Dispositional Factors Affecting Children's Early Numerical Development

Sophie Batchelor

A thesis submitted for the degree of

Doctor of Philosophy

Loughborough University

September 2014

Abstract

Children show large individual differences in numerical skills, even before they begin formal education. These early differences have significant and long-lasting effects, with numerical knowledge before school predicting mathematical achievement throughout the primary and secondary school years. Currently, little is known about the dispositional factors influencing children's numerical development. Why do some children engage with and succeed in mathematics from an early age, whilst others avoid mathematics and struggle to acquire even basic symbolic number skills?

This thesis examines the role of two dispositional factors: First, spontaneous focusing on numerosity (SFON), a recently developed construct which refers to an individual's tendency to focus on the numerical aspects of their environment; and second, mathematics anxiety (MA), a phenomenon long recognised by educators and researchers but one which is relatively unexplored in young children. These factors are found to have independent effects on children's numerical skills, thus the empirical work is presented in two separate parts.

The SFON studies start by addressing methodological issues. It is shown that the current measures used to assess children's SFON vary in their psychometric properties and subsequently a new and reliable picture-based task is introduced. Next, the studies turn to theoretical questions, investigating the causes, consequences and mechanisms of SFON. The findings give rise to three main conclusions. First, children's SFON shows little influence from parental SFON and home numeracy factors. Second, high SFON children show a symbolic number advantage. Third, the relationship between SFON and arithmetic can be explained, in part, by individual differences in children's ability to map between nonsymbolic and symbolic representations of number.

The MA studies focus primarily on gender issues. The results reveal no significant differences between boys' and girls' overall levels of MA; however, there are gender differences in the correlates of MA. Specifically, boys' (but not girls') MA is related to parents' MA. Moreover, the relationship between MA and mathematical outcomes is stronger for boys than it is for girls. Possible causal explanations for these gender differences are explored in two ways: First, by examining the reliability of the scales used to assess MA in boys and girls. Second, by investigating the relationship between girls' (and boys') mathematics anxiety and their societal math–gender stereotypes.

The findings from both sets of studies draw a link between children's emerging dispositions towards mathematics and their early numerical skills. Future research needs to examine how these dispositional factors interact with other (cognitive and non-cognitive) predictors of mathematics achievement.

Contents

11 Discussion of Mathematics Anxiety Studies   **216**

Part IV   General Discussion

12 Conclusions   **225**

Part V   References & Appendices

List of Figures

List of Tables

Acknowledgements

I would like to start by thanking my supervisors Camilla Gilmore and Matthew Inglis for their support, guidance and inspiration. Throughout my PhD they have been there with encouragement and provided me with opportunities, whilst giving me the freedom to develop as an independent researcher. I could not have asked for more from either of them.

Next, I would like to thank the members of the Midlands Mathematical Cognition Group. I am very grateful for the many suggestions, discussions and cakes(!) over the years. Also, thank you to the staff and students of the Mathematics Education Centre at Loughborough University for providing such a positive and friendly place to work.

Very special thanks to the 500+ children who took part in my studies, together with their schools/preschools and families. Each and every child made this research possible. A big thank you to Annie Wilkinson and the Foundation team at Round Hill Primary School. Their enthusiasm and continued support has been invaluable. Also, many thanks to my colleagues at the University of Nottingham for giving me the opportunity to run studies each year at the 'Summer Scientist Week' event.

Finally, I would like to thank my family and friends for their love and understanding. I would not have got this far without them. Special thanks to my mum, Annette, and my brothers, Jamie and Robbie, for believing in me and helping in so many ways. Also, thank you to my gran, Phyllis, who sadly was unable to see the completion of this thesis, but whose confidence in me I will always remember.

Declaration

I, the author, declare that the work presented in this thesis is my own and has not been submitted for a degree at any other institution. None of the work has previously been published in this form.

Part I

General Introduction

Chapter 1

Introduction

The aim of this first chapter is to introduce the overall topic and purpose of this research. I start by presenting the rationale for investigating children's early numerical development. Next, I review the literature on the development of symbolic number skills and, in particular, the factors that give rise to individual differences in numerical abilities. Finally, I provide an overview of the scope and structure of the current thesis.

1.1 Statement of the Problem

Understanding how numbers work is an essential life skill. It is important not only for the individual negotiating life's daily demands, but for modern society as a whole. In recent years the need for a numerate population has become increasingly recognised (Walport, 2010). Several reports have shown that poor numeracy is associated with a range of negative outcomes, from low income and unemployment, to higher rates of physical illness, depression and delinquency (e.g. Bynner & Parsons, 2006). As such, numeracy has become a national and international priority. It has risen to the forefront of economic policy as well as educational programs in many countries worldwide.

Here in the UK, in particular, there have been growing concerns about numeracy standards. The latest worldwide education assessments, such as the 'Programme for International Student Assessment' (PISA), have highlighted persistent problems of underachievement in UK students' mathematics (National Numeracy, 2014). Moreover, government-commissioned

inquiries have estimated that roughly a quarter of adults are "functionally innumerate" (Vorderman, Budd, Dunne, Hart, & Porkess, 2011, p.3). This means that many students leave school with only a basic level of arithmetic. They may struggle with important everyday activities such as managing personal finances and interpreting data.

There have been a string of government reports and policies aimed at improving numeracy levels in the UK (Marsh, 2011; Norris, 2012; Smith, 2004; Vorderman et al., 2011). Perhaps most important is the shift in attention to numeracy in the early years (Munn, 2006; Williams, 2008). This follows recent findings that gaps in numeracy begin to emerge as early as preschool (e.g. Ginsberg, Lee, & Boyd, 2008), and that numerical knowledge before school predicts mathematical achievement throughout the primary and secondary school years (Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Baroody, 2000; Duncan et al., 2007; Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; Melhuish et al., 2008b; Reeve, Reynolds, Humberstone, & Butterworth, 2012; Sammons et al., 2012).

Research suggests that early interventions are more effective and economically efficient (Heckman, 2004; Heckman & Masterov, 2007). As argued by one of the world's leading economists, James Heckman, the earlier we invest in children, the greater our returns in education, health and productivity. In a paper published in the Wall Street Journal back in 2004, Heckman and Wax made the following influential statement:

> "Like it or not, the most important mental and behavioural patterns, once established, are difficult to change once children enter school" (p. 14).

Many similar claims have also been made:

> "If the race is already halfway run even before children begin school, then we clearly need to examine what happens in the earliest years" (Esping-Andersen, 2004, p. 133).

These statements help to underline the urgency of early years research. Note that traditionally, the emphasis in preschool education has been on literacy (Davis, 2009). With a wealth of studies showing that early linguistic input is essential for later reading development (see Galloway & Richards,

1994, for a review), educators and policy-makers have tended to focus on promoting preschool children's need for language-rich environments (e.g. Raban & Ure, 2000). Research into early numeracy, whilst it may have a history just as long, is much less comprehensive than its literacy counterpart (LeFevre, 2000). It is only now that we see it coming to the forefront.

To summarise, recent years have seen widespread efforts to raise numeracy standards across many Western societies. The importance of early numerical experiences have started to be acknowledged and preschool numeracy practices have become a major part of educational reform. This shift to the early years marks significant progression; however, there is much that needs to be done, particularly in terms of research. As it stands, it is not clear what type of interventions will be most effective at supporting children's early numerical development. Crucially, there is a need for further investigations into the factors that give rise to individual differences in symbolic number skills. The current thesis seeks to address this need.

1.2   Review of the Literature

The following review comprises three main sections. The first section provides some brief historical perspectives. The second section highlights and defines some key concepts within the literature on children's symbolic number development. Finally, the third section discusses the current state of research on individual differences in early number skills.

1.2.1   Historical Context

In 1925, in one of the first publications on children's number development, Harl Douglass asked a question that went on to intrigue psychologists for many years:

> "Of what numbers from 1 to 10 does the kindergarten child (from ages four and one half years to six years) have reasonably adequate concepts? In what degree does he have concepts of these numbers?" (p. 445).

Traditional theories of number learning granted children very little in terms of conceptual understanding. Early counting behaviours were widely

studied to begin with (Buckingham & MacLatchy, 1930; Douglass, 1925; Freeman, 1912) but they became increasingly dismissed. They were thought to have no operational value (Sarama & Clements, 2009). In one of the most influential theories of learning, Piaget (1952) argued that verbal and object counting were simply skills learnt by rote; skills which had little impact on constructing a concept of number. He believed that a concept of number was built up from the development of logic and reasoning, emerging at around 6 years of age.

Interestingly, later research painted a markedly different picture. The seminal work of Gelman and Gallistel (1978) sparked a resurgence of interest in children's early counting skills. In contrast to Piaget, Gelman and Gallistel saw counting as a fundamental part of children's natural number concept. They argued that any child who could count in accordance with a set of five "counting principles" could represent number. (Note that these counting principles, described in Section 1.2.2.4, have been widely adopted.) In their neo-nativist account, they further maintained that children hold an innate understanding of these counting principles.

In the decades that followed Gelman and Gallistel's (1978) influential book, *The Child's Understanding of Number*, a series of theoretical debates ensued. It was generally agreed that Piaget had underestimated young children's numerical competencies (and overlooked the importance of early counting behaviour); however, the notion of a preverbal innate counting system was contested.

Numerous studies were conducted to explore children's acquisition of the counting principles. These studies used a range of experimental methods, testing not only children's counting production skills, but their ability to detect errors in the counting routines of others. Broadly speaking, two theoretical camps emerged. There were those who believed that the counting principles came first, guiding the development of counting skills (e.g. Gelman & Meck, 1983; Gelman, Meck, & Merkin, 1986; Greeno, Riley, & Gelman, 1984), and those who believed that the counting principles came later, acquired in conjunction with, or after the counting routine (e.g. Briars & Siegler, 1984; Frye, Braisby, Lowe, Maroudas, & Nicholls, 1989; Fuson, 1988; Fuson & Hall, 1983).

As these epistemological discussions continued (Gallistel, 2007; Gallistel & Gelman, 1992; Le Corre & Carey, 2007, 2008; Le Corre, Van de Walle,

Brannon, & Carey, 2006), many scholars changed focus. In 2002, in an award winning article in Psychological Science, Newcombe aptly stated:

> "Rather than endlessly replaying the empiricist-nativist debate, researchers need to get on with the detailed work of proposing exactly how starting points in infancy–stronger than those postulated by Piaget–are transformed into mature competence– perhaps not quite in the way Piaget imagined, but nonetheless in generally interactional ways" (p. 400).

In line with Newcombe's (2002) suggestion, recent research has centered less on where children's counting skills originate from, and more on how these early counting skills develop into more advanced numerical competencies.

1.2.2   The Development of Symbolic Number Skills

1.2.2.1   Symbolic number representations

As children acquire knowledge of the counting sequence, they learn to use external symbols, such as number words and Arabic digits, to represent number. These external symbols, or symbolic number representations, enable exact number comparison and manipulation. They from the basis of formal arithmetic, allowing us to perform all kinds of precise mathematical operations.

Studies have shown that before children acquire this symbolic number system, they are able to represent quantities nonverbally or nonsymbolically. There is evidence for two nonsymbolic number systems: (i) an approximate number system which generates noisy representations of large numerical sets, and (ii) a parallel individuation system which generates precise representations of individual elements in small arrays (Feigenson, 2005; Feigenson, Carey, & Hauser, 2002; Feigenson, Dehaene, & Spelke, 2004; Pica, Lemer, Izard, & Dehaene, 2004).

1.2.2.2   Nonsymbolic number representations

In 1980, Starkey and Cooper suggested that infants were capable of representing small quantities nonsymbolically. Using a habituation paradigm they demonstrated that $4 - 6$ month old infants could discriminate between small arrays of dot stimuli. Infants were repeatedly exposed, or habituated,

to a given number of dots (e.g. 2). Then, after some time, they were presented with a different number of dots (e.g. 3). Interestingly, the infants looked longer at the novel array of dots (in this case 3), and thus they were assumed to have detected the change in numerosity.

This experiment sparked several investigations into infants' capacity to represent number. Researchers questioned whether infants were attending to numerosity, or to other perceptual aspects of the arrays, e.g. shape, size or density (Clearfield & Mix, 1999; Feigenson, Carey, & Spelke, 2002; Mix, Levine, & Huttenlocher, 1997). In attempts to tease apart these possibilities a variety of controls were employed. Strauss and Curtis (1981) used coloured photographs of objects that varied in size and alignments so that their only constant was their number. Other researchers used auditory tones (Lipton & Spelke, 2003, 2004) and moving displays of objects (Van Loosbroek & Smitsman, 1992; Wynn, Bloom, & Chiang, 2002). The consensus from these studies was that infants do indeed make discriminations based on numerosity.

As this area of research grew, it was demonstrated that infants form nonsymbolic representations of large numerical sets (greater than 4) as well as small numerical sets (less than 4). These representations were found to be approximate. The accuracy of discrimination depended on the ratio between the numerosities being compared, rather than the absolute set sizes (Feigenson et al., 2004). Lipton and Spelke (2003), for example, showed that 6 month old infants could discriminate between 4 vs. 8 and 8 vs. 16 elements, but not 4 vs. 6 or 8 vs. 12. In a subsequent experiment they found that older infants could discriminate between numerosities with smaller ratios, thus suggesting that the acuity of these nonsymbolic representations increases with age (Lipton & Spelke, 2004).

In further extensions of this work, researchers showed that infants can not only represent number nonsymbolically, but they can manipulate these representations too. Using the habituation paradigm, Wynn (1992a) demonstrated that 5 month old infants could perform basic arithmetic computations. They were able to add and subtract small quantities nonsymbolically, e.g. $1 + 1$ or $2 - 1$ (see also, Koechlin, Dehaene, & Mehler, 1997). Follow-up studies yielded similar findings with large quantities (e.g. McCrink & Wynn, 2004).

Overall, research suggests that before children learn to count, they can

represent number nonsymbolically. Infants have been shown to compare, add and subtract both small and large numerical sets presented in multiple sensory modalities. In view of these nonsymbolic abilities, children are often described as having an intuitive understanding of quantities, or 'number sense' (Dehaene, 1992, 2001). A question that naturally follows is, what role do nonsymbolic representations play in the development of formal symbolic number knowledge?

