials with a great physical distance where physical comparison is easiest, the performance on word trials approached that of the performance on digit trials. No significant effects were found in the task-irrelevant dimension (numerical distance) suggesting that when physical comparison takes place, the exact numerical distance of the stimuli does not interfere with the process.

Numerical Comparison. Figure 3.4 presents the distance effects in the numerical comparison task across task-relevant and -irrelevant distance dimensions for a) the Low Maths group and b) the High Maths group (mean RTs in the physical and numerical comparison tasks across distance levels are presented in Tables 3.2 and 3.3).

Table 3.2. *Means and Standard Deviations of RTs on incongruent trials in the Physical Comparison Task for the Low (L) and High (H) Maths groups.*

		Physical Distance											
		1				3				5			
		Digit		Word		Digit		Word		Digit		Word	
		L	**H**	**L**	**H**	**L**	**H**	**L**	**H**	**L**	**H**	**L**	**H**
Numerical Distance	1	1675.67	1662.36	1925.53	1942.12	1500.87	1476.92	1581.79	1559.7	1484.18	1466.43	1530.92	1494.64
		(134.2)	(192)	(294.3)	(272.9)	(73.12)	(71.8)	(96.67)	(91.36)	(69.65)	(87.95)	(90.85)	(71.37)
	3	1702.05	1672.49	1949.23	1877.2	1540.40	1504.15	1603.46	1569.6	1503.72	1475.23	1526.33	1497.98
		(147.59)	(180.28)	(291.39)	(259.6)	(111.03)	(89.09)	(184.13)	(106.89)	(106.06)	(80.88)	(85.61)	(76.43)
	5	1707.74	1756.84	1978.27	1824.25	1484.81	1490.28	1604.47	1582.35	1494.89	1499.51	1534.79	1499.02
		(165.8)	(169.91)	(496.58)	(249.2)	(89.16)	(92.92)	(116.9)	(109.76)	(90.07)	(87.88)	(84.31)	(80.15)

Table 3.3. *Means and Standard Deviations of RTs on incongruent trials in the Numerical Comparison Task for the Low (L) and High (H) Maths groups.*

		Numerical Distance											
		1				3				5			
		Digit		Word		Digit		Word		Digit		Word	
		L	**H**	**L**	**H**	**L**	**H**	**L**	**H**	**L**	**H**	**L**	**H**
Physical Distance	1	1835.8	1687.1	1970.23	1823.22	1712.79	1615.6	1900.77	1751.35	1695.12	1594.48	1841.87	1747.16
		(349.82)	(119.05)	(225.5)	(138.04)	(129.89)	(105.06)	(275.24)	(130.27)	(124.72)	(110.02)	(153.97)	(124.01)
	3	1821.18	1690.69	2024.26	1831.57	1720.73	1640.89	1862.2	1762.62	1734.98	1635.17	1842.1528	1739.33
		(229.04)	(126.08)	(240.50)	(131.73)	(151.27)	(108.78)	(213.26)	(120.09)	(176.11)	(122.87)	(154.88)	(155.29)
	5	1799.34	1689.58	2007.73	1827.19	1752.92	1641.68	1925.20	1753.23	1718.21	1626	1844.6	1731.45
		(223.86)	(136.39)	(306.89)	(127.34)	(185.45)	(115.01)	(246.05)	(104.24)	(97.59)	(115.50)	(195.18)	(119.71)

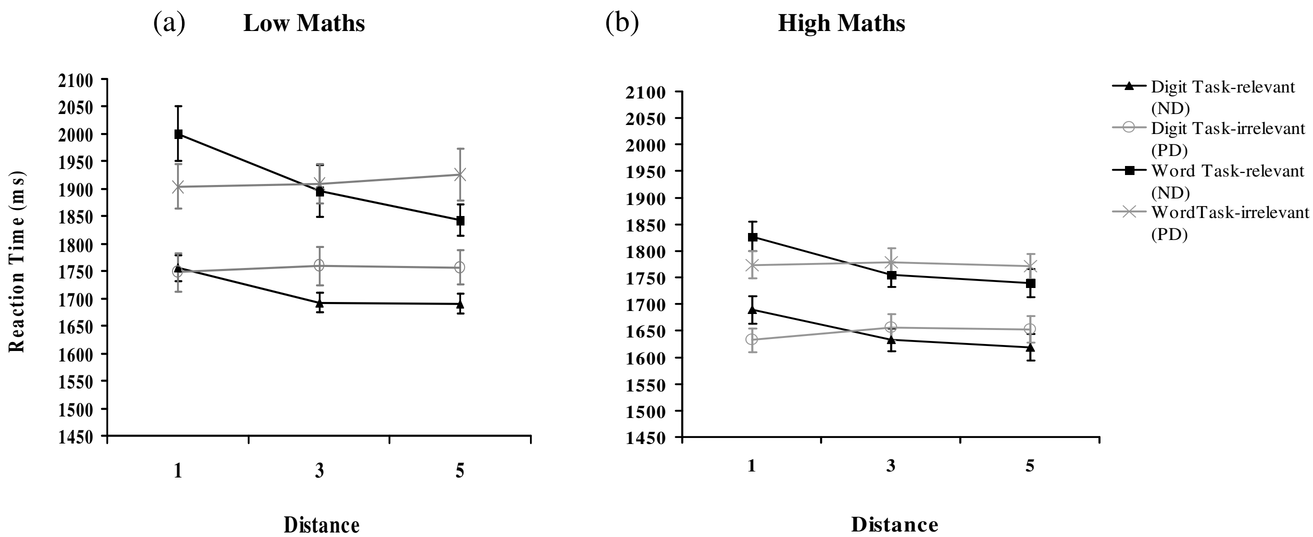

Figure 3.4. *Distance effects in the numerical comparison task across task-relevant (numerical distance) and -irrelevant (physical distance) dimensions for (a) the Low Maths and (b) the High Maths group (± SEM).*

The same analysis was conducted for the numerical comparison task where participants were asked to attend to numerical magnitude and to ignore physical size. Faster RTs were found on trials with a greater numerical distance (e.g. **2 7**) than on trials with a smaller numerical distance (e.g. **2 3**), whereas the physical distance of the stimuli did not seem to influence RTs. An analysis of variance was conducted with the factors of format, numerical distance, physical distance and Maths group. A main effect was found for format, $F(1, 42) = 195.24$, $p < .001$ (partial eta squared = 0.82), reflecting overall faster responses for digit than for word stimuli. Responses were overall significantly slower on trials with a small numerical distance (ND1), $F(2, 84) = 31.12$, $p < .001$, (partial eta squared = 0.43). Overall, the High Maths group responded faster than the Low Maths group, $F(1, 42) = 7.74$, $p = .008$ (partial eta squared = 0.16) suggesting an advantage for numerical comparison.

Paired samples t-tests with Bonferroni corrections were conducted to investigate the significance of the distance effects. RTs for digit stimuli decreased from ND 1 (*High M* = 1689.12, *SD* = 115.61; *Low M* = 1755.91, *SD* = 115.34) to ND 3 (*High M* = 1632.73, *SD* = 98.23; *Low M* = 1692.80, *SD* = 87.61), but not significantly from ND 3 to ND 5 (*High M* = 1618.55, *SD* = 112.03; *Low M* = 1690.41, *SD* = 87.83). The decrease from ND 1 to ND 3 was significant for both High, $t(19) = 4.92$, $p < .001$, and Low, $t(23) = 3.85$, $p = .001$, Maths groups.

While performance on word stimuli was slower overall, a similar numerical distance effect was found as for digit stimuli. RTs decreased from ND 1 (*High M* = 1827.33, *SD* = 120.95; *Low M* = 2000.74, *SD* = 244.28) to ND 3 (*High M* = 1755.73, *SD* = 105.17; *Low M* = 1896.06, *SD* = 227.63), but not significantly from ND 3 to ND 5 (*High M* = 1739.31, *SD* = 119.25; *Low M* = 1842.86, *SD* = 138.12). The decrease from ND 1 to ND 3 was significant for both High, $t(19) = 4.82$, $p < .001$, and Low, $t(23) = 5.75$, $p < .001$, maths groups.

To summarise, while responses on digit trials were faster overall, similar numerical distance effects were found for digits and words: on trials where numerical distance was small (ND1) participants found it more difficult to make a numerical comparison than on trials where numerical distance was greater (ND3 and ND5). No significant effects were found in the task-irrelevant dimension (physical distance) suggesting that although the physical size of digit stimuli interfered with RT (the size congruity effect presented in figure 3.2), the processing of physical size was not as refined as the processing of numerical magnitude when physical size is the task-irrelevant dimension. The data also show that for numerical comparison, the High Maths participants were faster responders on incongruent trials compared to Low Maths participants.

3.4. Discussion

Experiment 2 compared the processing of physical size and numerical magnitude for arabic digits and number words at different levels of mathematics experience. By adapting the task developed by Tang et al. (2006) the two dimensions of physical size and numerical magnitude were methodically manipulated in order to control the amount of interference for each of these dimensions. Considering arabic digits, the size congruity findings are in line with the Stroop literature (e.g. Girelli et al., 2000; Tang et al., 2006; Tzelgov et al., 1992), namely that participants found it difficult to ignore number meanings in physical comparison and difficult to ignore the physical sizes of stimuli in numerical comparison.

For number words however, the size congruity findings suggest differential processing of the two formats. Firstly, in the physical comparison task, the influence of congruency did not follow the expected Stroop pattern marked by faster responses for congruent and slower response for incongruent, relative to neutral word trials. Instead, for number words no influence of congruency was found for High Maths participants and Low Maths participants seemed to find it easier to respond to neutral word trials than to congruent or incongruent word trials. As response times for word stimuli were relatively slow overall, this effect seems to be related to the difficulty in responding to two different words, compared to responding to two identical words, in which case the size difference between the two words is more visually obvious. While care was taken to match letter sizes to digit sizes, this effect could also be related to the fact that different number words differ in physical length (e.g. ‘two’ vs. ‘three’). This could explain why a similar effect was not observed for digit format trials in the physical comparison task, in which case the Stroop effect was found.

Task-relevant distance effects were found in both physical and numerical comparison at task-relevant level, namely a physical distance effect in physical comparison and a numerical distance effect in numerical comparison, in accordance with the Stroop literature (Fias et al., 2003; Pinel et al., 2004; Tang et al., 2006). The presence of a distance effect suggests that processing has gone deeper than small/large classifications and that the magnitude meanings of each number have been encoded (Tzelgov et al., 1992). As distance effects were found for digits and number words, it suggests that this automatic process takes place for both formats at task-relevant level.

For word stimuli, the presence of a numerical distance effect suggests that automatic access to number meanings is gained, but the absence of a size congruity effect suggests that participants found task-irrelevant information easier to ignore. Cohen-Kadosh et al. (2007) interprets distance effects from a purely verbal point of view and argue that, verbally, numbers are connected to each other in a similar way to semantic relations between different words (e.g. **DOG** and **CAT**). Numbers that are closer together are thus more difficult to compare as they are verbally more connected to the same category than numbers that are numerically further apart. The further apart the numbers are, the less the verbal code would interfere. If this interpretation is followed, magnitude representations need not even come into effect when two number words are being compared. A purely symbolic distance effect could thus take place for word stimuli, thus explaining the diminished size congruity effects.

High Maths participants showed an overall advantage on numerical comparison, whereas no such group difference was found in the physical comparison task. This suggests that the High Maths participants were not merely faster responders overall, but that the advantage lies in the processing of numerical magnitude *per se*. Physical size and numerical magnitude processing could thus diverge at a cognitive level in relation to an individual's experience with numbers. This finding is in line with Zorzi and Butterworth's (1999) classification of numbers as either "discrete numerosities" or physical sizes presented analogically. The former is thought to require higher processing. Tang et al. (2006) also found greater parietal activation for numerical distance processing relative to physical distance processing indicating a quantitative difference in processing between numerical and physical distance. Experience with numbers could therefore confer an advantage for numerical comparison, whereas no advantage is gained for physical comparison which takes place at a perceptual level and does not depend on higher level processing (Tang et al., 2006).

To conclude, the current study showed differences in processing between digits and number words in terms of the size congruity and distance effects. This suggests that the process of gaining access to underlying number meanings occurs more automatically for arabic digits than for number words. High Maths participants were also faster at numerical comparison in arabic digit as well as number word format, suggesting an advantage for numerical processing from various formats. Since High Maths participants showed this advantage for digits and number words, such an effect might also occur for extracting number meaning from language more generally. To investigate this possibility, Experiment 3 employed quantifier words that do not convey number meanings as explicitly as arabic digits or words.

Chapter 4

Experiment 3: Congruency Effects in a Quantifier Word Counting Task as a Function of Mathematical Experience

4.1. Introduction

Evidence from neuropsychology suggests that the degree to which a digit or number word activates underlying number representations reflects the degree to which the semantic referent of the format is preserved (see Chapter 1, p. 22). Quantifier words, such as “both” or “each” have also been investigated in this regard (e.g. Cappelletti, Butterworth & Kopelman, 2006; McMillan, Clarke, Moore & Grossman, 2006; Troiani, Peelle, Clark & Grossman, 2009). Lexically, quantifier words operate similarly to other words; however, semantically these words refer to quantities (e.g. Cappelletti et al., 2006). As these words do not convey number meaning explicitly, the question of whether or not these words mainly reflect a numerical or linguistic representation remains uncertain, with little existing research on this question (Cappelletti et al., 2006).

Neuroscientific evidence has provided some insights, suggesting that the processing of quantifier words operate more numerically than linguistically. Such studies show that the same brain areas are activated during numeral and quantifier word processing. McMillan and colleagues (2005), for example, found right intraparietal activation during a true/false quantifier word judgement task (McMillan, Clark & Moore et al., 2005). During this task participants saw an array of objects and had to judge whether the sentence presented with the array (e.g. “some of the balls are blue") was 'true' or 'false'. The observed right intraparietal activation was consistent with the literature that highlighted its involvement in number knowledge (e.g. Chochon, Cohen, van de Moortele & Dehaene, 1999; Simon, Mangin & Cohen et al., 2002). Patients with selective impairment of numerical comprehension, such as Corticobasal degeneration (CBD) also generally display both a numeral and quantifier word processing deficit, despite being unimpaired in other language functions (McMillan et al., 2006). Cipolotti, Butterworth and Denes (1991) also presented a patient that suffered a stroke whose severe numerical deficits could not be attributed to impairments in language or memory. Number knowledge thus seems to play a central role in quantifier word comprehension.

Similarly, Troiani et al. (2009), based on evidence from both healthy and dyscalculic adults, argue that abstract number knowledge is central to quantifier comprehension. In brain organisation, a dissociation seems to be evident between numerical quantifiers (e.g. "at least three") and logical quantifiers (e.g. "some"), with the former depending on areas typically involved in number processing and the latter depending on areas involved in focusing attention on specific elements in a distribution ("conceptual logic"). Specifically, quantifiers that are explicitly related to cardinal knowledge activate lateral parietal-dorsolateral prefrontal regions, which are also involved in numeral comprehension. Logical quantifiers, on the other hand, activate a rostral medial prefrontal-posterior cingulate network, suggesting that processing of general quantifiers is more 'logically' than numerically based (Troiani et al., 2009).

However, Cappelletti et al. (2006) argued that even quantifiers that refer to approximate (e.g. “some” and “every”) rather than exact quantities are organised in a more numerical than linguistic pattern in the brain. This argument was based on the observation that patient AM who suffered from semantic dementia, had preserved comprehension of quantifier words, despite being impaired in the understanding of non-quantifier words of the same frequency. The patient’s numerical knowledge was also preserved, while the meanings of words, objects and linguistic concepts were impaired. Cappelletti et al. (2006) thus argued that the reason that quantifier as well as numeral knowledge was preserved is that the semantic referent of quantifier words, namely the number domain, was preserved.

In the development of quantifier understanding, it seems that learning and experience play a fundamental role in whether exact or general meanings of quantifier words are understood (Barner et al., 2009). However, whereas quantifier and numeral acquisition in infancy is significantly correlated, they do not seem to facilitate the development of one another (see Chapter 1, p. 12). Extensive training and experience is needed in order for children to learn the specific numerical meanings of quantifier words. The understanding of quantifier words thus seems to become more and more ‘numerical’ (meaning based) and less ‘linguistic’ with development as specific number knowledge increases. These findings strongly suggest that quantifier processing does not operate on a purely linguistic basis and also support the argument that semantic memory is organised in different specific domains, an example of which is number (Caramazza & Shelton, 1998; McMillan et al., 2005).

CHAPTER 4

Most previous research regarding the numerical vs. linguistic processing of quantifier words has been neuropsychological in nature and has not taken individual differences related to mathematics into account. However, the processing of quantifier words could also be investigated by means of the counting Stroop task. As discussed in Chapters 1 and 2, during counting Stroop tasks the two highly automatic processes of counting small numbers (subitizing) and reading are placed in competition with one another and the degree to which one of these processes slows down the processing of the other is indicative of the degree to which the two systems overlap (e.g. Brugger, Pietzsch, Weidmann & Biro, 1995). The original counting Stroop task (Bush et al., 1998; 2006; see Chapter 2) found congruency effects for number words such that congruent trials resulted in a speeded response (Stroop facilitation) and incongruent trials resulted in a slowed response (Stroop interference). Such effects arise from the automatic spontaneous processing of certain stimulus features (e.g. number of words on-screen) even if these features are to be ignored under task demands. Thus, if quantifier knowledge is largely dependent on number knowledge we might predict greater effects of congruency, if any, for individuals with more mathematical education and experience.

Since the previous experiments showed an advantage for High Maths participants in number encoding from arabic digits and number words, the current experiment investigated if this effect might also emerge for quantifier words. The stimuli used in the current counting task were quantifier words that could either refer to specific quantities (e.g. “both” or “each”) or general quantities (e.g. “some” or “much”) and trials were also manipulated to be congruent, neutral or incongruent in meaning.

4.2. Method

4.2.1. Participants

Thirty participants took part in the experiment (age 18 – 29; *M* = 21.7; *SD* = 3.42). The sample consisted of 15 High Maths (9 men; 6 women) and 15 Low Maths (8 men; 7 women) participants based on the same criteria as Experiments 1 and 2. The same exclusion criteria as was used in Experiments 1 and 2 were also used here.

4.2.2. Apparatus and Materials

Quantifier word counting task. Fifteen quantifier words and 5 neutral words were selected as stimuli in the experiment. These consisted of 10 quantifier words referring to the numbers ‘one’ or ‘two’ (5 each), 5 ‘General Quantifier’ words and 5 neutral words. The ‘One Quantifier’ words were: ‘first’, ‘unit’, ‘single’, ‘once’ and ‘each’. The ‘Two Quantifier’ words were: ‘second’, ‘pair’, ‘double’, ‘twice’ and ‘both’. These two stimuli groups were matched in terms of common part of speech, such that, for example, the equivalent of the word ‘first’ in the One Quantifier word list would correspond to the word ‘second’ in the Two Quantifier word list. The General Quantifier words were: ‘few’, ‘little’, ‘some’, ‘much’, and ‘many’.

Five neutral words were chosen on the basis that they do not semantically relate to quantity or measurement. The neutral stimulus group consisted of the words: ‘still’, ‘lady’, ‘busy’, ‘soon’ and ‘able’. The neutral stimuli were matched to the critical stimuli for number of letters, number of syllables, Kucera-Francis written frequency and Brown verbal frequency using the MRC Psycholinguistic database (Coltheart, 1981).

The One, Two and General Quantifier words were presented as either congruent or incongruent trials. For the One and Two Quantifier words, congruent trials involved the number that the word corresponded to (one or two) matching the number of identical words presented on-screen (e.g. '**twice twice**' or '**first**'; respond '2' and '1' respectively), whereas incongruent trials involved the number that the word corresponded to mismatching the number of identical words presented on screen ('**twice**' or '**first first**'; respond '1' and '2'). Congruency was also manipulated for the General Quantifier word trials such that when the words corresponding to smaller quantities (e.g. 'little' or 'few') were presented once on-screen these acted as congruent trials (e.g. '**little**'; respond '1'). Trials where these words were presented twice on-screen (e.g. '**little little'**; respond '2') acted as incongruent trials.

Similarly, the General Quantifier words corresponding to greater quantities (e.g. 'much' or 'many') presented twice on-screen (e.g. '**many many**'; respond '2') acted as congruent trials; and presented once on-screen (e.g. '**many**'; respond '1') acted as incongruent trials. Neutral trials consisted of the neutral words (e.g. still, soon etc.) presented either once or twice at a time on-screen.

Each stimulus was presented three times in a test block, except for the neutral trials where each stimulus was presented 6 times (3 times as one word on-screen and 3 times as two words on-screen) resulting in a total of 120 trials. There were 45 congruent, 45 incongruent and 30 neutral trials overall. Trials were presented in a quasi-random order on a laptop screen. Each stimulus was positioned centrally on the screen and subtended between approximately 1 to 1.9 degrees of visual angle, presented in black print against a white background. Programming for the task was done in *Superlab ®,* which recorded all participant input and reported reaction times (RTs) in milliseconds and accuracy.

Participants also completed the same numeracy test, subjective numeracy scale (SNS) and working memory span tasks as described in Chapters 2 and 3.

4.2.3. Procedure

The experiment took place in a small windowless cubicle. Each participant was told that the study would investigate the processing of numerical stimuli and was then asked to sign an informed consent form (Appendix 1). The participant then completed the 17-item numeracy test (Lipkus et al., 2001). The same procedure was followed for the numeracy test as is described in Chapters 2 and 3.

The experimenter then asked the participant their age, whether they had studied ordinary or higher level Leaving Certificate Mathematics and the grade they obtained. Once the demographic information had been collected and participants had completed the numeracy test. The following message appeared on-screen:

YOU ARE ABOUT TO SEE SOME WORDS ON SCREEN.

EACH TIME THERE WILL BE EITHER ONE OR TWO WORDS PRESENT. YOU HAVE TO INDICATE AS FAST, BUT AS ACCURATELY AS YOU CAN, HOW MANY WORDS ARE PRESENT EACH TIME.

IF THERE IS ONE WORD ON THE SCREEN PRESS THE ‘D’ KEY ON THE KEYBOARD.

IF THERE ARE TWO WORDS, PRESS THE ‘K’ KEY ON THE KEYBOARD.

PRESS THE SPACE BAR TO TRY SOME PRACTICE TRIALS.

The experimenter also told the participant to try and ignore the meaning of the words and to just indicate the number of words by either pressing the ‘d’ key (left of keyboard) or the ‘k’ key (right of keyboard). Participants were clearly instructed that the ‘d’ key should be pressed to indicate that one word is on-screen and that the ‘k’ key should be pressed to indicate that two words are on-screen. The index fingers of both hands were recommended for pressing the corresponding keys and the keys were labelled. It was also emphasized that both speed and accuracy were important in the task.

Once the task instructions were read and explained in more detail by the experimenter, the participant pressed the space bar and two practice trials followed. The practice trials involved one stimulus from the incongruent ‘One Quantifier’ condition (e.g. ‘**each each**’) and one stimulus from the incongruent ‘Two Quantifier’ condition (e.g. ‘**double**’). Bush et al. (1998) noticed a practice effect in the counting Stroop task, with improved performance emerging after a few minutes. Practice trials were therefore limited to two.

Once it was clear that the task instructions were understood, the experimenter left the room and the participant commenced the experiment by pressing the space bar when ready. Each stimulus remained on-screen until the participant responded by pressing either the ‘d’ or ‘k’ key. An inter-stimulus interval of 1000 milliseconds blank white screen was used and trials were presented in a quasi-random order. After participants completed the quantifier task, they completed the subjective numeracy scale (SNS) and working memory span tasks as described in Chapter 2.

4.3. Results

Reaction times (RTs) were recorded as response latencies to press the d ('1') or k ('2') key on the keyboard to indicate the number of words on-screen after each stimulus onset. Mean correct RTs were calculated for each participant on each of the seven stimulus categories. These were congruent (One, Two and General Quantifier words), incongruent (One, Two and General Quantifier words) and neutral (neutral words) trials.

After assessing the normality of the sample, one outlier was removed from the High Maths group as most of these scores were extreme data points. An independent samples t-test indicated that on average the High Maths participants (M = 12.357; SD = 3.692) outperformed Low Maths participants (M = 9.533; SD = 2.997) on the numeracy test, $t(27) = 2.269$, $p = .032$. Men also outperformed women on the numeracy test (*Men* M = 12.24, SD = 3.49; *Women* M = 9, SD = 2.89), $t(27) = 2.63$, $p = .014$. There were no significant working memory or self-perceived numeracy differences between the High and Low Maths groups.

4.3.1. Accuracy

Errors were minimal and were excluded from the reaction times (RTs) analysis. Overall error for the High Maths group was 2.8 % and for the Low Maths group 3 %. Table 4.1 presents the mean error percentages for the different stimulus categories for the High and Low Maths groups. Overall, congruency did not seem to have a strong effect on error rates. The only significant increase in errors on incongruent conditions was found for the High Maths group in the One Quantifier condition, $t(13) = -3.12$, $p = .008$ (see Table 4.1). Overall, 2.9 % of the data was excluded from any subsequent RT analyses due to errors made.

Table 4.1. *Mean percentages of errors for the Low and High Maths groups across congruent, neutral and incongruent quantifier word trials.*

Group	Neutral	Congruent General	Incongruent General	Congruent One	Incongruent One	Congruent Two	Incongruent Two
Low	3.33	2.67	3.11	1.78	1.78	3.11	4.89
High	2.61	3.33	3.81	0.48	3.33	2.86	3.33

4.3.2. Reaction Time

Figure 4.1 presents the mean correct RTs across the different stimulus categories for a) the Low Maths and b) the High Maths groups. On average, there seemed to be a small increase in RT on incongruent relative to congruent trials. However, the RTs were relatively slow overall across One, Two and General Quantifier trials. The performance of the High Maths group was also faster overall than the performance of the Low Maths group. A 2 x 3 x 2 mixed between–within groups ANOVA was conducted to analyse the RTs differences between congruent and incongruent trials. The factors were congruency (congruent and incongruent), and word type (One, Two and General Quantifier) and Maths group. A significant main effect was found for congruency, $F(1, 27) = 11.02$, $p = .003$, indicating that RTs on congruent trials were faster overall than RTs on incongruent trials, however the size of the effect was small (partial eta squared = .29). A main effect was also found for Maths group, $F(1, 27) = 12.6$, $p = .001$, reflecting the faster overall performance of the High Maths group (partial eta squared = .32). No further main or interaction effects were found.

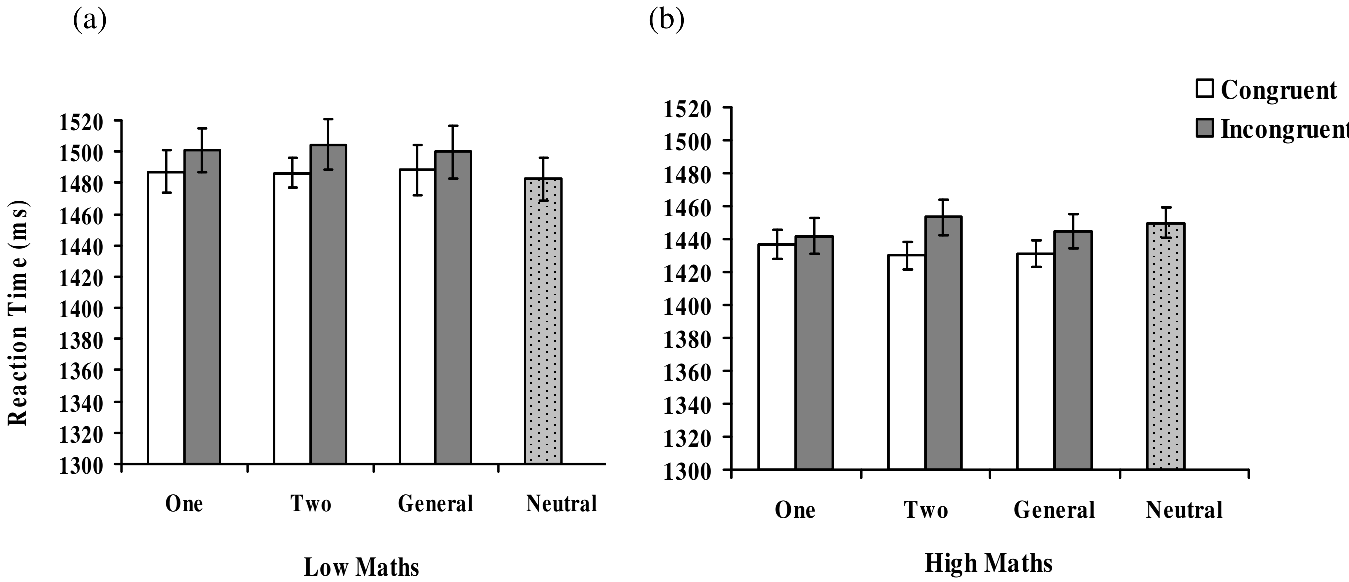

Figure 4.1. *Mean RTs (± SEM) across congruent, neutral and incongruent trials for One, Two and General Quantifier words for (a) the Low Maths (n = 15) and (b) the High Maths (n = 14) group.*

On average, High Maths participants responded 23.33 ms faster on congruent Two Quantifier trials (M = 1430.02, SD = 30.88) than on incongruent Two Quantifier trials (M = 1453, SD = 40.16), $t(13) = -2.925$, $p = .012$ (Bonferroni corrected). No significant effect of congruency was found for General Quantifier (*congruent* M = 1431, SD = 29.34 and *incongruent* M = 1444.77, SD = 38.22) and One Quantifier words (*congruent* M = 1436.62, SD = 32.21 and *incongruent* M = 1441.75, SD = 40.91; $p = .55$).

For Low Maths participants, no significant RT differences were found between congruent and incongruent trials for One (*Congruent* M = 1487.19, SD = 52.02 and *Incongruent* M = 1500.65, SD = 54.77), Two (*Congruent* M = 1486.38, SD

= 37.6 and *Incongruent M* = 1504.56, *SD* = 62.94) or General (*Congruent M* = 1488.27, *SD* = 62.83 and *Incongruent M* = 1499.81, *SD* = 66) Quantifier words (all $p > .15$).

To summarise, differences in RT between congruent and incongruent trials were only found in the High Maths group. High Maths participants responded slightly faster on trials where quantifier word meaning matched the number of items on-screen (e.g. **both both**; correct response ‘2’) compared to where quantifier meaning did not match the number of items on-screen (e.g. **both**; correct response ‘1’). The analysis showed that this congruency effect was only due to the Two Quantifier words (e.g. both, second, double etc.). The Low Maths group showed no advantage for congruent over incongruent trials. To investigate Stroop interference and facilitation effects, the RTs on congruent and incongruent trials were compared with RTs on neutral trials and are discussed in the following section.

4.3.3. Interference and Facilitation

Regarding Stroop facilitation and interference (the RTs discrepancy between congruent/incongruent and neutral trials), paired samples t-tests (Bonferroni corrected) showed a facilitation effect in the High Maths group, with significantly faster RTs on congruent relative to neutral trials in the Two Quantifier condition, $t(13) = 3.94$, $p = .002$. However, no facilitation effect occurred for One and General Quantifier words. No interference effect was found as RTs on incongruent trials did not differ significantly from RTs on neutral trials in any of the conditions. In the Low Maths group, RTs on neither congruent nor incongruent trials differed significantly from RTs on neutral trials.

The RT data showed that the High Maths group responded faster than the Low Maths group on congruent and incongruent trials (all $p \leq .015$) whereas the two groups did not differ significantly in RT on neutral trials (*Low M* = 1482.559, *SD* = 53.28 and *High M* = 1449.708, *SD* = 34.33; $p = .06$). This suggests that the High Maths group's observed advantage for congruent Two Quantifier trials was not just due to these participants merely being faster responders, in which case an advantage would have been expected on neutral trials as well.

To summarise, congruency did not seem to have a strong influence on RTs with relatively slow RTs overall. However, any significant effects related to the congruency of the stimuli were only found in the High Maths group; a facilitation effect was found, showing faster RTs on congruent than neutral trials for quantifier words relating to the number 'two'. Specifically, High Maths participants were slightly faster to respond on trials where the number or words matched the quantifier word (e.g. **both both**) than on trials where a neutral word was presented (e.g. **still still**). No interference effect was found, however, since RTs on incongruent conditions did not differ significantly from RTs on neutral conditions. The Low Maths group showed no significant effects of congruency.

4.4. Discussion

During the counting Stroop task the two highly automatic processes of subitizing and reading are placed into competition (as discussed in Chapter 2), and effects related to the congruency of the stimuli are thought to be indicative of the degree to which the words are processed automatically (e.g. Bush et al., 1998; 2006). The current task investigated the processing of quantifier words by means of a counting Stroop task. Unlike the original counting Stroop task with number words, congruency did not seem to influence overall response times to a great extent. However, the small, but significant congruency effects that did emerge were only found for High Maths participants. As previous research suggested that the reading of quantifier words do not operate on a purely linguistic basis (e.g. Cappelletti et al., 2006; McMillan, Clarke, Moore & Grossman, 2006), the current study further supports this argument for individuals with greater mathematics experience. Although a minimal effect overall, the facilitation effect found for High Maths participants is similar to the facilitation effect found in the traditional counting Stroop task (Bush et al., 1998; 2006), namely that a congruent number word (e.g. **two two**) speeds up the counting process. As quantifier words semantically refer to number and lexically operate similarly to other words (Cappelletti et al., 2006), this finding could be interpreted as more spontaneous access to the underlying number meanings of some quantifier words as a result of experience with numbers.

The observed congruency effect was related to Stroop facilitation rather than Stroop interference, as is generally observed in numerical Stroop tasks (e.g. Bush et al., 2006; Girelli et al., 2000). The processes that give rise to facilitation effects, however, are not as clear to account for as those that give rise to interference effects (the latter occurring due to parallel processing of two conflicting stimulus features). MacLeod and Macdonald (2000) caution against the view of facilitation as the advantage of congruence mirroring the disadvantage of incongruence and argue that the two effects should rather be seen as reflecting different processing mechanisms or bases of responding. In traditional counting Stroop tasks, for example, the main difficulty in interpreting facilitation effects is that on congruent trials a response to either the meaning *or* the number of the items (e.g. two two; respond ‘2’) would result in the correct response (MacLeod & MacDonald, 2000), despite the fact that the task requires only responding to the number of items. In other words, on congruent trials there is a ‘double chance’ of responding correctly. This is not the case for incongruent trials, where only a response to the task-relevant dimension (number of words) would be correct. In light of this, it is possible that undetected responses based on reading are included in the overall RT of congruent trials. MacLeod and Dunbar (1988), for example, provided evidence for this in an experiment where reading errors were filtered. In the current task, facilitation could thus be based on reading response times, which circumvents the counting process.

However, even if facilitation is partly based on reading responses, the fact that this effect only occurred in the High Maths group points towards the involvement of number knowledge in quantifier word processing for this group. Also, while High Maths participants responded faster overall, the two groups did not differ significantly in reaction time on neutral stimuli. The possibility that the facilitation effect is due to the High Maths participants being faster readers was thus ruled out. Instead, the effect could be accounted for as follows: an arbitrary response number line was created from one (left) to two (right), where quantifier meanings could correspond to either one of these numbers. Given that quantifier words do not convey number meanings as explicitly as number words (e.g. in the original counting Stroop), if responses were based on reading of the quantifier words, participants had to map the number meaning of the quantifier word onto this arbitrary number line (e.g. the word ‘both’ corresponds to the number ‘two’ in the number line). Performance on congruent trials could thus still, at least to some extent, reflect the automaticity with which quantifier words are transcoded to underlying number meanings. This facilitation effect was, however, only found for quantifier words relating to the number ‘two’ and not for ‘one’ or ‘general’ quantifier words, which limits strong conclusions being drawn regarding individual differences in the processing of quantifier words.

A number of methodological issues relating to the stimuli employed in the current study are also worth mentioning, which could account for the limited effects of congruency that were noted. Although responses on congruent Two Quantifier trials were faster than responses on neutral and incongruent trials in the High Maths group, no significant congruency effects were found in any of the other quantifier conditions. For One and General Quantifiers, RTs on congruent trials did not differ significantly from incongruent trials. A possible explanation for why the facilitation effect only occurred in the Two Quantifier condition could be that the words chosen to represent the number ‘one’ in the task do not do so as explicitly as the words chosen to represent the number two. While the two stimulus groups were matched in terms of word frequency and part of speech, those in the Two Quantifier word category always relate to the number ‘two’, whereas the meanings of those in the One Quantifier word category could be more dependent on sentence context . In language usage, the word ‘both’, for example, always refers to the number ‘two’, whereas the corresponding word ‘each’ could be used when describing a whole collection of objects and referring to ‘each’ object in the whole collection. The number ‘one’ might thus not be the only or most immediate representation that is evoked when the word ‘each’ is presented. Similarly, the word ‘every’ could refer to any number of individual objects in a collection of items. Processing of other words, such as ‘first’ might be more explicitly related to cardinal number knowledge and might thus operate similarly to the Two Quantifier words.

Regarding the neutral stimuli, the overall slow RTs and lack of an interference effect, could have been due to the fact that the neutral words used in this study were not all in a single semantic category, unlike the animal names used in the original counting Stroop task (Bush et al., 1998). However, the neutral stimuli were matched to the critical stimuli for number of letters, number of syllables, Kucera-Francis written frequency and Brown verbal frequency using the MRC Psycholinguistic database (Coltheart, 1981), which is why animal names were not suitable. Care was also taken to ensure that the neutral words did not semantically relate to quantity or measurement in any way (although the word ‘lady’ like the animal names in the original counting Stroop might suggest ‘one’). Furthermore, the words in each of the critical stimulus categories were not in a single semantic category either. The main differences between critical and neutral stimuli were thus that the critical stimuli had numerical meanings whereas the neutral stimuli (arguably) did not. Group differences in RTs are thus believed to reflect this, as the two groups only differed in quantifier word RTs and not neutral word RTs.

Finally, the small, but significant, congruency differences found between the two groups could be related to the sampling method used. A more robust measure of assigning participants to groups of mathematics experience might yield greater differences. As High Maths participants in this study were those who obtained a grade of a C3 or higher for higher level Leaving certificate mathematics, whereas the rest of the participants were placed in the Low Maths group, highly mathematics experienced individuals could have been included in the Low Maths group. Also, unlike Experiments 1 and 2, the two groups did not differ in working memory or self-reported numerical efficacy, suggesting that the differences between the High and Low Maths groups could have been more robust in Experiment 3. The sample could also have been more balanced in terms of gender, with women being relatively unrepresented in the High Maths group (9 men and 5 women), which could have contributed to the gender difference in numeracy scores. Nonetheless, in accordance with the findings of Experiments 1 and 2, the group differences obtained in Experiment 3 shows that the advantage that individuals with greater mathematics experience show for extracting number meaning from stimuli could also extend to quantifier words.

Overall, the findings from Experiments 1 to 3 suggest that processing differences between individuals with differing levels of mathematics experience are evident when considering basic numerical processes such as number comparison and subitizing. While no response time differences occurred on neutral stimuli, individuals in the High Maths group were generally faster responders on trials that contained numerical stimuli. For the High Maths individuals, a heightened appreciation for numerical information seems to emerge, in particular for digit format, but also to a lesser extent for number words and quantifier words.

Chapter 5

Experiment 4: Format-specific Effects in Arithmetic at Different Levels of Mathematics Experience: Evidence from Eye-tracking

5.1. Introduction

Stimulus format seems to affect early stages of number processing (e.g. Ischebeck, 2003), such as subitizing (Experiment 2) or number size comparison (Experiment 3). However, mixed reports exist regarding the influence of numerical surface format on more advanced numerical functions such as arithmetic (e.g. Bernardo, 2001; Rickard et al., 1994). As discussed in Chapter 1 (p. 15 – 18), some theorists argue that number representations are independent from the input format (e.g. Dehaene & Cohen, 1995; McCloskey & Macaruso, 1995) and that performance is not expected to differ with different surface formats. Others argue that surface format influences calculation per se (e.g. Campbell, 1994; Campbell & Epp, 2005) and that arabic digit operands would result in better performance than word operands.

Some recent studies have suggested that effects of surface format can arise due to different formats promoting or hindering the use of different strategies in arithmetic, such as counting or directly retrieving the answer from memory (Campbell & Alberts, 2009; Campbell et al., 2004; Szücs & Csépe, 2004). Different effects of format seem to emerge for each of the four arithmetic operations (addition, subtraction, multiplication and division) reflecting the different strategies of problem solving that are promoted by digit/word format in each. In an arithmetic task that compared addition, subtraction, multiplication and division, Campbell and Alberts (2009) presented problems in digit and word format (e.g. ‘2 + 3’ or ‘two + three’) and asked participants to report the strategies that they used to solve each problem. The strategies used were direct memory retrieval, reference to another operation (e.g. if ‘2 + 3 = 5’ then ‘5 – 3 = 2’), using knowledge of a related problem (e.g. if ‘2 + 2 = 4’ then ‘2 + 3 = 5’) and counting one by one (e.g. 1 + 1 + 1 + 1 + 1 = 5). In addition and subtraction, direct retrieval was the most common strategy reported, followed by addition/subtraction reference and counting. Retrieval was used less in subtraction and on larger number problems. Participants also reported that word format problems promoted counting strategies over direct retrieval, which corresponded with slower RTs on word format problems. In multiplication and division however, the cost of word format on retrieval was much less evident than in addition and subtraction.

Following Campbell and Alberts (2009), the lack of, or less prominent, format effects related to performance and strategy reports could be an indication that participants use similar strategies for digit and word stimuli (e.g. memory based strategies such as direct retrieval or multiplication-reference). Operations that utilise one dominant strategy for problem solving (e.g. memory retrieval in multiplication; Campbell & Xue, 2001; LeFevre & Morris, 1999) do not show much difference in performance between digit and word format problems since both formats promote the use of the same strategy, namely retrieval. Other operations, that can either be solved through retrieval or procedural strategies (e.g. addition), for example, show clearer effects of format, as word format seems to promote procedural strategies and digit format seems to promote direct memory retrieval (e.g. Campbell & Alberts, 2009).

Since retrieval is generally a faster process than calculation, the cost of word format on performance is very evident here.

Effects of format in arithmetic can thus highlight the degree to which problem solving processes operate separately from input format. One line of evidence that sheds some light on the debate on the influence of surface format comes from the problem size effect (PSE; see Chapter 1, p. 14), which shows that response time and errors in an arithmetic problem usually increase when the operands in the problem increase in magnitude (e.g. 2 + 3 vs. 8 + 9; see Ashcraft & Christy, 1995; Geary, 1996). Smaller numbers have stronger memory retrieval strength due to more extensive exposure, which make small number problems easier to solve (Zbrodoff & Logan, 2005). Larger problems are more likely to be solved by a strategy other than retrieval (e.g. calculation or reference to another operation), which takes longer and can be more error prone (Campbell & Xue, 2001). Studies consistently show that the problem size effect is greater for numbers written in number word format (e.g. two + three) compared to arabic digit format (e.g. 2 + 3; e.g. Campbell et al, 1999; Campbell & Alberts, 2009). Campbell and colleagues suggest that the slower performance on large word format problems is because retrieval processes are less efficient with number words than with arabic digits (see Campbell & Epp, 2005, for review). Campbell and Fugelsang (2001), for example, found that in a simple addition true–false verification task, participants reported the use of procedures (e.g. counting vs. retrieving answer from memory) much more with words (41 %) than with digits (26 %) and that this effect was even greater for large number problems (see also Campbell & Penner-Wilger, 2006). As both format and problem size seem to influence strategy choice, a switch to an alternative strategy occurs when retrieval strength is low. The cost of word format on retrieval is thus more evident on large number problems with already low retrieval strength.

Apart from Campbell and Alberts's (2009) study, the influence of format on strategy use in arithmetic has only been demonstrated for addition. More evidence is thus needed to link format effects with arithmetic strategies in subtraction, multiplication and division. Campbell and Alberts (2009) suggested that the use of retrieval is related to the efficiency and accuracy with which an answer can be accessed from long-term memory. It follows then that individual differences related to mathematics should influence the use of retrieval in arithmetic and that this might differ between formats. The current study considered mathematics experience and how it can further regulate the interactions of format, problem size and operation in arithmetic. As Campbell and Alberts (2009) suggested, operation and problem size effects in arithmetic should reflect long-term learning and experience. If this is the case, individuals with high mathematics experience could show an advantage for a) transcoding between number formats, b) solving large number problems and c) arithmetic fact retrieval in general. Effects of format, problem size and operation should thus reflect these advantages of high mathematics experience.

In addition to behavioural measures, the current experiment employed eye-tracking to explore processing differences that can occur as a result of surface format. As Zhang et al. (2010) pointed out, reaction time data is insufficient for highlighting processing differences between formats, since different formats might still yield similar behavioural responses, despite being processed along separate pathways.

Eye-tracking offers a more precise measure of processing than reaction time or accuracy (e.g. Desroches et al., 2006) as it provides an index of the location as well as the duration of fixation on certain stimuli (Merkley & Ansari, 2010). Eye-tracking has proven a useful technique in reading tasks by, for example, illustrating the process of integrating information read with information stored in memory and the pattern in which the information is processed (see Liversedge & Findlay, 2000 for a review). In reading, gaze duration has been argued to be an indication of access to an internal lexicon and integrating text information with existing knowledge in memory (e.g. Inhoff, 1984, 1985; Rayner & Pollatsek, 1987). Eye-tracking thus seems to be particularly useful for investigating strategies such as direct memory retrieval of arithmetic facts versus calculation. With regards to numerical cognition specifically, Merkley and Ansari (2010) have also recently employed measures of fixation count and fixation duration to study numerical magnitude processing and showed that both measures revealed additional effects that were not evident from behavioural data alone.

Eye-tracking might also be a useful alternative to self-reports in studies of strategy use in arithmetic. Whereas most research on strategy use in arithmetic have employed self-reports of participants (e.g. Campbell & Alberts, 2009; Campbell & Penner-Wilger, 2006), some shortcomings have been noted, which calls the validity and reliability of this approach into question. As cognitive processes that have become automatic are generally not readily available to self-report (e.g. Crutcher, 1994; Ericsson & Simon, 1993; Wilson, 1994), it could be argued that only information in short-term memory that is attended to could be consciously reported (e.g. Kirk & Ashcraft, 2001). Arithmetic strategies which involve working memory, such as counting, might also be more accurately reported than those that occur more automatically, such as direct memory retrieval (Kirk & Ashcraft, 2001). In the case of the latter, the answer is reportable, but not the strategy, which is believed to be a largely automatic process (e.g. Ashcraft, 1992; Campbell & Graham, 1985; Lebiere & Anderson, 1998). Strategy self-reports might thus not always be an accurate reflection of the actual strategies employed in a task.

Furthermore, strategy self-reports may influence strategy choice in a task. If participants are aware that they would have to report strategies after performing each operation, they might deliberately engage in the use of certain strategies in order to arrive at an answer (Kirk & Ashcraft, 2001). While self-reports in such a task might be highly accurate, they might not be indicative of the typical strategies that an individual would use in arithmetic. Indeed, Kirk and Ashcraft (2001) biased instructions towards either direct retrieval or non-retrieval based strategies in a simple arithmetic task and found that participants' strategy reports were highly influenced by these instructions. However, self-reports have still been shown to converge with RT findings based on the assumption that procedural strategies are generally slower than direct retrieval (Campbell & Alberts, 2009; Campbell & Penner-Wilger, 2006) suggesting some level of validity to strategy self-reports. Nonetheless, self-reports should still be interpreted with caution since instruction can strongly influence strategy choice (Kirk & Ashcraft, 2001).

In light of these shortcomings, the current experiment did not record participants' self-reports of strategies used. Instead, the current study is the first to use eye-tracking measures (average fixation count and duration) to study effects of format and problem size in simple arithmetic. The degree to which format and problem size effects differed across conditions and with mathematics experience was presumed to reflect the use of different strategies. Eye-tracking patterns were expected to indicate differences in cognitive processes between the different conditions, and were assumed to eliminate any strategy report biases.

The use of eye-tracking technology in the study of arithmetic problem solving has been limited and such studies have mostly investigated arithmetic word problems (e.g. De Corte, Verschaffel & Pauwels, 1990; Hegarty, Mayer & Monk, 1995; Verschaffel, De Corte & Pawels, 1992). Merkley and Ansari (2010) recently noted that, surprisingly, eye movement patterns in number processing have to date not been systematically investigated. While behavioural measures are useful, we can devise more precise and testable hypotheses of the underlying mental processes involved in a task if we know where, when and for how long participants looked at a certain stimulus (Merkley & Ansari, 2010). Furthermore, if eye-tracking patterns, like self-reports, also converge with reaction time data, it would suggest that strategy self-reports are valid and that eye-tracking measures provide another index for investigating strategy use in arithmetic. However, if eye-tracking measures diverge from RT and accuracy data, it could suggest that eye-tracking measures pick up on subtle underlying cognitive processes that are not evident from reaction time and accuracy data alone.

Very few studies have employed this technique to investigate arithmetic in arabic digit format. However, the few studies that have done this have shown that eye movements differ with the steps required in the different conditions. Suppes (1990) recorded eye movements of three participants while performing single-digit subtraction and addition problems in column format (the procedure taught in schools). Eye fixations varied according to structural features of each problem, such as the number of columns and whether or not the operation required a carry or borrow action. Verschaffel and colleagues presented 8 and 9 year olds with addition problems with three addends presented in a horizontal line (e.g. 2 + 5 + 6; Verschaffel, De Corte, Gielen & Struyf, 1994). Fixations of at least 100ms were identified and the final gaze that lasted the longest (at least 180 ms) was assumed to be the number that participants added to the other two operands. Importantly, eye movements concurred with verbal reports of strategies: participants rearranged items so as to first add two complimentary numbers that equalled 10 or two identical numbers together.

The current study closely followed the method of Campbell and Alberts (2009). The first part compared the performance of addition and subtraction problems across different levels of problem size and across digit and word format equations. Following Campbell and Alberts (2009), greater format X problem size interactions were expected for addition problems, as problems written in word-format were expected to be more taxing on memory retrieval (and thus performance) for larger, more difficult addition problems, relative to problems in digit-format (e.g. Campbell & Penner-Wilger, 2006). This effect was not expected for subtraction problems, based on the argument that in education subtraction is introduced after addition and taught as inverse addition, rendering subtraction subordinate to addition

(Campbell & Alberts, 2009). Larger subtraction problems are thus often solved by strategies that make use of the addition-reference (e.g. Le Fevre et al., 2006). Small subtraction problems are solved primarily through retrieval (however this retrieval strength is relatively weak compared with addition) with small subtraction performance closely matching large addition performance (Cambell & Xue, 2001). Small subtraction problems were thus expected to display similar word-format performance costs to addition problems. On large subtraction problems, on the other hand, due to their retrieval strength being too low to promote retrieval strategies even in digit format, word-format effects were not predicted to be particularly prominent, as was found in the study of Campbell and Alberts (2009).

Part 2 of the study compared format and problem size effects for multiplication and division across different levels of mathematics experience. In multiplication and division, counting-based strategies are rarely used due to their relative inefficiency as a strategy: counting would involve repeated addition and subtraction to solve multiplication and division problems respectively (Campbell & Xue, 2001). Instead, the main strategy for multiplication is direct memory retrieval (e.g. Campbell & Xue, 2001; LeFevre, Bisanz & Daley et al., 1996; LeFevre & Morris, 1999) and for division either direct retrieval or multiplication-reference (e.g. 2 x 3 = 6 therefore 6 ÷ 3 = 2; LeFevre & Morris, 1999; Mauro, LeFevre, & Morris, 2003). Regarding format effects, Campbell and Alberts (2009) found that similar to addition and subtraction, word format hindered retrieval, but that this effect was greater for division than multiplication. The reason for this is that the multiplication-reference strategy is efficient enough for division problem solving to afford a rapid shift away from retrieval. Multiplication, on the other hand, does not afford a shift away from retrieval as another strategy would be too inefficient (division-reference or counting is unlikely) and both digit and word formats are thus mainly solved via retrieval. Format and problem size effects are still expected in multiplication, as the retrieval strength of word format or large problems are generally weaker than digit format or small problems. However, these effects are still unlikely to encourage strategies other than retrieval and would thus be less evident in multiplication (Campbell & Alberts, 2009). To investigate effects of format and problem size a number of dependent measures were used in Experiment 4. The four dependent measures were 1) accuracy 2) response latency in milliseconds 3) total number of fixations across each problem and 4) the average fixation duration across each problem. In all four operations, performance was expected to be generally poorer for Low Maths participants. Participants across each level of Maths group were expected to answer simple arithmetic problems (e.g. 2 + 2) accurately and relatively quickly. However, the eye-movement patterns were expected to reflect subtle between–groups differences in arithmetic fact retrieval that might not be evident from behavioural measures alone. Stimulus format and problem size was also expected to interact with problem solving strategies, which were expected to be reflected in the eye-tracking measures.

5.2. Method

5.2.1. Participants

Eighteen women and 23 men participated in the study with ages ranging from 18 to 30 (M = 23.1, SD = 4.25). The current study aimed to improve on the previous method (see Chapters 2 – 4) of assigning participants to groups of mathematics experience, by including a 'Middle' Maths group. Participants were recruited so as to have three groups of differing mathematics experience ('High', 'Middle' and 'Low' Maths) based on self-reported Leaving Certificate performance, numeracy test results and third level mathematics education. Twelve participants were recruited from a university department of mathematics and comprised the High Maths group (9 men and 3 women). These were individuals who are currently completing advanced mathematics courses (e.g. abstract mathematics) at degree level or who had completed a degree course in mathematics. Of these 12 participants, 9 had studied Higher Level Leaving Certificate mathematics and had obtained a grade in the A (N = 4) or B (N = 5) range (one participant had studied Ordinary Level Mathematics with an obtained A grade). Two participants in the High Maths group had also studied Higher level Leaving Certificate mathematics with an obtained grade of a D. Both these participants had obtained the maximum score in the numeracy test (17) and had also studied mathematics to degree level and were therefore included in the High Maths group. Seventeen participants were assigned to the 'Middle Maths' group (8 men and 9 women). These participants were those who had studied Ordinary Level mathematics with an obtained grade A or Higher Level mathematics with obtained grade A – D at Leaving Certificate level, but did not hold/were not pursuing a degree in mathematics. Of these participants, 11 had studied Higher Level Leaving Certificate mathematics (grades A = 1, B = 9 and C = 1) and 6 had studied Ordinary Level Leaving Certificate mathematics (A = 6). The twelve participants in the 'Low Maths' group (6 men and 6 women) were those who had studied Ordinary Level mathematics at Leaving Certificate level with an obtained B grade or lower and who did not hold/were not pursuing a degree in mathematics. Two participants (both from the Middle Maths group) had to be excluded from the analysis due to missing eye-tracking data. One participant from the Low Maths group was excluded from the multiplication and division analyses (Part 2) due to too many errors made.

5.2.2. Apparatus and Materials

Participants completed the17-item numeracy test adapted from Lipkus et al.'s (2001) Numeracy Scale as well as on-screen arithmetic tasks of addition, subtraction, multiplication and division.

Eye-tracking apparatus. Participants' eye movements were recorded at 50 Hz with a remote Tobii 1750 eye tracker manufactured by Tobii Technology AB (Tobii, Stockholm, Sweden). A chin rest was used in order to ensure that all participants were seated the same distance from the computer screen (approximately 60cm) and to minimise any head movements. The Tobii system's analysis software, ClearView was used to identify participants' total fixation count and fixation durations per stimulus.

Arithmetic Stimuli. The stimuli, based on the study of Campbell and Alberts (2009) were addition problems ranging from 2 + 2 to 9 + 9 and corresponding subtraction problems (4 – 2 to 18 – 9) presented in arabic digit or word format. For subtraction pairs the second number in the addition problem became the subtrahend (e.g. 3 + 8 became 11 – 8). For each operation, 36 pairings of the numbers 2 to 9 were used, ignoring operand order (e.g. 3 + 4 and 4 + 3) resulting in a total of 72 stimuli per operation and stimulus format (144 problems per block). The stimuli included 8 tie problems (e.g. 2 + 2 or 8 + 8). These were included in the test blocks, but were excluded from the analysis due to unique encoding characteristics (e.g. Campbell et al., 2004). Operand order for each of the non-tie pairs were selected quasi-randomly and constrained such that the same operand order was used for digit and word format versions of the same pair. Pairs with a product of less than or equal to 25 were classified as small problems and those with a product of more than 25 were classified as large problems. Campbell and Alberts (2009) used this method to define problem size in order to have two balanced sets of 18 problems and to aid comparison of results across operations and with previous research (e.g. Campbell & Xue, 2001). In Part 2 (multiplication and division) problems ranged from 2 x 2 to 9 x 9 in multiplication and from 4 ÷ 2 to 81 ÷ 9 in division. Apart from the different stimuli, the same procedure and analyses were followed as for the addition and subtraction task (Part 1). Participants took part in both Part 1 and 2 of the study, with order of Parts 1 and 2 counterbalanced across participants. Each trial began with a fixation dot which flashed twice over a 2 second interval. The problem appeared on what would have been the third fixation flash, with the operation sign (+ or –) appearing in the space where the fixation dot would have been (Campbell & Alberts, 2009). Each problem remained on-screen until the participant responded by pressing the space bar.

Stimuli were presented centrally on a computer screen and subtended between 1 and 1.9 degrees of visual angle. Each equation appeared horizontally on the computer screen in either arabic digit or lower case written number word format in white ink against a black background. Digit operands were separated by a ‘+’ or ‘–’

sign with three character spaces on each side of the sign. Digit problems occupied 9 or 10 character spaces and word problems occupied 10 – 17 spaces in length (each character space was approximately 3mm wide and 5 mm high). Figure 5.1 shows examples of the digit and word format stimuli employed.