1.2.2.3   Mapping between symbolic and nonsymbolic numerical representations

Research suggests that when children learn to represent numbers symbolically, the nonsymbolic number system is not overridden. Instead, nonsymbolic and symbolic representations become mapped onto one another. There are at least three pieces of evidence to support this suggestion. Firstly, when children and adults are asked to compare two symbolic numbers (e.g. Arabic digits) their performance is slower and less accurate when the symbols are numerically close together. It takes longer to compare 7 vs. 9 than it does to compare 5 vs. 9. This distance effect is a robust effect which shows that symbolic numbers automatically activate approximate (nonsymbolic) representations (Moyer & Landauer, 1967; Dehaene, Dupoux, & Mehler, 1990; Temple & Posner, 1998).

Secondly, evidence for these mappings comes from the fact that children and adults can perform rapid approximate calculations on Arabic digits (Gilmore, McCarthy, & Spelke, 2007; Xenidou-Dervou, De Smedt, van der Schoot, & van Lieshout, 2013). According to neuroscientific findings, this approximate symbolic arithmetic activates a separate neural system to that which is engaged during exact symbolic calculations (Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999). Thirdly, there is evidence from neuropsychological patients to support a link between the nonsymbolic and symbolic number systems (Dehaene, Dehaene-Lambertz, & Cohen, 1998). Individuals with impaired nonsymbolic processing show impaired symbolic processing, and vice versa.

Whilst it is widely agreed that the nonsymbolic and symbolic systems become mapped onto one another, it is not clear whether children harness their nonsymbolic abilities to master the symbolic system. Researchers disagree over whether the formation of these mappings is part of the acquisition

of the counting principles (Carey, 2004; Gallistel & Gelman, 2000; Le Corre & Carey, 2007; Lipton & Spelke, 2005).

1.2.2.4   The counting principles

The development of symbolic number knowledge is typically a long and arduous process. Learning the number sequence by rote tends to happen very early on –children often begin counting around their second birthday– but it takes years for a child to grasp the true meaning of these words. Gelman and Gallistel (1978) identified five principles that must be adhered to in order to count successfully:

1. The one-to-one principle: Each word in our count list must be assigned to one, and only one, of the items to be counted. No word must be used more than once and all items must be counted.

2. The stable-order principle: Counting involves more than the ability to assign arbitrary tags to items in an array. Each word in our count list must be recited in the same order. An array of 3 items must be counted as "1, 2, 3" as opposed to "2, 3, 1" or "3, 2, 1".

3. The cardinal principle: This refers to our understanding that the last word in our count list is special; it not only tags the final item, signalling the end of the count, but also tells us how many objects have been counted. It represents the set as a whole, indicating what we call the numerosity of the array.

4. The abstraction principle: The realisation that any array of discrete items can be counted. The principles listed above can be applied to all kinds of arrays (e.g. heterogeneous vs. homogeneous and tangible vs. intangible).

5. The order-irrelevance principle: The items in an array can be counted in any order. It doesn't matter whether we count from left to right, from right to left or from somewhere in the middle, providing that all items in the array are counted once and only once.

These counting principles have been widely acknowledged. However, as noted in Section 1.2.1, there has been considerable debate surrounding how

and when they are acquired. The details of these debates are beyond the scope of this thesis but are ongoing and well-documented in the literature (e.g. Gallistel, 2007; Le Corre & Carey, 2007, 2008).

1.2.2.5   Summary

The aim of this section was to introduce some of the key concepts within research on children's symbolic number development. This should provide the necessary background for discussions that follow later. The concepts of symbolic representations, nonsymbolic representations and the mapping between nonsymbolic and symbolic representations are particularly relevant to Studies 4 and 5 presented in Chapters 5 and 8. The counting principles are not discussed in any greater depth but note that the concept of cardinality (the cardinal principle) will be referred to within the studies that look at children's counting skills (Studies 1 and 2 presented in Chapters 2 and 3).

1.2.3   Individual Differences in Symbolic Number Skills

As noted in Section 1.1, individual differences in numeracy start to emerge before children begin formal education. Whilst many children enter school with a wide range of numerical skills –from counting, matching and ordering sets, to basic calculations such as adding and subtracting small numbers– others have yet to grasp the meanings of the words in their count list and the concept of cardinality (Klibanoff et al., 2006). These individual differences appear to have long-lasting consequences. Studies have shown that the relationship between early numerical knowledge and later mathematical achievement is remarkably strong (Aunola et al., 2004; Baroody, 2000; Klibanoff et al., 2006; Melhuish et al., 2008b; Reeve et al., 2012; Sammons et al., 2012), roughly twice as strong as that between early and later reading achievement (Duncan et al., 2007).

Researchers in psychology and mathematics education have become increasingly concerned with understanding where these individual differences come from. Thus far, four factors have been highlighted: (i) domain-general cognitive factors, (ii) domain-specific cognitive factors, (iii) dispositional factors, and (iv) environmental factors. An overview of research relating to each of these factors is provided below.

1.2.3.1   Domain-general cognitive factors

Domain-general factors are those cognitive skills which predict achievement in many subject areas. They include general intelligence (or IQ) and executive functions such as working memory, planning and inhibitory control. Executive functions can been defined as the "online" processes which enable one to self-regulate complex cognitive activities (Clark, Pritchard, & Woodward, 2010).

A number of studies have documented the role of these general skills in children's early mathematical development (Bull, Espy, & Wiebe, 2008; Clark et al., 2010; Epsy et al., 2004; Kroesbergen, Van Luit, Van Lieshout, Van Loosbroek, & Van de Rijt, 2009; Monette, Bigras, & Guay, 2011; Verdine, Irwin, Golinkoff, & Hirsh-Pasek, 2014). Clark et al. (2010) questioned whether preschool executive function abilities predict early mathematics achievement in school. In their longitudinal study they followed 104 children from age 2 to 6 years. At age 4 children were given a battery of executive function tasks which measured planning, shifting[1] and inhibitory control[2]. They also completed the Wechsler Preschool and Primary Scale of Intelligence (WPPSI). At age 6 children received a similar battery of tasks with the addition of the Woodcock-Johnson Mathematics Fluency and Passage Comprehension tests. The findings revealed that children's performance on all three measures of executive functioning at age 4 was positively related to their performance on the Woodcock-Johnson Mathematics Fluency test at age 6. These relationships remained significant even after controlling for individual differences in general intelligence and reading achievement.

It is important to note that in Clark et al.'s (2010) study, children's numerical competence was assessed via a single mathematics fluency task which measured speeded arithmetic only. Mathematics fluency is not synonymous with untimed arithmetic (Hart, Petrill, & Thompson, 2010), therefore follow-up work has investigated the relationship between domain-general skills and other components of mathematics. Passolunghi and Lanfranchi (2012), for example, conducted a similar longitudinal study in which children

---

[1]The ability to disengage from an irrelevant task set or strategy and subsequently activate a more appropriate one (Bull et al., 2008).

[2]The ability to suppress dominant action tendencies in favor of more goal-appropriate behaviours (Bull et al., 2008).

completed a wider range of mathematics tasks[3] from verbal counting and numerical comparison, to more standardised tests of logic, arithmetic and geometry. Children also completed general measures of intelligence, processing speed, verbal and visual-spatial short-term memory, working memory and phonological abilities. Passolunghi and Lanfranchi found that working memory and processing speed were directly related to numerical skills in kindergarten children. By the end of the children's first year of school the influence of working memory was mediated by early numerical competence.

These results add to the literature on domain-general skills in children with mathematical learning difficulties (D'Amico & Guarnela, 2005; Geary, Hamson, & Hoard, 2000; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; McLean & Hitch, 1999; Passolunghi & Siegel, 2001; Toll, Van der Ven, Kroesbergen, & Van Luit, 2011). Research has demonstrated that children who experience difficulties in mathematics show impairments in executive function, and in particular, visual-spatial processing (e.g. Geary et al., 2000). The link between visual-spatial processing and numerical skills has also been emphasised in studies with typically developing children (Alloway & Passolunghi, 2011; Cheng & Mix, 2014; Kyttala, Aunio, Lehto, Van Luit, & Hautamaki, 2003; Raghubar, Barnes, & Hecht, 2010). Kyttala et al. (2003), for example, found that visual-spatial working memory was related to preschool children's early counting ability.

Overall, there is a large body of cross-sectional and longitudinal research showing an association between children's general cognitive abilities and their early numerical development. This suggests that early numerical instruction may benefit from focusing not only on the acquisition of numerical knowledge, but on foundational competencies such as executive function skills.

1.2.3.2 Domain-specific cognitive factors

The most widely studied domain-specific cognitive predictors of children's numerical skills have been nonsymbolic and symbolic numerical magnitude representations (see Section 1.2.2). Studies have demonstrated that children (and adults) show variations in the precision of these magnitude represen-

---

[3]Both speeded and unspeeded.

tations and thus researchers have questioned whether they are related to individual differences in formal mathematics. Current data on this topic is vast (see De Smedt, Noël, Gilmore, & Ansari, 2013 for a review).

Halberda and colleagues have shown that the precision of children's nonsymbolic representations is positively related to their current and prior mathematics performance (Halberda, Mazzocco, & Feigenson, 2008; Libertus, Feigenson, & Halberda, 2011). Halberda et al. (2008) used a nonsymbolic dot comparison task to measure the acuity of children's nonsymbolic numerical magnitude representations at 14 years. They found that this nonsymbolic acuity was related to previous mathematics performance measured between $5 - 11$ years. In a later study, Libertus et al. (2011) found that nonsymbolic comparison performance was related to concurrent mathematical ability in preschoolers, even after controlling for age and verbal skills.

These findings are compelling, but they are by no means conclusive. Whilst some studies have replicated these effects in children and adults (Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Libertus, Odic, & Halberda, 2012; Lourenco, Bonny, Fernandez, & Rao, 2012), others have failed to find a relationship in either population (Castronovo & Göbel, 2012; Iuculano, Tang, Hall, & Butterworth, 2008; Price, Palmer, Battista, & Ansari, 2012; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013; Sasanguie, Van den Bussche, & Reynvoet, 2012; Vanbinst, Ghesquiere, & De Smedt, 2012). Interestingly, there is evidence to suggest that children's ability to integrate (or map between) nonsymbolic and symbolic number representations may be more important for formal mathematics than nonsymbolic abilities per se (e.g. Holloway & Ansari, 2009; Mundy & Gilmore, 2009).

Holloway and Ansari (2009) found a relationship between children's symbolic (but not nonsymbolic) numerical comparison performance and their performance on the Mathematics Fluency and Calculation subtests of the Woodcock Johnson. Similarly, De Smedt, Verschaffel, and Ghesquière (2009) demonstrated that children's performance on a symbolic number comparison task (completed at the beginning of formal schooling) predicted their symbolic number skills one year later. In another study, Mundy and Gilmore (2009) developed a novel mapping task in which children had to choose which of two nonsymbolic numerosities matched a symbolic numerosity (and vice versa). They found that children's performance on this mapping task was related to their achievement on tests of school mathematics, over and above the

variance explained by symbolic and nonsymbolic comparison performance.

In sum, research suggests that the precision of children's numerical magnitude representations and in particular, the mapping between nonsymbolic and symbolic representations, may play an important role in the acquisition of formal symbolic number knowledge. Given these findings, one way to support children's numerical development may be to provide games and activities that help them to make the connections between different representations of number (i.e. number words, Arabic digits and nonsymbolic numerosities). The effects of some initial interventions have been small, but promising. Whyte and Bull (2008), for example, found that number card games involving nonsymbolic magnitude judgments (and mapping between nonsymbolic and symbolic numerosities) led to some improvements in preschool children's basic number skills.

1.2.3.3 Dispositional factors

Dispositional factors can be defined as the attitudes, beliefs and motivations that a child brings with them to the mathematics classroom (Buckley & Reid, 2013). As emphasised by Mazzocco, Hanich, and Noeder (2012), "a productive disposition towards mathematics is an essential component of mathematics proficiency" (p. 1). It is thus important to investigate the emergence of these dispositions, and the role that they play in the early school years.

One dispositional factor to consider is spontaneous focusing on numerosity (SFON). This is a recently developed construct which captures an individual's tendency to focus on the numerical aspects of their environment (Hannula, 2005; Hannula & Lehtinen, 2005; Hannula, Lepola, & Lehtinen, 2010; Hannula, Mattinen, & Lehtinen, 2005; Hannula, Räsänen, & Lehtinen, 2007; Hannula-Sormunen, 2014). The term "spontaneous" is used to refer to the fact that the process of "focusing attention on numerosity" is self-initiated or non-guided. That is, attention is not explicitly guided towards the aspect of number or the process of enumeration. The idea is that "SFON tendency indicates the amount of a child's spontaneous practice in using exact enumeration in her or his natural surroundings" (Hannula et al., 2010, p. 395).

Hannula and colleagues have demonstrated that preschool children show individual differences in SFON, and that these individual differences are re-

lated to current counting skills (Hannula et al., 2007) and later arithmetical success (Hannula et al., 2010). They suggest that SFON is an important sub-process of enumeration, one which gives children more practice using their exact symbolic number skills (Hannula & Lehtinen, 2005). Findings from a recent study highlight that children with mathematical learning difficulties, such as developmental dyscalculia, may show reduced SFON tendency compared to their typically developing peers (Kucian et al., 2012).

Another dispositional factor that may affect children's numerical development is mathematics anxiety. Mathematics anxiety refers to the syndrome of negative emotions that many individuals experience when engaging in tasks demanding numerical or mathematical skills. In contrast to SFON, which is a new construct, the notion of mathematics anxiety dates back many years (Dreger & Aiken Jr, 1957). Lang (1968) asserted that mathematics anxiety is a phobia which, like any other phobia, influences individuals on three dimensions: (i) physiological, e.g. increased heart rate, (ii) cognitive, e.g. worrisome thoughts, and (iii) behavioural, e.g. avoiding mathematics classes. Other researchers have since emphasised that mathematics anxiety comprises a range of emotional responses, from mild states of apprehension to intense feelings of uneasiness and distress (e.g. Richardson & Suinn, 1972). Importantly, mathematics anxiety is related to, but nonetheless distinct from, other forms of anxiety such as test anxiety (e.g. Dew & Galassi, 1983).

Numerous studies have shown that mathematics anxiety is negatively associated with performance on a range of mathematical tasks (see, Hembree, 1990 and Ma, 1999, for reviews). Up until recently, this work focused on older children and adults. It was traditionally believed that mathematics anxiety arose in the context of complex mathematics, surfacing at the end of elementary school when more advanced mathematical concepts were introduced (Maloney & Beilock, 2012). In this tradition, the most dominant account to explain the relationship between mathematics anxiety and mathematics performance was the "working memory disruption" hypothesis (Ashcraft & Kirk, 2001; Ashcraft & Krause, 2007; Faust, 1996; Hopko, Ashcraft, Gute, Ruggiero, & Lewis, 1998). Ashcraft and colleagues proposed that mathematics anxiety results in worrisome thoughts and ruminations which consume vital working memory resources needed to solve complex mathematical problems.

Two recent strands of research have challenged this traditional view. Firstly, studies conducted with adults have shown that mathematics anxiety is associated with impairments in basic, low-level numerical processing as well as higher-level mathematics (Maloney, Ansari, & Fugelsang, 2011; Maloney, Risko, Ansari, & Fugelsang, 2010; Nùñez-Peña & Suàrez-Pellicioni, 2014). Maloney et al. (2011), for example, gave high mathematics-anxious and low mathematics-anxious adults a symbolic numerical comparison task, typically used to measure the precision of an individual's magnitude representations. They found that high mathematics-anxious individuals showed a larger numerical distance effect than their low mathematics-anxious peers. This suggests that their numerical magnitude representations (and in particular the mapping between nonsymbolic and symbolic representations) were less precise.

In addition to these findings, studies have demonstrated that mathematics anxiety is present in children much younger than previously thought, i.e. children who are in the early stages of formal schooling (Jameson, 2013; Krinzinger, Kaufmann, & Willmes, 2009; Ramirez, Gunderson, Levine, & Beilock, 2013; Thomas & Dowker, 2000; Wu, Barth, Amin, Malcarne, & Menon, 2012; Young, Wu, & Menon, 2012). In a neuroimaging study, Young et al. (2012) showed that this anxiety is associated with a distinct pattern of neural activity. High mathematics-anxious children (aged between $7-9$ years) showed hyperactivity in the right amygdala, a brain region that is important for the processing of negative emotions. In comparison to their low mathematics-anxious peers they also exhibited reduced activity in two brain regions (the posterior parietal cortex and the dorsolateral prefrontal cortex) that have been implicated in mathematical reasoning.

Importantly, the recent construction of some developmentally-appropriate mathematics anxiety scales[4] has allowed researchers to start exploring the effects of mathematics anxiety on younger children's numerical development. Thus far, the findings from some initial studies have been mixed (Eden,

---

[4]Thomas and Dowker (2000) developed the Mathematics Anxiety Questionnaire (MAQ) to assess mathematics anxiety in children aged $6-9$ years. Other recently developed scales include: Wu et al.'s (2012) Scale for Early Mathematics Anxiety (SEMA) for children aged $7-9$ years; Jameson's (2013) Children's Anxiety of Math Scale (CAMS) for children aged $6-9$ years; and Ramirez et al.'s (2013) Child Mathematics Anxiety Questionnaire (CMAQ) for children aged $5-9$ years.

Heine, & Jacobs, 2013). Further investigations are needed to help us understand not just the consequences but the causes of mathematics anxiety in the early years. There is growing evidence to suggest that both teacher and parental attitudes may be particularly important (Gunderson, Ramirez, Levine, & Beilock, 2012). However, note that there are inconsistencies in this literature, particularly regarding the role of gender (e.g. Beilock, Gunderson, Ramirez, & Levine, 2010).

To summarise, researchers have highlighted the role of dispositional factors on students' mathematical outcomes, but little is known about the development (and effects) of these dispositions in early childhood. There are two key factors that need to be investigated further. These are SFON and mathematics anxiety. SFON is a recently developed construct (Hannula & Lehtinen, 2005), which, over the course of the last year, has become more widely recognised (Bojorque, Torbeyns, Hannula-Sormunen, Van Nijlen, & Verschaffel, 2014; Bull, 2013; Gray & Reeve, 2013; Poltz, Wyschkon, Hannula-Sormunen, von Aster, & Esser, 2014; Torbeyns, Rathé, & Verschaffel, 2014). There are several avenues for further research both methodologically (e.g. developing reliable and valid measures of SFON) and theoretically (e.g. investigating the developmental roots and mechanisms of SFON). Mathematics anxiety is a phenomenon with a much longer history; however, it has only recently started to be examined in young children. It is not yet clear how mathematics anxiety affects numerical development in the initial stages of schooling.

1.2.3.4   Environmental factors

Research has shown that early environmental input plays a crucial role in the development of a wide range of cognitive skills (Reynolds & Temple, 1998; Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001). In the area of language acquisition, studies have long demonstrated that the linguistic input children receive is strongly associated with their general vocabulary growth (e.g. Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). Recently, it has become clear that early experiences play an important role in number development too. There are far fewer studies looking at home numeracy experiences compared to home literacy ones, but the evidence for a relationship between early numerical input and later numerical development is growing (e.g. Skwarchuk, Sowinski, & LeFevre, 2014).

One way in which researchers have studied children's early numerical experiences is through self-report measures. For example, LeFevre et al. (2009) developed a home numeracy questionnaire designed to measure parents' attitudes towards mathematics, their academic expectations for their child in mathematics, and the frequency of informal and formal number activities in the home. The results from this study revealed a significant positive correlation between the frequency of informal numeracy activities and children's standardised mathematics performance. This corroborated and extended previous work which had shown a relationship between numerical input and numerical development, but had not distinguished between different types of home numeracy experiences (Saxe, Guberman, & Gearhart, 1987; Blevins-Knabe & Musun-Miller, 1996; Starkey et al., 1999).

Other researchers have used more in-depth observational methods to assess children's early exposure to number. Levine, Suriyakham, Rowe, Huttenlocher, and Gunderson (2010) used a longitudinal design to uncover the relationship between parents' number-related talk and children's developing understanding of cardinality. They sought to obtain a measure of parental input that was not biased by self-reports or influenced by children's prior numerical knowledge, thus they used naturalistic observations with young children (aged $14 - 46$ months) who had yet to map number words onto their cardinal meanings. The results from this study revealed large variations in parents' number talk. These variations were associated with children's developing understanding of cardinality, even after controlling for children's number-related talk, parental talk in general and socioeconomic status (SES).

In addition to these investigations of parental input, studies have also looked at the role of preschool numeracy experiences. Klibanoff et al. (2006), for example, examined the effects of variations in preschool teachers' "math talk" on children's early numerical development. Children aged $4 - 5$ years were recruited from 26 classrooms; they were tested twice, once at the beginning and once at the end of the academic year. Interestingly, the results mirrored the findings from studies of parental mathematics input. Firstly, there were marked individual differences in preschool teachers' number-related talk, both in terms of frequency and diversity. (Over the course of one 1 hour classroom observation, the amount of "math talk" varied from just a single instance to 104 instances, with an average of 28 instances). Secondly,

the frequency and diversity of teachers' "math talk" –both of which were highly correlated– were significantly related to the growth of children's number knowledge over the academic year. Crucially, they were not related to children's numerical knowledge at the beginning of the preschool year.

Together, these results suggest that children's early exposure to quantitative activities is highly important for their symbolic number development. Specifically, informal numerical input (as opposed to direct numerical instruction) appears to play a particularly significant role. This is highlighted by the work of LeFevre and colleagues (LeFevre et al., 2009; Skwarchuk et al., 2014) and also by Siegler and Ramani, who have conducted a series of investigations into the developmental effects of playing number board games (Siegler & Ramani, 2008; Ramani & Siegler, 2008; Siegler & Ramani, 2009; Ramani & Siegler, 2011).

Siegler and Ramani emphasise the differences in numerical knowledge between preschoolers from low and high socioeconomic backgrounds. They refer to a number of studies showing that these differences exist (e.g. Ginsburg & Pappas, 2004) and that they likely stem from variations in environmental input (e.g. Clements & Sarama, 2007). Compared to low-income parents, middle- and high-income parents engage in a broader range of number-based activities with their children. The frequency of these activities has been shown to be related to children's early symbolic number knowledge (e.g. Siegler & Ramani, 2008).

In a three-week intervention study, Siegler and Ramani (2009) investigated the effects of linear vs. circular number board games. Preschool children from low socioeconomic backgrounds took part in five $15 - 20$ minute sessions with either a linear board game, a circular board game or numerical control activities (e.g. counting and number word identification). Results showed that playing the linear number board game promoted greater learning than playing the circular game or completing the control activities. Significant improvements were seen in numerical magnitude comparison and number line estimation. This suggests that linear board games may help to strengthen the mapping between children's symbolic and nonsymbolic representations of number, a finding that ties in with the research reviewed in Section 1.2.3.2.

Overall, a wide range of studies have shown that children's early numerical input is related to their later numerical development. In light of this,

researchers are beginning to develop preschool interventions that may help to reduce the large discrepancies found in children's early numerical understanding. Recent investigations suggest that informal numerical activities may play an important role in alleviating this gap.

1.2.3.5 Summary

Researchers have uncovered a number of factors that give rise to individual differences in symbolic number development. Currently, the literature on some of these factors is more extensive than others. It is widely documented that cognitive skills (both domain-general and domain-specific) are associated with children's early numerical abilities. Environmental influences (such as home numeracy experiences) have also become increasingly established. In comparison to these factors, little is known about the dispositional predictors of children's numerical development. As highlighted in Section 1.2.3.3, there are two dispositional factors that warrant further investigation. Firstly, SFON, a recently developed construct which refers to an individual's tendency to focus on the numerical aspects of their environment. Secondly, mathematics anxiety, a phenomenon long recognised by educators and researchers but one which is relatively unexplored in young children. These two factors form the focus of this thesis.

1.3   Overview of the Current Thesis

This thesis investigates the role of SFON and mathematics anxiety in the early childhood years. The empirical findings relating to each of these factors are presented in two separate parts. This is because SFON and mathematics anxiety were found to have independent effects on children's numerical processing skills. Interestingly, there were no differences in the mathematics anxiety levels reported by children of low, middle and high SFON tendency. This was revealed in two studies, both of which measured children's SFON and mathematics anxiety concurrently (see Figure 1.1).

Part II of this thesis focuses on the construct of SFON. First, Chapter 2 examines the current tasks used to assess children's SFON. Next, Chapter 3 develops and validates a new SFON measure. Finally, having addressed these methodological issues, Chapters 4 and 5 investigate the developmental

roots and mechanisms of SFON. Following these empirical chapters, Part II closes with a general discussion of SFON in Chapter 6.

Part III of this thesis focuses on mathematics anxiety. Chapter 7 examines the relationship between child and parent mathematics anxiety and Chapter 8 investigates the effects of children's mathematics anxiety on their numerical processing skills. The findings from this research reveal consistent gender differences in the correlates of children's mathematics anxiety. These gender differences are explored further in Chapters 9 and 10. Chapter 11 then reviews all of the empirical work presented in Part III.

In Part IV, Chapter 12 concludes this thesis by bringing together the findings from both strands of research.
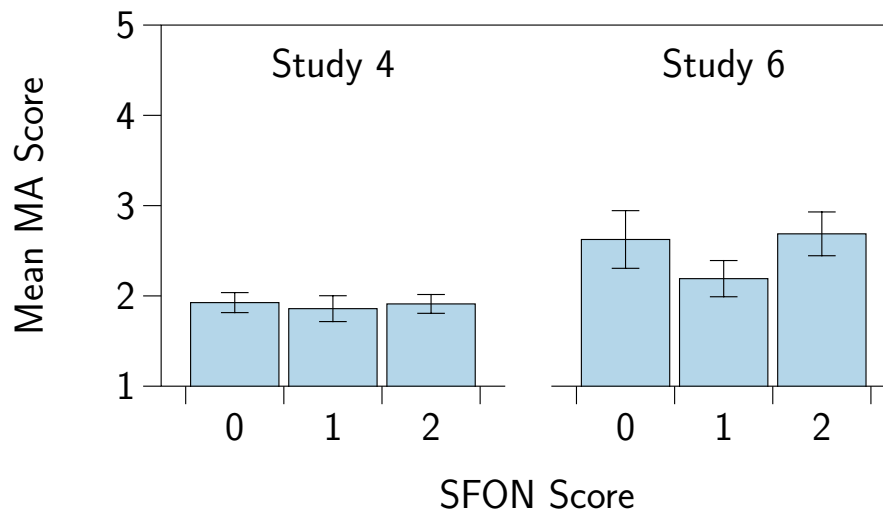
Figure 1.1: Mean mathematics anxiety score for children obtaining SFON scores from $0-2$ in Study 4 and Study 6 (error bars show ± 1 standard error of the mean).

Note. Mathematics anxiety was scored from $1-5$ (where $1 =$ low anxiety and $5 =$ high anxiety). SFON was scored from $0-2$ (where $0 =$ low SFON and $2 =$ high SFON). One-way ANOVAs were conducted to compare the mathematics anxiety scores of low, middle and high SFON children. These yielded no significant differences in mathematics anxiety between the different SFON groups: Study 4, $F(2, 99) = 0.08$, $p = .926$; Study 6, $F(2, 35) = 1.40$, $p = .260$.

Part II

Spontaneous Focusing on Numerosity

Chapter 2

Assessing Current SFON Measures (Study 1)

SFON is a recently developed construct which captures an individual's tendency to focus on the numerical aspects of their environment. As discussed in the literature review, researchers have demonstrated that preschool children show individual differences in SFON, and that these differences are associated with early counting and arithmetic skills. Researchers have used a variety of experimental tasks to assess children's SFON but, thus far, the psychometric properties of these tasks have yet to be established. In the current chapter I use a test-retest methodology to investigate the reliability and validity of three SFON measures in preschool and school-aged children.

2.1 Introduction

As reviewed in Chapter 1 (Section 1.2.3), recent research has highlighted the role of children's informal numerical experiences in the development of formal symbolic number skills. In particular, Hannula and her colleagues in Finland have demonstrated that preschool children show individual differences in their tendency to focus on numerical information in informal everyday contexts (Hannula, 2005; Hannula & Lehtinen, 2005; Hannula et al., 2010, 2005, 2007; Hannula-Sormunen, 2014). These individual differences in SFON have been shown to predict both current numerical skills (Hannula et al., 2007) and later mathematics success in school (Hannula et al., 2010).