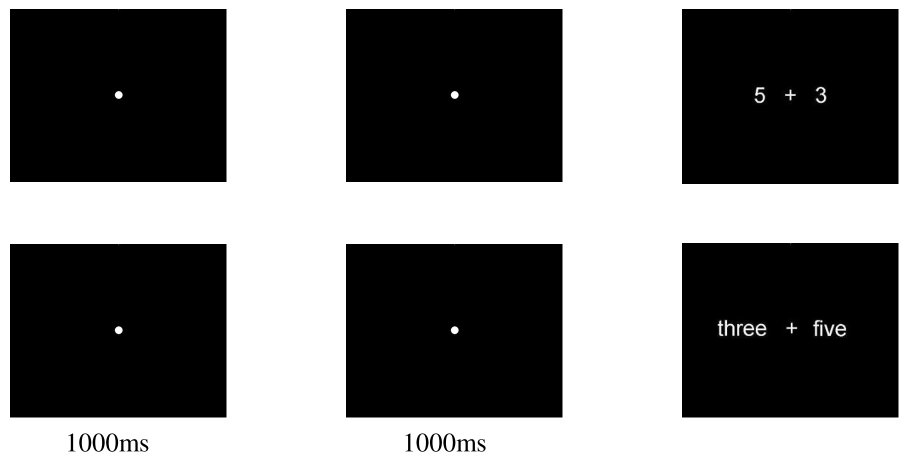

Figure 5.1. *Presentation examples of digit and word format equations*

Note. Fixation dots were presented twice over a 2 second interval. The problem appeared on what would have been the third fixation flash and stayed on-screen until the participant responded.

Numeracy Test. The same numeracy test adapted from Lipkus et al. (2001) as was used in the previous experiments was used in this task. However, given the more competent sample, participants were given six, as opposed to eight, minutes for the test to guard against possible ceiling effects. A blank sheet of paper was provided to work out the answers. The experimenter also noted demographic information regarding each participant’s age, gender and mathematics experience (Leaving Certificate and third level education).

5.2.3. Procedure

The experimenter told the participant that the experiment would involve a number of arithmetic tasks and a numeracy task. The task followed the same procedure as the Campbell and Alberts (2009) study. However, instead of using a voice recorder, which can be subject to microphone failure, participants were asked to press the space bar once they know the answer to the problem and then to report the answer verbally, which was noted by the experimenter. The following task instructions were given, both verbally and on-screen:

A NUMBER OF EQUATIONS WILL APPEAR ON THE COMPUTER SCREEN. YOUR TASK IS THIS: AFTER EACH EQUATION, PRESS THE SPACE BAR WHEN YOU KNOW THE ANSWER. THEN SAY THE ANSWER OUT LOUD. SPEED AND ACCURACY ARE IMPORTANT.

After the experimenter had given basic task instructions, the eye-tracker was calibrated to ensure that gaze direction could be accurately calculated for each individual. This was done for each participant prior to commencing the arithmetic tasks. A series of practice trials then followed. The practice trials involved four problems (two digit and two word problems) for each operation. Once the experimenter was confident that the participant understood the task instructions, the participant was told that the task would begin. Participants were told to try and answer the problems correctly, but to also try and answer as fast as they could. The experimenter emphasised that it was important not to sacrifice accuracy for speed.

The order of completing Parts 1 (addition and subtraction) and 2 (multiplication and division) was counterbalanced across participants.

Before each test block the word 'ADDITION' or 'SUBTRACTION' appeared on-screen to indicate which operation the block would test. Participants completed two test blocks of 72 trials (one addition and one subtraction block). Order of presentation of the test blocks were counterbalanced across participants. Each test block consisted of all 36 problems once presented as digits and once presented as words. Following Campbell and Alberts (2009), digits were presented on odd trials and words were presented on even trials. The problem order was quasi-random, but constrained such that digit and word versions of the same problem did not appear within at least 10 trials of one another. Each participant received the same order of trials.

The experimenter remained in the room during the arithmetic tasks, seated at a table behind the participant with an answer sheet to record participants' accuracy on the task. Trials on which the participant pressed the space bar, but did not promptly report an answer were noted and were excluded from the analyses (these were minimal overall).

After the participant had performed the addition and subtraction tasks, the experimenter offered the participant a 5 minute break if desired. The participant then performed the multiplication and division tasks (part 2). Apart from the different stimuli, Part 2 followed the same procedure as Part 1. After completing all the arithmetic tasks, the participant was then asked to complete the timed 17-item numeracy test. To control for ceiling effects, participants were only given 6, as opposed to 8, minutes to perform the test. Instructions for the numeracy test were the same as for the other experiments (see Chapter 2, p. 44). Participants were then debriefed and it was explained that group as opposed to individual data was of interest in this study.

5.2.4. Ethical Considerations

In addition to the ethical considerations that were set out in Chapter 2 (p. 50), participants were asked not to participate in Experiment 4 if they had prior head injury, suffered from epilepsy or had any reading or visual difficulties. None of the participants were excluded based on these criteria. Before commencing the tasks, each participant was given an information sheet with general information on what to expect from an eye-tracking study (Appendix 4) and each participant was required to read the information sheet before the experiment began.

5.3. Results

Reaction times (RTs) were recorded as time taken in milliseconds to press the space bar after each stimulus onset. The eye-movement measures were the total number of fixations and average fixation duration (ms) per stimulus. Errors were also recorded and were excluded from RT and eye-movement analyses. Overall, 2.08 % of the data was excluded from Part 1 (addition and subtraction) and 4.75 % from Part 2 (multiplication and subtraction) due to errors made. For each part (1 and 2) a separate 2 x 2 x 2 x 3 mixed between–within groups analysis of variance was conducted for each dependent measure. The dependent measures in each case were accuracy, RT, number of fixations and average fixation duration. In each case the influences of operation (e.g. addition and subtraction), format (digits and words), problem size (small and large) and Maths group (Low, Middle and High Maths groups) were investigated. Post-hoc comparisons used the Tukey HSD test.

Preliminary t-tests were conducted to assess the differences between the three Maths groups on the numeracy test. The Low Maths group (M = 8, SD = 3.22) was outperformed by both the High (M = 13.92, SD = 3.55), $t(22) = 4.27$, $p < .001$, and Middle Maths groups (M = 12.2, SD = 3.63), $t(25) = 3.14$, $p = .004$. The High Maths group did not perform significantly better than the Middle Maths group ($p = .23$). Overall, men outperformed women on the numeracy test (men M = 13.43, women M = 8.94), $t(38) = 3.97$, $p < .001$.

5.3.1. Part 1: Addition and Subtraction

Accuracy. Errors were minimal overall and did not seem to differ with mathematics experience or format. However, problem size seemed to have an influence on overall error rates. On average participants made more errors on large problems (addition 2.06 % and subtraction 4.48 %) relative to small problems (addition 0.78 % and subtraction 1 %). Table 5.1 presents the percentage of errors made in addition and subtraction problems for the High, Middle and Low Maths groups.

Table 5.1. *Mean percentages of errors in the addition and subtraction tasks.*

Operation	Problem Size	Maths Group	Percentage Errors	
Addition			Digit	Word
	Large	Low	1.85	0.93
		Middle	0.37	1.11
		High	1.39	0.69
		Total	1.14	0.93
	Small	Low	0.23	0.69
		Middle	0.19	0.56
		High	0.46	0.23
		Total	0.29	0.5
Subtraction				
	Large	Low	3.24	3.24
		Middle	1.67	1.85
		High	1.85	1.85
		Total	2.21	2.28
	Small	Low	0.69	0.93
		Middle	0.56	0.18
		High	0.69	0
		Total	0.64	0.36

The ANOVA found a main effect for operation, $F(1, 36) = 9.17$, $p = .005$, with a large associated effect size (partial eta squared = 0.2), indicating that participants made significantly more errors in subtraction (2.74 %) than in addition (1.42 %). A main effect was also found for problem size, $F(1, 36) = 30.84$, $p < .001$ (partial eta squared = .46). A significant problem size x operation interaction effect, $F(1, 36) = 9.59$, $p = .004$ (partial eta squared = .21), was also found. Dependent t-tests with Bonferroni corrections showed that the size effect was only significant in the subtraction task, for problems written in both digit, $t(38) = -3.75$, $p = .001$, and word format, $t(38) = -5.2$, $p < .001$. Problem size thus influenced error rates more in subtraction than addition, with most errors made on large subtraction problems in comparison with the other problem types. This is in line with the notion that large subtraction is usually solved through calculation (as opposed to retrieval), which is a more error prone strategy. Overall 2.08 % of the data was excluded from any subsequent analyses due to errors made.

Reaction times (RTs). Figure 5.2 presents the mean correct RTs in milliseconds in terms of problem size (small and large), format (digits and words) and Maths group in *(a)* the addition and *(b)* the subtraction task.

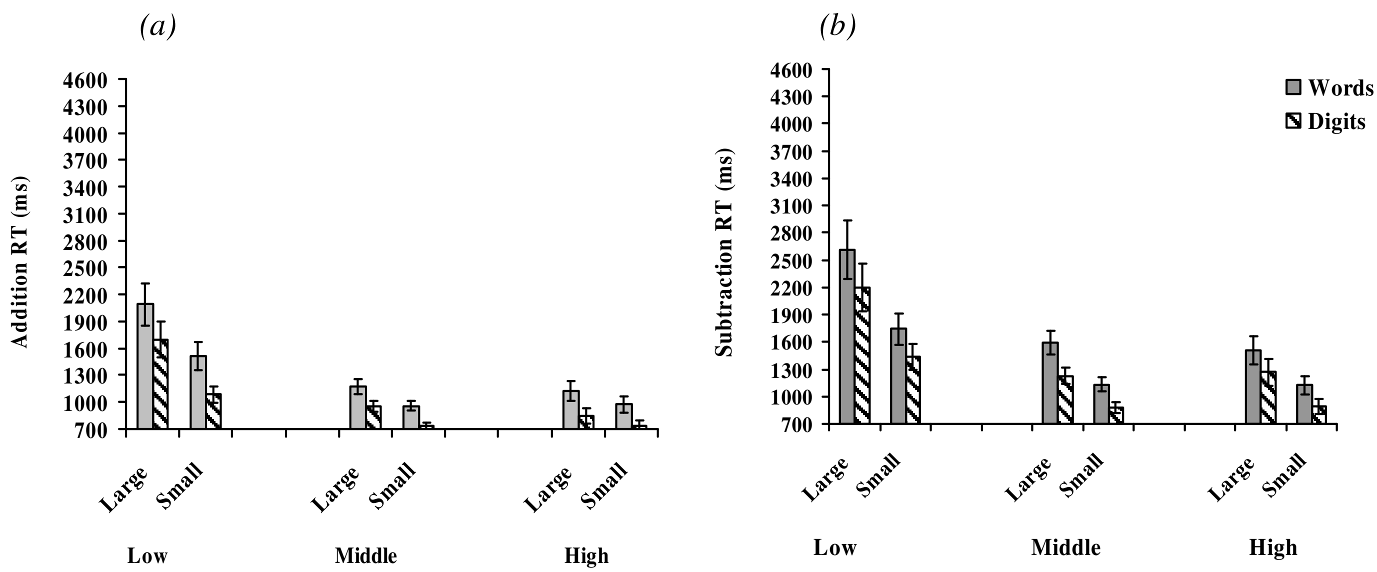

Figure 5.2. *Mean RTs across format, problem size and maths group (Low, Middle or High) in (a) the addition and (b) the subtraction task (± SEM).*

Note. The scale on the Y-axis is set at 700 – 4600 ms to allow comparison with RTs in the multiplication and division tasks (presented in Figure 5.3).

Overall, participants performed the addition task faster than the subtraction task (M = 468.28 ms and M = 1157.81 ms respectively). In both tasks participants answered problems written in digit format (M = 1163 ms) faster than problems written in word format (M = 1463.9 ms) and small problems (M = 1100.73 ms) faster than large problems (M = 1525.36 ms). The Low Maths group's performance (M =

1797.38 ms) was also overall slower than the Middle (M = 1081.81 ms) and High Maths groups' (M = 1059.93 ms) performance. Significant main effects were found for operation, $F(1, 36) = 86.2$, $p < .001$, format, $F(1, 36) = 80.4$, $p < .001$, problem size, $F(1, 36) = 80.63$, $p < .001$, and Maths group, $F(2, 36) = 11.86$, $p < .001$ (all partial eta squared ≥ .4). Post-hoc comparisons showed that the Low Maths group took significantly longer to answer than the Middle ($p < .001$) and High Maths groups ($p < .001$). On average, RTs for the Middle and High Maths groups did not differ significantly ($p = .99$). Problems size effects also differed with Maths group and operation as indicated by significant problem size x Maths group, $F(2, 36) = 8.78$, $p = .001$ (partial eta squared = 0.33) and problem size x operation, $F(2, 36) = 19.19$, $p = .001$ (partial eta squared = 0.35) interaction effects. No significant interaction effects were found for format.

To compare the size effects between the operations and maths groups, difference scores were calculated by subtracting the mean RT on small problems from the mean RT on large problems. Table 5.2 presents the mean difference scores in the addition and subtraction tasks. Overall, the problem size effect was greater in subtraction (M = 533.49 ms difference) than addition (M = 315.76 ms difference). Problems size also seemed to become less influential on performance as the level of Maths group increased (Low M = 705.82 ms difference, Middle M = 313.12 ms difference and High M = 254.93 ms difference).

Table 5.2. *Mean RT disparity (ms) between small and large problems (problem size effect) in the addition and subtraction tasks.*

Operation	Maths Group	Reaction Time difference (ms)	
Addition		Digit	Word
	Low	614.78 (450.82)	576.68 (299.16)
	Middle	222.41 (132.02)	219.04 (171.74)
	High	111.84 (121.78)	149.81 (114.76)
	Total	309.12 (338.09)	307.78 (272.78)
Subtraction	Low	762.35 (567.75)	869.47 (829.47)
	Middle	347.33 (249.45)	463.72 (274.12)
	High	373.33 (250.93)	384.73 (286.4)
	Total	483.03 (412.52)	564.26 (542.38)

A 2 x 2 x 3 mixed between–within groups ANOVA was conducted on the difference scores with operation, format and maths group as factors. A main effect was found for operation, $F(1, 36) = 19.19$, $p < .001$ (partial eta squared = 0.35), indicating that the size effect was overall greater for subtraction than addition. A main effect was also found for Maths group, $F(2, 36) = 8.78$, $p = .001$ (partial eta squared = 0.33). Post-hoc comparisons with Tukey HSD showed that the Low Maths group was significantly more susceptible to the problem size effect compared to the Middle ($p = .004$) and High Maths groups ($p = .001$). The size effect did not differ significantly between the Middle and High Maths groups ($p = .86$). No main or interaction effects were found for format.

To summarise, the RT problem size effect, namely a slowed response on large problems, varied with operation and Maths group. The size effect was greater in subtraction than addition in line with the notion that large subtraction usually involves counting or addition-reference strategies, whereas small subtraction can be solved by direct retrieval. For addition, the size effect was also significant, but smaller, as both small and large addition problems can be solved through retrieval (although larger problems have relatively weaker retrieval strength than small problems). Overall, the Low Maths group was also more susceptible to the problem size effect than the Middle and High Maths groups, who showed a relatively small RTs difference between small and large problems.

Regarding format effects, participants answered word format problems slower than digit format problems, consistent with the argument that retrieval strength is relatively low for word format problems, which causes a shift from retrieval to procedural strategies and, in turn, slower RTs. However, while RTs on word-format problems were slower in all conditions, the slowed response on large problems was relatively similar for digit and word formats. In other words, word format-costs were similar on small and large problems. Word-format costs on RTs were also relatively similar for addition and subtraction and across the three maths groups.

Number of Fixations. The mean number of fixations across each stimulus category was calculated for each participant and is presented in Table 5.3. The fixation count data reflected the overall patterns in the RT data: participants made more fixations in subtraction (M = 3.4) than addition (M = 2.91), more fixations on word format problems (M = 3.7) than digit format problems (M = 2.62), and more fixations on large (M = 3.6) than small problems (M = 2.71). On average, the Low Maths group also made more fixations (M = 4.1) than the Middle (M = 2.71) and High Maths groups (M = 2.66).

Table 5.3. *Means and standard deviations of number of fixations in the addition and subtraction tasks*

Operation	Problem Size	Maths Group	Number of Fixations	
Addition			Digit	Word
	Large	Low	3.99 (2.02)	4.99 (2.47)
		Middle	2.37 (0.54)	3.11 (0.81)
		High	2.25 (1.17)	3.04 (1.36)
		Total	2.83 (1.52)	3.66 (1.83)
	Small	Low	2.24 (0.92)	3.99 (1.76)
		Middle	1.63 (0.47)	2.77 (0.69)
		High	1.56 (0.71)	2.95 (1.25)
		Total	1.79 (0.75)	3.2 (1.35)
Subtraction				
	Large	Low	4.69 (2.64)	5.63 (2.53)
		Middle	2.76 (0.92)	3.81 (1.32)
		High	2.63 (1.38)	3.91 (1.37)
		Total	3.31 (1.94)	4.4 (1.93)
	Small	Low	3.18 (1.37)	4.1 (1.82)
		Middle	2.16 (0.61)	3.07 (0.85)
		High	1.93 (0.88)	2.99 (1.135)
		Total	2.4 (1.09)	3.36 (1.36)

The ANOVA showed that the influences of operation, $F(1, 36) = 16.83$, $p < .001$, format, $F(1, 36) = 129.88$, $p < .001$, problem size, $F(1, 36) = 62.24$, $p < .001$ and maths group, $F(2, 36) = 5.69$, $p = .007$ (all partial eta squared $\geq .3$) on number of fixations were significant. Post-hoc comparisons showed that the Low Maths group made significantly more fixations than the Middle ($p = .014$) and High Maths ($p = .016$) groups, whereas the Middle and High Maths groups did not differ significantly in this regard ($p = .99$). Similar to the RT data, a significant size x maths group interaction effect was also found, $F(2, 36) = 6.04$, $p = .005$ (partial eta squared = 0.25) suggesting that the influence of problem size on fixation count decreased with the level of Maths group. The fixation count analyses also showed a significant three-way operation x size x format interaction effect, $F(1, 36) = 19.78$, $p < .001$

(partial eta squared = 0.35). Differences in effects between the RT and fixation analyses emerged here, as no significant format x size interaction effect was found in the RT analyses.

To compare the fixation count size effects between the operations, Maths groups and formats, difference scores were calculated for each participant as the discrepancy in mean number of fixations between small and large problems (presented in Table 5.4). Similar to the RTs findings, problem size seemed to influence the number of fixations more for the Low Maths group (M = 1.44) than the Middle (M = 0.61) and High Maths (M = 0.6) groups. The size effect also seemed to be greater for digit than word format problems in addition, whereas in subtraction this effect was relatively similar for the two formats (see Table 5.4).

Table 5.4. *Mean disparity in number of fixations between small and large problems (fixation count problem size effect) in the addition and subtraction tasks.*

Operation	Maths Group	Difference in no. fixations	
Addition		Digit	Word
	Low	1.75 (1.2)	0.99 (0.85)
	Middle	0.74 (0.4)	0.34 (0.24)
	High	0.7 (0.55)	0.09 (0.36)
	Total	1.04 (0.89)	0.46 (0.64)
Subtraction	Low	1.5 (1.77)	1.53 (1.68)
	Middle	0.6 (0.5)	0.75 (0.58)
	High	0.7 (0.77)	0.92 (0.49)
	Total	0.91 (1.16)	1.04 (1.06)

A 2 x 2 x 3 ANOVA was conducted on the fixation count problem size difference scores with operation, format and Maths group as independent variables. A main effect was found for Maths group, $F(2, 36) = 6.03$, $p < .006$ (partial eta squared = 0.25). Post-hoc comparisons showed that the overall size effect was significantly greater for the Low Maths group relative to the Middle (p = .01) and High Maths (p = .014) groups. A significant operation effect, $F(1, 36) = 4.23$, $p = .047$ (partial eta squared = 0.1), and format x operation interaction was also found, $F(1, 36) = 19.79$, $p < .001$, (partial eta squared = 0.35). Paired samples t-tests (Bonferroni corrected) showed that in the addition task the size effect was significantly greater for problems in digit relative to problems in word format. This was found for the High, $t(11) = 4.43$, $p = .001$, Middle, $t(14) = 3.44$, $p = .004$, and Low Maths groups, $t(11) = 3.03$, $p = .011$. In the subtraction task, the increase in the number of fixations on large problems was relatively similar for digit ($M = 0.91$) and word ($M = 1.04$) format problems (all $p > 0.4$). Thus, on large subtraction problems, word format did not result in significantly more fixations than digit format.

Overall, the fixation count analysis found a format x size interaction in addition, but not in subtraction, which showed that the increase in fixations that accompanies large problem size was relatively similar for digit and word format problems in subtraction. This pattern was found for all three groups, however, the overall influence of problem size on fixations was greater for the Low Maths group.

Fixation Duration. Table 5.5 presents the mean fixation duration in milliseconds in the addition and subtraction tasks. Participants fixated longer on digit ($M = 275.54$ ms) than word format problems ($M = 226.65$), and slightly longer on large ($M = 256.94$ ms) than small problems ($M = 245.25$). Overall, fixation durations were relatively similar across addition ($M = 252.1$ ms) and subtraction ($M = 250.09$ ms).

Table 5.5. *Means and standard deviations of fixation duration in the addition and subtraction tasks.*

Operation	Problem Size	Maths Group	Fixation Duration (ms)	
Addition			Digit	Word
	Large	Low	256.48 (84.96)	229.85 (57.92)
		Middle	308.45 (125.01)	244.28 (78.16)
		High	284.05 (156.08)	214.03 (56.64)
		Total	284.95 (123.99)	230.54 (65.66)
	Small	Low	248.51 (79.91)	213.81 (47.65)
		Middle	281.34 (99.64)	238.57 (63.64)
		High	296.85 (139.75)	209 (48.7)
		Total	276.01 (107.45)	221.85 (54.94)
Subtraction				
	Large	Low	248.67 (95.55)	239.16 (87.43)
		Middle	302.05 (122.19)	262.68 (88.06)
		High	287.77 (188.86)	205.85 (42.75)
		Total	281.23 (137.8)	237.96 (78.52)
	Small	Low	242.66 (84.07)	216.18 (70.26)
		Middle	285.32 (104.67)	242.1 (75.52)
		High	264.32 (157.38)	204.35 (64.42)
		Total	265.73 (116.5)	222.51 (70.71)

The effects of format, $F(1, 36) = 18.37$, $p < .001$ (partial eta squared = 0.39) and problem size, $F(1, 36) = 8.073$, $p = .007$ (partial eta squared = 0.183), were significant and a four way interaction effect was found for operation x format x problem size x Maths group, $F(2, 36) = 5.13$, $p = .011$ (partial eta squared = 0.22). However, effect sizes were relatively small overall.

To investigate the format x size interactions, difference scores were calculated as the discrepancy in mean fixation duration between small and large problems and are presented in Table 5.6. Overall, the increase in fixation duration on large problems was greater in subtraction than in addition, in accordance with the RT and fixation count data.

Table 5.6. *Mean disparity in fixation duration between small and large problems (problem size effect) in the addition and subtraction tasks.*

Operation	Maths Group	Difference (ms) in mean fixation duration	
Addition		Digit	Word
	Low	7.97 (31.21)	16.04 (28.6)
	Middle	27.11 (43.41)	5.72 (26.1)
	High	-12.79 (33.29)	5.03 (17.56)
	Total	8.94 (39.71)	8.68 (24.53)
Subtraction	Low	6.01 (28.1)	22.98 (31.12)
	Middle	16.72 (65.73)	20.58 (30.35)
	High	23.45 (49.99)	1.49 (38.63)
	Total	15.5 (50.87)	15.44 (33.79)

Paired samples t-tests suggested that fixation duration was not particularly sensitive to effects of problem size or format: no significant differences in fixation duration were found between small and large problems for any of the maths groups (see Table 5.6). With regards to format effects, paired samples t-tests (Bonferroni corrected) showed that the only significant decrease in fixations on word problems relative to digit problems was found in the Middle maths group, $t(14) = 3.82$, $p = .002$. Overall, fixation durations were relatively similar across maths groups, operations, formats and problem sizes.

5.3.2. Part 2: Multiplication and Division

Similar analyses as for Part 1 were conducted to compare performance in the multiplication and division tasks in terms of Maths group, format and problem size. One participant from the Low Maths group was excluded from the Multiplication and Division data due to excessive errors made.

Accuracy. Table 5.7 presents the error rates in the multiplication and division tasks. Errors were minimal overall and were excluded from any subsequent analyses. Participants made more errors in division (4.75%) than in multiplication (3.98%) and made more errors on large (1.25 %) than small problems (0.31 %). Error rates were relatively similar between digit (4.38 %) and word format (4.35 %) problems. Overall, the Low Maths group made more errors (5.43 %) than the Middle (4.77 %) and High (2.89 %) maths groups.

Table 5.7. *Mean percentages of errors in the multiplication and division tasks.*

Operation	Problem Size	Maths Group	Percentage Errors	
Multiplication			Digit	Word
	Large	Low	4.55	5.56
		Middle	4.63	4.63
		High	0.69	1.62
		Total	3.36	3.95
	Small	Low	0.51	0.51
		Middle	0.74	0
		High	0	0.23
		Total	0.44	0.22
Division				
	Large	Low	3.79	3.28
		Middle	2.59	3.89
		High	3.7	3.01
		Total	3.29	3.44
	Small	Low	2.27	1.26
		Middle	1.3	1.3
		High	1.62	0.69
		Total	1.68	1.1

A main effect was found for problem size, $F(1, 35) = 19.01$, $p < .001$ (partial eta squared = .35) as well as a significant problem size x operation interaction, $F(1, 35) = 6.56$, $p = .015$ (partial eta squared = .16), with no further main or interaction effects. In multiplication, problem size influenced errors more (1.18 % difference between small and large problems) than in division (0.7 % difference). This reflects the fact that participants made very few errors on small multiplication problems, compared to small division problems, resulting in a clear difference in errors between small and large multiplication problems. Overall 4.75 % of the data was excluded from any subsequent analyses due to errors made.

Reaction Times. Figure 5.3 presents the mean correct RTs in the multiplication and division tasks in terms of format, problem size and Maths group.

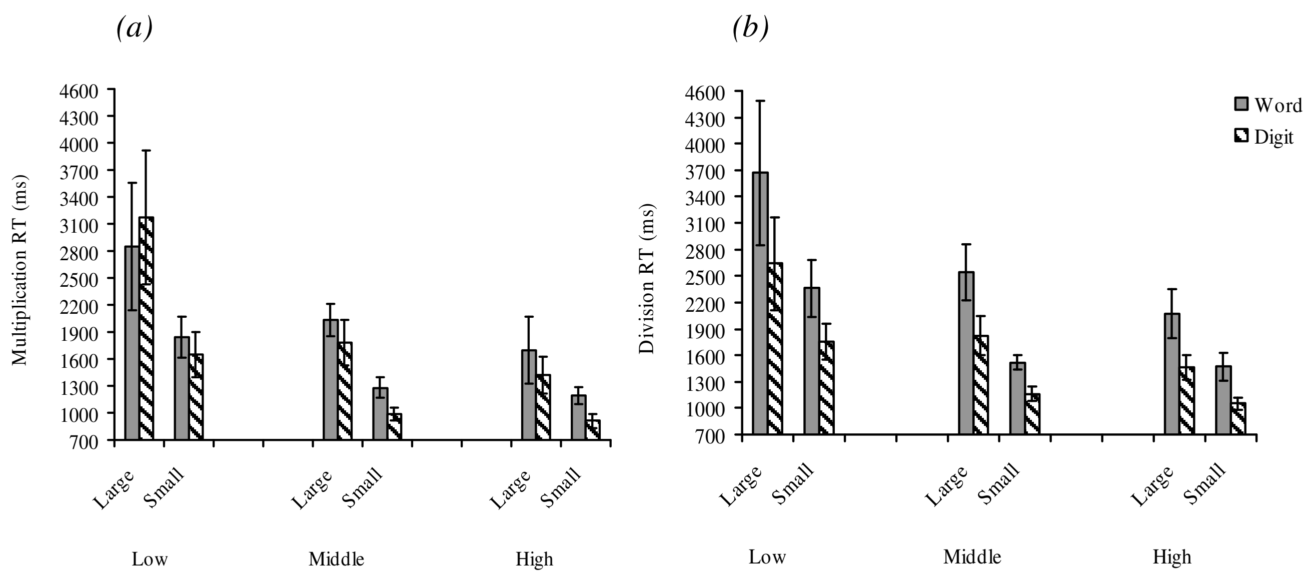

Figure 5.3. *Mean RTs across format, problem size and maths group (Low, Middle or High) in (a) the multiplication and (b) the division task (± SEM).*

Participants performed the multiplication task (M = 1733.76 ms) faster on average than the division task (M = 1958.91 ms) and small problems were answered faster (M = 1430.53 ms) than large problems (M = 2262.14 ms). Regarding stimulus format, problems in digit format (M = 1650.53 ms) were answered faster than problems in word format (M = 2042.14 ms), however, this effect seemed to be much more prominent for division than multiplication (see Figure 5.3 for comparison). On average, RTs increased from the High Maths (M = 1409 ms) to the Middle Maths (M = 1640.24 ms) to the Low Maths (M = 2489.77 ms) group. An ANOVA showed that participants answered multiplication problems significantly faster than division problems, $F(1, 35) = 5.12$, $p = .03$, and RTs on word problems were overall slower than RTs on digit problems, $F(1, 35) = 22.67$, $p < .001$. The effects of problem size, $F(1, 35) = 29.99$, $p < .001$, and Maths group, $F(2, 35) = 4.7$, $p = .016$, were also significant (all partial eta squared = ≥ 0.13). Post-hoc comparisons showed that the Low Maths group's performance was significantly slower than the High Maths group ($p = .017$), whereas RTs for the Middle Maths group did not differ significantly from the High ($p = .78$) or Low ($p = .055$) Maths groups. As expected, a number of significant interaction effects were found for format. These were for operation x format, $F(1, 35) = 27.32$, $p < .001$, operation x format x Maths group, $F(2, 35) = 5.32$, $p = .01$, and operation x format x size, $F(1, 35) = 8.41$, $p = .006$ (all partial eta squared ≥ 0.19). No further interaction effects were found.