In a three-year longitudinal study, Hannula and Lehtinen tracked the development of preschool children's counting skills, together with their SFON. SFON was assessed with a series of structured pretend play activities. For example, children completed a variety of imitation tasks in which they watched an experimenter post a given number of items (e.g. plastic berries) into a concealed object (e.g. a toy parrot) and then they were asked to do exactly the same. Children were scored as focusing on numerosity if they posted the same number of items as the experimenter and/or if they demonstrated any verbal or nonverbal quantitative acts (such as utterances including number words or counting with their fingers).

Results showed that children varied in the extent to which they focused on numerosity during the pretend play activities. At Time 1, 46% of children attended to numerosity on all possible occasions whilst 26% did not attend to numerosity at all. Interestingly, these individual differences in SFON, measured at 4, 5, and 6 years, were significantly related to children's counting skills. High SFON children performed better than low SFON children on measures of number word sequence production, object counting and cardinality understanding. Path analyses revealed a reciprocal relationship suggesting that SFON both precedes and follows the development of early counting skills.

In a follow-up experiment, Hannula and Lehtinen (2005) demonstrated that children's lack of SFON did not stem from inadequate procedural counting skills. They showed that children who did not spontaneously focus on numerosity could successfully complete the SFON tasks when they were guided towards the numerical aspects of the activities; guided focusing on numerosity (GFON). This suggests that SFON is not simply a proxy for counting knowledge itself; rather, it can be distinguished as a "separate sub-process in enumeration" (Hannula & Lehtinen, 2005, p. 237).

Further to this, Hannula et al. (2010) investigated the domain specificity of SFON as a predictor of children's numerical skills. In another longitudinal study they measured children's SFON together with their spontaneous focusing on a non-numerical aspect of the environment, namely, spontaneous focusing on spatial locations (SFOL). The findings showed that SFON in preschool predicted arithmetical skills, but not reading skills, two years later in school. This relationship could not be explained by individual differences in nonverbal IQ, verbal comprehension or SFOL.

SFON is thus emerging as a key factor for explaining variations in children's counting and arithmetical development, opening up several avenues for further research. To advance our theoretical understanding, future studies need to investigate why SFON is associated with a numerical advantage. In other words, what are the mechanisms behind this relationship? Moreover, from a more applied perspective, research needs to explore whether SFON is something that can be encouraged, or trained. Can we increase children's tendency to recognise and use numbers in informal everyday contexts? And, if so, do increases in SFON lead to better mathematical outcomes?

Before we address these theoretical and practical questions, there are some important methodological issues that need to be raised. First and foremost, we need to establish the extent to which SFON is a stable trait that can be reliably measured. Hannula and Lehtinen (2005) developed a variety of pretend play activities designed to assess preschool children's SFON. These activities (or tasks) have been used in a few different studies, mentioned above; however, it is not yet clear how reliable they are, or whether they are all measuring the same SFON construct. It is important to consider these aspects of SFON tasks when interpreting and comparing findings across studies. In the section below I outline key criteria within the context of measurement reliability and validity, before returning to review the current SFON tasks.

2.1.1   Reliability & Validity

Reliability and validity are two interrelated concepts that are important for evaluating the measures we use in research. Broadly speaking, reliability can be defined as the extent to which a research tool yields a consistent measure of a construct, and validity can be defined as the extent to which a research tool measures the specific construct it is intended to measure. There are several types of reliability and validity that need to be considered when assessing a research tool. These are discussed in the following sections, with particular reference to the measurement of SFON.

2.1.1.1   Reliability

Test-retest reliability   One way of testing the reliability of a measure is to examine its temporal stability using the test-retest method. The idea here is that if we administer the same measure to the same sample on two separate occasions, then the results obtained should be consistent (Bryman, 2008). If we find a high positive correlation between the results at Time 1 and Time 2 then the measure would appear to show good test-retest reliability.

Thinking about this in terms of children's SFON, if we assume that SFON is a stable trait, then children's tendency to focus on numerosity on a task designed to assess their SFON should show little variation over time. A critical issue to consider is the length of this time period. The test-retest reliability of a task will partly depend on the duration of the interval between the two measurement points. Generally, shorter gaps (i.e. days or weeks) will result in higher test-retest correlations than longer gaps (i.e. months or years) because factors relating to error are more similar. As the gap increases it becomes harder to distinguish between a lack of reliability in a measure and 'actual' changes in the construct (Bryman, 2008).

While shorter gaps are therefore preferable for assessing test-retest reliability, it is important to note that if the administrations are too close together in time then the results may be biased by memory effects or practice or boredom effects. Memory effects refer to the fact that participants may remember many of their responses on the first administration of the measure and respond identically. This can result in spuriously high test-retest reliabilities. In contrast, practice or boredom effects can result in lower test-retest reliabilities because participants have either learned from the first administration or they have become disinterested or fatigued (Segal & Coolidge, 2004). Given these issues, researchers examining the test-retest reliability of a measure should carefully consider an appropriate time interval and check for any systematic variations (i.e. practice or boredom effects) by comparing scores across the two testing timepoints.

Internal reliability   Another way of assessing the reliability of a measure is to examine its internal consistency. Essentially, if all items within a measure are related to each other, then they can be seen to be reliably measuring the same construct (Bryman, 2008). Consider, for example, a task designed to measure children's SFON. All of the items on this task should

be measuring children's spontaneous tendency to focus on numerosity. If this is the case then they will be highly correlated with each other and the task will show good internal reliability. If, however, there are some items on the task measuring children's spontaneous tendency to focus on another continuous quantity variable (e.g. size or density), then these items will be less correlated with the items assessing SFON, and consequently the task will show poorer internal reliability.

Internal reliability is often estimated using a split-half analysis or a Cronbach's alpha. Using the split-half method, the items on a measure are randomly divided into two halves and the correlation between participants' scores on these two sets of items is calculated. Cronbach's alpha is a more computationally complex application of the split-half method that averages across every possible split-half combination (Bryman, 2008). Typically, a measure is seen to have good internal reliability if it yields an alpha coefficient of 0.7 or above (Field, 2005), although lower thresholds are often seen as acceptable (R. J. Cohen & Swerdlik, 1999).

It is important to note that internal reliability estimates increase with the number of items on a measure. This means that measures with few items (e.g. fewer than seven) may be consistently measuring the desired construct but their alpha coefficients may suggest low internal reliability (Segal & Coolidge, 2004). This issue is particularly relevant in relation to the measurement of SFON. As emphasised by Hannula et al. (2010), tasks designed to assess SFON aim to capture behaviour that is spontaneous (rather than primed) therefore trial numbers are limited. In view of this, the reliability coefficients yielded by SFON tasks can be expected to be lower than those for other psychological constructs.

Inter-rater reliability   Reliability can also be indexed by the extent to which two or more raters (or observers) agree in their measurements. This type of reliability is important when the measurements obtained are subjective; for example in observational studies where data may be coded into categories (Bryman, 2008). If two or more raters are consistent in their coding then there will be a high percentage of agreement in the categorisations and thus, a good degree of inter-rater reliability. Note that if the measurements obtained are continuous then a correlation coefficient can be computed to compare the ratings of two observers. Inter-rater reliability is

particularly relevant to the measurement of SFON using observational methods rather than experimental tasks, as is the case in Study 3 (see Chapter 3, Section 3.4).

2.1.1.2   Validity

As with reliability there are a number of ways to assess the validity of the measures we use in our research. Measurement validity is concerned with whether a measure of a construct is really measuring the construct it is intended to measure, and it is commonly referred to as 'construct validity' (Bryman, 2008). Below I define four types of construct validity: predictive, concurrent, convergent and discriminant. Note that this is not an exhaustive list and it is sometimes organised in different ways; here I take the approach of Trochim (2006).

As well as construct validity, there are other types of validity relevant to scientific research. Validity is a broad term which includes both internal validity (the extent to which the results of a study are due to the variable the researcher intended to study and not some other confounding variables) and external validity (the extent to which the results of a study can be generalised to situations beyond the research context). The following discussion is restricted to construct (or measurement validity), but issues surrounding external validity are elaborated upon in Chapter 3, Section 3.4.

Predictive validity   One way of assessing the construct validity of a measure is to examine how well it predicts something that, in theory, it should be able to predict (Trochim, 2006). With regards to SFON, children's attention to number during a SFON task should be related to real-world observations of their recognition and use of numbers. Moreover, it is theorised that children who are more inclined to focus on the numerical aspects of their environment should show better numerical skills than children who are less inclined to focus on numerosity. In view of this theory, children's performance on a task designed to measure SFON should be positively correlated with their numerical achievement in school. This relationship between SFON scores and numerical skills has been previously demonstrated in the literature (e.g. Hannula & Lehtinen, 2005), thus the current tasks used to assess SFON can be seen to exhibit some degree of predictive validity.

Concurrent validity   Another way of assessing the construct validity of a measure is to compare a group of participants' scores on this measure with their scores on a theoretically-related measure (when both measures are taken concurrently). If a measure has concurrent validity then it should correlate positively with the theoretically-related measure (Segal & Coolidge, 2004). A difficulty with this type of validity is that it relies on there being an existing theoretically-related measure that has previously been validated. If there is no known existing measure, such as is the case with SFON (a recently developed construct), then concurrent validity is not easily demonstrated and thus, researchers must use other forms of construct validity.

Convergent validity   The construct validity of a measure can also be indexed by the extent to which measures of the same construct converge (Trochim, 2006). For example, in terms of SFON, the different tasks developed by Hannula and Lehtinen (2005) to assess children's SFON should, in theory, all be measuring (or converging on) the same SFON construct. If this is the case then children's performance on these tasks should be positively correlated. Note that convergent validity is similar to concurrent validity but it involves the comparison of multiple measures of the same construct.

Discriminant validity   As well as establishing the convergent validity of a measure it is necessary to assess its discriminant validity; that is, the extent to which it can discriminate between constructs that are theoretically different. To illustrate this in terms of SFON, it is important that the different tasks designed to assess children's SFON not only all converge on the same SFON construct, but that they diverge from dissimilar constructs such as spontaneous focusing on a non-numerical aspect of the environment (e.g. SFOL). This discriminant validity has been previously demonstrated by Hannula et al. (2010) who found little correlation between children's performance on measures of SFON and SFOL ($r = .15$).

2.1.2   Current SFON Tasks

Children's SFON has been measured with a variety of experimental tasks involving pretend-play activities.[1] In the longitudinal study conducted by Hannula and Lehtinen (2005) different tasks were employed across the three testing timepoints. At 4 and 5 years children were presented with a single SFON task involving quantities that were not visible. On each of three trials the experimenter posted a given number of items into a concealed object and then the child was asked to do exactly the same. The objects involved were plastic berries and a toy parrot at 4 years, and money (Finnish Marks) and a savings-box at 5 years. At the final point of testing (6 years) children were presented with a wider range of SFON tasks involving both visible and invisible quantities. These tasks are briefly described below:

1. Posting Task: This was similar to that presented at 4 and 5 years. The experimenter posted some letters into a toy postbox and then asked the child to do the same.

2. Model Task: The experimenter stamped some geometric shapes (a circular "node" and triangular "spikes") onto a drawing of a dinosaur. They then asked the child to make their dinosaur look the same.

3. Finding Task: The experimenter placed an object under 1 of 27 toy hats and then asked the child to remember under which hat the object had been hidden.

As stated above, little is known about the psychometric properties of these tasks. It is not clear whether they show test-retest reliability and construct validity. Note that although they have demonstrated some predictive validity (predicting children's numerical skills) and discriminant validity (distinguishing between children's SFON and SFOL) their convergent validity has yet to be assessed. We don't know whether all of these tasks measure, or converge on, the same SFON construct.

Intuitively, the tasks seem to differ in their demands. While the Model Task involves quantities that are visible to the child, the Posting and Finding Tasks involve quantities that are hidden. They therefore vary in terms

---

[1] See Hannula (2005) for full details of the development of SFON assessments.

of their working memory demand. Furthermore, the Posting and Finding Tasks can be seen to differ from each other because the Posting Task draws on children's understanding of cardinality (i.e. understanding that the last number in a count represents the numerosity of the set as a whole), whereas the Finding Task is more about the concept of ordinality (i.e. understanding that the last number in a count represents the position in a sequence).

Adding to these issues, it remains to be seen whether the reliability and validity of these tasks change with (age-related) development and with exposure to formal schooling. Note that thus far, these tasks have only been used with preschool samples, since children in Finland start school later than they do in other parts of the world. The compulsory age for children starting school in Finland is 7 years whereas in the UK the legal age is 5, and children generally start school before this point (in the year of their fifth birthday).

Overall, the results of Hannula and Lehtinen (2005) provided mixed evidence in terms of the stability of children's SFON. Path analyses revealed a significant correlation between SFON tendency at 4 and 5 years ($r = .53$), but there was no relationship between SFON at 4 and 5 years and SFON at 6 years. This suggests one of two possibilities: either (i) SFON is not a stable trait, or (ii) different measures of SFON do not measure the same SFON construct. Study 1 sought to distinguish between these possibilities.

2.1.3   Aims of Study 1

The aims of this study were threefold:

1. To investigate the test-retest reliability and convergent validity of current SFON measures;

2. To explore the effects of age and formal education on children's SFON tendency;

3. To replicate and extend previous findings on the relationship between children's SFON and early numerical skills.

To achieve these objectives, a test-retest experiment was carried out in which preschool and school-aged children in the UK (aged $3 - 6$ years) completed three SFON tasks (developed by Hannula & Lehtinen, 2005),

administered one week apart. This one week interval was chosen on the basis that it would be long enough to avoid memory effects, or practice or boredom effects, but not so long that it would be difficult to distinguish between a lack of reliability in the tasks and actual changes in children's SFON (as discussed in Section 2.1.1.1). Following this experiment, preschool practitioners and school teachers completed an assessment of each child's numeracy skills.

Based on previous studies it was hypothesised that SFON would be related to children's numerical development (Study Aim 3). Given the exploratory nature of Aims 1 and 2, no predictions were made regarding the psychometric properties of the SFON tasks or the effects of age and formal schooling.

2.1.4 Ethical Considerations

There are a number of ethical issues relevant to this topic of research. In the sections below I digress from the current study to consider these issues in relation to all of the studies presented in this thesis. Each ethical issue is discussed with reference to the BPS Code of Human Research Ethics (The British Psychological Society, 2010).

2.1.4.1 Valid Consent

All participants involved in research studies should consent freely to take part. In accordance with the BPS Code of Human Research Ethics, this consent should be obtained in a manner that is suitable to the participant's age and level of competence. In cases where the research involves children (those under the age of 16 years) consent should be sought from children's parents or legal guardians prior to the start of the study. Where possible, children should also provide verbal assent before their participation begins. If this is not possible, for example if the child is very young, then "their assent should be regularly monitored by sensitive attention to any signs, verbal or non-verbal, that they are not wholly willing to continue with the data collection" (The British Psychological Society, 2010, p. 17).

All of the research presented in this thesis involved young children of preschool and primary-school age, thus consent to take part was obtained from children's parents or legal guardians. The children were deemed too

young to provide their own written consent; therefore, their assent was monitored closely through their verbal and non-verbal behaviour during all testing sessions. Children were recruited in one of two ways: (1) through local schools and nurseries, and (2) through the University of Nottingham's 'Summer Scientist Week' scheme. This is an annual research event which takes place during the school holidays. Further details can be found in Section 3.2.1.1.

For studies where data collection took place in local schools and nurseries, permission was first sought from the headteacher or manager of the relevant institution. Once this permission had been received, parents and guardians of children in the participating school or nursery received letters detailing the proposed research activities and inviting their child to take part. If they were happy for their child to take part then they were asked to sign and return an attached form to their child's school or nursery teacher.

For studies carried out as part of the University of Nottingham's 'Summer Scientist Week', permission was sought from children's parents and guardians directly. Parents and guardians who registered for their child to take part in the event were sent an information pack detailing the research activities involved in each of the proposed studies. If they were happy for their child to take part in all (or some) of the research studies then they were asked to sign and return a consent form to the university (indicating which of the studies they were happy for their child to take part in).

As part of the consent process, participants should be given all relevant information about (i) the nature of the research, (ii) the type and length of data collection involved, (iii) the confidentiality associated with the data collected, (iv) how the research findings will be made available, (v) the right to withdraw from the study at anytime without giving a reason and without any adverse consequences, and (vi) who to contact should they wish to talk to anyone about the research.

In line with these guidelines, all study information sheets given to schools, nurseries and parents/guardians contained details relating to each point (i – vi) listed above, except for some minor withholding of information about the numerical nature of the research (see Section 2.1.4.2 below).

2.1.4.2 Withholding Study Information

In order to measure an individual's spontaneous, self-initiated, attention to numbers (or SFON) it is essential that the individual does not know that their engagement with numbers is being observed. Therefore, all SFON studies reported in this thesis were presented to participants (schools/nurseries, parents/guardians and children) as studies of general thinking skills as opposed to early numeracy. With regards to this withholding of information the BPS Code of Human Research Ethics specifies the following:

> "Where an essential element of the research design would be compromised by full disclosure to participants, the withholding of information should be specified in the project protocol that is subjected to ethics review and explicit procedures should be stated to obviate any potential harm arising from such withholding" (p. 24).

While it was crucial for the participants to be uninformed about the numerical focus of this research, it was possible to provide full details of the SFON activities themselves. Thus, the withholding of information presented no risk to participants. All SFON research presented in this thesis was reviewed and approved by the Loughborough University Ethics Approvals (Human Participants) Subcommittee, or, in the case of those studies conducted as part of the University of Nottingham's 'Summer Scientist Week' scheme, the University of Nottingham School of Psychology Ethics Committee.

2.1.4.3 Recruiting & Testing Child Participants

As outlined in Section 2.1.4.1, children were recruited either through local schools and nurseries or through the University of Nottingham's 'Summer Scientist Week' scheme. Data collection that took place within schools and nurseries was arranged to fit in around everyday classroom routines. Each testing session lasted a maximum of 30 minutes in order to minimise the time taken out of the child's school or nursery day. The researcher (myself) had an enhanced CRB check together with appropriate training and experience in working with young children.

2.2   Method

I now return to the first study which sought to examine the current tasks used to measure children's SFON.

2.2.1   Participants

Sixty-five children were recruited from two nurseries and one primary school in Nottinghamshire, UK. The school was of mid to high SES based on the proportion of pupils eligible for free school meals compared to the national average. The nurseries were in the same catchment area as the school, and thus all children are assumed to be from similar socioeconomic backgrounds.

Thirty-one children (17 girls) aged $2.8-4.5$ years ($M = 3.6$ years, $SD = 0.5$ years) were in preschool nursery classes. Thirty-four children (24 boys) aged $4.4-5.7$ years ($M = 4.9$ years, $SD = 0.3$ years) were in the second term of their first year of school. At this early stage of schooling learning is play-based and child-led, following the UK government's Early Years Foundation Stage (EYFS) framework for children aged 0 - 5 years.

Participation was voluntary and the children received stickers to thank them for taking part. All parents provided written consent prior to the start of the study and children's assent was monitored closely during each testing session. Study procedures were approved by the Loughborough University Ethics Approvals (Human Participants) Subcommittee (see Section 2.1.4 for further details).

Two children from the primary school sample were excluded from the analyses; one of the children was not a native speaker of English and the other was identified by the school as having learning difficulties. A further six children from the nursery sample had missing data (and were excluded from the individual analyses that included those data) due to being absent at the second time of testing ($N = 5$) or being unable to complete all of the tasks ($N = 1$). In total there were 57 complete datasets.

2.2.2   Design

A test-retest design was employed with an exact time interval of one week (± 60 minutes). During each of two sessions children completed three SFON tasks (developed by Hannula & Lehtinen, 2005), the order of which was counterbalanced across participants. Children were presented with the tasks

in the same counterbalanced order in each session. Following the SFON tasks, in Session 1 only, children completed the Block Design subtest from the WPPSI (Wechsler, 1967).

In addition to these measures, school and nursery teachers were asked to complete an assessment of each child's numeracy skills. This data was collected after all testing was complete in order to keep the numerical aspect of the study concealed (see Section 2.1.4.2).

2.2.3   Materials & Procedure

Testing took place during two sessions, each lasting $20 - 30$ minutes. Children were tested individually, in a corridor outside their classroom or a quiet area of their nursery. The researcher was careful to ensure that the testing area was free from any numerical displays (e.g. noticeboards displaying number work) that might have prompted the children to focus on numerosity. Children were not told that the tasks were in anyway numerical or quantitative. Likewise, the children's parents and teachers were not informed of the numerical aspect of the study; rather, they were told that the study was looking at children's general thinking skills (see Section 2.1.4.2).

Throughout all tasks children received general praise (e.g. "You're watching really nicely") but no specific feedback was given. After each task children were allowed to choose a sticker.

2.2.3.1   SFON

Each of three SFON tasks were administered in accordance with the procedure of Hannula and Lehtinen (2005). The researcher recorded all verbal and nonverbal quantitative acts. These included the following:

A. Utterances including number words
   (E.g. "I did yellow three times. . . did you do yellow three times?")

B. Counting acts
   (E.g. "one, two, three, four. . . I want four")

C. Use of fingers to denote numbers
   (E.g. pointing)

D. Utterances referring to quantities or counting
   (E.g. "I think I did too many")

E. Interpretation of the goal of the task as quantitative
   (E.g. "I don't know how many you did")

Based on the previous literature, all of the tasks involved small numerosities (between 2 and 7) to ensure that they were within the children's counting range.

It is important to note that during the examination of this thesis, it became evident that there were some differences between the SFON tasks used here and those used by Hannula and Lehtinen (2005). Whilst the procedures reported in Hannula and Lehtinen (2005) were followed directly, unfortunately some aspects of the tasks were interpreted in a different way to what had been intended. A series of footnotes are included in the remainder of this chapter to highlight these differences.

Posting Task The materials used in this task were a pop-up toy postbox (diameter = 28.7cm, height = 42.0cm), a pile of 20 blue letters (9.5cm × 6.0cm), a pile of 20 yellow letters (9.5cm × 6.0cm) and a toy postman. The researcher sat opposite the child and the materials were placed in front of them on a table as shown in Figure 2.1. The postbox was positioned on the left hand side of the researcher and on the right hand side of child.

The researcher introduced the materials by saying: "Here is Pete the Postman's postbox, and here are some letters. We have some blue letters and some yellow letters. Now, watch carefully what I do, and then you do just the same". The researcher posted two yellow letters, one at a time, into the postbox followed by one blue letter. She then prompted the child: "Now you do just the same". On the second trial the researcher posted one blue letter and one yellow letter and on the third trial she posted two blue letters and three yellow letters. The researcher progressed from one trial to the next by saying: "Okay, let's go again".

For each trial children were scored as spontaneously focusing on numerosity if they posted the correct number of letters into the postbox and/or if they presented any quantifying acts listed above (A – E). Each child received a total score out of three.

Model Task The materials used in this task were six identical line drawings of a dinosaur (printed and laminated on A4 sized paper), three ge-

Figure 2.1: The experimental materials used in the SFON Posting Task.

ometric shaped rubber stamps (one circle, one isosceles triangle and one equilateral triangle) and one green ink-pad.

The researcher sat opposite the child and began by placing one dinosaur picture in front of the child and one in front of themselves. She set out the stamps and the ink-pad and said: "Here are some stamps and here is some green ink. Now watch carefully while I make my picture into a dinosaur". The researcher then picked up the circular stamp, dabbed it in the ink and stamped five circles over the dinosaur's body. When complete, she rotated the picture to face the child.[2] Next, the researcher dabbed the stamp in the ink-pad, gave the stamp to the child and said: "Now, make your picture look just like my dinosaur". On the second and third trials the procedure was repeated with three triangles and six triangles (see Figure 2.2). The researcher introduced the next trial by saying: "Okay, let's put these pictures away and make a different dinosaur".

---

[2]Note that in Hannula and Lehtinen's (2005) studies the dinosaur was turned over so that the child could not see the researcher's model. The number of stamps was therefore concealed rather than visible. This difference between the task procedures stems from a misinterpretation of the instruction "turned upside down".

(a) Trial 1   (b) Trial 2

(c) Trial 3
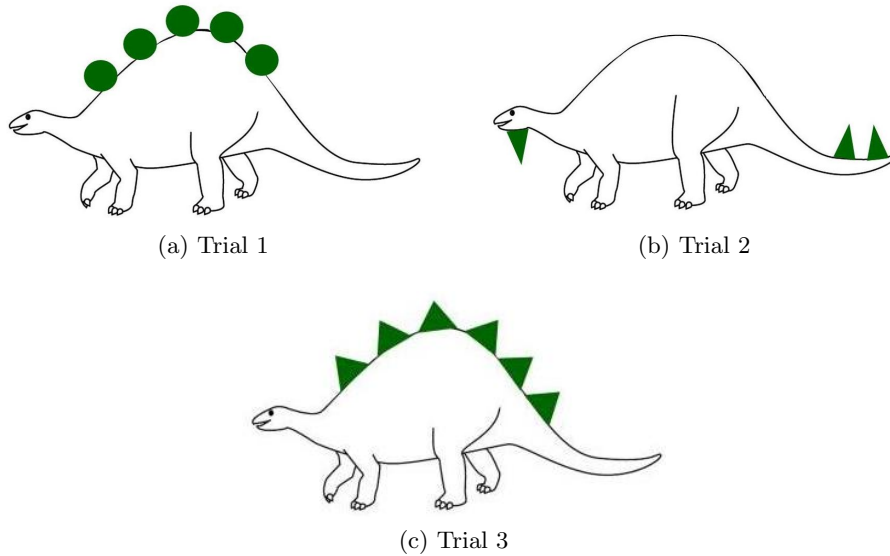
Figure 2.2: The models of the first, second and third trials of the SFON Model Task.

For each trial children were scored as spontaneously focusing on numerosity if they stamped the correct number of circular "nodes" or triangular "spikes" onto their dinosaur and/or if they presented any quantifying acts listed above (A – E). Each child received a total score out of three.

Finding Task The materials used in this task were 27 blue plastic pots (diameter = 2.9cm, height = 5.1cm), a stamp (2.5cm × 2.3cm) and a toy postman. The pots were arranged side by side in a semi circle in front of the child (see Figure 2.3).[3] All pots were placed upside down so that the stamp could be hidden under one of them.

The researcher sat opposite the child and introduced the materials by saying: "Here is Pete the Postman and here are his pots". Next she said to the child: "Now, you may watch while I hide this stamp under one of Pete the Postman's pots. Then you can tell Pete the Postman where the stamp is hidden. Now watch under which pot I hide the stamp". At this point the researcher lifted the pot that was sixth on her right and placed

---

[3]Note that Hannula and Lehtinen (2005) used wooden hats as opposed to plastic pots. These hats were arranged side by side in a semi-circle with no gaps between them.

the stamp inside the pot. She counted silently to four, in order to give the child time to remember the location of the stamp, before placing the pot back down. The researcher continued by saying: "Now, can you close your eyes for me (the researcher counted silently to four). And reach up to the sky (again, the researcher counted silently to four). And look back at me". These instructions were given in order to prevent the child from pointing or keeping their eyes on the target pot. The researcher then asked the child: "Can you tell Pete the Postman where the stamp is hidden?" If the child lifted up the wrong pot then the researcher lifted up the target pot to show the child where it was.

This procedure was repeated for the second and third trials. On the second trial the stamp was hidden under the seventh pot on the researcher's left and on the third trial it was hidden under the fifth pot on the researcher's right. Following all three trials the researcher asked the child: "How did you try to remember where the stamp was hidden?"

For each trial children were scored as spontaneously focusing on numerosity if, and only if, they found the stamp using counting procedures (as indicated by their verbal and nonverbal behaviour during each trial and their response to the researcher's question: "How did you try to remember where the stamp was hidden?"). If a child found the stamp by chance, rather than on the basis of number knowledge, then they were not scored as focusing on numerosity. Moreover, if a child stated that they had used counting procedures, but got some trials wrong, then they were only scored as focusing on numerosity for the trials on which they correctly found the stamp. Each child received a total score out of three.