Since word format costs on RT appeared much more prominently in division (M = 611.82 ms difference) than in multiplication (M = 174.44 ms difference), difference scores were calculated as the discrepancy in RT between digit and word format problems to compare the format effects between the two operations. Format difference scores are presented in Table 5.8.

Table 5.8. *Mean disparity in RT between digit and word format problems in the multiplication and division tasks.*

Operation	Maths Group	Reaction Time difference (ms)	
Multiplication		Small	Large
	Low	191.05 (397.69)	-326.56 (1424.63)
	Middle	290.7 (263.79)	245.44 (658.62)
	High	279.65 (149.95)	278.38 (809.49)
	Total	258.63 (278.71)	90.26 (990.14)
Division	Low	607.27 (635.13)	1029.35 (1109.14)
	Middle	359.6 (340.61)	722.19 (536.29)
	High	415.08 (419.47)	606.42 (726.63)
	Total	448.82 (465.08)	774.83 (792.36)

A 2 x 2 x 3 ANOVA was conducted on the format difference scores to investigate the influences of operation, problem size and maths group. A main effect was found for operation, $F(1, 35) = 27.32$, $p < .001$, indicating the greater influence of format in division than in multiplication. Interaction effects were found for operation x size, $F(1, 35) = 8.413$, $p = .006$, and operation x Maths group, $F(2, 35) = 5.32$, $p = .01$. Paired samples t-tests (Bonferroni corrected) showed that in multiplication, the only advantage gained for digit stimuli was found in the Middle, $t(14) = -4.27$, $p = .001$, and High Maths groups, $t(11) = -4.27$, $p = .001$, and only on small number problems (see Table 5.8). This reflects the High and Middle Maths groups' relatively fast responses on small digit multiplication problems. Considering division, the High Maths group showed an advantage for digit stimuli on small problems, $t(11) = -3.43$, $p = .006$, but not large problems. The Middle Maths group responded faster on digit problems on small, $t(14) = -4.09$, $p = .001$, as well as large problems, $t(14) = -5.22$, $p < .001$. For the Low Maths group, format effects in division were short of significance ($p \leq .07$).

To summarise, the RT data showed overall greater word format costs for division than multiplication. However, the Low Maths group did not show significant format effects, with relatively slow RTs across digit as well as word format problems. The problem size effect did not differ significantly with Maths group. Thus while the Low Maths group's performance was slower overall, the disadvantage on large problems was relatively similar for all three maths groups.

Number of Fixations. Table 5.9 presents the mean number of fixations in the multiplication and division tasks.

Table 5.9. *Means and standard deviations of number of fixations in the multiplication and division tasks.*

Operation	Problem Size	Maths Group	Number of Fixations	
Multiplication			Digit	Word
	Large	Low	5.96 (5.12)	5.67 (5.03)
		Middle	2.84 (1.26)	3.95 (1.38)
		High	2.87 (1.92)	4.27 (2.98)
		Total	3.75 (3.29)	4.55 (3.28)
	Small	Low	3.44 (2.13)	4.02 (2.29)
		Middle	2.19 (0.54)	3.26 (0.88)
		High	2.02 (1.13)	3.13 (1.67)
		Total	2.5 (1.45)	3.44 (1.64)
Division				
	Large	Low	5.18 (3.12)	8.3 (5.6)
		Middle	3.4 (0.95)	5.97 (1.89)
		High	3.13 (1.33)	5.81 (2.93)
		Total	3.83 (2.07)	6.59 (3.69)
	Small	Low	4.34 (2.04)	5.44 (2.55)
		Middle	2.97 (0.64)	4.24 (1.04)
		High	2.87 (1.15)	4.26 (2.07)
		Total	3.33 (1.45)	4.59 (1.94)

Overall, the fixation count patterns of operation, problem size and format were in accordance with the RT data. Participants made more fixations in the division task (M = 4.66) than the multiplication task (M = 3.64) and more fixations on word (M = 4.86) relative to digit stimuli (M = 3.43). More fixations were also made on large number problems (M = 4.78) than small number problems (M = 3.52). The overall number of fixations seemed to increase as the level of Maths group decreased (*High* M = 3.55, *Middle* M = 3.6 and *Low* M = 5.29). The increase in fixations in the division task relative to the multiplication task was significant, $F(1, 35) = 19.29$, $p < .001$. Main effects were also found for size, $F(1, 35) = 21.29$, $p < .001$, and format, $F(1, 35) = 67.19$, $p < .001$, showing that more fixations were made on large problems and problems in word format (all partial eta squared ≥ .38). However, the fixation count increase for the Low Maths group was short of significance ($p = .07$). Similar to the RT data, a number of interaction effects were found for format. The difference in fixations between digit and word problems was greater in division, $F(1, 35) = 50.35$, $p < .001$, and this effect seemed to be greater for the High Maths group, $F(2, 35) = 4.92$, $p = .013$. In division, the word format cost on fixation count was also greater on large problems as indicated by significant format x size, $F(1, 35) = 10.1$, $p = .003$, and operation x format x size, $F(1, 35) = 40.17$, $p < .001$, interactions (all partial eta squared = ≥ .22).

To compare the format effects across the operations and maths groups, difference scores were calculated as the discrepancy in number of fixations between digit and word format problems (Table 5.10).

Table 5.10. *Mean discrepancy in number of fixations between digit and word format problems in the multiplication and division tasks.*

Operation	Maths Group	Fixation Count difference	
Multiplication		Small	Large
	Low	1.10 (0.86)	-.28 (2.1)
	Middle	1.06 (0.6)	1.12 (0.89)
	High	1.12 (0.86)	1.4 (1.32)
	Total	0.94 (0.71)	0.8 (1.59)
Division	Low	1.10 (0.81)	2.68 (2.14)
	Middle	1.26 (0.72)	2.58 (1.25)
	High	1.39 (1.3)	3.12 (2.88)
	Total	1.26 (0.94)	2.76 (2.06)

A 2 x 2 x 3 ANOVA was conducted on the format difference scores to investigate the format effects in terms of problem size, operation and Maths group. A main effect was found for operation, $F(1, 35) = 50.27$, $p < .001$ (partial eta squared = .59), indicating the overall greater influence of format in division than in multiplication. A main effect was also found for size, $F(1, 35) = 10.13$, $p = .003$ (partial eta squared = .224), as well as an interaction for operation x size, $F(1, 35) = 40.3$, $p < .001$ (partial eta squared = .535), reflecting the greater influence of problem size on fixation in division than in multiplication (see Table 5.10). Interactions were also found for operation x Maths group, $F(2, 35) = 4.93$, $p = .013$, and operation x size x Maths group, $F(2, 35) = 4.45$, $p = .019$, showing that on large division problems, the format of the operands was most influential on fixation count in the High Maths group (both partial eta squared ≥ .2). This reflects the relatively few fixations made by the High Maths group on digit format problems in division, which made the increase in fixations on word format problems more evident.

Bonferroni corrected paired samples t-tests showed that in multiplication, the increase in fixations on word format problems was relatively similar for small (M = 0.94 difference) and large (M = 0.8 difference) problems (all $p > .2$). However, in division, the increase in fixations on word format problems was greater on large (M = 2.76 difference) than small (M = 1.26 difference) problems. This was the case for the High, $t(11) = -3.51$, $p = .005$, and Middle, $t(14) = -5.35$, $p < .001$, Maths groups, but not the Low Maths group (more fixations overall). For the High and Middle maths groups, word format costs on fixation count were thus more evident on large than small division problems.

Fixation duration. Table 5.11 presents the means and standard deviations of the fixation durations in the multiplication and division tasks. Average fixations were longer on digit (M = 280.32) format than on word format (M = 235.9) problems. Overall, participants' fixations were slightly longer in the multiplication task (M = 264.88) than in the division task (M = 250.07). Fixations were also longer on large (M = 262.91) than on small problems (M = 252.05). The ANOVA showed significant influences of operation, $F(1, 35) = 4.12$, $p = .05$ (partial eta squared = .105), format, $F(1, 35) = 13.04$, $p = .001$ (partial eta squared = .27) and problem size, $F(1, 35) = 5.75$, $p = .022$ (partial eta squared = .14) on fixation duration. The fixation duration data showed no interaction of operation x format x size. Instead, for both small and large problems, fixations were longer in multiplication than division and longer on digit than on word problems (see Table 5.11).

Table 5.11 *Means and standard deviations of fixation durations in the multiplication and division tasks.*

Operation	Problem Size	Maths Group	Fixation Duration (ms)	
Multiplication			Digit	Word
	Large	Low	270.51 (108.31)	243.89 (77.43)
		Middle	344.71 (180)	288.81 (106.58)
		High	285.63 (174.03)	212.34 (62.57)
		Total	304.57 (159.84)	251.66 (90.34)
	Small	Low	257.77 (94.88)	230.78 (65.4)
		Middle	290.42 (94.48)	254.44 (77.13)
		High	288.49 (175.91)	210.79 (55.76)
		Total	280.36 (123.39)	233.81 (68.4)
Division				
	Large	Low	251.73 (94.29)	225.79 (67.31)
		Middle	298.61 (141.69)	257.97 (78.07)
		High	267.92 (140.98)	206.96 (37.11)
		Total	275.35 (127.72)	232.55 (66.54)
	Small	Low	245.02 (76.13)	227.41 (71.46)
		Middle	296.57 (121.09)	248.79 (69.32)
		High	257.96 (138.6)	216.15 (42.03)
		Total	269.46 (115.51)	232.3 (62.65)

Key findings in Part 1 and Part 2: To summarise, taking the main findings from Part 1 and 2, behavioural data and fixation patterns showed effects of operation, format and problem size. Taking addition, a clear advantage was found for digit format, whereas in subtraction, the increase in fixations on large problems was similar for word and digit format problems. The Low Maths group was also more susceptible to the problem size effect in Part 1. In Part 2, format effects were much more prominent in division than in multiplication. Word format costs were also more evident on large than small division problems. Problem size seemed to influence the three groups similarly in Part 2, however, the Middle and High Maths groups showed an advantage for digit problems in multiplication.

5.4. Discussion

The current study investigated the influence of surface format on arithmetic across different operations, problem sizes and individual differences related to mathematics experience. Evidence from RT and eye-tracking showed that input format (digits vs number words) affected the problem solving processes in arithmetic and supports the argument that arithmetic performance is not abstracted away from the format of the operands (e.g. Campbell, 1994; Campbell & Alberts, 2009; Campbell & Epp, 2005, Cohen-Kadosh et al., 2008). The RT and fixation patterns in the current study have modelled similar effects as was found for self-reports of strategies used in arithmetic (Campbell & Alberts, 2009). In support of these previous findings, word format (e.g. two + three) seemed to hinder retrieval of arithmetic facts compared to digit format (e.g. 2 + 3).

Furthermore, eye-tracking patterns did not just support RT patterns, but also showed subtle effects that were not evident from RT findings alone. In Part 1, for example, the fixation patterns showed a format x size interaction for addition, but not subtraction, which supported the argument that retrieval strength is lowest on large subtraction problems, regardless of format. Since this finding is in line with Campbell and Alberts's (2009) reports of strategy use in arithmetic, it suggests: a) that fixation patterns can give an index of strategies used in arithmetic and b) that self-reports of strategies are reasonably valid. Accordingly, on large subtraction problems, both formats promote procedural strategies (as opposed to retrieval) and word format would thus not necessarily hinder retrieval more than digit format (Campbell & Alberts, 2009). The fixation data supported this argument by showing that the fixation increase on large problems was similar for digit and word formats in subtraction.

Overall, individuals in the Low Maths group seemed to be more influenced by the magnitude of the operands in a problem as indicated by behavioural as well as fixation data. This was found to be the case in addition and subtraction, since large problems are thought to promote counting-based strategies, as opposed to direct retrieval (e.g. Campbell & Fugelsang, 2001). It is thus likely that an individual's experience with mathematics influences the retrieval strength of arithmetic facts, which makes those in the High Maths group less prone to resort to counting strategies, and gives them an advantage on large problems, regardless of format. High Maths individuals could thus have an advantage for transcoding numerical information from different formats. Overall, these findings make the case for considering individual differences in mathematics when investigating interactions of format and problem size in arithmetic.

In Part 2 of the study, format effects were much more evident in division than in multiplication. In terms of reaction time, the cost of word format was greater for division than multiplication (except for the Low Maths group whose response times were relatively slow overall). The fixation data also suggested that in division the cost of word format on retrieval was greater on large than small problems, a finding that was not evident from the behavioural data alone. This is in line with the strategy report findings of Campbell and Alberts (2009): since direct retrieval is the preferred strategy for multiplication regardless of format, the increase in RT on large multiplication problems (with relatively weak retrieval strength) is thus quite similar for digit and word format problems. However in division, problem solving can sustain a shift in strategy towards multiplication-reference, which is frequently the preferred strategy on large word format problems where retrieval is less likely. This was supported by the fixation count analysis that showed similar format x size interactions.

Unlike addition and subtraction, the problem size effect did not differ significantly with Maths group in multiplication and division. Thus while the Low Maths group's performance was slower overall, the disadvantage on large problems was relatively similar for all three Maths groups. This is likely to be due to the fact that in multiplication and division, with counting being unlikely, strategies are more retrieval-based regardless of problem size. If similar strategies are used on both small and large problems, it is thus not surprising that the size effect affected the three groups quite similarly. In addition and subtraction, large problems, due to their relatively weak retrieval strength in comparison with small problems, could promote counting, an effect that the Low Maths group were more susceptible to. However, in multiplication and division this is not the case since strategies are overall more retrieval based.

Overall, fixation count seemed to be a particularly sensitive measure of interactions between number format, operation, problem size and Maths group. The fixation count findings were largely in support of the strategy reports of Campbell and Alberts (2009) and highlighted effects beyond the accuracy or RT data. On the other hand, fixation duration did not seem to be particularly sensitive to interactions between format, operation and problem size. Since arithmetic operations such as subtraction and division might require more eye-movements back and forth than addition and multiplication, future research on strategy use in arithmetic could benefit from employing measures of gaze direction rather than fixation duration. Although fixation duration has proven to be a useful measure of participants' strategies in solving a parity judgement task (Merkley & Ansari, 2010), arithmetic is likely to involve different problem solving strategies, which do not seem to be reflected in fixation duration patterns to a great extent.

To conclude, the current study added to the current literature on arithmetic performance by showing that eye-tracking is sensitive to effects of operation, format, problem size and individual differences related to mathematics across each of the four operations. The current findings support previous findings that encoding format affects calculation processes per se (e.g. Campbell, 2004) and that operands are not necessarily abstracted away from input format (as suggested by McCloskey & Macaruso, 1995, for example). The findings are in line with the format-specific view of number processing postulated by models such as Dehaene's Triple Code Model (Dehaene, 1992; Dehaene & Cohen, 1995) and Campbell and Clark's Encoding Complex View (Campbell & Clark, 1988; 1992; Campbell, 1994; see Chapter 1, p. 16 – 17) which assume interactions of format with problem size and operation. Since clear differences in performance was found between arabic digit and number word formats on both behavioural and eye-tracking measures, the findings also suggest that eye-tracking provides an additional level of analysis for studying calculation processes in arithmetic. In previous studies of arithmetic, the main indices of strategy use have been self-reports, accuracy and RT (e.g. Campbell & Alberts, 2009;

Campbell & Penner-Wilger, 2006). While the current fixation count patterns are largely in line with the self-reports of Campbell and Alberts (2009), eye-tracking may provide a more reliable measure of strategies than self-reports, which are prone to bias. Also, as Zhang (2010) pointed out, stimulus features that might follow different processing routes (e.g. digits and words) might still yield similar RT and accuracy patterns, an argument that has been overlooked in some studies that supported the format-independent view of number processing (e.g. Ganor-Stern & Tzelgov, 2008). The current eye-tracking findings that diverged from behavioural findings support this argument by showing interactions of format and problem size that were not evident from behavioural data alone, and suggests that format-specific processing takes place in arithmetic. Overall, the format-specific view of number processing suggests the close interaction of the encoding conditions (such as format) with the answer-retrieval stage in arithmetic. In Chapter 6 (Experiment 5) this was explored further by investigating the event-related potentials that occur during encoding and retrieval separately.

Chapter 6

Experiment 5: The Interaction between Encoding and Answer-retrieval Stages of Arithmetic: Format and Operation Effects at Different Levels of Mathematics Experience

6.1. Introduction

Campbell and Epp (2005) suggested that solving of arithmetic occurs across three stages. These stages involve the encoding of the operands, retrieving or calculating the answer, and reporting the answer. The results from cognitive interference tasks (Chapters 2 and 3) and eye-tracking (Chapter 5) suggest that encoding features, such as format, have an influence on basic number encoding, as well as answer-retrieval processes. However, considerable debate still exists in the literature on the relationship between the encoding and retrieval conditions in arithmetic. The two main viewpoints are the additive and interactive views of arithmetic. The former view argues that encoding and retrieval operate independently of one another and is related to the format-independent view of number representation, namely that numerical information from various surface formats is translated into a uniform abstract code (e.g. Dehaene, 1997; Dehaene & Cohen, 1995; McCloskey, 1992; McCloskey & Macaruso, 1995). After the operands in an equation have been translated to an internal abstract code, answers can be retrieved or calculated. The results are then sent to arabic, written or verbal number output codes, depending on task requirements (Dehaene's Triple Code model, Dehaene, 1992; McCloskey's Abstract Code model, McCloskey, 1992, McCloskey & Macaruso, 1995). Since the retrieval/calculation phase operates from the abstract numerical code, input conditions such as operand format (e.g. arabic digits or number words), should have no influence on subsequent retrieval or calculation. In support of this view, neuropsychological evidence has shown that brain-injured patients have operation-specific deficits regardless of input format (e.g. McCloskey & Macaruso, 1995). In contrast with the argument of Campbell and Alberts (2009; see Chapter 5), the additive viewpoint argues that different formats should not hinder or promote the use of different retrieval processes in arithmetic, but these processes should rather differ between operations.

The opposite viewpoint, namely that the encoding and retrieval/calculation phases interact with one another, proposes that the encoding conditions have a direct influence on the subsequent calculation processes and is in line with the format-specific view of number representation. Campbell's Encoding Complex model is the main supporter of this view (e.g. Campbell & Clark, 1989; 1992; Campbell & Alberts, 2009; Campbell et al., 2004). As discussed in Chapter 5, this view does not assume an analogue number code, but rather modality-specific mental number representations, which promote the use of different strategies (Campbell & Alberts, 2009). Regarding format, arabic digits and number words are thus each represented in a separate code, rather than a uniform abstract representation. Support came from Campbell and Colleagues' studies that showed interactions of arithmetic format with operation and problem size. Format effects differed, for example, between small and large number problems, such that word format costs on performance were more prevalent on large problems (e.g. Campbell et al., 1999; Campbell & Alberts, 2009) and more prevalent in division than in multiplication (See Chapter 5; Campbell & Alberts, 2009). Campbell and Epp (2005) argue that these interaction effects show that differences in performance between digit and word format problems do not just arise due to differences in encoding of the formats, but rather as a result of format-specific influences which occur during the calculation/retrieval stage per se.

The current behavioural and eye-tracking results (Experiment 5) are mainly in support of the interactive viewpoint of arithmetic, showing similar interactions of operation, problem size and format to Campbell and colleagues' findings. Specifically, format seemed to affect the strategies used in arithmetic and this varied with problem size. The findings also show that format-specific influences can be further regulated by individual differences related to mathematics experience. The eye-tracking data suggest, for example, that while participants found large number equations in word format the most difficult to process, this effect was less pronounced for High Maths individuals (e.g. in subtraction).

In the study of arithmetic, event-related potential technology is particularly useful as it can highlight effects related to different operations, which might not be evident from behavioural data alone. Studies have shown, for example, that during addition and subtraction, visuospatial processing takes place suggesting that these facts are stored in visuospatial memory areas (Zhou et al., 2006, 2007). On the other hand, knowledge of multiplication is usually represented in verbal memory areas (e.g. left anterior activation; Rickard, Romero & Basso et al., 2000; Zhou et al., 2006, 2007) in accordance with the argument that multiplication is usually solved through direct memory retrieval (e.g. Campbell & Alberts, 2009).

CHAPTER 6

The debate on the interaction between the stages of arithmetic still continues in the literature, with recent evidence supporting the additive rather than the interactive viewpoint (e.g. Zhou, 2011; Zhou, Chen & Qiao et al., 2009). Aiming to resolve the conflict between the two viewpoints, Zhou (2011) presented arithmetic equations in two parts to create 'pure' encoding and retrieval/calculation phases of presentation. Equations (e.g. 3 + 3 = 6) were presented on-screen in two parts during a true/false verification study using event-related potentials (ERPs). On each trial, the first operand appeared initially (e.g. 3) and remained on-screen for 400 ms. The rest of the equation followed on a subsequent screen (e.g. + 3 = 6) and remained for 600 ms. This presentation method allowed the operation effects that take place at each stage to be investigated separately. To isolate the effects of operation, equations were also presented in separate blocks of addition and multiplication in order for participants to anticipate the operations to be solved. In support of the additive viewpoint, Zhou (2011) found that participants encoded arabic digit operands differently for addition and multiplication, namely a verbal code was activated for multiplication, but an analogue code for addition, from which the appropriate multiplication and addition facts could be retrieved. In other words, an operation effect, (reflected as larger left anterior ERP responses for multiplication) was found during both the encoding and retrieval stages of presentation. If the interactive viewpoint were supported, no left anterior operation effect should have emerged during encoding, since addition and multiplication operands presented in the same format (arabic digits in this case) should be encoded in a similar way regardless of operation, with operation-specific effects only emerging during the retrieval phase (Zhou, 2011).

The current study aimed to further investigate the relationship between the encoding and the retrieval/calculation stage in arithmetic, while also controlling for surface format and mathematics experience. Whereas the previous chapter suggested that digit and word operands influence the strategies used in arithmetic per se, separating the encoding and retrieval phases can provide a more in-depth analysis of the format and operation effects at each stage of arithmetic problem solving. In Experiment 5, the presentation method of Zhou (2011) was closely followed, but the experiment included word as well as digit format problems in order to directly investigate the operation effects between the two formats at the encoding and retrieval/calculation phase. While the results of Zhou (2011) supported the additive instead of the interactive viewpoint, arithmetic problems were only presented in arabic digit format in that study and individual differences in mathematics were not taken into account. The interactive viewpoint of arithmetic was thus rejected based on operation-specific, rather than format-specific effects. Format effects are important to consider, however, since they can be informative as to whether or not calculation procedures differ with the format of the operands (Campbell & Alberts, 2009), and in turn how encoding conditions such as format influence subsequent retrieval. By including digit and word versions of arithmetic problems, a direct comparison can be made between the operation-specific effects that emerge for each format and at each stage of arithmetic problem solving. Similar results to Zhou's (2011) for both digit and word format equations, namely operation effects at both the encoding and retrieval phases, would lend more support to the additive viewpoint of arithmetic, suggesting that input format does not influence subsequent calculation or retrieval. However, if operation effects only emerge during the retrieval/calculation phase for equations presented in the same format, the format-specific view can be supported suggesting that surface format directly influences retrieval and calculation. As the evidence in Chapter 5 suggested an advantage for arithmetic fact retrieval for the High/Middle Maths groups, the location of the operation effects as presented by Zhou and colleagues (Zhou et al., 2006, 2007; Zhou, 2011) could also differ between individuals. With regards to whether or not direct retrieval or visual spatial/magnitude processing take place, such effects can be informative of how the groups differ in problem solving strategies and how the mental representation of arithmetic facts differs across formats.

6.2. Method

6.2.1. Participants

Eighteen right-handed participants took part in the study with ages ranging from 18 to 30 (M = 22.06, SD = 3.02). All participants reported having normal or corrected-to-normal vision. In Chapter 5, the numeracy test confirmed differences between the High and Low Maths groups, whereas the Middle group did not differ significantly from the High group (see Chapter 5, p. 124). Therefore, in the current experiment, participants were recruited and allocated so as to only have a ‘High’ and ‘Low’ Maths group. Participants who had studied higher level Leaving Certificate mathematics with an obtained grade in the A (n = 2), B (n = 5) or C (n = 3) region were assigned to the High Maths group (N = 10; 7 men and 3 women). Participants who had studied ordinary level Leaving Certificate mathematics with an obtained grade of a B or lower were assigned to the Low Maths group (*N* = 8; 5 men and 3 women). Four participants in the Low Maths group indicated an obtained grade of a B and four indicated an obtained grade of a D.

6.2.2. Materials and Apparatus

Stimuli. Similar to the method employed by Zhou (2011), 28 single digit addition problems and 28 single digit multiplication problems were used as stimuli. The same number of word format problems was also included as stimuli. Problems ranged from 2 + 3 to 8 + 9 in addition and from 2 x 3 to 8 x 9 in multiplication. The first operand in each problem was the smaller of the two. Tie problems (e.g. 3 x 3) or problems containing ‘1’ or ‘0’ were excluded from the stimulus set due to their unique or rule-based encoding characteristics (e.g. Blankenberger, 2001; Campbell & Gunter, 2002, LeFevre et al., 1996). Four of the 28 problems were randomly selected to form false arithmetic problems. False problems were formed by adding or subtracting one number from one of the operands as in the study of Zhou (2011). The resulting answer was added to the original problem to form a false answer. False answers still had the same number of digits as the true answers (one or two) to ensure that true and false answers were closely matched. Equations were presented in either arabic digit (e.g. 2 + 3) or number word (e.g. two + three) format. Each test block contained 64 stimuli in total. This included 28 true digit format problems, 28 true word format problems, 4 false digit format problems and 4 false word format problems. Stimuli were presented centrally on a computer screen in white print against a black background and subtended between 1 and 1.9 degrees of visual angle.

Digit and letter sizes corresponded to one another (each character space was approximately 3mm wide and 5 mm high). Each operation contained two test blocks resulting in each stimulus presented twice. Trials were presented in a pseudorandom order and each participant received the same order of trials. Stimuli were programmed and presented with E-Prime©, which recorded all participant input and calculated average response times and accuracy. The electroencephalography (EEG) materials and procedure are described in section 6.2.3.

Numeracy Test. The same numeracy test that was modified from Lipkus et al. (2001) was used and the same procedure was followed as for Chapter 5 (see p. 120).

6.2.3. Electroencephalography (EEG) materials and procedure

EEG was recorded using silver/silver–chloride (Ag/AgCl) electrodes mounted on a 32-channel elastic electrode cap. The extended version of the International 10-20 system for electrode placement (American Encephalographic Society, 1994) was used to collect EEG data from 32 scalp sites. One electrode was placed on the nasion as a reference. Electrooculography (EOG) was used to record horizontal (HEOG) and vertical (VEOG) eye movements. HEOG electrodes were placed on the outer canthus of each eye and VEOG electrodes were placed above and below the left eye. The impedance level was kept to below 10kΩ. A BrainVision© amplifier with a band-pass of 0.16-100Hz and a gain of 1000 was used to amplify EEG activity.

Stimulus presentations and participant input were logged in real time on the EEG recordings. This was achieved as the E-prime software logged participant responses and sent TTL voltage triggers to the EEG acquisition PC representing the different stimulus categories. For analyses, EEG recordings were notch-filtered off-line at 50 Hz. Eye-blinks were filtered off-line and a blink reduction algorithm was applied to the data, which involved automatic artefact correction (Berg & Scherg, 1991; Ille, Berg & Scherg, 2002). The EEG data were digitised at a sampling rate of 500 Hz.

6.2.4. Procedure

Participants signed an informed consent form which stated that the experiment aims to investigate processing differences between digits and words (Appendix 1). Before commencing the arithmetic task participants completed the numeracy test. Scoring and instructions were similar to the other experiments. Participants were given six minutes to complete as many of the answers as possible.