2.2.3.2   Block Design

The Block Design subtest of the WPPSI was administered in accordance with the standardised procedure (Wechsler, 1967). Children were presented with a series of three-dimensional geometric patterns and they were asked to replicate them, within a given time limit, using a set of two to four blocks. If they succeeded on the three-dimensional items then they progressed onto a series of two-dimensional patterns. Each child received a score out of 20.

The purpose of this task was to obtain a measure of nonverbal IQ; specifically, it measured spatial processing and visual motor integration. The Block Design subtest is well regarded as a measure of nonverbal reasoning ability.
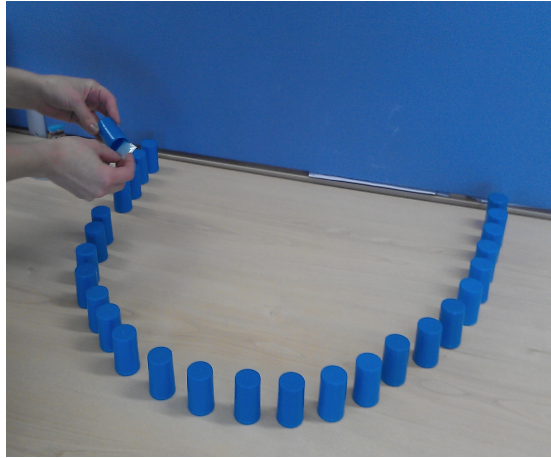
Figure 2.3: The experimental materials used in the SFON Finding Task.

It is known for its high reliability coefficients and strong correlation with the full Performance IQ scale (LoBello, 1991).

2.2.3.3   Numeracy Skills

School teachers and nursery staff were asked to complete an assessment of each child's numerical abilities. Several studies have shown that teacher-based judgements correspond highly with children's scores on standardised achievement tests (see Südkamp, Kaiser, & Möller, 2012, for a review). This teacher assessment covered the following skills: number word sequence production, numerical ordering, cardinality understanding and simple addition and subtraction. The following six questions were presented:

1. Can this child reliably count up to 5?

2. Can this child reliably count up to 10?

3. Can this child order the numbers $1 - 10$?

4. If presented with a group of items, can this child tell you how many there are altogether? For what numbers $1 - 10$ can they do this?

5. Can this child add one or more items?

6. Can this child subtract one or more items?

This type of information is regularly recorded by school teachers and nursery staff in order to monitor children's progress along the EYFS framework. Based on this assessment, each child received a numeracy score $1-5$ (where $1=$ the lowest numeracy score and $5=$ the highest numeracy score) as outlined in Table 2.1.

	Count to 5	Count to 10	Order numbers	How many?	Add & subtract
1	-	-	-	-	-
	Yes	-	-	-	-
2	Yes	Yes	-	-	-
	Yes	-	Yes	-	-
	Yes	-	-	Yes	-
3	Yes	Yes	Yes	-	-
	Yes	Yes	-	Yes	-
4	Yes	Yes	Yes	Yes	-
5	Yes	Yes	Yes	Yes	Yes

Table 2.1: Criteria for each numeracy score $1-5$.

2.3   Results

2.3.1   Data Preparation

In line with Hannula and Lehtinen's (2005) analysis, children received a score from 0 to 3 for each of the three SFON tasks. For the Posting and Model Tasks children were deemed to be focusing on numerosity if they produced the correct number and/or if they displayed any verbal or nonverbal quantitative acts. For the Finding Task children were deemed to be focusing on numerosity if, and only if, they found the stamp using counting procedures (as indicated by their verbal and nonverbal behaviour during each trial and their response to the researcher's question: "How did you try to remember where the stamp was hidden?"). If a child found the stamp by chance, rather than on the basis of number knowledge, then they were not scored as focusing on numerosity.

This method of scoring is somewhat broad. Consider, for example, two children scoring three out of three on the Posting Task. One child may have

posted the correct number of letters on all three trials, while the other child may have been unsuccessful at posting the correct number of letters but have made verbal references to quantities (e.g. "I think I did too many"). In order to ensure that this broad method of scoring was not concealing any patterns in the data, children's responses were coded in three additional ways. Each method of scoring is described below.[4]

1. Hannula & Lehtinen score   The first method of scoring was based on Hannula and Lehtinen's (2005) coding system. As described above, for the Posting and Model Tasks children were scored as focusing on numerosity if they produced the correct number and/or if they displayed any verbal or nonverbal quantitative acts. For the Finding Task children were scored as focusing on numerosity if, and only if, they found the stamp using counting procedures. If they found the stamp by chance, rather than on the basis of number knowledge, then they were not scored as focusing on numerosity.

2. Pure accuracy score   The second method of scoring was based purely on accuracy. Children were scored as focusing on numerosity if, and only if, they produced the correct number. This meant posting the correct number of letters on the Posting Task, stamping the correct number of "nodes" or "spikes" on the Model Task and identifying the correct pot on the Finding Task.

3. Pure quantitative acts score   The third method of scoring was based purely on quantitative acts. Children were scored as focusing on numerosity if, and only if, they displayed any verbal or nonverbal quantitative acts. Examples of these quantitative acts are given in Section 2.2.3.1.

4. Combined score   The fourth method of scoring was based on accuracy and quantitative acts. Children were scored as focusing on numerosity if they produced the correct number and/or if they displayed any verbal or nonverbal quantitative acts. Note that the only difference between this method and Hannula and Lehtinen's method was the way in which the Finding Task

---

[4]Note that in Hannula and Lehtinen's (2005) study, the observational data served as an additional method which aimed to increase the reliability of the SFON assessments on trials when the child made an enumeration mistake.

was scored. Here, children were scored as focusing on numerosity on the Finding Task if they identified the correct pot and/or if they used counting procedures (as indicated by their verbal and nonverbal behaviour during each trial and their response to the researcher's question: "How did you try to remember where the stamp was hidden?"). Note that if a child found the correct pot, but showed no evidence of counting or number knowledge, then they were still scored as focusing on numerosity under this method of scoring.

2.3.2   Data Analysis

First, descriptive statistics are presented for children's performance on each of the SFON tasks. The sections that follow address each of the three research aims outlined in Section 2.1.3. Group-level comparisons are made across the preschool and school samples to explore the effects of age and formal education on children's SFON tendency (Aim 2). Correlations are performed to investigate the test-retest reliability and convergent validity of each SFON task (Aim 1) as well as the relationship between children's SFON and numerical skills (Aim 3).

2.3.2.1   Descriptive Statistics

Figures 2.4, 2.5 and 2.6 show the distribution of children's SFON scores (at Time 1) on the Posting Task, the Model Task and the Finding Task, respectively. For each task, separate graphs are presented for preschool and school-aged children, and for each of the four methods of scoring. Together, these graphs demonstrate that both preschool and school-aged children showed individual differences in SFON. The datasets for each task were of a normal or skewed distribution as opposed to a bimodal distribution. This suggests that SFON is continuous rather than discrete.

Visual inspection of the data reveals that very few children displayed pure quantitative acts across any of the tasks (scoring method 3). Children's SFON scores were driven primarily by accuracy (i.e. producing the correct number) as opposed to verbal or nonverbal quantitative acts (i.e. utterances or gestures that indicated counting).

In light of this, the statistical analyses that follow are presented for the combined method of scoring only. Recall that this method of scoring was the
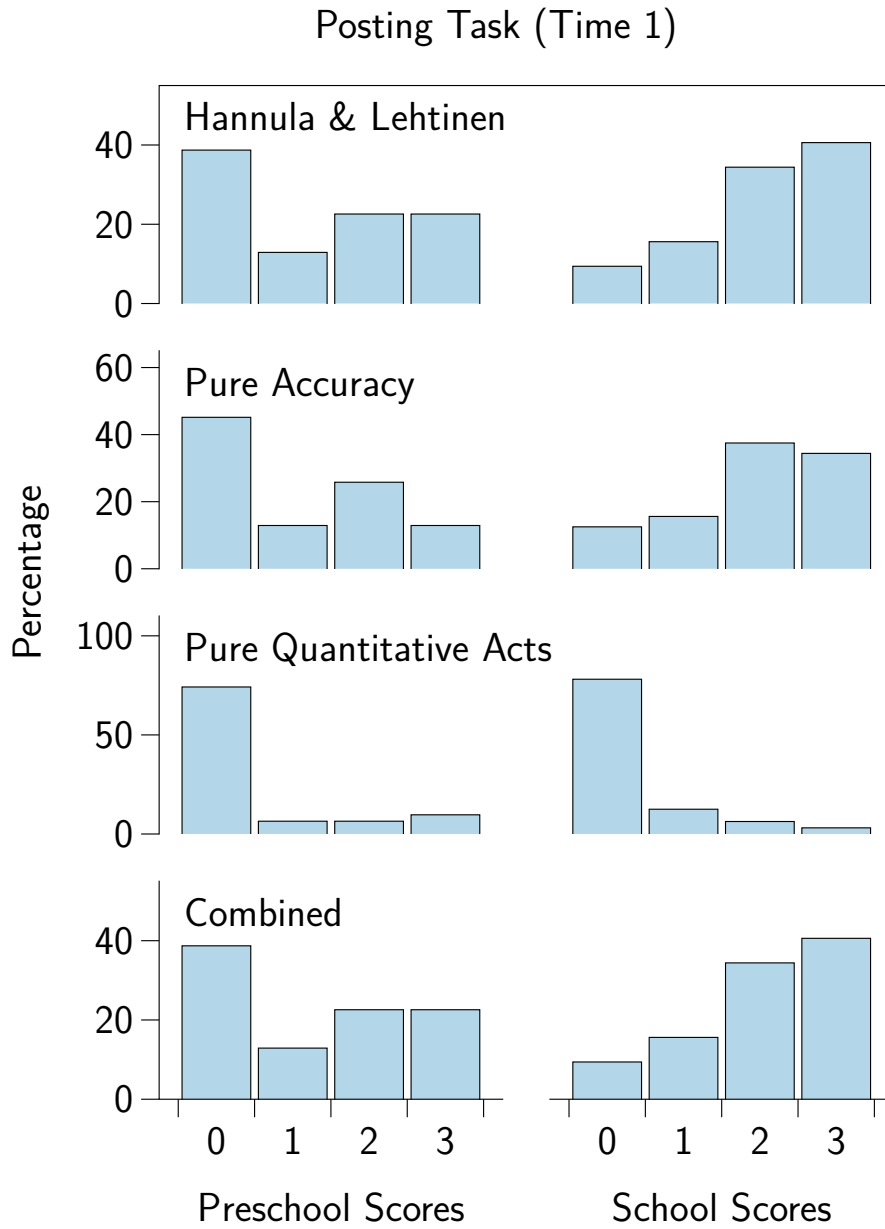
Figure 2.4: The percentage of preschool children (left) and school children (right) obtaining SFON scores from $0-3$ on the Posting Task (at Time 1) for each method of scoring.
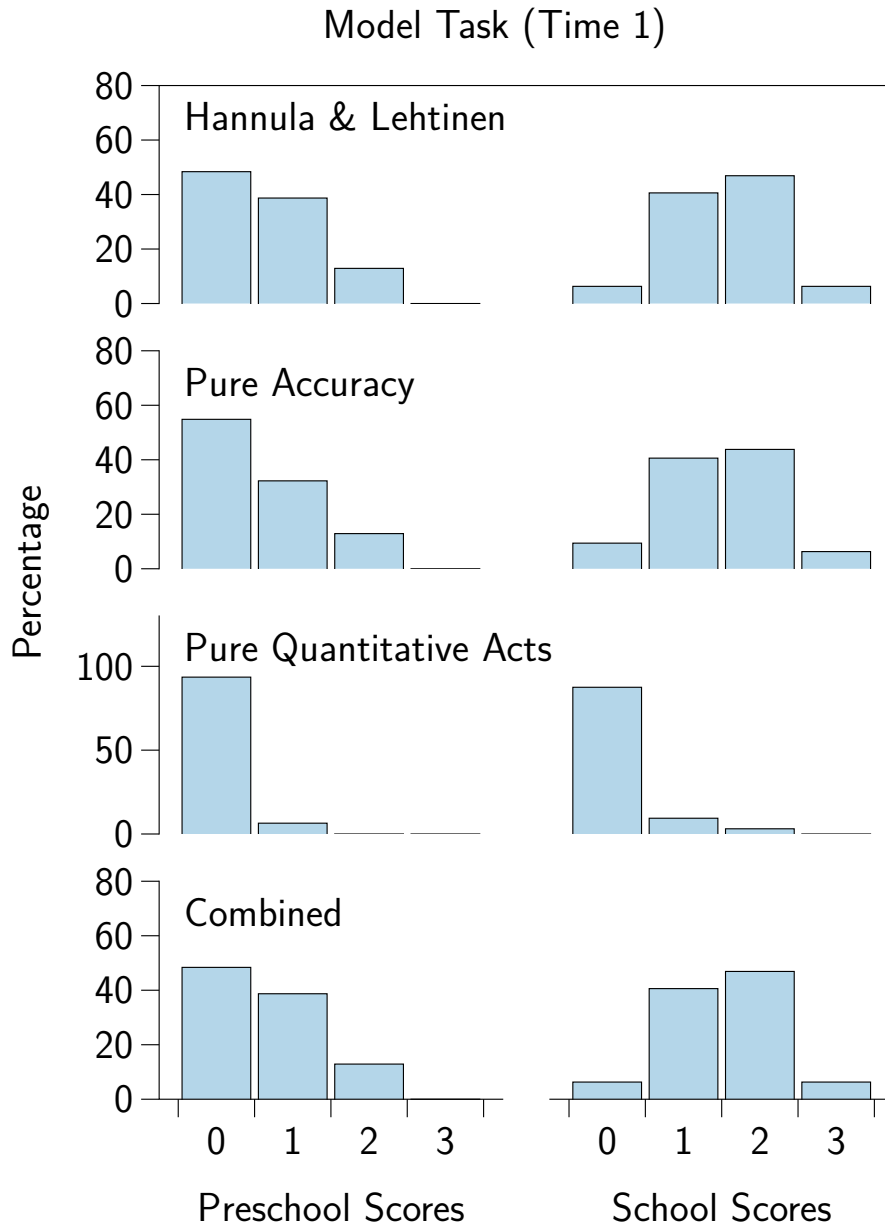
Figure 2.5: The percentage of preschool children (left) and school children (right) obtaining SFON scores from $0 - 3$ on the Model Task (at Time 1) for each method of scoring.
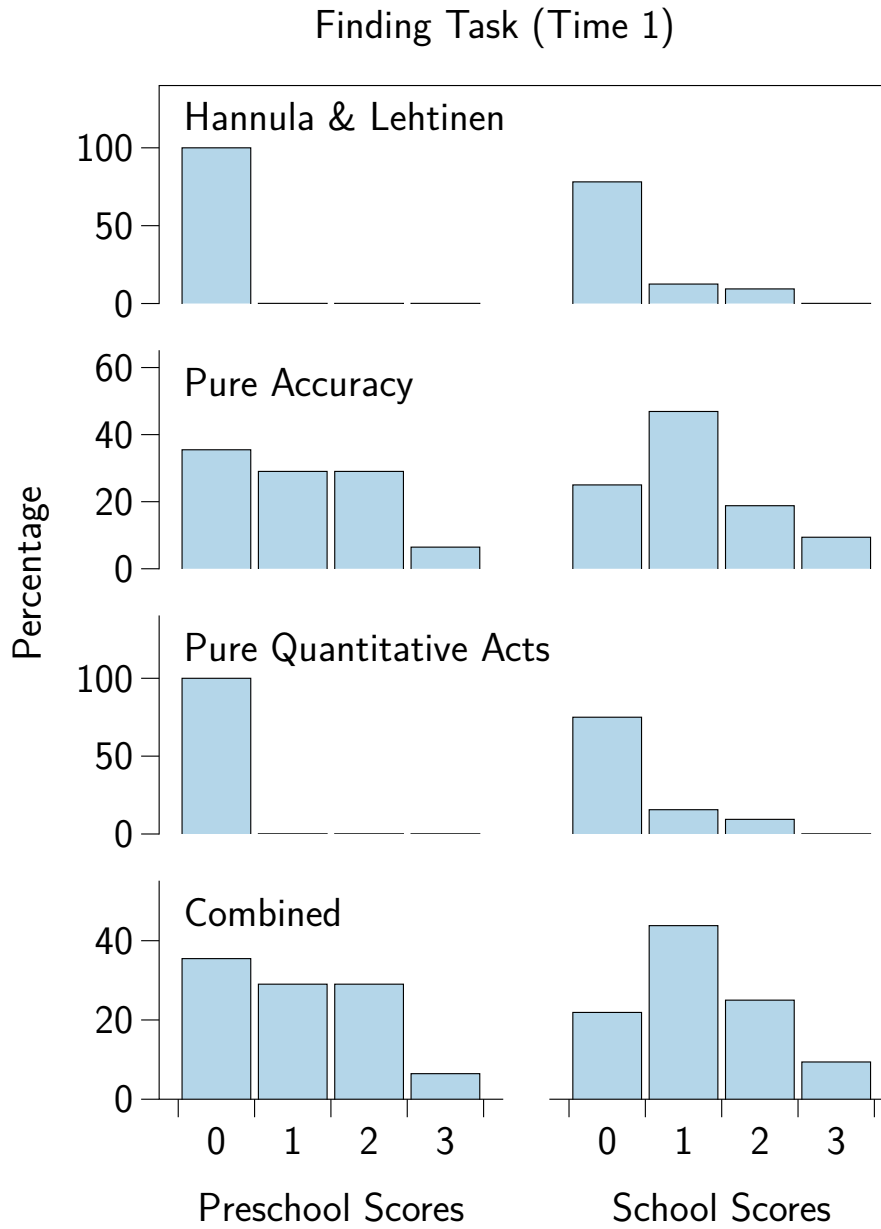
Figure 2.6: The percentage of preschool children (left) and school children (right) obtaining SFON scores from $0-3$ on the Finding Task (at Time 1) for each method of scoring.