After attaching the electrodes and connecting the EEG equipment, participants were seated in a darkened cubicle (150cm X 180cm) approximately half a meter from the LCD computer monitor. The cubicle was copper-plated and electrically shielded. The experimenter explained that the computerised task that was to follow would involve answering 'true' or 'false' to arithmetic problems presented on-screen. The experimenter remained in the room while the participant read the on-screen instructions and completed practice trials. Half of the participants were instructed to use the 'd' key (left of keyboard) to respond 'true' and the 'k' key to respond 'false'. The rest of the participants were instructed to use the 'k' key (right of keyboard) to respond 'true' and the 'd' key to respond 'false'. After the participant had read the instructions and completed the practice trials, a screen appeared which indicated that the following test block would either be addition or multiplication. Test blocks were counterbalanced across participants. Each operation contained two test blocks of 64 stimuli each. After each block, participants could take a one minute break if desired. The following instructions appeared on-screen before the participant commenced the tasks:

A NUMBER OF EQUATIONS WILL APPEAR ON THE SCREEN SOMETIMES THE EQUATION WILL BE TRUE: E.G. 2 + 2 = 4 SOMETIMES THE EQUATION WILL BE FALSE: E.G. 2 + 3 = 4 IF YOU THINK THE EQUATION IS TRUE, PRESS THE 'D' KEY ON THE KEYBOARD IF YOU THINK THE EQUATION IS FALSE, PRESS THE 'K' KEY ON THE KEYBOARD	EACH EQUATION WILL APPEAR IN TWO PARTS E.G. '2' AND '+ 2 = 4' THE FIRST NUMBER WILL APPEAR FIRST AND THEN DISAPPEAR. THE REST OF THE EQUATION WILL APPEAR THEN. AFTER EACH EQUATION A QUESTION MARK WILL APPEAR ON THE SCREEN - THIS IS WHEN YOU SHOULD RESPOND EACH EQUATION WILL ONLY BE ON-SCREEN FOR A VERY SHORT TIME PRESS THE SPACEBAR TO SEE SOME PRACTICE TRIALS

Stimuli were presented in white font against a black background. The first operand was presented centrally on-screen for 400 ms. The second part of the problem, which included the operation sign, second operand, equal sign and answer, was then presented on-screen for 800 ms (each problem was on-screen for 1200 ms in total). To accommodate word format problems, which might take longer to read than digit format problems, the second part of the equation was presented for 800ms and not 600ms as was the case in Zhou’s study (2011). After this, a question mark was presented centrally, which remained on-screen until the participant responded by pressing the ‘d’ or ‘k’ key. Trials were presented in a pseudorandom order and constrained such that consecutive problems did not contain the same operand or answer as was the case in the study of Zhou (2011). Participants were instructed that most of the answers would be true, but that some of them would be false and that it was their task to try to respond as quickly, but as accurately as possible. Figure 6.1 presents examples of the presentation of true digit and word format problems.

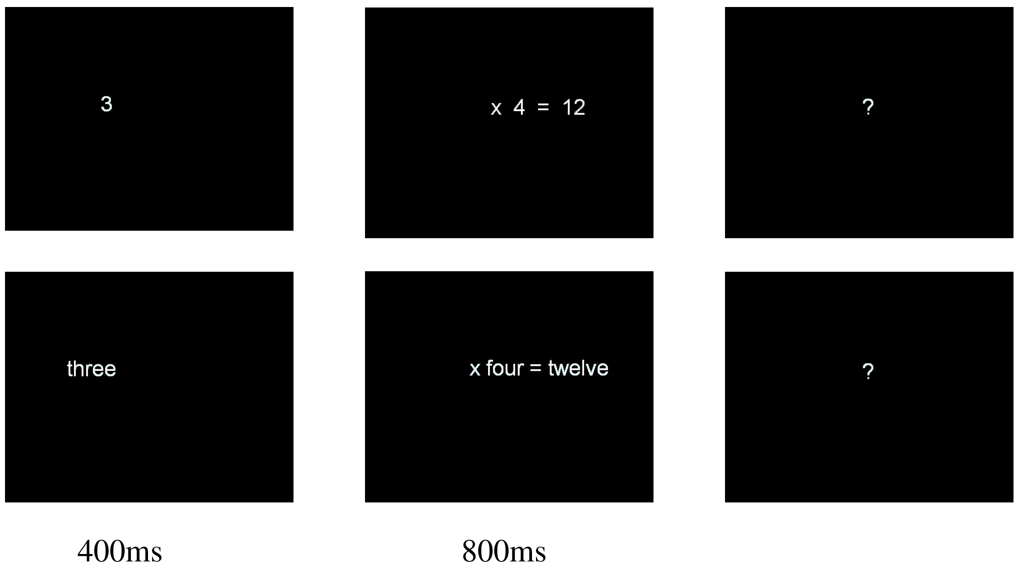

Figure 6.1. *Examples of true digit and word format equations.*

6.2.5. Ethical Considerations

In addition to the ethical considerations set out in Chapter 2, participants were asked not to participate if they had sustained prior head injury, suffered from epilepsy, any neurological disorders or claustrophobia. Participants with reading or visual difficulties were also excluded from participating. Prior to the experiment each participant received an information sheet on what to expect from the EEG experiment (see Appendix 5) and were told that the experiment would also involve a short numeracy scale.

6.3. Results

6.3.1. Behavioural data

Reaction times were measured as the time taken in milliseconds for the participant to indicate whether each on-screen equation was true or false after the question mark appeared on-screen. E-prime© calculated the average response latency and number of errors across each stimulus category for each participant. Analyses focused on responses to true arithmetic equations following Zhou's (2011) study. An independent t-test showed that on average, High Maths participants (*M* = 12.9, *SD* = 3.45) outperformed Low Maths participants (*M* = 7.88, *SD* = 2.95) on the numeracy test, $t(16) = 3.27$, $p = .005$. There were no gender differences in numeracy performance.

6.3.1.1. Accuracy

One participant from the Low Maths group was excluded from the accuracy and RTs analyses due to too many errors made. The overall error rate for the High maths group on true addition and true multiplication trials were 2.68 % and 5.89 % respectively. The overall error rate for the Low Maths group on true addition and true multiplication trials were 4.08 % and 5.36 % respectively. A 2 x 2 x 2 ANOVA was conducted on error rates on true trials with operation, format and maths group as factors. A main effect was found for format, $F(1, 15) = 21.74$, $p < .001$, indicating that participants made significantly more errors on word format (6.05 %) than on digit format (2.38 %) problems. A main effect was also found for operation $F(1, 15) = 6.85$, $p = .019$, indicating that more errors were made in multiplication (5.67%)

than addition (3.26 %). Error rates did not differ as a function of maths group. Overall, 4.22 % of the data was excluded due to errors made.

6.3.1.2. Reaction Time (RT)

The RTs analyses focused only on correct responses on true arithmetic equations. Figure 6.2 presents the mean RTs in milliseconds across digit and word format equations in the addition and multiplication tasks for the High and Low Maths groups. In the Addition task the High Maths group showed overall faster performance than the Low Maths group for both digit (*High M* = 467.04, *SD* = 243.34 and *Low M* = 814.95, *SD* = 333.56) and word format equations (*High M* = 605.62, *SD* = 279.92 and *Low M* = 1048.29, *SD* = 474.35). The same pattern was found for multiplication with the High Maths group showing faster performance across digit (*High M* = 507.55, *SD* = 297.064 and *Low M* = 896.8, *SD* = 442.97) and word format equations (*High M* = 669.18, *SD* = 356.41 and *Low M* = 1140.74, *SD* = 634.2). Overall faster performance was found for digit than word format equations, whereas no clear RT differences seemed to occur between addition and multiplication.

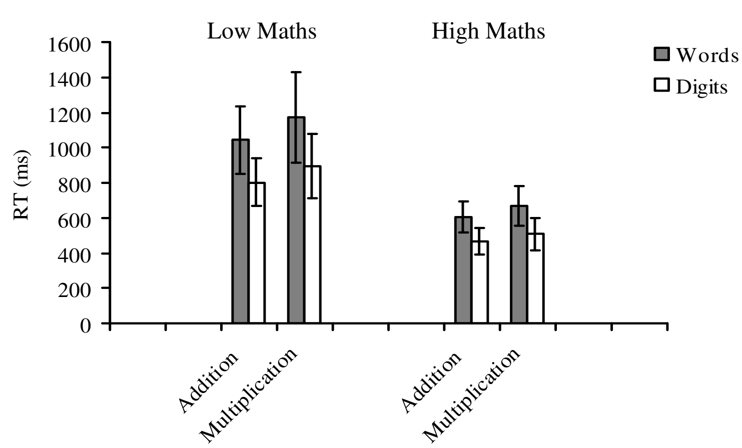

Figure 6.2. *Mean RTs (± SEM) across format and operation for the Low and High Maths groups.*

A 2 x 2 x 2 mixed between–within groups ANOVA was conducted on the mean RTs on true arithmetic trials. Main effects were found for format, $F(1, 15) = 31.16$, $p < .001$, and Maths group, $F(1, 15) = 4.99$, $p = .041$, with no further main or interaction effects.

Overall, the reaction time and accuracy data showed similar patterns for addition and multiplication, with performance differing more with format and Maths group than with operation. While overall performance improved with the High Maths group, the pattern of performance was similar for both groups and for both operations, namely faster performance for digit format problems.

6.3.2. Event-related Potentials (ERPs)

EEG data were averaged using Brain Electrical Source Analysis (BESA©) software. Similar to Zhou (2011), the ERP analysis focused on the encoding (first operand presentation) and retrieval (second part of equation) stages of arithmetic. ERPs were time-locked to the onset of the second operand. Zhou et al. (2006, 2009) found operation effects for digit stimuli emerging in left anterior and right posterior electrodes. The electrodes F3 over the left anterior scalp and P4 over the right posterior scalp were thus selected for analyses. The analyses also focused on corresponding electrodes F4 and P3 over the right and left hemisphere, respectively. Figure 6.3 presents the 32-channel montage and the scalp locations of the four selected electrodes.

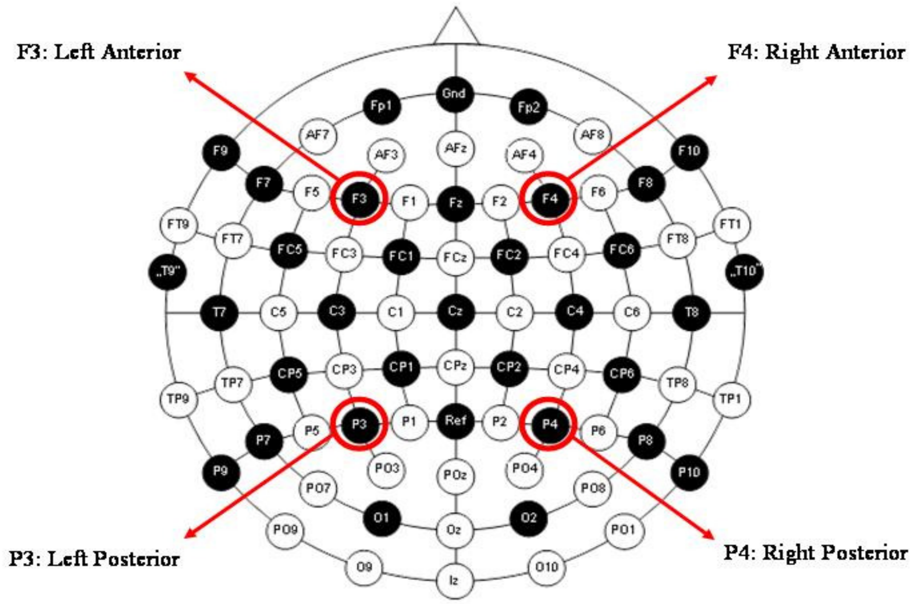

Figure 6.3. *The 32-channel ERP montage showing the scalp locations of electrodes F3, F4, P3 and P4.*

6.3.2.1. Amplitude

Encoding. Figures 6.4 and 6.5 present the grand mean waveforms over the left and right anterior and posterior scalp for the presentation of the first operand of digit and word format equations respectively for the Low Maths group. Figures 6.6 and 6.7 present these waveforms for the High Maths group. The scale is set at ± 3 μV to allow comparison with components elicited in response to the retrieval/calculation phase (presented in Figures 6.8 to 6.11). The two shaded regions represent the time-windows 0 to 60 ms and 70 to 140 ms, which were selected for analysis based on visible variations in amplitude here from inspecting the grand mean waveforms. However, from visually inspecting the grand mean waveforms, no clear components seemed to emerge for the presentation of the first operand of each equation which remained on-screen for 400 ms.

To investigate operation effects, analyses of variance were conducted on the mean amplitudes for the two selected time windows. In each ANOVA the within-groups variables were operation (addition or multiplication), electrode position (F3, F4, P3 or P4) and format (digits or words), and the between-groups variable was Maths group (High or Low). Post-hoc dependent t-tests (Bonferroni corrected) focused on operation effects over the left anterior and right posterior scalp, following Zhou (2011).

In the interval between 0 and 60ms (first shaded region presented in the figures) a significant main effect was found for electrode position, $F(3, 48) = 7.35$, $p < .01$ (partial eta squared = 0.315). No further main or interaction effects were found. Dependent t-tests showed that there was no operation effect for word or digit problems.

The same analysis was conducted for the time window between 70 and 140ms. A main effect was found for electrode position, $F(3, 48) = 7.35$, $p < .001$ (partial eta squared = 0.34), and an interaction effect was found for operation x format x maths group, $F(1, 16) = 4.62$, $p = .047$, however, the effect size was small (partial eta squared = 0.224). Dependent t-tests (Bonferroni corrected) showed that there was no operation effect for digit or word format problems for High or Low Maths groups. Overall, no clear components were found during the encoding phase where participants only saw a single operand on-screen.

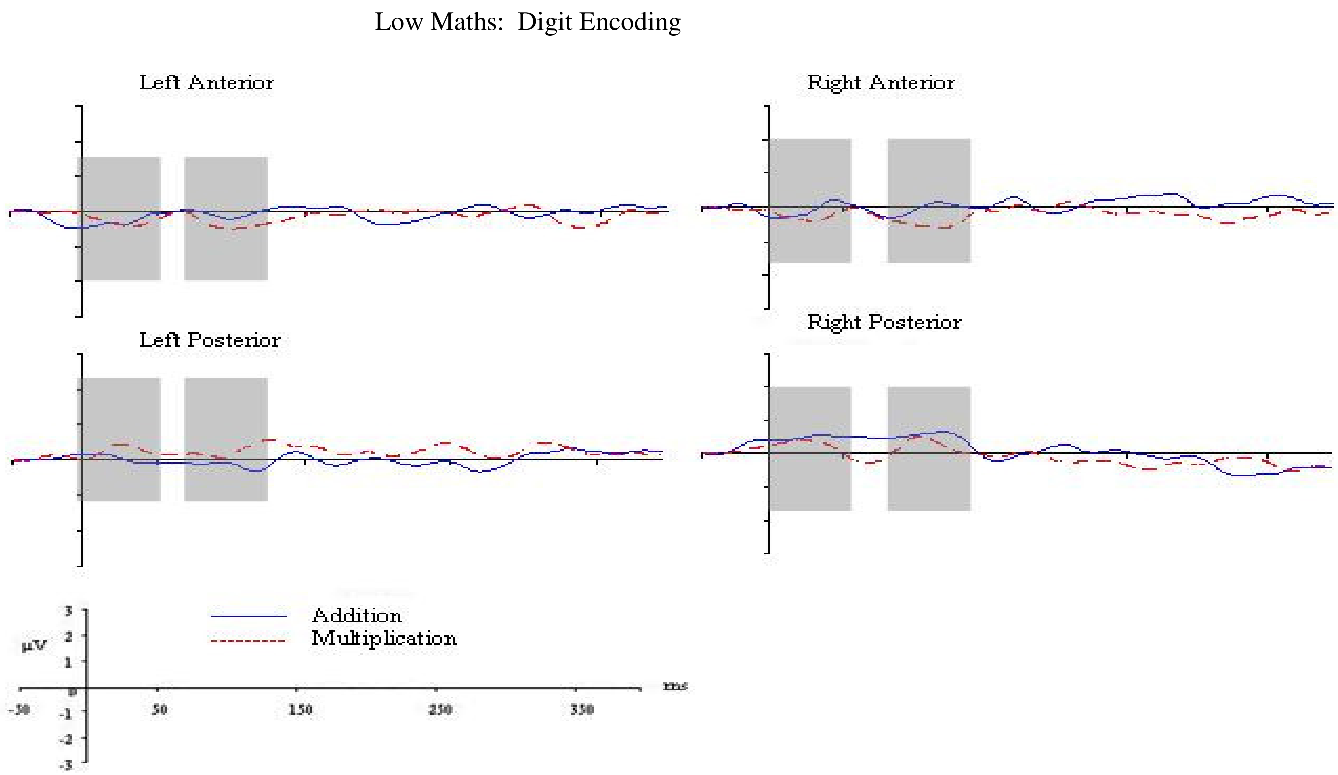

Figure 6.4. *Grand mean waveforms of the Low Maths group over the left and right anterior and posterior scalp in the time window 0 – 60 ms and 70 – 140 ms post-stimulus for the first operand in digit format equations.*

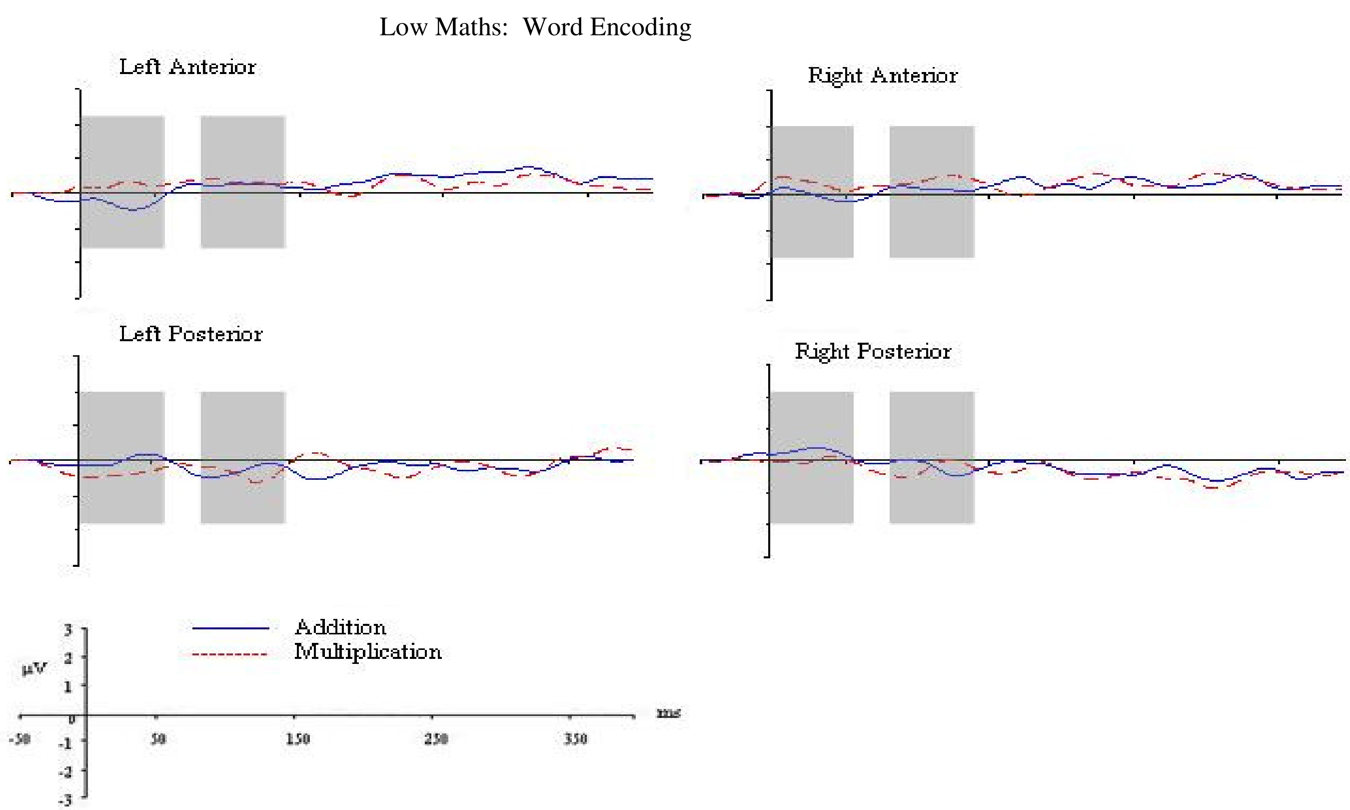

Figure 6.5. *Grand mean waveforms of the Low Maths group over the left and right anterior and posterior scalp in the time window 0 – 60 ms and 70 – 140 ms post-stimulus for the first operand in word format equations.*

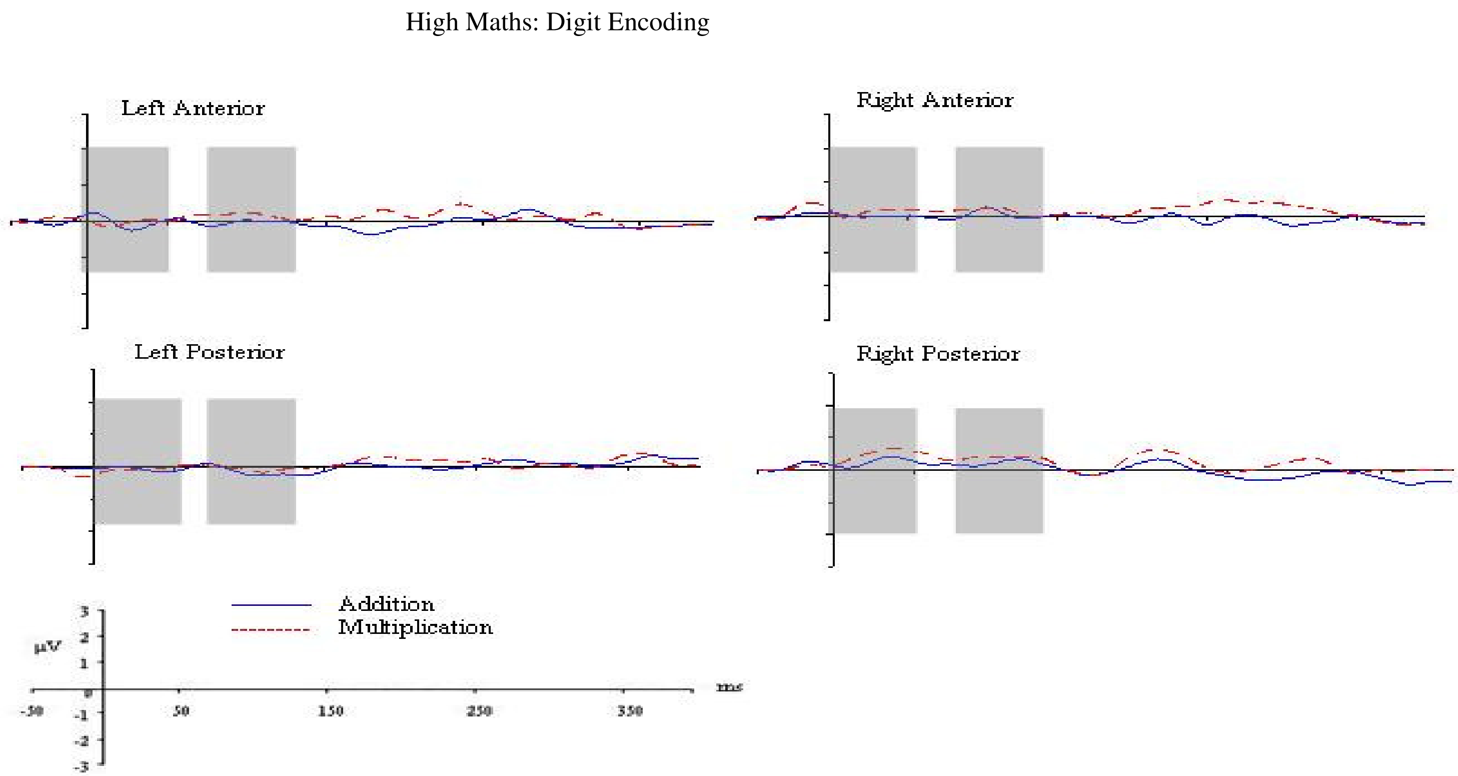

Figure 6.6. *Grand mean waveforms of the High Maths group over the left and right anterior and posterior scalp in the time window 0 – 60 ms and 70 – 140 ms post-stimulus for the first operand in digit format equations.*

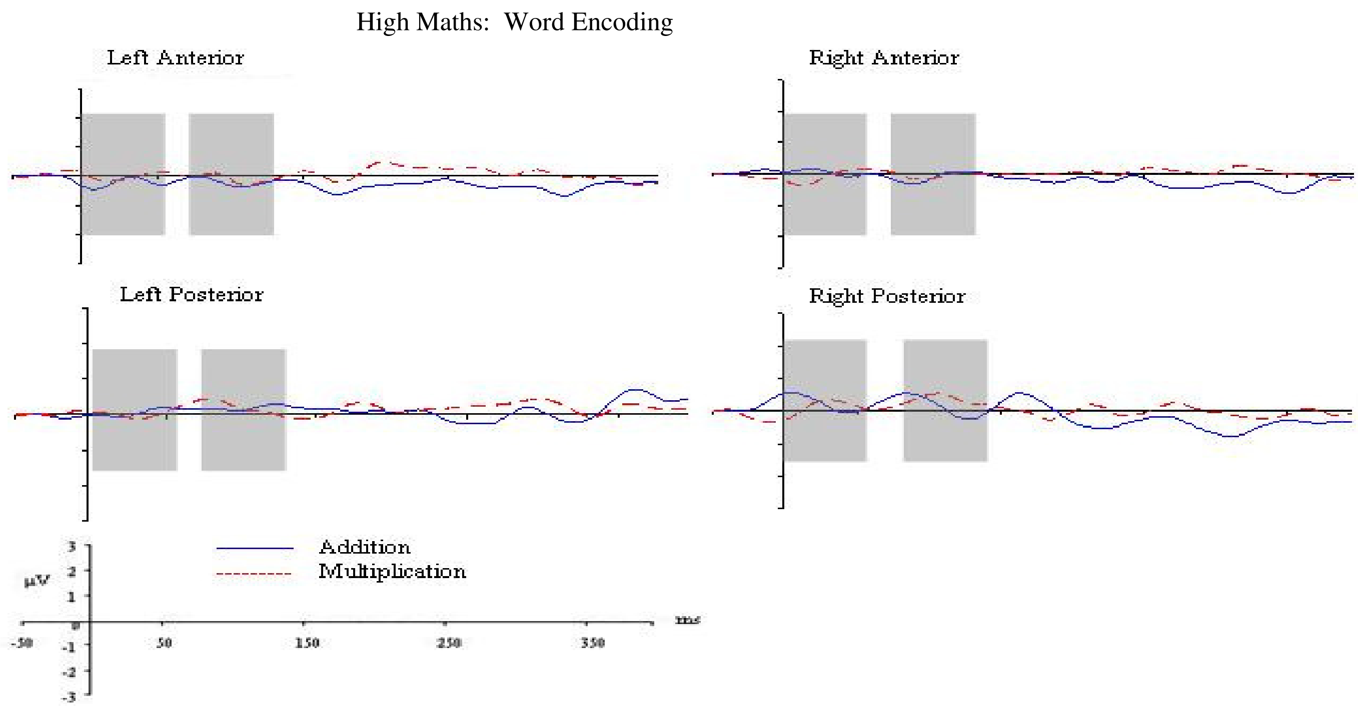

Figure 6.7. *Grand mean waveforms of the High Maths group over the left and right anterior and posterior scalp in the time window 0 – 60 ms and 70 – 140 ms post-stimulus for the first operand in word format equations.*

Retrieval/Calculation. Similar analyses were conducted for the second part of the equation presentation (operation sign, second operand, equals sign and answer) focusing on electrodes F3, F4, P3 and P4. Figures 6.8 to 6.11 present the event-related potentials for the retrieval/calculation phase of the digit and word format problems over the left and right anterior and posterior scalp for the Low and High Maths groups. The two shaded regions represent the time-windows 100 to 180 ms and 270 to 440 ms post-stimulus, which were selected for analysis based on clear variations in amplitude based on the grand mean waveforms. In comparison with the encoding phase, clear components emerged for the retrieval/calculation phase. From visually inspecting the grand mean waveforms, the High Maths group seemed to show a left anterior operation effect in the second time window for digit and word format problems, whereas the Low Maths group did not seem to show any clear amplitude differences between addition and multiplication over the left anterior region.

An ANOVA was conducted on the mean amplitudes in the time window between 100 – 180 ms post-stimulus (first shaded region in Figures 6.8 – 6.11) based on the presence of large negativity in this time window from observing the grand mean waveforms. A main effect was found for electrode position, $F(3, 48) = 24.73$, $p < .01$ (partial eta squared = 0.61). Interaction effects were found for format x electrode, $F(3, 48) = 7$, $p = .001$ (partial eta squared = 0.3) as well as format x operation x electrode, $F(3, 48) = 3.67$, $p = .018$ (partial eta squared = 0.19) showing that the location of the operation effect differed between digit and word format equations. No further main or interaction effects were found.

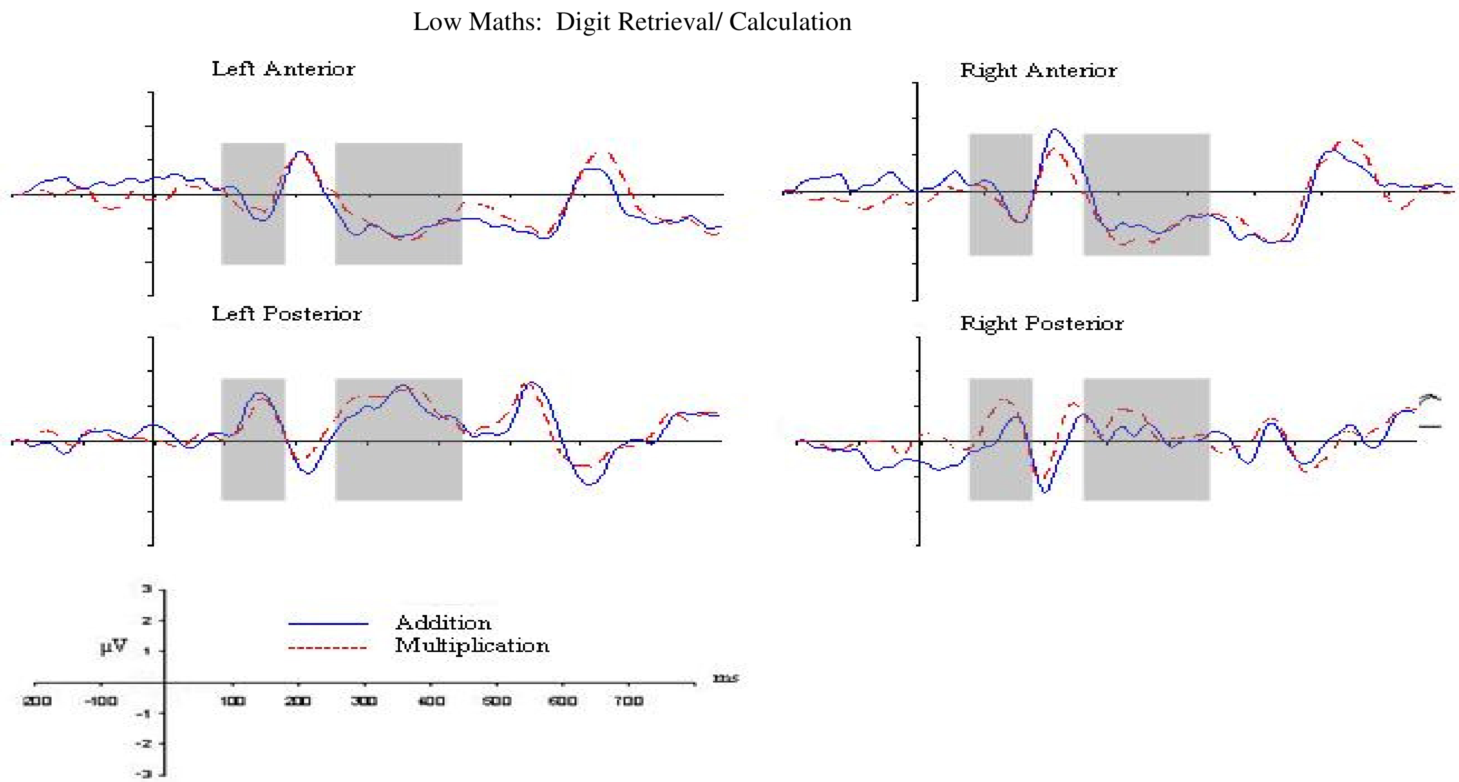

Figure 6.8. *Event-related potentials of the Low Maths group over the left and right anterior and posterior scalp in the time windows 100 – 180 ms and 270 – 440 ms post-stimulus for the second part of digit format equations.*

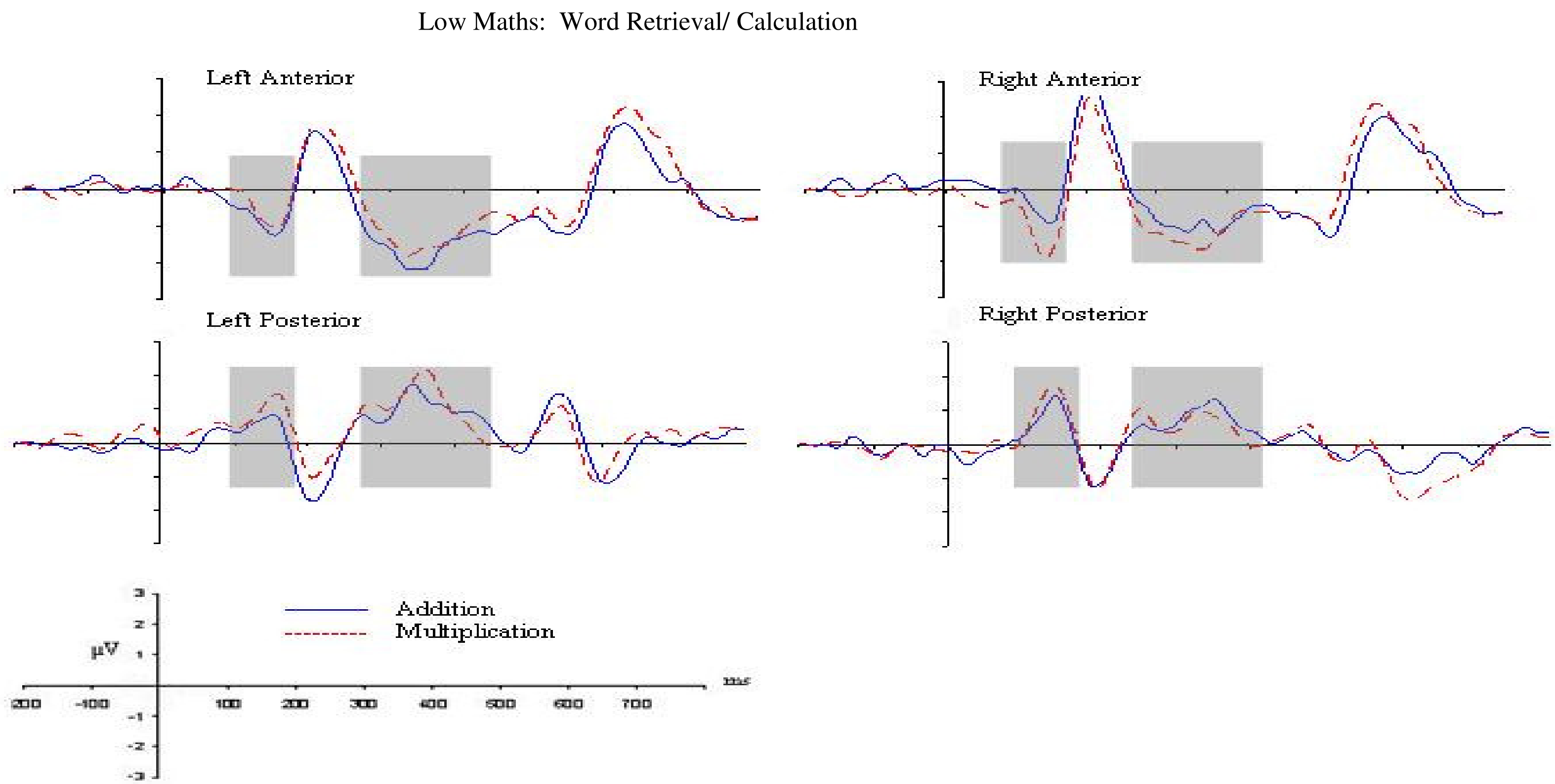

Figure 6.9. *Event-related potentials of the Low Maths group over the left and right anterior and posterior scalp in the time windows 100 – 180 ms and 270 – 440 ms post-stimulus for the second part of word format equations.*

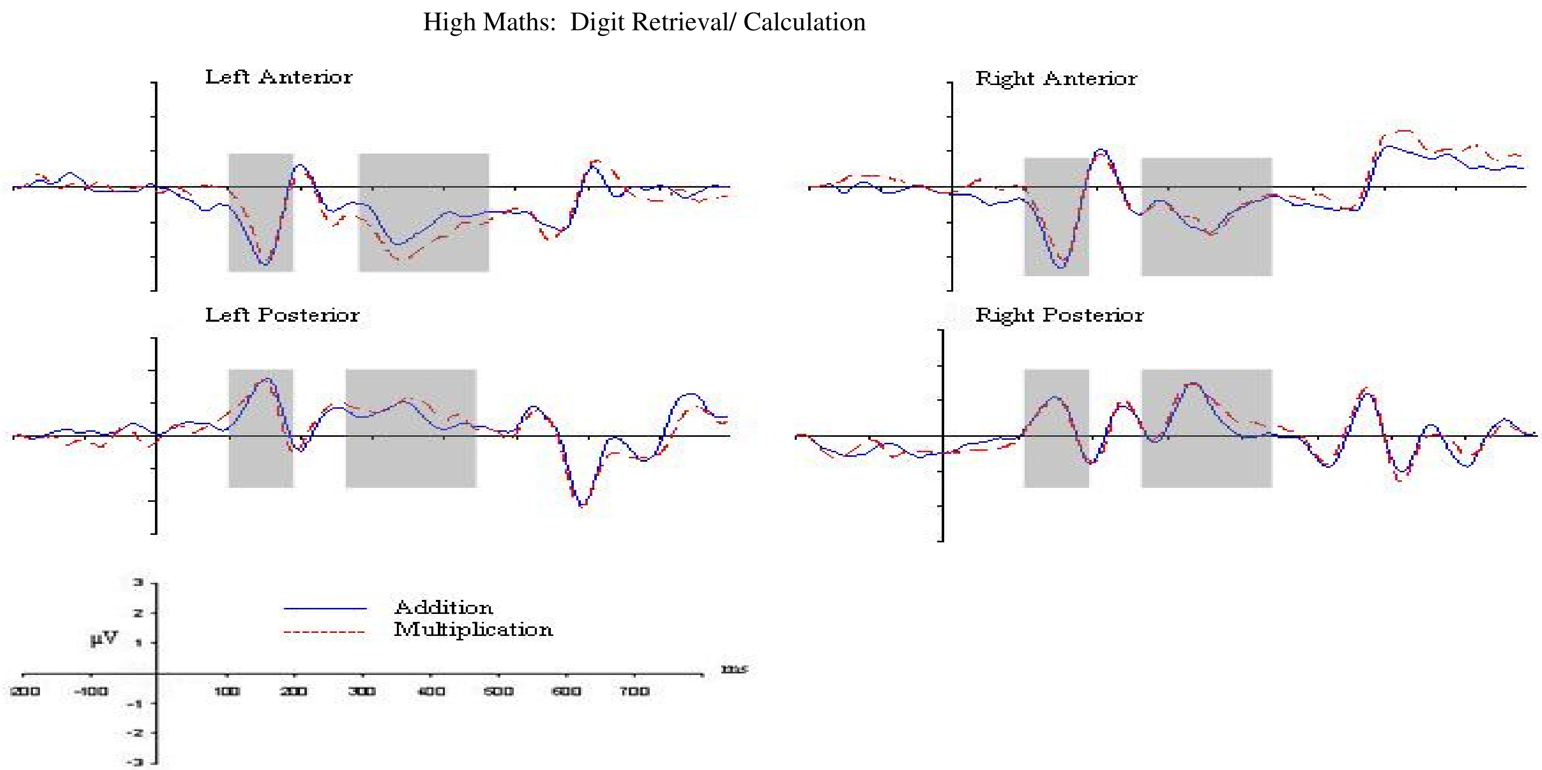

Figure 6.10. *Event-related potentials of the High Maths group over the left and right anterior and posterior scalp in the time windows 100 – 180 ms and 270 – 440 ms post-stimulus for the second part of digit format equations.*

High Maths: Word Retrieval/ Calculation

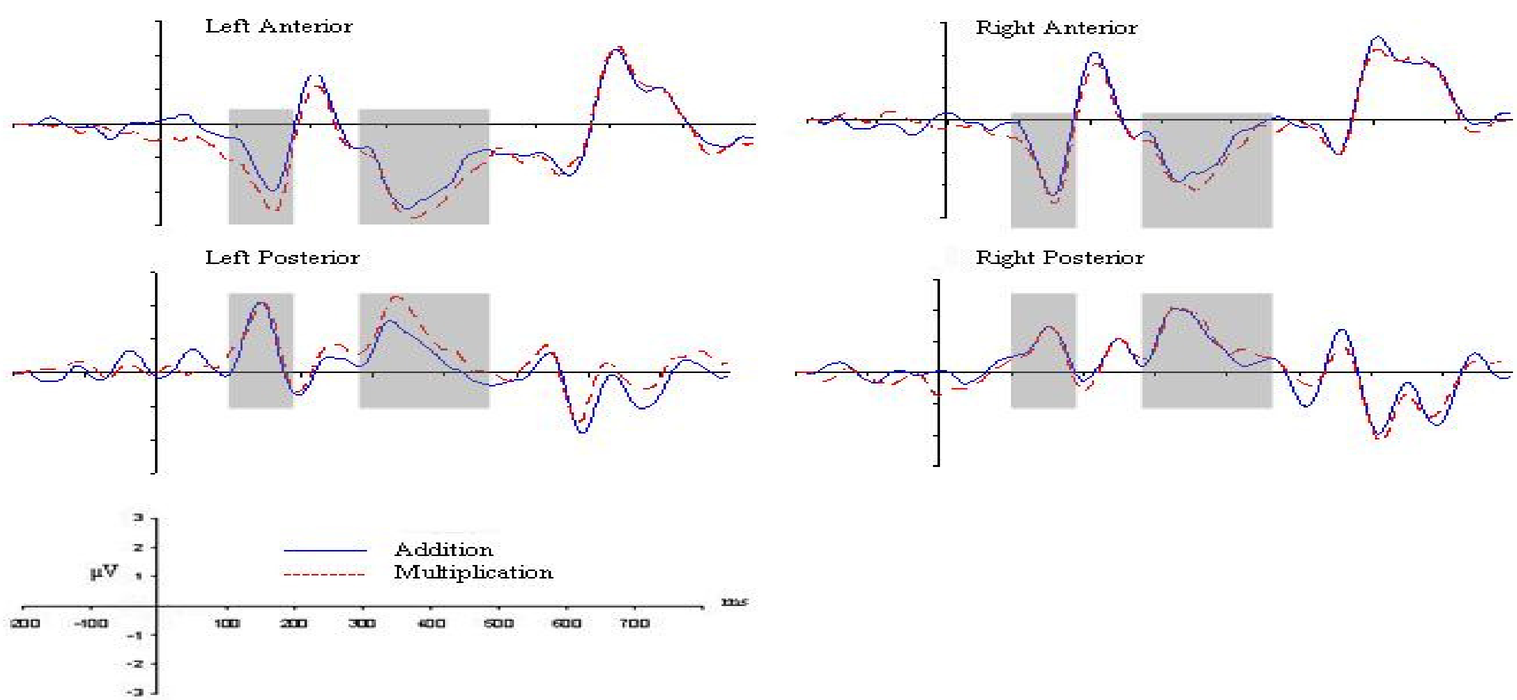

Figure 6.11. *Event-related potentials of the High Maths group over the left and right anterior and posterior scalp in the time windows 100 – 180 ms and 270 – 440 ms post-stimulus for the second part of word format equations.*

Dependent t-tests showed that for digit stimuli the Low Maths group showed an operation effect in the first time window with slightly larger right posterior ERP responses for multiplication than for addition, $t(7) = -2.74$, $p = .029$. The High Maths group showed no operation effect in this time window, with no significant difference in amplitudes between multiplication and addition.

To summarise, in the time-window between 100 – 180 ms post-stimulus, no left anterior operation effect occurred. However, the Low Maths group showed a small, but significant, operation effect for digit stimuli over the right posterior region.

A second analysis of variance was conducted for the time window between 270 and 440 ms based on visible amplitude variations in the grand mean waveforms. A main effect was found for electrode position, $F(3, 48) = 28.403$, $p < .001$ (partial eta squared = 0.64). Significant interaction effects were found for operation x Maths group, $F(1, 16) = 4.83$, $p = .043$ (partial eta squared = 0.23), and format x Maths group $F(1, 16) = 5.7$, $p = .03$ (partial eta squared = 0.26) suggesting that the format and operation effects differed between the two groups. Overall, the location of the operation effect differed with the format of the problems as indicated by significant interaction effects for operation x format, $F(1, 16) = 9.54$, $p = .007$ (partial eta squared = 0.37), operation x electrode, $F(3, 48) = 19.46$, $p < .01$ (partial eta squared = 0.55) and operation x format x electrode, $F(3, 48) = 8.73$, $p < .01$ (partial eta squared = 0.35). No further main or interaction effects were found.

Taking the High Maths group, an operation effect emerged over the left anterior scalp for digit, $t(9) = 2.99$, $p = .015$, as well as word format problems, $t(9) =$ -

3.97, $p = .003$, reflected as greater amplitude responses for multiplication than addition. The Low Maths group showed no left anterior operation effect.

Overall, clear operation effects emerged during the retrieval/calculation phase of arithmetic, reflected as greater amplitude responses for multiplication than addition, an effect that was not found during the encoding phase. The operation effect, found in the time window of 270 – 440 ms, was only found for the High Maths group, whereas the Low maths group showed no difference in amplitude in this regard. Furthermore, the High Maths group showed this operation effect for digit as well as number word equations.

6.4. Discussion

The current study utilised a true/false verification task to investigate operation effects in simple arithmetic for digit and word format equations. To investigate the processing that takes place at the encoding and retrieval/calculation stages of arithmetic, event-related potentials were investigated separately for the presentation of the first operand and then for the presentation of the rest of the equation (Zhou, 2011). The results showed that clear differences in processing between addition and multiplication only emerged during the answer retrieval phase and not during the presentation of the first operand as was found by Zhou (2011). In this study Zhou (2011) argued that in anticipation of the operation that is to follow, addition and multiplication problems presented in the same format (digits) are already encoded differently during the presentation of only the first operand. Addition and multiplication operands were thus thought to be represented in separate codes from which the answer could be retrieved. Zhou (2011) argued that encoding conditions (such as format) should have no direct influence on retrieval conditions based on the observation that digit operands are encoded as either addition or multiplication operands. The encoding and retrieval stages were thought to be additive rather than interactive, in which case operation effects should have been more evident during the retrieval/calculation phase and not necessarily during encoding (Dehaene, 1992; Dehaene & Cohen, 1995; McCloskey, 1992; McCloskey & Macaruso, 1995). The latter view is supported by Campbell and colleagues (e.g. Campbell, 1992, 1994, 1999; Campbell & Clark, 2009) suggesting that the format of the operands (encoding conditions) have a direct influence on the strategies that are used in arithmetic. This view assumes that different formats, rather than different operations are represented in separate representational codes (encoding complex view). According to Zhou (2011), if this view (interactive view) were supported, digit operands in either addition or multiplication should have been encoded similarly.

By including digit as well as word format problems, the current findings are thus more in support of the interactive and format-specific view of arithmetic with behavioural and ERP patterns suggesting that operands presented in the same format are encoded relatively similarly for addition and multiplication. This view assumes that instead of a uniform abstract number code for different numerical formats, each format is presented in a separate code, which can hinder or promote the use of different strategies in arithmetic (e.g. Campbell & Alberts, 2009). The basis for this argument is that operation effects were generally absent from the overall reaction times and event-related potential patterns observed during the encoding phase (presentation of the first operand). Clear operation effects in the event-related potentials only emerged during the retrieval/calculation phase of presentation after participants had seen the part of the equation containing the operation sign and answer. Zhou (2011) similarly did not find any operation effects in error rates or reaction times, with these effects only emerging in the EEG analyses, which demonstrates the usefulness of the ERP technique for highlighting effects that are not evident from behavioural analyses alone.

Overall, no clear components were evident for the encoding of the first operand. Clearer effects of operation were evident for the retrieval/calculation stage and the results also suggested that these effects differed between High and Low Maths participants. Taking digit format equations, the operation effects which emerged during retrieval/calculation are in support of previous results (Zhou, 2011; Zhou et al., 2007). However, the current findings suggest that the left anterior operation effect might only hold for High Maths individuals. Zhou (2011) interpreted larger negative left anterior potentials for multiplication to be an indication that greater verbal processing takes place for multiplication than addition. This is in line with fMRI findings showing greater activation of language areas for multiplication (Rickard et al., 2000; Zhou et al., 2007).

The High Maths group also showed this operation effect over the left anterior region for word format problems, whereas the Low Maths group showed no difference in left anterior amplitudes between addition and multiplication equations. This could suggest that for the High Maths participants, multiplication facts were represented phonologically regardless of format, which is also in line with the argument that in multiplication direct memory retrieval takes place for both digit and word formats (e.g. Campbell & Alberts, 2009; see Chapter 5). Furthermore, this finding suggests that previous findings of an operation effect over left anterior scalp might disappear for Low Maths individuals.

For Low Maths participants, a modest operation effect occurred for digit format equations over the right posterior region in the earlier time window (100 – 180 ms post-stimulus). In line with evidence from neuropsychology, an operation effect over the right posterior region could suggest that addition involves more visual spatial processing and activation of numerical magnitude representations than multiplication (e.g. Dehaene et al., 2009; Zhou, 2011). The fact that this effect was only found for the Low Maths group could therefore suggest that, with more mathematics experience, visual spatial processing might not necessarily need to occur since multiplication as well as addition facts could be readily retrieved from memory, as opposed to utilising a different strategy (e.g. counting) to arrive at addition answers. The faster RTs of the High Maths group, and the presence of a left anterior operation effect, also support the argument that, in general, individuals with greater mathematics experience might rely more on direct memory retrieval as a strategy. Overall, these findings highlight the importance of investigating effects in arithmetic problem solving at different levels of mathematics experience.

A possible reason why no clear operation effects were observed during the encoding phase in the current experiment could reflect the fact that digit and word format problems were presented in the same block. When the format of the operands stays the same, as was the case in Zhou's (2011) experiment, participants might pay more attention to operation than format, since operation switched between blocks, but format stayed the same. However, if the format of the operands switches unpredictably between digits and words as was the case in the current experiment (the order was pseudorandom), participants might pay more attention to format, than to operation. Attention to operation is thus only necessitated during the second stage of presentation where participants see the operation sign and not necessarily during the presentation of a single digit/word operand. It might thus be useful for future studies to present digit and word format problems in separate blocks in order to observe the presence or absence of operation effects during encoding. What the current results do seem to suggest is a transcoding advantage that comes with greater mathematics experience, with left anterior operation effects during retrieval, found for High Maths individuals regardless of operand format.

To conclude, by examining the event-related potentials that occur for digit and word format equations the current study suggests that operand format influences the answer retrieval stages of arithmetic. The study also highlights group differences, which is in support of the previous chapters which suggested that with high mathematics experience an advantage is gained for arithmetic fact retrieval, regardless of format.

Overall, the cost of word format on performance is consistent with the previous chapters and also in line with the argument that word format hinders retrieval of arithmetic facts (Campbell & Alberts, 2009). Also, in support of the findings of Chapter 5, the findings from Chapter 6 show that High Maths participants performed better on the arithmetic tasks in digit, as well as word format; a finding that can be interpreted as a transcoding advantage. On the basis that Experiments 1 –

3 (Chapters 2 – 4) suggested that with high mathematics experience, a general advantage is gained for extracting numerical information from various formats, the eye-tracking and ERP data show that this early advantage could aid these individuals in more complex numerical functions such as arithmetic.

Chapter 7

General Discussion

7.1. Overview

Considering the wide range of opinion regarding the mental representation and manipulation of numbers, the current thesis aimed to investigate how symbolic numerical notation influences numerical cognition, and how mathematics experience might regulate this process further. As discussed in the earlier chapters, the current debates concerning the representation of numerical information from different formats assume format-specific (e.g. Campbell & Clark, 1988; 1992; Campbell, 1994) or format-independent processing (e.g. Gallistel & Gelman, 1992; McCloskey, 1992; McCloskey & Macaruso, 1995). Yet other views postulate the co-existence of format-independent and format-dependent processing pathways depending on the numerical function that is required in a task (e.g. Dehaene, 1992; Dehaene & Cohen, 1995; Nieder et al., 2006). With regards to calculation, for example, it is yet uncertain if format effects can merely be attributed to differential encoding processes that take place for different formats, or if number format plays a role at all levels of numerical processing, including number manipulation which occurs after encoding. In addition to this, individual differences related to mathematics experience have generally not been considered in studies of adult numerical cognition, even though format-specific experience with numbers has been suggested to play a key role in processing differences between formats (e.g. Campbell & Alberts, 2009). Effects of format were thus investigated in the current research for individuals of differing levels of mathematics experience across a wide range of tasks to include basic numerical processing, as well as more advanced functions such as arithmetic. The following sections highlight the key findings related firstly to format effects and secondly to mathematics experience in terms of basic number encoding and arithmetic.

7.2. Format-specific Encoding

The first two experiments considered processing differences between digits and number words in terms of basic numeral encoding. During these simple tasks, two stimulus features were placed into competition with one another and participants were required to respond to one feature and to try to ignore the other stimulus feature. The degree to which the processing of the task-irrelevant stimulus feature interfered with attending to the task-relevant stimulus feature gave a measure of cognitive interference, and in turn a measure of the automaticity of processing of the task-irrelevant feature. It was thus possible to compare the processes by which underlying number meanings were accessed from digits and number words.

Number format was found to be more influential on number comparison (Experiment 2) than on subitizing (Experiment 1). In Experiment 1 (Counting Stroop task), overall, similar Stroop interference effects seemed to occur for digit and word formats; format effects only emerged when mathematics experience was considered, as discussed in section 7.4 below. When the two highly automatic processes of reading and subitizing were placed into competition with one another, incongruent number word conditions (e.g. '**two two two**'; respond '3') slowed down the counting process to much the same extent as incongruent arabic digits (e.g. '**2 2 2**';

respond ‘3’). In other words, the automaticity of processing of number words and arabic digits was relatively similar. This effect can be accounted for as follows: The numerosity (number of items on-screen) of the items on each trial was represented analogically rather than symbolically and format thus had no effect on the numerosity of the items (i.e. it is just as easy to count four digits as it is to count four words). Participants automatically counted the number of items on-screen regardless of whether or not they were a number of identical digits, symbols, or words, demonstrating the automaticity of the subitizing process. The level of ‘competition’ imposed by the numerosity of the items was thus similar for digit and word format trials.

With regards to reading, a similar process also seemed to take place for digits and number words. As mentioned in Chapter 1, highly practised number words, especially small numbers, can also often follow a conceptual reading route similar to digit reading, without the need for letter-sound mapping. Therefore, with similar subitizing and reading processes taking place for the two formats, no significant advantage was gained for digit processing in the counting Stroop task. As will be discussed in section 7.4, subtle differences in digit processing did emerge between High and Low Maths participants. However, overall, number words showed similar cognitive interference to digits suggesting that both formats can be processed along a similar route.

In Experiment 2 (Chapter 3), the process of comparing two numbers in terms of numerical magnitude seemed to be much more difficult for number words (e.g. ‘**two** **seven**’; respond ‘7’) than for arabic digits (e.g. ‘**2** **7**’; respond ‘7’). This was also the case for comparing numbers in terms of physical size (e.g. '**two** five'; respond '2' vs. '**2** 5'; respond '2'). The asymmetry in the observed format effects in Experiments 1 and 2 is likely to reflect the fact that the numbers used in Experiment 1 were smaller than those in Experiment 2. Both experiments used numbers in arabic digit and number word format, however, in Experiment 1, only numbers within the subitizing range (1 – 4) were used, whereas in Experiment 2, the numbers ranged from 2 – 9. Following the argument from Dehaene (1997) that in language the numbers 1 – 3 are widely used and therefore highly practised, in Experiment 1, digits as well as number words could thus have been read through a meaning-mediated reading route, with letter-sound mapping not necessarily taking place for word stimuli. In Experiment 2, however, it could be the case that for numbers up to 4, arabic digits and number words are still processed relatively similarly, whereas for numbers that are less practised and that are outside the subitizing range, an advantage is found for digits, as this is a more familiar numerical representation. The digit '3' and the word 'three' might thus be processed similarly, but the magnitude meaning of the digit '9' might be accessed more automatically than that of the word 'nine', which could account for the advantage on digit trials. Under time-pressure, the time taken to encode two different number words should also be expected to take longer than the time to encode two different digits, unlike in Experiment 1 where the number words or digits were identical. When number words were read, accurate numerical and size comparisons were still made suggesting that number meaning was still spontaneously encoded. The distance effects also demonstrated that the number meaning of each word was encoded distinctly. However, the process of accessing underlying number meanings seemed to be much slower for number words than for digits, which could relate to time-consuming letter-sound mapping that take place for large number words (less practised) especially.

Overall, these early experiments show that the task used to study format effects, can strongly influence inferences made regarding the abstractness of symbolic numerical representations, which could account for the wide theoretical disagreement on this issue in the literature, with evidence concerning format effects predominantly coming from studies of arithmetic (e.g. Campbell, 1994; Campbell & Alberts, 2009; Noël et al., 1997). By taking the findings from the cognitive interference tasks, the observed advantage for digits on number comparison is thus not necessarily an indication that digits and words follow separate representational pathways, but rather that the time taken to transcode from number symbols to number meanings can take longer for number words than for arabic digits. Following Dehaene (1997), in order to carry out accurate number comparison, numbers from various formats need to be transcoded to the underlying abstract numerical code, a process that might take longer for number words than for digits.