same as Hannula and Lehtinen's (2005) method of scoring for the Posting and Model Tasks. The only difference concerned the Finding Task whereby children were scored as focusing on numerosity if they identified the correct pot and/or if they used counting procedures. Using the combined method of scoring, rather than Hannula and Lehtinen's method, ensured that results for the Finding Task were not biased by floor effects.

2.3.2.2   Group-level Analyses

Do preschool and school-aged children show different levels of SFON?

A series of Mann-Whitney U tests were performed to compare preschool and school-aged children's SFON scores on each of the three tasks at Time 1 and Time 2. The alpha level was Bonferroni corrected to .008 (.05/6) to control for the effect of multiple group comparisons. Results revealed a significant difference between preschool and school-aged children's SFON scores on the Model Task at Time 1 ($z = -4.21$, $p < .001$) and Time 2 ($z = -2.78$, $p = .005$), reflecting the fact that school-aged children showed higher levels of SFON than preschool children (Time 1: school median = 2.0, preschool median = 1.0; Time 2: school median = 1.5, preschool median = 1.0). No significant differences between preschool and school-aged children's SFON scores were found on the Posting Task ($ps > .014$) or the Finding Task ($ps > .057$) at either timepoint.

These comparisons provide mixed evidence as to whether children's SFON tendency increases with (age-related) development and exposure to formal schooling. It seems that school-aged children demonstrate higher levels of SFON than preschool children, but only when SFON is measured with certain tasks, i.e. the Model Task (and to some extent the Posting Task). Two possible explanations for these mixed findings are: (1) that the three tasks used to measure SFON vary in their suitability for use with different age groups, and (2) that the different SFON tasks are not all tapping into the same SFON construct. The second of these possibilities is examined in Section 2.3.2.3 below.

Do boys and girls show different levels of SFON?

In an exploratory manner, group-level comparisons were also made between boys and girls to examine the effect of gender on children's SFON. A series of Bonferonni-corrected Mann-Whitney U tests indicated that boys and girls showed no significant differences in their SFON scores on any of the SFON tasks at Time 1 or Time 2 (all $ps > .135$). This suggests that there are no gender differences in terms of the extent to which children spontaneously attend to the numerical aspects of their environment. Note that these gender comparisons are included here, and in all subsequent SFON chapters, because the null effects will contrast with later gender effects reported in Part III of this thesis.

2.3.2.3   Correlational Analyses

Spearman correlational analyses were performed to investigate: (1) the stability of children's SFON over time, (2) the extent to which different tasks used to assess children's SFON measure the same SFON construct, and (3) the relationship between children's SFON and numerical skills.

Prior to conducting these correlations, a series of Wilcoxon signed ranks tests were carried out to check whether there were any changes in children's SFON scores across the two testing timepoints. The results revealed no significant differences in SFON scores on any of the tasks between Time 1 and Time 2: Posting Task, $z = -1.11$, $p = .266$; Model Task, $z = -1.52$, $p = .129$; Finding Task, $z = -1.51$, $p = .130$. This demonstrates that there were no systematic practice (or boredom) effects.

How stable is SFON over time?

The test-retest reliability of each SFON task was calculated using Spearman rank-order correlation coefficients. Table 2.2 shows the correlation coefficients for each task, for preschool children, school children and both preschool and school children combined. Figure 2.7 presents bubbleplots of children's SFON scores at Time 1 and Time 2, on each SFON task. Here, the preschool and school children's data is presented together. Recall that these analyses are based on the combined method of scoring.
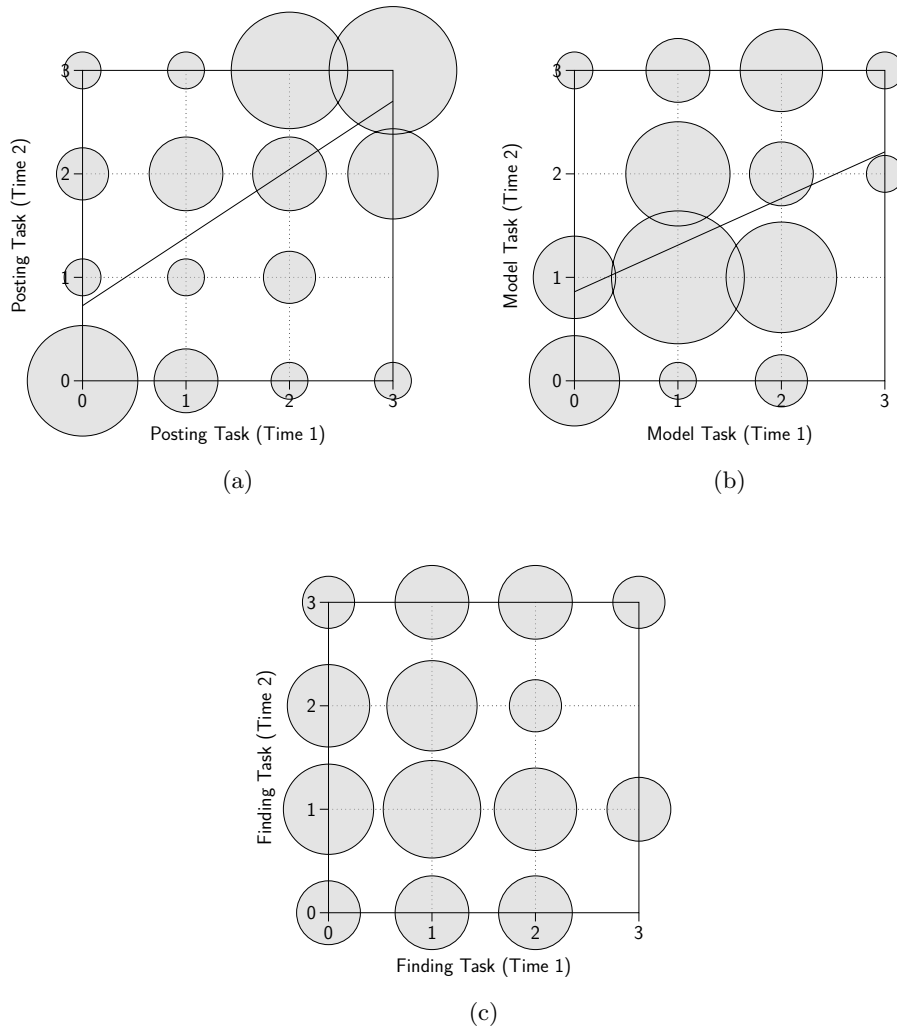
Figure 2.7: Bubbleplots of children's SFON scores at Time 1 and Time 2, for (a) the Posting Task, (b) the Model Task, and (c) the Finding Task.

Note. The area of each point indicates the relative number of children for each score.

	Preschool data	School data	All data
Posting Task	.578**	.500**	.590***
Model Task	.437*	.196	.365**
Finding Task	−.039	.076	.047

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

Table 2.2: Spearman (one-week) test-retest correlation coefficients for each SFON task.

Firstly, it is important to note that there were no significant differences between the correlation coefficients yielded by the school data and the preschool data, for any of the tasks. Following statistical guidelines laid out by Myers and Sirois (2006), the Spearman coefficients were treated as if they were Pearson coefficients and the standard Fisher's z-transformation was used to compare the two datasets (Posting Task, $z = −0.39$, $p = .70$; Model Task, $z = −0.97$, $p = .33$; Finding Task, $z = −0.41$, $p = .68$). This suggests that the test-retest reliability of the tasks does not vary with age and formal education.

Combining the school and preschool data together, the Spearman correlation coefficients ranged between .047 and .590, indicating that the tasks have zero to moderate test-retest reliability. The SFON task that showed the highest test-retest reliability was the Posting Task, followed by the Model Task. The correlation coefficients for these tasks were highly significant at $p < .001$ and $p = .005$, respectively. The Finding Task yielded the lowest correlation coefficient and thus the lowest SFON score stability.

Do the different SFON tasks measure the same SFON construct?

The convergent validity of the three SFON tasks was calculated using Spearman rank-order correlation coefficients. Table 2.3 presents the correlation coefficients between all tasks, at Time 1 and Time 2. There were no significant differences between the correlations yielded by the school and preschool datasets,[5] thus only the combined correlations are reported.

---

[5]A Fisher's z test was used to compare the correlations yielded by the school and preschool datasets at Time 1 (all $ps > .15$) and Time 2 (all $ps > .65$).

	Time 1	Time 2
Posting Task & Model Task	.255*	.360**
Posting Task & Finding Task	.164	.298*
Model Task & Finding Task	.117	.255

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

Table 2.3: Spearman correlation coefficients between all SFON tasks at Time 1 and Time 2.

Findings revealed a low, but nonetheless significant correlation between children's scores on the Posting Task and the Model Task, both at Time 1 ($r_s = .255$, $p = .046$) and Time 2 ($r_s = .360$, $p = .005$). These correlations were corrected for attenuation to control for the measurement error, or unreliability, associated with each task (by dividing them by the square root of the product of the two test-retest reliabilities)[6]. The resulting disattenuated correlations were considerably stronger: Time 1, $r_s = .550$; Time 2, $r_s = .776$. This suggests that these tasks were, certainly to some extent, measuring the same SFON construct. Children's scores on the Finding Task were not significantly correlated with their scores on the other two tasks (apart from a weak correlation with scores on the Posting Task at Time 2). This was not surprising given its poor test-retest reliability.

How does SFON relate to children's numerical skills?

Further nonparametric correlations were conducted to investigate the relationship between children's SFON and their numerical skills, as measured via teacher ratings (see Table 2.1). In line with the findings from Hannula and Lehtinen (2005) it was predicted that high SFON children would demonstrate more advanced numerical skills.

To control for age and nonverbal IQ, Spearman partial rank-order correlations were carried out with age and Block Design score as covariates. Table 2.4 presents the partial correlation coefficients between numeracy and SFON score on the Posting Task and the Model Task. The Finding Task was excluded from this analysis given its lack of test-retest reliability and

---

[6]See J. Cohen, Cohen, West, & Aiken, 2013, p. 56.

convergent validity. Again there were no significant differences between the correlations yielded by the school and preschool datasets,[7] therefore only the combined correlations are reported. Figure 2.8 presents bubbleplots depicting the relationship between children's SFON scores (on the Posting & Model Tasks) and their numeracy scores at Time 1.