7.3. Format-specific Arithmetic

The format-independent viewpoint argues that once numbers have been transcoded from different formats to underlying number meanings, any subsequent processes that take place should not differ between formats, since both operate from the same identical, amodal number representation (e.g. Fias et al., 1996; Zhou, 2011). In this view, any performance differences between digits and words are thought to reflect differences in encoding processes rather than differences in processing that take place subsequent to encoding, such as calculation strategies. The processing pathways of digits and words might thus differ initially (e.g. digit processing might be faster), but once numbers have been translated to an amodal number code, the processing for the two formats remains similar. While the findings from Experiment 2 could be taken to suggest that different encoding processes for digits and words underlie the observed advantage for digits in number comparison, the evidence from the studies on arithmetic (Experiments 4 and 5) suggest that the advantage for digit format in arithmetic is not merely related to faster encoding, but rather that different calculation procedures might take place for digits and number words. In line with the suggestion from Campbell and colleagues (Campbell & Alberts, 2009; Campbell & Epp, 2005), the findings from Experiments 4 and 5 suggest that performance differences between formats can be related to the use of different calculation strategies, which occur subsequent to encoding.

In support of the findings of Campbell and Alberts (2009) on strategy use in arithmetic, the eye-tracking data presented in Chapter 5 suggested that word format discouraged, but digit format *en*couraged, the use of direct retrieval as a strategy. Specifically, in Experiment 4, the fixation count data showed a format x problem size interaction for addition, but not for subtraction. For addition, a greater size effect was found for digit than for word format problems, reflecting the fact that very few fixations were made on small digit format problems (participants found these problems the easiest). However, for subtraction a similar pattern was found for the two formats: the increase on large number problems was similar for digits and number words. Following Campbell and Alberts (2009), a diminished format effect for large subtraction suggests that the same strategy was used for both formats. Since large subtraction problems are usually more difficult, participants resort to counting-based strategies rather than direct retrieval, regardless of format, which makes format effects less prominent here. If the format-independent viewpoint had been supported, one would not expect such interactions of format, operation and problem size in the observed fixation patterns. Instead, if the two formats followed similar processing pathways, the use of a similar strategy should be expected for both formats, and the observed differences in terms of the problem size effects should not be expected for addition.

This argument was further supported by the findings for multiplication and division in Experiment 4 Part 2, namely similar behavioural and eye-tracking patterns for digits and words in the former, but clear word-format costs on performance in the latter. In line with the format-specific view of Campbell and colleagues, the reason for a smaller format effect in multiplication is that the same strategy is used for both formats, namely direct retrieval. Under time pressure, multiplication necessitates direct retrieval, since another strategy (e.g. repeated addition, division-reference, counting etc.) would be too inefficient. Word-format is thus not particularly costly in multiplication since both digit and word format problems are solved via retrieval. In division, however, there is opportunity for a switch to an alternative efficient strategy, namely the multiplication-reference strategy (e.g. $2 \times 3 = 6$, therefore, $6 \div 3 = 2$), which seems to be the case for division problems in word format. Similar to the findings for addition, if format differences were merely related to encoding differences, similar word format costs should be expected in both operations. In other words, one might expect overall weaker performance for word-format problems due to slower encoding/weaker retrieval strength, but not that word format should interact with operation and problem size to promote the use of different strategies in comparison to that of digit format problems, as was suggested by the eye-tracking patterns.

As a whole, these findings highlighted the importance of investigating format effects across a wide range of numerical tasks including all arithmetic operations. Prior to the study of Campbell and Alberts (2009), interactions of format, operation and problem size had not been investigated for operations other than addition. Experiment 4 was also the first to investigate such effects by means of eye-tracking in an attempt to eliminate self-report bias in relation to strategy use and suggested that it is a useful means of studying calculation processes in arithmetic. In relation to the format-specific view of arithmetic, the observed operation x format x problem size interactions are in support of the interactive view of arithmetic, which argues for the close interaction of encoding conditions with answer-retrieval conditions in arithmetic.

The final experiment (Experiment 5) provided further support from event-related potentials for the interactive view of arithmetic by showing that format and operation effects were mostly attributable to the answer-retrieval stage of arithmetic, suggesting that format influences retrieval strategies specifically. By isolating a pure encoding stage (presentation of the first operand) from the presentation of the rest of the equation (e.g. '**2**' and '**+ 3 = 5**'), the aim of the experiment was to investigate if the two stages of encoding and retrieval were additive (e.g. McCloskey & Macaruso, 1995; Zhou, 2011) or interactive (e.g. Campbell & Clark, 1989; 1992; Campbell & Alberts, 2009). During this true/false verification task, addition and multiplication equations were presented in separate blocks in order to investigate if participants could anticipate the addition/multiplication operation that was to follow even during the phase of encoding the first operand when only a single digit or word was presented on-screen. In the case of the additive viewpoint, similar effects of operation and format are expected to occur during encoding and retrieval. In this case, numbers (e.g. digits) are thought to be encoded as distinctly addition or distinctly multiplication operands (regardless of format), from which the appropriate arithmetic facts can be retrieved or calculated. If this view were to be supported, amplitude responses during the encoding phase (when only the first operand is seen) should vary with operation, but not necessarily with format, since numerical information from all formats is thought to be transcoded to an underlying uniform code. Thus, if addition and multiplication operands (presented in the same format) are encoded separately, this should be evident during both the encoding and retrieval stages, suggesting that the two stages operate serially, rather than interactively.

In the case of the alternative view, namely that the encoding and retrieval stages interact with one another, the differences in performance between addition and multiplication should rather relate to the answer retrieval stage specifically, and not necessarily to the ‘pure’ encoding stage when only a single digit or word is presented on-screen (Zhou, 2011). Encoding features, such as format, are also thought to play an important role on any subsequent retrieval processes with operand encoding relating more to format than to operation (e.g. Campbell’s Encoding Complex model). In this view, addition *or* multiplication operands presented in the same format are thus encoded similarly and operation effects are only expected to emerge during answer retrieval.

The findings from Experiment 5 are in support of this view and suggested that the encoding and retrieval stages operate interactively rather than serially. The ERP data showed that effects of operation were generally absent from the encoding phase (where participants only saw a single digit or number word on-screen), but only emerged during retrieval (when participants saw the rest of the equation including the operation sign). This suggests that the observed differences in performance between digits and words can be attributable to effects that occur during the answer-retrieval stages specifically, such as strategy choice as was suggested from the eye-tracking evidence in Experiment 4.

The benefit of more sensitive measures, such as eye-tracking and ERP technology, is that they can highlight effects that are not evident from behavioural measures (e.g. accuracy and RT) alone. As pointed out by Zhang et al. (2010) stimuli that result in similar behavioural effects might still be processed along separate representational pathways, suggesting that more sensitive measures are needed. In Experiment 4, Part 1, for example, format x size interactions were not clear from RT or accuracy data, but the fixation patterns suggested interactions of operation, format and problem size that are in accordance with the reports of Campbell and Alberts (2009) on strategy use in arithmetic. Similarly, in Experiment 5, no clear operation effects were evident from the RT data (as was the case in the study of Zhou, 2011), but the ERP patterns suggested that for High Maths participants solving multiplication relies more on verbal processing than addition (e.g. Zhou, 2011). The ERP data have also shown that the effects of format and operation can occur at the retrieval stage specifically and not necessarily due to mere encoding differences.

It is worth mentioning that the lack of operation or format effects during the encoding condition could have related to the fact that digit and word format equations were presented in the same test block, as discussed in Chapter 6 (p. 178 – 179), and similarly for Chapter 5, which followed the procedure of Campbell and Alberts (2009) who presented digit equations on odd and word equations on even trials. In situations where format stays the same throughout the experiment, as was the case for Zhou's (2011) study, arithmetic operands might well be encoded as distinctly multiplication or addition, since less attention needs to be paid to other encoding features such as format. Also, the degree to which the encoding and retrieval stages could be separately investigated in this task is also questionable. While the presentation of a single numeral (first operand) can be thought of as a purely encoding stage, effects related to number encoding could also be attributed to the 'retrieval' stage when participants see the rest of the equation. It is thus difficult to isolate effects related to the encoding of the second operand from effects related to answer retrieval/calculation specifically. What the current results do show, however, is that effects of format and operation were minimal for the presentation of the first operand, suggesting that encoding and retrieval does not necessarily operate serially. In the current tasks, format also seemed to interact with operation and problem size in arithmetic, suggesting that performance differences between digits and number words reflect more than just differential encoding processes.

With regards to the debate on the influence of format in numerical cognition, the current findings emphasise the importance of investigating format effects across a wide range of tasks in order to gain a comprehensive view of how numbers from different formats are represented. Overall, format seemed to be more salient for some numerical functions than others, suggesting that numerical processing is sometimes abstracted away from input format, but sometimes different formats follow separate processing routes, depending on the numerical function that is required under task-demands.

7.4. The Influence of Mathematics Experience on Format-specific Numerical Cognition

The studies used in the current research were designed to also consider individual differences related to experience with numbers and how this might regulate the influence of surface format in numerical cognition, a factor that has not yet been considered in previous accounts. Since the influence of format has been suggested to reflect an individual's experience with the specific format in question (e.g. Campbell & Alberts, 2009), and since recent reports show a concern over adult numeracy (e.g. Jukes & Gilchrist, 2006; Lipkus et al., 2001), mathematics experience seemed an important variable to consider. The main objective was to compare the numerical cognition of individuals with relatively more experience with numbers to that of individuals with less experience with numbers, in order to see if individuals with greater mathematics experience show an advantage for accessing number meanings from various formats.

It is worth mentioning at the outset of this discussion on mathematics experience, the difficulty that was encountered, firstly, in defining mathematics experience and secondly, in recruiting individuals to participate in mathematical research. While efforts were made to divide participants into groups of ‘high’ and ‘low’ levels of experience with mathematics, it must be kept in mind that numerical competency can be a reflection of many interwoven factors (e.g. Mazzocco, 2008). The definition of ‘mathematics experience’ was thus kept broad, to take into account factors such as an individual’s numerical ability, numerical self-efficacy, mathematics education history and working memory capacity. The main differences between the groups of mathematics experience was that those in the High Maths group had reported better performance in the Irish Leaving Certificate mathematics examination and also performed better on the numeracy measure that was administered in each experiment. However, Experiments 1 and 2 showed that the High Maths group also generally reported greater numerical self-efficacy and better working memory performance. The observed group differences in the current experiments could thus relate to an individual’s exposure to numbers through education or to an individual’s numerical aptitude in general, or to a combination of both factors.

Secondly, in cases where group differences were less clear, such as in Experiment 3, it could be a reflection of the sampling method. The measure of assigning participants to High and Low Maths groups could have been more robust in some cases due to the difficulty in recruiting participants. Individuals were often reluctant to participate in numerical cognition research, which seemed to be related to the fact that participants knew the experiments would include a time-restricted numeracy assessment or arithmetic equations, in some cases. In Experiment 3, for example, it could thus be the case that highly numerically competent individuals were included in the Low Maths group, since less numerically competent individuals were less likely to participate. Unlike Experiments 1 and 2, the two groups in Experiment 3 also did not show significant self-efficacy or working memory differences, further suggesting that group differences could have been more robust here. Also, in Experiments 1, 3, and 4 men showed an advantage for numeracy performance compared to women, whereas no gender differences were found in the other experiments, which suggest that this effect could be related to a sampling issue.

Notwithstanding this difficulty in defining mathematics experience, a number of findings from the experiments on basic numerical encoding (Experiments 1 – 3) suggest that the observed performance of individuals in the High Maths group was related to an advantage for processing numerical information specifically and that it was not merely a reflection of general memory efficiency or aptitude. Firstly, in the earlier experiments, the performance of the High and Low Maths groups did not differ in reaction time on neutral trials, but only on trials that featured numerical information. Secondly, with high mathematics experience, an advantage was found for numerical comparison (e.g. '**3** 5'; which number is higher?), but not for physical comparison (e.g. '**3** 5'; which number is physically bigger?), in which case the two groups performed relatively similarly. Finally, the High Maths group showed an advantage for accessing numerical meanings from language in general, as was found for quantifier words (Experiment 3). Again, no difference in performance was found for responding to neutral word stimuli, whereas the High Maths group responded faster on trials that contained quantifier words. Overall, these findings are in support of the accounts that postulate the existence of our number concept as a specific semantic domain that is independent of other abilities (e.g. Butterworth et al., 2001; Caramazza & Shelton, 1998; Koechlin et al., 1998). The fact that the High Maths group showed such a specific processing advantage for numerical stimuli from various formats supports this claim and suggests that if the semantic referent (number) is strongly represented in memory, it aids the transcoding of numerical information from various formats, and not just the well-practised digit format. Thus, while performance on word format trials was generally weaker than on digit format trials, the High Maths participants showed better performance for both formats, in comparison with the Low Maths participants.

However, format did seem to interact with Maths group on conditions where format effects were generally diminished, with no clear overall performance differences between arabic digits and number words. In such cases, where digit and number word processing was relatively similar, arguably reflecting the use of similar strategies for both formats (e.g. Experiment 4), the High Maths group showed an advantage for processing arabic digits. This was the case, for example, in Experiment 1, where no overall effect of format was noted, but the High Maths group showed more automatic processing, evidenced by cognitive interference, for digit stimuli compared to the Low Maths group. Similarly, in Chapter 5 Part 2, where overall performance on multiplication was relatively similar for digit and word format problems, individuals in the High Maths group showed an advantage for digit format. By considering mathematics experience, these findings show that the influence of format in numerical cognition can, at least to some extent, be regulated by an individual's experience with numbers.

Overall, Experiments 1 – 3 suggest that with greater mathematics experience an advantage can emerge for accessing number meanings, especially from digit format, but also to a lesser extent for number words and quantifiers. As expected, differences in performance were also noted in arithmetic performance, which showed that effects of operation, problem size and format differed across the groups. The evidence from reaction time and eye-tracking in Chapter 5 Part 1 (addition and subtraction) suggest that the Low Maths group were more influenced by the magnitude of the operands in an arithmetic problem than the High Maths group. A greater problem size effect thus emerged for individuals with less mathematics experience, showing that these individuals found large problems much more difficult than individuals with more mathematics experience. This could suggest that individuals with greater mathematics experience have an overall arithmetic fact retrieval advantage, which makes them less prone to resort to counting based strategies on large, more difficult problems. The performance for individuals with less mathematics experience on small problems overlapped with that of individuals with more mathematics experience on large problems, demonstrating the problem-solving efficiency of those with more mathematics experience. With regards to format, this advantage emerged for digit as well as word format equations, in accordance with the findings from the first three experiments, which suggested a general transcoding advantage that accompanies high mathematics experience.

For multiplication and division (Experiment 4 Part 2), the influence of problem size was relatively similar across each the groups. While individuals with high mathematics experience answered faster overall, the increase in RT and fixations on large equations was relatively similar for all groups. In line with the suggestions put forward in section 7.3 regarding strategy use, this suggests that in solving multiplication and division, strategies generally are more retrieval-based overall, in comparison with addition and multiplication. Thus, since counting-based strategies would be too inefficient, especially under time constraints, even individuals with less mathematics experience seemed to use retrieval based strategies, and not calculation, on large problems.

The final experiment (Experiment 5) further highlighted the importance of considering individual differences when investigating event-related potential effects of operation and format in arithmetic. The findings from Experiment 5 showed that the left anterior operation effect (e.g. Zhou, 2011; Zhou et. al. 2006, 2007) can differ with mathematics experience, arguably reflecting the use of different calculation strategies. Zhou and colleagues (2006, 2007; Zhou, 2011) demonstrated the operation effect for digit format equations as greater left anterior and right posterior amplitude responses for multiplication than addition. However, the findings from Experiment 5 showed that the operation effect over the left anterior region, thought to reflect the greater involvement of verbal memory in multiplication, was only found for the High Maths group. This effect was also found for digit as well as word format problems, suggesting similar processes for the two formats, namely verbal memory retrieval. On the other hand, the operation effect over the right posterior scalp, thought to be a reflection of more visual spatial and magnitude processing for addition than multiplication, was only found for the Low Maths group (although this component emerged earlier than that reported by Zhou and colleagues). In accordance with the evidence from Experiment 4, this can be interpreted as an advantage for arithmetic fact retrieval that accompanies high mathematics experience. The lack of a right posterior operation effect suggests that greater magnitude processing might not necessarily take place for addition than multiplication for individuals with high mathematics experience. In light of faster overall RTs, it was thus likely that these individuals solved both addition and multiplication problems through retrieval strategies, without the need for greater magnitude processing in addition.

7.5. Implications of the Current Research

With regards to number representation and manipulation, the current thesis highlighted a number of characteristics that seem to underlie proficient adult numeracy. Firstly, individuals who were more numerate seemed to show an overall transcoding advantage, namely more automatic access to underlying number meanings from different symbolic notations, but especially arabic digits. Secondly, these individuals seemed to show an overall advantage for arithmetic fact retrieval, regardless of format and problem size. In arithmetic, processing thus seemed to be generally more memory- than calculation-based, in comparison with individuals with less mathematics experience. Thirdly, the advantage of individuals with high mathematics experience seemed to be specifically numerical and not related to other advantages (e.g. response speed).

CHAPTER 7

In light of the recent concerns over declining adult numeracy the current findings link adult numerical proficiency to simple numerical functions such as translating from symbols to quantities. While all the participants in the current research were educated, numerate and literate adults, the fact that the benefits of mathematics experience was already evident at such a basic level suggests that mathematics experience creates a strong symbol–number concept which forms the basis for more complex numerical functions such as arithmetic.

In support of the recent findings on adult numeracy (e.g. Lipkus et al., 2009), the current findings also showed that even highly educated individuals often struggle with basic numerical and probability concepts (numeracy test). The relatively simple numeracy test used here proved to be challenging for some individuals, which showed that in the absence of explicit instruction, some individuals do not know which strategy or operation to use in order to solve such basic problems. Some individuals might thus not have acquired the necessary mathematical skills in education, which helps them apply numerical concepts to novel situations in adulthood. Overall, the evidence could suggest that in education, more attention should be focused early on translating between different symbolic formats and generalising to novel situations as these basic numerical functions seem to play an important role in more complex numerical functions such as arithmetic.

Lipkus et al. (2001) suggested that in everyday numeracy, which requires applying learned number concepts to practical situations, errors usually occur when individuals are required to switch between metrics. This occurs, for example, when participants switch from proportions (e.g. ½) to percentages (e.g. 50 %), similar to the difficulty that is encountered when transcoding between arabic digits and number words in arithmetic. In light of this, in training individuals to become more numerate, education methods should gain from recognising a) the importance of being able to switch between different symbolic formats and b) the importance of being able to generalise mathematical concepts to novel situations.

7.6. Limitations of the Current Research and Outlook for Future Research

While the advantage for digit format problems in arithmetic have been demonstrated across a number of languages, including French, Dutch, English and Chinese (Noël et al., 1997; Campbell et al., 1999), the observed format differences related to more basic number processing (Experiments 1 and 2) may only be applicable to the English language. As Zhang et al. (2010) pointed out, the visual-verbal number form that is usually compared with arabic digits is relatively rare in languages such as Chinese, for example. In this case, the advantage for digit format might thus be more obvious than in English. Cultural variables should also be expected to play a role in format-specific processing. In Chinese education, for example, emphasis is placed on extensive reciting of multiplication tables, which renders verbal-numerical number forms more salient than visual-verbal number forms (Zhang et al., 2010). Future research should thus take into account cross-language format effects that might relate more to modality of input (e.g. auditory or visual) than to surface format.

Leading theorists in the field of mathematical research emphasise the defining role of practice and memory in numerical proficiency and highlight it as the best predictor of mathematics achievement (e.g. Butterworth, 1999; Dehaene, 1997).

However, with regards to mathematics experience, it is beyond the scope of the thesis to say whether or not the observed advantage of the High Maths participants relates to practice and learning or whether it reflects numerical aptitude. While automaticity of processing (for numbers in this case), as was observed in Experiments 1 – 3, is thought to stem from extensive practice and memory for specific stimuli (e.g. Ashcraft, 2006), the advantage of the High Maths participants could also have reflected superior working memory efficiency or motivational variables, such as self-efficacy beliefs, as suggested by Experiments 1 and 2. Referring back to the issues related to defining mathematics experience and recruiting participants for participation in mathematics research (p. 193), it seems difficult to isolate practice and memory for numbers from other variables such as aptitude or general memory efficiency.

What is lacking in the adult mathematical cognition research is a consideration of individual differences related to mathematics experience, especially in basic number processing such as subitizing and number comparison. The current research has highlighted the importance of considering individual differences by showing that some of the effects observed in the literature might only hold for some individuals depending on their level of mathematics experience. The direction of future research should thus aim to employ more robust measures of adult mathematics experience and also consider other individual differences variables and how this might influence format-specific number processing.

7.7. Conclusions

In order to accurately understand and manipulate numbers, it is necessary to read, write and transcode numbers from various different symbolic formats (Dehaene, 1994). The importance of these basic functions are noted in early number processes such as number comparison, as well as more complex processes such as arithmetic, with theoretical accounts differing on how numbers from different formats are mentally represented. The current results are in support of views that assume the co-existence of format-specific and format independent numerical processing pathways (e.g. Dehaene's Triple Code Model). Arabic digits and number words might thus be represented in distinct pathways in the brain; however, for some numerical functions, information from both formats might require a similar processing route in order to arrive at a solution.

While most mathematically educated individuals can perform these basic transcoding functions, individuals who are more numerate show an advantage for accessing number meanings from not just the highly familiar digit format, but also from word format, and to a lesser extent quantifier words. This is in line with the argument for a language-independent number domain in semantic memory, and also supports the argument that accessing numerical information from different symbolic inputs (transcoding) is essential for numerical proficiency in adulthood, a property that seems to underlie the enhanced performance of highly mathematics experienced individuals. As a whole, the current research highlights the diversity of adult numeracy, and how this is evident from very basic numerical tasks, such as number comparison and subitizing, to more complex numerical tasks, such as addition, subtraction, multiplication and division. The findings show that individual differences related to mathematics are an important consideration for theoretical accounts of number processing.

References

Alario, F. X., Segui, J., & Ferrand, L. (2000). Semantic and associative priming in picture naming. *Quarterly Journal of Experimental Psychology, 53*, 741-764.

American Electroencephalographic Society. (1994a). Guidelines on evoked potentials. *Journal of Clinical Neurophysiology*, *11*, 40-73.

American Electroencephalographic Society. (1994b). Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *11*, 111-113.

Anderson, S.W., Damasio, A.R., & Damasio, H. (1990). Troubled letters but not numbers: Domain specific cognitive impairments following focal damage in frontal cortex. *Brain, 113*, 749-766.

Ansari, D., & Karmiloff-Smith, A. (2002). Atypical trajectories of number development: A neuroconstructivist perspective. *Trends in Cognitive Science, 6*, 511-516.

Antell, S., & Keating, D.P. (1983). Perception of numerical invariance in neonates. *Child Development, 54*, 695-701.

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

Ashcraft, M. H. (2006). *Cognition* (4th ed.). New Jersey: Pearson.

Ashcraft, M. H., & Christy, K. S. (1995). The frequency of arithmetic facts in elementary texts: Addition and multiplication in grades 1– 6. *Journal for Research in Mathematics Education*, *26*, 396- 421.

Baddeley, A., & Hitch, G. (1974). Working memory. In Bower, G. (ed.), *The Psychology of Learning and Motivation*. (pp. 47-89). Academic Press.

Baillargeon, R. (1994). Physical reasoning in young infants: seeking explanations for impossible events. *British Journal of Developmental Psychology, 12,* 9-33.

Banks, W. P., Fujii, M., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 435-447.

Barner, D., Chow, K., & Yang, S. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology, 58,* 195-219.

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005) Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences USA 102,* 14116-14121.

Bassok, M. (2001). Semantic alignments in mathematical word problems. In Gentner, D., Holyoak, K. J., & Kokinov, B. N. (Eds.) *The analogical mind: Perspectives from cognitive science.* Cambridge, MA: MIT Press.

Berent, I., Pinker, S., Tzelgov, J., Bibi, U., & Goldfarb, L. (2005). Computation of semantic number from morphological information. *Journal of Memory and Language, 53*, 342-358.

Berg, P., & Scherg, M. (1991). Dipole models of eye movements and blinks. *Electroencephalography and Clinical Neurophysiology*, *79*, 36-44.

Bernardo, A. B. I. (2001). Asymmetric activation of number codes in bilinguals: Further evidence for the encoding-complex model of number processing. *Memory and Cognition, 29,* 968-976.

Besner, D., & Coltheart, M. (1979). Ideographic and alphabetic processing in skilled reading of English. *Neuropsychologia, 17*, 467-472.

Blankenberger, S., & Vorberg, D. (1997). The single-format assumption in arithmetic fact retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23,* 721-738.

Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in english number agreement. *Language and Cognitive Processes, 8,* 57-99.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23*, 45- 93.

Boysen, S.T., & Capaldi, E.J. (1993). *The development of numerical competence:animal and human models.* Hillsdale, N.J: Erlbaum.

Brugger, P., Pietzsch, S., Weidmann, G., & Biro, P. (1995). Stroop-type Interference in random-number generation. *Psychological Reports, 77,* 387-390.

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology, 103,* 1131-1136.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting Stroop: An interference task specialized for functional neuroimaging – validation study with functional MRI. *Human Brain Mapping, 6,* 270-282.

Bush, G., Whalen, P.J., Shin, L.M., & Rauch, S.L. (2006). The counting Stroop: a cognitive interference task. *Nature Protocols, 1,* 230-233.

Butterworth, B. (1999). *The Mathematical Brain.* London: Macmillan.

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, 46*, 3-18.

Butterworth, B., Cappelletti, M., & Kopelman, M. (2001). Category specificity in reading and writing: the case of number words. *Nature: Neuroscience, 4(8),* 784-786.

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

Campbell, J. I. D., & Alberts, N. M. (2009). Operation-specific effects of numerical surface form on arithmetic strategy. *Journal of Experimental Psychology: Learning, Memory and Cognition, 35*, 999- 1011.

Campbell, J.I.D., & Clark, J.M. (1988). An encoding-complex view of cognitive number processing: Comment on McCloskey, Sokol & Goodman (1986). *Journal of Experimental Psychology: General, 117,* 204-214.

Campbell, J.I.D., & Clark, J.M. (1992). Cognitive number processing: an encoding complex perspective. In J. I. D. Campbell (Ed.). *The nature and origins of numerical skills* (pp. 457-492). Amsterdam: Elsevier.

Campbell, J. I. D., & Clark, J. M. (1989). Time course of error priming in number-fact retrieval: Evidence for excitatory and inhibitory mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 920-929.

Campbell, J. I. D., & Epp, L., J. (2005). Architectures for arithmetic. In J. I. D. Campbell (Ed.), *The handbook of mathematical cognition* (pp. 347-360). New York: Psychology Press.

Campbell, J. I. D., & Fugelsang, J. (2001). Strategy choice for arithmetic verification: Effects of numerical surface form. *Cognition, 80*, 21-30.

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

Campbell, J. I. D., Kanz, C. L., & Xue, Q. (1999). Number processing in Chinese-English bilinguals. *Mathematical Cognition, 5,* 1-39.

Campbell, J. I. D., Parker, H. R., & Doetzel, N. L. (2004). Interactive effects of numerical surface form and operand parity in cognitive arithmetic. *Journal of Experimental Psychology: Learning, Memory and Cognition, 30,* 51-64.

Campbell, J. I. D., & Penner-Wilger, M. (2006). Calculation latency: The mu of memory and the tau of transformation. *Memory & Cognition, 34,* 217-226.

Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130*, 299-31.

Cappelletti, M., Butterworth, B., & Kopelman, M. (2006). The understanding of Quantifiers in semantic dementia: a single case study. *Neurocase, 12*, 136-145.

Caramazza, A., & Shelton, R.S. (1998). Domain-specific knowledge systems in the brain: The animate–inanimate distinction. *Journal of Cognitive Neuroscience, 10*, 1-34.

Chochon, F., Cohen, L., van de Moortele, P. F., & Dehaene, S. (1999). Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience*, *11*, 617-630.

Church, R.M., & Meck, W.H. (1984). The numerical attribute of stimuli. In H.L. Roitblat, T.G. Bever & H.S. Terrace (Eds.), *Animal Cognition*. Hillsdale, NJ: Erlbaum.

Cipolotti, L., Butterworth, B., & Denes,G . (1991). A specific deficit for numbers in a case of dense acalculia. *Brain, 114,* 2619-2637.

Cipolotti, L., Warrington, E.K., & Butterworth, B. (1995). Selective impairment in manipulating Arabic numerals. *Cortex, 31*,73-86.

Cohen, L., & Dehaene, S. (1994). Amnesia for arithmetic facts: A single case study. *Brain and Language, 47,* 214-232.

Cohen, L., Dehaene, S., & Verstichel, P. (1994). Number words and number non-words: a case of deep dyslexia extending to Arabic numerals. *Brain, 117*, 267-269.

Cohen-Kadosh, R., Henik, A., & Rubinsten, O., Mohr, H., Dori, H., et al. (2005). Are numbers special? The comparison systems of the human brain investigated by fMRI. *Neuropsychologia, 43,* 1238-1248.

Cohen-Kadosh, R., Henik, A., & Rubinstein, A. (2007). The effect of orientation on number word processing. *Acta Psychologica, 124*, 370-381.

Cohen-Kadosh, R., Henik, A., & Rubinsten, O. (2008). Are arabic and verbal numbers processed in different ways? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1377-1391.

Coltheart, M. (2005). Modelling reading: the dual-route approach. In Snowling M.J., Hulme, C., (Eds). *The Science of Reading: A Handbook*. (pp. 6-23), Malden, MA: Blackwell.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology, 33A*, 497-505.

Cox, W. M., Fadardi, J. S., & Pothos, E. M. (2006). The addiction-Stroop test: Theoretical considerations and procedural recommendations. *Psychological Bulletin, 132,* 443-476.

Crutcher, R. J. (1994). Telling what we know: The use of verbal report methodologies in psychological research. *Psychological Science, 5,* 241-244.

Damian, M. F. (2004). Asymmetries in the processing of Arabic digits and number words. *Memory & Cognition 32*, 164-171.

Deacon, T. (1997). *The symbolic species.* London: Penguin.

Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Perception and Psychophysics, 45,* 557-566.

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

Dehaene, S. (1997). *The Number Sense*. Oxford: Oxford University Press.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1,* 83-120.

Dehaene S, Molko N, Cohen L, & Wilson A.J. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology, 14*, 218-224.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioural and brain-imaging evidence. *Science, 284*, 970-974.

Desroches, A. S., Joanisse, M. F., & Robertson, E. K. (2006). Specific phonological impairments in dyslexia revealed by eye-tracking. *Cognition, 100,* B32-B42.

De Corte, E., Verschaffel, L., & Pauwels, A. (1990). Influence of the semantic structure of word problems on second graders' eye movements. *Journal of Educational Psychology, 82,* 359-365.

Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and Language, 36,* 147-164.

Edwards, M. S., Burt, J. S., & Lipp, O. V. (2006). Selective processing of masked and unmasked verbal threat material in anxiety: Influence of an immediate acute stressor. *Cognition and Emotion, 20,* 812-835.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (rev. ed.). Cambridge, MA: Bradford/MIT Press.

Fias, W., Brysbaert, M., Geypens, F., & d'Ydewalle, G. (1996). The importance of magnitude information in numeric processing: Evidence from the SNARC-effect. *Mathematical Cognition*, *2*, 95-100.

Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., & Orban, G.A. (2003). Parietal representation of symbolic and nonsymbolic magnitude. *Journal of Cognitive Neuroscience, 15,* 47-56.

Fias, W., Reynvoet, B., & Brysbaert, M. (2001). Are Arabic numerals processed as pictures in a Stroop interference task? *Psychological Research, 65,* 242-249.

Flowers, J. H., Warner, J. L., & Polansky (1979). Response and encoding factors in 'ignoring' irrelevant information. *Memory and Cognition, 7,* 86-94.

Foltz, G. S., Poltrock, S. E., & Potts, G. R. (1984). Mental comparisons of size and magnitude. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 442-453.

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition, 108*, 819-824.

Frege (1974). The foundations of arithmetic: a logic-mathematical enquiry into the concept of number (2nd ed. Rev., J.L. Austin, Trans.). Oxford, England: Blackwell. (Original work published 1884).

Friedman, A. (1978). Memorial comparisons without the "mind's eye". *Journal of Verbal Learning and Verbal Behaviour, 17*, 427-444.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43-74.

Ganor-Stern, D., & Tzelgov, J. (2008). Across-notation automatic numerical processing. *Journal of Experimental Psychology: Learning, Memory,and Cognition, 34*, 430-437.

Geary, D. C., Frensch, P. A., & Wiley, J. G. (1993). Simple and complex mental subtraction: Strategy choice and speed-of-processing differences in younger and older adults. *Psychology and Aging, 8,* 242-256.

Gelman, R., & Butterworth, B. (2005). Number and language: How are they related? *Trends in Cognitive Sciences, 9*, 6-10.

Gelman, R., & Gallistel, C.R. (1978) *The Child's Understanding of Number.* Harvard University Press.

Gilmore, C.K., McCarthy, S.E., & Spelke, E. (2007). Symbolic arithmetic knowledge without instruction. *Nature, 447*, 589-591.

Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology*, *76*,104-22.

Gleitman, L., & Papafragou, A. (2005). Language and thought. In K. Holyoak & R. Heathcote, D. (1994). The role of visuo-spatial working memory in the mental addition of multi-digit addends. *Current Psychology*, *13*, 207-245.

Greenblatt, S.H. (1973). Alexia without agraphia or hemianopsia. Anatomical analysis of an autopsied case. *Brain, 96*, 307-316.

Holloway, I.D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: the distance effect and children's mathematical competence. *Journal of Experimental Child Psychology, 103,* 17-29.

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of Arithmetic Word Problems: A Comparison of Successful and Unsuccessful Problem Solvers. *Journal of Educational Psychology, 87*, 18-32.

Hauser, M. D., & Spelke, E. S. (2004). Evolutionary and developmental foundations of human knowledge: A case study of mathematics. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences, Vol. 3.* Cambridge: MIT Press.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory and Cognition, 10*, 389-395.

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language learning and development, 2*, 77-96.

Ille, N., Berg, P., & Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology, 19*, 113-124.

Inhoff, A. W. (1985). Phonological effects in the visual processing of words: Some methodological considerations. *Behavioral and Brain Sciences, 8*, 714-715.

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behaviour, 23,* 612-624.

Ischebeck, A. (2003). Differences between digit naming and number word reading in a flanker task. *Memory & Cognition 31*, 529-537.

Jukes, L., & Gilchrist, M. (2006). Concerns about numeracy skills of nursing students. *Nurse Education in Practice, 6,* 192–198.

Kadosh, R.C., Kadosh, K.C., Linden, D.E.J., Gevers, W., Berger, A., & Henik, A. (2007). The Brain Locus of Interaction between Number and Size: A Combined Functional Magnetic Resonance Imaging and Event-Related Potential Study. *Journal of Cognitive Neuroscience, 19,* 957-970.

Kashiwagi, A., Kashiwagi, T., & Hasegawa, T. (1987). Improvement of deficits in mnemonic rhyme for multiplication in Japanese aphasics. *Neuropsychologica, 25,* 443-447.

Kaufmann, L., Handl, P., & Thoeny, B. (2003). Evaluation of a numeracy intervention program focusing on basic numerical knowledge and conceptual knowledge. A pilot study. *Journal of Learning Disabilities*, *36*, 564-573.

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, & Cognition, 27,* 157-175.

Koechlin, E., Dehaene, S., & Mehler, J. (1997). Numerical transformations in five month-old human infants. *Mathematical Cognition, 3,* 89-104.

Landy, D. & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 720-733.

Lebiere, C, & Anderson, J. R. (1998). Cognitive arithmetic. In J. R. Anderson & C. Lebiere (Eds.), *The atomic components of thought* (pp. 297-342). Mahwah, NJ: Erlbaum.

LeFevre, J. (2000). Research on the development of academic skills: Introduction to the special issue on early literacy and early numeracy. *Canadian Journal of Experimental Psychology, 54*, 57-60.

LeFevre, J. A., Bisanz, J., Daley, K. E., Buffone, L., Greenham, S. L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General, 125,* 284-306.

LeFevre, J. A., DeStefano, D., Penner-Wilger, M., & Daley, K. E. (2006). Selection of procedures in mental subtraction. *Canadian Journal of Experimental Psychology, 60*, 209-220.

LeFevre, J. A., & Morris, J. (1999). More on the relation between division and multiplication in simple arithmetic: Evidence for mediation of division solutions via multiplication. *Memory & Cognition, 27*, 803-812.

Lewis, C., Hitch, G. J., & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- to 10-year-old boys and girls. *Journal of Child Psychology and Psychiatry, 35,* 283-292.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, *21*, 37-44.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*, 6-14.

Logan, G. (1985). Skill and automaticity: Relations, implications and future directions. *Canadian Journal of Psychology, 39***,** 367-386.

MacLeod, C. M. (1991). Half a Century of Research on the Stroop Effect: An Integrative View. *Psychological Bulletin, 109,* 163-203.

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14,* 126-135.

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences, 4,* 383-391.

Marangolo, P., Nasti, M., & Zorzi, M. (2004). Selective impairment for reading numbers and number words: A single case study. *Neuropsychologia, 42,* 997-1006.

Mauro, D. G., LeFevre, J. A., & Morris, J. (2003). Effects of problem format on division and multiplication performance: Division facts are mediated via multiplication-based representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 163-17.

Mazzocco, M.M.M. (2008). Introduction: Mathematics ability, performance and achievement. *Developmental Neuropsychology*, *33*, 197-204.

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

McCloskey, M. & Macaruso, P. (1995). Representing and using numerical information. *American Psychologist, 50*, 351-363.

McCloskey, M., Macaruso, P., & Whetstone, T. (1992). The functional architecture of numerical processing mechanisms: Defending the modular model. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 493-537). Elsevier: Amsterdam.

McMillan, C., Clark, R., Moore, P., Devita, C., & Grossman, M. (2005). Neural basis for generalized quantifier comprehension. *Neuropsychologia, 43*, 1729–1737.

McMillan, C., Clark, R., Moore, P., & Grossman, M. (2006). Quantifiers comprehension in corticobasal degeneration. *Brain and Cognition, 65*, 250-260.

McNeil, J., & Warrington, E. K. (1994). A dissociation between addition and subtraction with written calculation. *Neuropsychologica, 32,* 717-728.

Merkley, R., & Ansari, D. (2010). Using eye tracking to study numerical cognition: the case of the ration effect. *Experimental Brain Research, 206*, 455-460.

Moeller, K., Neuburger, S., Kaufmann, L., Landerl, K., & Nuerk, H.C. (2009). Basic number processing deficits in developmental dyscalculia Evidence from Eye-tracking. *Cognitive Development, 24,* 371-386.

Muroi, M., & MacLeod, C. M. (2004). Integration versus separation in Stroop-like counting interference tasks. *Japanese Psychological Research, 46,* 56-64.

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

Nieder, A., Diester, I., & Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. *Science, 313*, 1431-1435.

Noël, M. P., Fias, W., & Brysbaert, M. (1997). About the influence of the presentation format and arithmetical-fact retrieval process. *Cognition, 63,* 335-374.

Noël, M. P., & Seron, X. (1992). Notational constraints and number processing: A reappraisal of the Gonzalez and Kolers (1982) study. *The Quarterly Journal of Experimental Psychology, 45,* 451-478.

Ostad, S. A. (1998). Developmental differences in solving simple arithmetic word problems and simple number fact problems: A comparison of mathematically normal and mathematically disabled children. *Mathematical Cognition, 4*, 1-19.

Parkman, J. M. (1971). Temporal aspects of digit and letters inequality judgment. *Journal of Experimental Psychology, 33*, 131-162.

Paterson, S. J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome. *Journal of Child Psychology and Psychiatry, 47*, 190-204.

Pavese, A., & Umiltà, C. (1998). Symbolic Distance Between Numerosity and Identity Modulates Stroop Interference. *Journal of Experimental Psychology, 24,* 1535-1545.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17,* 407-413.

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size and luminance during comparative judgments. *Neuron, 41*, 938-993.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103(1),* 56-115.

Rayner, K., & Pollatsek, A. (1987). *Eye movements in reading: a tutorial review.* In K. Rayner, Eye movements in reading: perceptual and language processes (pp. 327-362). New York: Academic.

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *Quarterly Journal of Experimental Psychology, 55*, 1127-1139.

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

Rickard, T. C., Romero, S. G., Basso, G., Wharton, C., Flitman, S., & Grafman, J. (2000). The calculating brain: an fMRI study. *Neuropsychologia, 38,* 325-35.

Roelofs, A. (2006). Functional Architecture of Naming Dice, Digits and Number Words. *Language and Cognitive Processes, 21,,* 78-111.

Rouselle, L., & Noël (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs. non-symbolic number magnitude processing. *Cognition, 102*, 361-395.

Rubinsten, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their association with Arabic numerals. *Journal of Experimental Child Psychology, 81,* 74-92.

Saxe, G. B. (1981). Body parts as numerals: A developmental analysis of numeration among the Oksapmin in Papua New Guinea. *Child Development, 52,* 306-316.

Schiffrin, R. M. (1989). Attention. In R. C. Atkinson, R. J. Herrnestein, G. Lindzey, & R. D. Luce (Eds.), *Steven's handbook of experimental psychology* (2nd ed.). New York: Wiley.

Schmithorst, V. J., & Brown, R. D. (2004). Empirical validation of the triple-code model of numerical processing for complex math operations using functional MRI and group Independent Component Analysis of the mental addition and subtraction of fractions. *Nueroimage, 22,* 1414-1420.

Schwarz, W., & Heinze, H. J. (1998). On the interaction of numerical and size information in digit comparison: A behavioural and event-related potential study. *Neuropsychologia, 36*, 1167-1179.

Shor, R. E. (1971). Symbol processing speed differences and symbol interference effects in a variety of concept domains. *Journal of General Psychology, 85,* 187-205.

Simon, O., Mangin, J. F., Cohen, L., Le Bihan, D., & Dehaene, S. (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron*, *33*, 475-487.

Simon, T.J. (1997) Reconceptualizing the origins of number knowledge: a 'nonnumerical' account. *Cognitive Development, 12*, 349-372.

Simon, T.J., Hespos, S.J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development, 10,* 253-269.

Sokol, S. M., McCloskey, M., Cohen, N. J. & Aliminosa, D. (1991). Cognitive representations and processes in arithmetic: Inferences from the performance of brain-damaged patients. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17,* 355-376.

Spelke, E., & Dehaene, S. (1999). Biological foundations of numerical thinking. Response to T.J. Simon. *Trends in Cognitive Sciences, 3,* 365-366.

Stanescu-Cosson, R., Pinel, P., Van de Moortele, P.F., Le Bihan, D., Cohen, L., & Dehaene, S. (2000). Understanding dissociations in dyscalculia. A brain imaging study of the impact of number size on the cerebral network for exact and approximate calculation. *Brain*, *123,* 2240-55.

Starkey, P., & Cooper, R.G., Jr. (1980). Perception of numbers by human infants. *Science, 210,* 1033-1035.

Starkey, P., Spelke, E.S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36,* 97-128.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*(4), 869-918.

Stevenson, H. W., & Stigler, J. M. (1992). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education.* New York: Summit Books.

Strauss, M.S., & Curtis, L.E. (1981). Infant perception of numerosity. *Child Development, 52,* 1146-1152.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psyhcology, 18,* 643-662.

Suppes, P. (1990). Eye-movement models for arithmetic and reading performance. In E. Kowler (ed.), *Reviews* of *Oculomotor Research, Vol. IV. Eye Movements and Their Role in Visual and Cognitive Processes.* New York: Elsevier.

Szücs, D., & Csépe, V. (2004). Similarities and differences in the coding of numerical and alphabetical order using acoustic stimulation as revealed by event-related potentials in humans. *Neuroscience Letters.* 360 (1 – 2), 65-68.

Tang, J., Critchley, H. D., Glaser, D. E., Dolan, R. J., & Butterworth, B. (2006). Imaging Informational Conflict: An fMRI Study of Numerical Stroop. *Journal of Cognitive Neuroscience, 18,* 2049-62.

Troiani, V., Peelle, J.E., McMillan, C., Clark, R., & Grossman, M. (2009). Magnitude and parity as complementary attributes of quantifier statements. *Neuropsychologica, 47*, 2684-2685.

Tzelgov, J., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 166-179.

Udwin, O., Davies, M., & Howlin, P. (1996). A longitudinal study of cognitive abilities and educational attainment in Williams syndrome. *Developmental Medicine & Child Neurology, 38*(11), 1020-1029.

Verschaffel, L., De Corte, E., Gielen, I., & Struyf, E. (1994). Clever rearrangement strategies in children's mental arithmetic: A confrontation of eye movement data and verbal protocols. In J. E. H. van Luit (Ed.). *Research on Learning and Instruction of Mathematics in Kindergarten and Primary School* (pp. 153-180).

Verschaffel, L., De Corte, E., & Pauwels, A. (1992). Solving compare problems: An eye movement test of Lewis and Mayer's consistency hypothesis. *Journal of Educational Psychology, 84,* 85-94.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject–verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition, 61*, 261-298.

Wassmann, J., & Dasen, P.R. (1994) Yupno number system and counting. *Journal of Cross-Cultural Psychology, 25*, 78-94.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996): The emotional stroop task and psychopathology. *Psychological Bulletin, 120*, 3-24.

Wilson, T. D. (1994). The proper protocol: Validity and completeness of verbal reports. Psychological Science, 5, 249-252.

Windes, J. D. (1968). Reaction time for numerical coding and naming of numerals. *Journal of Experimental Psychology, 78(2),* 318-322.

Wynn, K. (1992a). Addition and subtraction by human infants. *Nature, 358*, 749-750.

Wynn, K. (1996). Infants' individuation and enumeration of sequential actions. *Psychological Science, 7,* 164-169.

Wynn, K., Bloom, P., & Chiang, W.C. (2002). Enumeration of collective entities by 5-month-old infants. *Cognition, 83*, 55-62.

Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds: The problem size effect. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 331-346). New York: Psychology Press.

Zhang, M., Si, J., Zhu, X., & Xu, X. (2010). Controversy on the essence of numerical surface form processing. *Journal of Mathematics Education, 3*, 158-166.

Zhou, X., Chen, C., Qiao, S., Chen, C., Chen, L. & Lu, N. et al. (2009). Event-related potentials for simple arithmetic in Arabic digits and Chinese number words: A study of the mental representation of arithmetic facts through notation and operation effects. *Brain Research, 1302*, 212-224.

Zhou, X., Chen, C., Zang, Y., Dong, Q., Chen, C., Qiao, S. et al. (2007). Dissociated brain organizations for single-digit addition and multiplication. *NeuroImage, 35,* 871-880.

Zhou, X., Chen, C., Zhang, H., Zhou, R., Zhao, H., Chen, C., et al. (2006). Event-related potentials of single-digit addition, subtraction, and multiplication. *Nueropsychologia, 44*, 2500-2507.

Zhou, X. (2011). Operation-specific encoding in single-digit arithmetic. *Brain and Cognition, 76,* 400-406.

Zorzi, M. & Butterworth, B. (1999). A computational model of number comparison. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the twenty first annual conference of the Cognitive Science Society* (pp. 778-783). Mahwah, NJ: Erlbaum.

Appendices

Appendix 1: Participant Consent Form

PARTICIPANT CONSENT FORM

Examining Cognitive Biases in Numerical Information Processing

Researcher
Justé Koller
[contact details]

Supervisor
Dr Fiona Lyddy
[contact details]

This study will be conducted by Justé Koller, a postgraduate student at the Department of Psychology, NUI Maynooth. The purpose of this research is to investigate processing differences for different numerical surface formats, such as digits (e.g. 3) and number words (e.g. 'three'). The study involves completing a number of simple computerised tasks.

All data will be kept entirely confidential and held on a secure computer. Data will be immediately coded and identifiable only by a participant code number.

The findings of the study may be published in the form of a research report. No individual responses will be reported and only group findings will be described.

Participation in the study is voluntary and you may refuse to participate or withdraw at any stage during the experiment. You may also withdraw your data from the study up until the report is published.

Any questions or concerns you may have about the study will be addressed by the experimenter.

Please note that the study does not involve any counselling or medical treatment and no form of medical diagnosis will be made.

By signing this consent form you indicate that you are 18 years of age or older and you have read the consent form and all your concerns have been addressed. You understand that you may withdraw from the study at any time, that you may withdraw your data and that all data will be kept confidential.

If during your participation in this study you feel the information and guidelines that you were given have been neglected or disregarded in any way, or if you are unhappy about the process please contact the Secretary of the National University of Ireland Maynooth Ethics Committee at research.ethics@nuim.ie. Please be assured that you concerns will be dealt with in a sensitive manner.

If the experiment caused you any discomfort or stress please contact Dr. Fiona Lyddy, head of the psychology department, NUI Maynooth.

____________________________ ____________________________

Participant Researcher

____

Date

Appendix 2: Numeracy Test (adapted from Lipkus et al., 2001)

Please attempt the following questions as accurately and as quickly as you can. The experimenter will contact you in a couple of minutes.

1) Imagine that we rolled a fair, six-sided die 1 000 times. Out of 1 000 rolls, how many times do you think the die would come up even (2, 4, or 6)?

2) Which of the following numbers represents the biggest risk of getting a disease?
____ 1 %
____ 10%
____ 5%

3) Which of the following numbers represents the biggest risk of getting a disease?
____ 1 in 100
____ 1 in 1000
____ 1 in 10

4) If Person A's risk of getting a disease is 1% in ten years, and person B's risk is double that of A's, what is B's risk?

5) If Person A's chance of getting a disease is 1 in 100 in ten years, and person B's risk is double that of A's , what is B's risk?

6) If the chance of getting a disease is 20 out of 100, this would be the same as having a ______% chance of getting the disease.

7) If the chance of getting a disease is 10%, how many people would be expected to get the disease:

A: out of 100?
B: out of 1000?

8) In a competition, the chance of winning a car is 1 in 1000. What percent of tickets win a car?

9) In the lottery, the chances of winning a € 10 000 prize is 1 %. How many people would win a € 10 000 prize if 1 000 people each buy a single lottery ticket?

10) The chance of getting a viral infection is .0005. Out of 10 000 people, about how many of them are expected to get infected?

11) If I get 6 hours sleep a night, what percentage of the week am I asleep?

12) If a product has been marked down from € 144 to € 132 in a sale, how much money is saved, presented as a fraction?

13) If I leave a waiter a € 4 tip for a € 25 bill, what percentage tip did I give?

14) Travel insurance for a holiday will cost me € 21.00 with company A. Travel insurance with company B is 25% cheaper than this company A price. How much will it cost me to take out insurance with company B?

15) At a restaurant the bill came to € 42. I gave the waiter € 50 as payment. After deducting a 10% tip, how much change will I have?

16) I buy a concert ticket online for € 49.00. This however does not include taxes and charges for mailing the ticket. If 2/7 of the original price goes to these charges, how much will I pay in the end?

Appendix 3: The Subjective Numeracy Scale (Fagerlin et al., 2007)

For each of the following questions, please check the box that best reflects **how good you are at doing the following things**:

1. How good are you at working with fractions?

☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6**

Not at all good **Extremely good**

2. How good are you at working with percentages?

☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6**

Not at all good **Extremely good**

3. How good are you at calculating a 15% tip?

☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6**

Not at all good **Extremely good**

4. How good are you at figuring out how much a shirt will cost if it is 25% off?

☐**1** ☐**2** ☐**3** ☐**4** ☐**5** ☐**6**

Not at all good **Extremely good**

For each of the following questions, please check the box that **best reflects your answer**:

1. When reading the newspaper, how helpful do you find tables and graphs that are parts of a story?

□**1** □**2** □**3** □**4** □**5** □**6**

Not at all helpful **Extremely helpful**

2. When people tell you the chance of something happening, do you prefer that they use words (e.g. "iit rarely happens") or numbers (e.g. " there is a 1% chance")? (1 = always prefer words; 6 = always prefer numbers)?

□**1** □**2** □**3** □**4** □**5** □**6**

Not at all good **Extremely good**

3. When people tell you the chance of something happening, do you prefer that they use words (e.g. "iit rarely happens") or numbers (e.g. " there is a 1% chance")?

□**1** □**2** □**3** □**4** □**5** □**6**

Always prefer words **Always prefer numbers**

4. How often do you find numerical information to be useful?

□**1** □**2** □**3** □**4** □**5** □**6**

Never **Very often**

Appendix 4: Eye Tracking Information Sheet

What is eye tracking?
People pay attention to some parts of a visual scene more than others. By examining the movements of their eyes we can tell which parts of the visual scene are of most interest and we can infer how people extract information from a visual scene. For example, when we present an equation on a computer screen, by examining which parts of the equation the eyes are drawn to, and for how long, we can infer how people are extracting information from the equation in their attempt to solve it.

What will I have to do during the experiment?
During the experiment, you will be seated in a comfortable chair with your chin on a chin-rest placed in front of a computer screen. The computer screen contains a small camera which records the eyes' movements. It is a reasonably comfortable procedure and it does not hurt or feel uncomfortable. The computer screen will show words or digits and you will respond by pressing a key on the keyboard as instructed. As it is important to keep relatively still, you will be also advised to refrain from blinks and head and body movements to facilitate accurate recording. You will receive full instructions before the start of the recording and have a trial run to familiarise you with the task. This experiment will investigate eye movements and duration of looking when certain numerical stimuli are presented. An eye tracker, a camera and computer apparatus that records eye movements will be used in this experiment. This research aims to inform us of the differences in the processing of different numerical stimuli. The apparatus will record your eye movements and the duration of your eyes' gaze as you look at numbers and words presented on a computer screen.

How long will the eye tracking sessions last?
The task and recording itself will usually last about 20-30 min. The whole experiment including a practice session will generally last no more than two hours.

Is the eye tracker safe?
This type of recording is considered completely safe and is non-invasive.

Are there any reasons why I should not participate?
Given the nature of the study, you should not participate if you have had a prior head injury or neurological illness, if you have epilepsy, or any difficulties with reading or vision. If you feel you may have a concern affecting whether you should participate please bring this to the attention of the researcher, who will be able to advise you.

Will I be rewarded for taking part?
There is no fee attached to participation in the study.

What if I change my mind during the study?
Your participation in the research is entirely voluntary. You have the right to withdraw from the study at any point, without having to give a reason and without your future study being affected in any way.

What will happen to the information from the study?
The recorded data and information for all the participants will be analysed and the results will be published in a postgraduate thesis and we hope in research literature in peer-reviewed journals. Any results about you personally will be held in the strictest confidence and not disclosed to anyone outside the project. The results will be described completely anonymously and no participant is named.

What if I have further questions?
Please do not hesitate to contact us at any time. You will be provided with full details of whom you may contact in relation to the study. We will be happy to answer any questions you may have.

Additional information:
Prior to participating in this study, you will be required to have read and fully understood this 'Eye Tracker Participant Information Sheet', and to complete and sign the attached 'Research Consent Form'. The recordings will take place in the eye tracker lab of the Psychology department (unless you are informed otherwise) at a time convenient for participants.

Department of Psychology, NUI Maynooth

Note: This Information Sheet is adapted from one on EEG recording from the Department of Psychology, University of Sheffield.

What are EEG and ERPs?
EEG (electroencephalogram) is a simple way of measuring the electrical activity of the brain or so-called 'brain waves'. Using it can tell us where, when and how the brain responds to a stimulus that we present. If it is recorded when a particular event occurs, then averaging a number of recordings will allow us to obtain the typical response of the brain related to the occurrence of this particular event. This is called the evoked (or event-related) potential (EP/ERP) technique.

What will the EEG recording involve?
This technique involves placing on the head a 'net' of small non-intrusive pads or electrodes. The Department uses a 32-channel cap of electrodes that are placed on your head by the experimenter. The pads are coated in a gel to improve the conductivity of brain signals. It is a reasonably comfortable procedure and it does not hurt or feel uncomfortable, although you will have to sit relatively still for about an hour while the pads are being attached in the correct places. Your hair will get slightly wet while applying the pads and you will be allowed to wash your hair in a sink in the Lab after the experiment. Once the pads are applied, you will be positioned in front of a computer screen. The experimenter will monitor the progress of the task in the adjacent room. You will always be able to communicate to the experimenter if you need to.

What will I have to do during EEG recording?
During recording, you will be seated in a comfortable chair and placed in front of a computer screen. The computer screen will show words or digits and you will respond by pressing a key on the keyboard as instructed. As it is important to keep relatively still, you will be also advised to refrain from blinks and head and body movements to facilitate accurate recording. You will receive full instructions before the start of the recording and have a trial run to familiarise you with the task. This experiment will investigate the changes in electrical activity in the brain when certain numerical stimuli are presented. It is expected that this research will help our understanding of the processing differences of different numerical formats such as digits (e.g. **3**) and words (e.g. **THREE**).

How long will the EEG sessions last?
The task and recording itself will usually last about 20-30 min. The whole experiment including the appliance of the pads and the practice session will generally last no more than two hours.

Is the EEG safe?
This type of recording is considered completely safe. It does not involve exposure to radiation; neither does it involve any injections. The pads attached to the scalp only record the ongoing activity of the brain. You will be in a normally illuminated room and able to speak to us throughout. The study can be stopped at any time if you wish. The type of EEG to be employed is in routine use in the Department of Psychology, NUI Maynooth, as it is in many other universities worldwide.

Will I be rewarded for taking part?
There is no fee attached to participation in the study.

Are there any reasons why I should not participate?
Given the nature of the study, you should not participate if you have had a prior head injury or neurological illness, if you have epilepsy, or any difficulties with reading or vision. If you feel you may have a concern affecting whether you should participate please bring this to the attention of the researcher, who will be able to advise you.

What if I change my mind during the study?
Your participation in the research is entirely voluntary. You have the right to withdraw from the study at any point, without having to give a reason and without your future study being affected in any way.

What will happen to the information from the study?
The EEG recorded data and information for all the participants will be analysed and the results will be published in a postgraduate thesis and we hope in research literature in peer-reviewed journals. Any results about you personally will be held in the strictest confidence and not disclosed to anyone outside the project. The results will be described completely anonymously and no participant is named.

What if I have further questions?
Please do not hesitate to contact us at any time. You will be provided with full details of whom you may contact in relation to the study. We will be happy to answer any questions you may have.

Additional information:
Prior to participating in this EEG study, you will be required to have read and fully understood this 'EEG Participant Information Sheet', and to complete and sign the attached 'Research Consent Form'. The EEG recordings will take place in the EEG lab of Psychology department at a time convenient for participants.

Department of Psychology, NUI Maynooth

Note: This Information Sheet was adapted from an information sheet on EEG of the Department of Psychology, University of Sheffield.