Results showed some moderate positive correlations between SFON and numeracy, indicating that the more children spontaneously focused on numerosity, the better their numerical skills. Crucially, these correlations remained significant even after controlling for children's age and nonverbal IQ.

	Time 1	Time 2
Numeracy & Posting Task	.277*	.400**
Numeracy & Model Task	−.133	.290*

Note. Covariates: Age and Block Design score, *$p <$ .05, **$p <$ .01.

Table 2.4: Spearman partial correlation coefficients between numeracy and SFON (on the Posting & Model Tasks) at Time 1 and Time 2.

	Time 1	Time 2
Numeracy & Posting Task	.370*	.354*
Numeracy & Model Task	−.075	.321*

Note. Covariates: Age and Block Design score, *$p <$ .05, **$p <$ .01.

Table 2.5: Spearman partial correlation coefficients between numeracy and SFON (on the Posting & Model Tasks) at Time 1 and Time 2, excluding all children who could not reliably count to 10.

One could argue that because children's SFON scores were driven largely

---

[7]A Fisher's z test was used to compare the correlations yielded by the school and preschool datasets at Time 1 (all $ps >$ .11) and Time 2 (all $ps >$ .62).
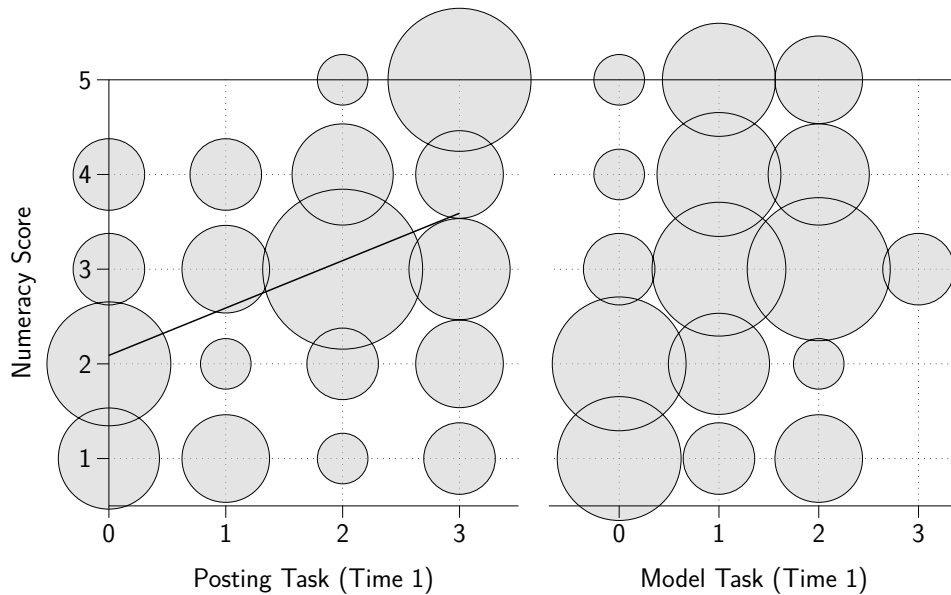
Figure 2.8: Bubbleplots depicting the relationship between numeracy scores and SFON scores (on the Posting & Model Tasks) at Time 1.

Note. The area of each point indicates the relative number of children for each score.

by their accuracy on the SFON tasks (as opposed to quantitative acts), SFON is simply a proxy for counting skills. To investigate this possibility, the partial correlations were re-run excluding those children who could not reliably count to 10 ($N = 14$). Note that all of the numerosities involved in the SFON tasks were less than 10; on the Posting and Model Tasks they ranged from 2 to 6.

Table 2.5 shows that the positive correlations between SFON and numeracy persisted, even after controlling for age and nonverbal IQ, and excluding those children with low counting abilities. This suggests that the SFON tasks were not simply measuring children's counting skills.

It is important to note some differences in the correlations yielded by the Posting Task and the Model Task. Children's SFON scores obtained from the Posting Task were significantly correlated with numeracy scores, both at Time 1 and Time 2, whereas children's SFON scores obtained from the Model Task were only significantly correlated with numeracy at Time 2; that is, after partialling out age and nonverbal IQ. Follow-up tests revealed

that these differences between the correlations at Time 1 and Time 2 were statistically significant; $t(23) = -8.28$, $p < .001$. There are two possible explanations for these differences: (1) the Model task may have been measuring something different across the two timepoints, or (2) there may have been noise in the data.

To try and distinguish between these two possibilities, further investigations into the types of errors that children made on the Model Task at Time 1 and Time 2 were carried out. It was reasoned that if this task was measuring something different at Time 1 and Time 2, then the types of errors children made (i.e. how far away they were from the correct numerosity) might have differed across the two timepoints.

An informal observation that the researcher made during the data collection process was that the Model Task was highly appealing to the children; as such, it appeared to demand a degree of inhibition. A number of children declared that they wanted to create their own dinosaur, rather than copying the researcher's model (in a couple of cases the child stamped the correct number of "nodes"/"spikes", paused, and then carried on stamping "nodes"/"spikes" all over the dinosaur's body). With this in mind, it was hypothesised that the differences in SFON scores obtained from the Model Task at Time 1 and Time 2 may have resulted from the novelty of the task at Time 1. If this was indeed the case, then we would expect to find that the children made larger (i.e. further away from the correct numerosity) errors at Time 1 than Time 2.

A chi-square test was performed to examine children's errors (near versus far) across the two time points. This test revealed no significant difference in the types of errors made, $\chi^2(1) = .81, p = .369$, suggesting that the differences in the Model Task at Time 1 and Time 2 were not due to varying levels of motivation and inhibition.

2.4   Discussion

The study presented in this chapter had three research aims: Firstly, to investigate the test-retest reliability and convergent validity of current SFON measures; secondly, to explore the effects of formal education on children's SFON tendency; and thirdly, to replicate and extend previous findings on the relationship between children's SFON and early numerical skills. Below,

I review the findings in relation to each of these aims.

With regard to the first aim, the results showed that the three SFON tasks varied in their test-retest reliability. The Posting Task demonstrated the highest SFON score stability ($r_s = .590$), followed by the Model Task which yielded a smaller, yet still significant, correlation between SFON scores at Time 1 and Time 2 ($r_s = .365$). In contrast to these two measures, the Finding Task showed poor test-retest reliability ($r_s = .047$). This indicates that the currently used version of this task is not an adequate measure of SFON, at least not in children aged $3 - 6$ years.

While the test-retest correlations of the Posting Task and the Model Task were both statistically significant (at $p < .01$), they were below the level usually considered acceptable. Typically, a measure which yields a test-retest correlation of at least 0.7 is deemed to be suitable for research purposes (Field, 2005). Despite this convention, it is widely acknowledged that when dealing with psychological constructs – as opposed to standardised intelligence tests, for example – we can expect to see correlations below 0.7. Many researchers have stressed the need to consider the specifics of each task. Cortina (1993) urged cautious interpretation of reliability figures given that reliability is often dependent on the number of items on a scale. On each of the SFON measures examined in this study children were presented with three trials and received a score from zero to three. The tasks sought to capture behaviour that was spontaneous therefore the number of possible trials was limited. With such a small scale we are less likely to see test-rest reliabilities of 0.7 and above. Arguably, the reliability figures observed here (on the Posting & Model Tasks) are adequate for these types of task. Certainly the Posting Task, which yielded the highest correlation between Time 1 and Time 2, can be considered to show sufficient SFON score stability.

After examining the test-retest reliability of each measure, the convergent validity between the tasks was assessed. Results revealed a significant correlation between children's scores on the Posting Task and the Model Task, both at Time 1 ($r_s = .255$) and Time 2 ($r_s = .360$). This suggests that these two tasks were, at least to some extent, measuring the same SFON construct. Again, the small magnitude of these correlations should be viewed in light of the small (4-point) scale of each of the measures. They should also be considered in view of the test-retest reliability figures; note

that if we correct for attenuation then these cross-task correlations are noticeably stronger. Overall, the Finding Task showed little correlation with either the Posting Task or the Model Task and thus it appears to provide neither a reliable nor a valid measure of children's SFON.

These findings may help to explain why previous research has found an inconsistent relationship between children's SFON scores over time. Recall that while Hannula and Lehtinen (2005) demonstrated a significant correlation between children's SFON scores at 4 and 5 years, they found no relationship between SFON at 4 and 5 years and SFON at 6 years. In this longitudinal study SFON was measured with different tasks across the various timepoints. Interestingly, at 4 and 5 years (where children's SFON showed significant stability) SFON was measured with the Posting Task, the task which yielded the highest score stability in the present study. At 6 years (where children's SFON showed a lack of stability) SFON scores were combined from the Posting Task, the Model Task and the Finding Task, three tasks that are shown here to vary in their reliability and validity. It is possible, therefore, that the stability of children's SFON at 6 years was masked by a lack of convergent validity between the tasks.

There is one important caveat to the interpretation of these findings. That is, there were some unforeseen differences between the tasks administered in the current study and the tasks administered by Hannula and Lehtinen (2005). In particular, the current version of the Model Task involved quantities that were visible to the child (the researcher's dinosaur was turned to face the child), whereas Hannula and Lehtinen's (2005) version of the Model Task involved quantities that were invisible or concealed (the researcher's dinosaur was turned over so that the child could not see it). Another difference to note is the age range of the children. In Hannula and Lehtinen's (2005) studies, the Model Task and the Finding Task were only used with children aged 6 years, whereas in the current study these tasks were used with younger children from 3 to 6 years. These differences mean that the methodological conclusions drawn the current study relate to the current versions of the tasks and not (necessarily) to the original versions developed by Hannula and Lehtinen (2005).

Turning now to the second aim, the results showed that the test-retest and intraclass correlations of the SFON measures were comparable across preschool and school-aged children. Previously, these SFON measures have

only been given to preschool samples. This is because children start school 3 years later in Finland where earlier research has been conducted (compared to the UK where the present study was carried out). The fact that there were no significant differences observed between the preschool and school datasets suggests that the psychometric properties of these SFON tasks do not vary with (age-related) development and exposure to formal schooling. SFON can be reliably measured – at least to some extent with two out of the three SFON tasks – and is equally stable in preschool and primary school children aged $3 - 6$ years. Note, however, that children's learning is mainly play-based at this early stage of schooling and therefore differences may emerge when classes become more formal in the later primary school years.

As well as comparing the correlations yielded by the preschool and school datasets, group-level analyses tested whether overall SFON tendency varied across these samples. These analyses produced mixed results. School children showed higher levels of SFON than preschool children on the Model Task, but not on the Posting Task or the Finding Task. In view of these mixed results, it is difficult to determine whether children's SFON tendency increases as they get older and start to receive mathematics education in school.

The inconsistencies observed here may reflect the fact that the three SFON tasks were not measuring exactly the same construct. We know that their intraclass correlations were relatively low, suggesting only limited convergent validity. Alternatively, it is possible that the tasks do tap into the same SFON construct, but that their extraneous task demands differentially affect the ease with which children can demonstrate this SFON. Recall that the tasks varied in terms of their working memory demands. The tasks involving invisible quantities (the Posting Task and the Finding Task) had a higher working memory demand than the task involving visible quantities (the Model Task). It is possible that the tasks also demanded varying levels of inhibition. The researcher observed that on the Model Task, the children had to inhibit their desire to make their own dinosaur. Given these extraneous factors, the tasks may vary in their suitability for children of different age groups.

Finally, in relation to the third aim, the results align with previous studies showing a positive association between children's SFON and their nu-

merical skills. The more children spontaneously focused on numerosity (on the Posting & Model Tasks), the higher their score on a teacher-based assessment of object counting, cardinality and simple arithmetic skills. Importantly, this relationship was significant, even after controlling for age and nonverbal IQ.

The relationship between SFON and numeracy appeared to depend on the SFON measure used. The Posting Task was significantly correlated with numeracy at both testing timepoints. Meanwhile, the Model Task was correlated with numeracy at Time 2, but not at Time 1; that is, after partialling out age and nonverbal IQ. It is difficult to ascertain whether this difference between Time 1 and Time 2 is a cause for concern. In other words, it is not clear whether the Model Task was measuring something different at Time 1, or whether this difference was due to statistical noise. Additional analyses sought to distinguish between these possibilities by comparing the types of errors children made on the Model Task at Time 1 and Time 2. These findings revealed no significantly different patterns, thus suggesting that the observed differences were likely a result of noise in the data. Either way, it seems reasonable to conclude that the Posting Task provided a better, more reliable measure of SFON than the Model Task.

A limitation of all of the SFON tasks examined here is that they failed to elicit many verbal or nonverbal quantitative acts. Children's SFON scores were driven largely by their accuracy (posting the correct number of letters in the case of the Posting Task, and stamping the correct number of "nodes" or "spikes" in the case of the Model Task) as opposed to their counting gestures and/or utterances. On the Posting Task at Time 1, for example, only 22% of children demonstrated any verbal or nonverbal quantitative acts. This figure was even lower for the Model Task at Time 1, with only 13% of children displaying overt numerical references.

This presents an issue because it means that children's SFON scores were dependent on them counting successfully. One could argue that some of the children with zero to low SFON scores were focusing on numerosity, but that they were making counting errors. One could further contend that SFON is merely a proxy for counting skills and therefore that explains its correlation with numeracy.

There are currently two lines of evidence to refute this account. Firstly, the present results demonstrated that the relationship between SFON and

numeracy persisted, even after those children with low counting skills (below the level needed to successfully complete the tasks) were excluded from the analysis. Secondly, previous results have shown that children's lack of SFON does not stem from a lack of procedural counting knowledge. Hannula and Lehtinen (2005) found that low SFON children could successfully complete the SFON tasks when they were guided to focus on numerosity (GFON). Together these findings suggest that SFON is not just a measure of counting knowledge itself.

Whilst this counter-evidence is compelling, the case in favour of SFON as a separate construct (as opposed to a proxy for numerical skill) would be even more convincing if children's SFON scores were independent of their counting accuracy. As such, there is a need to develop further SFON measures on which children receive SFON scores based on their quantitative acts rather than their production of the correct numerosity.

One way of doing this would be through observational methods. Another, much less time-intensive approach, would be to design some experimental tasks that require children to verbalise more. For example, a picture description task in which children are asked to describe a visual scene containing numerical and non-numerical aspects. This type of task has been used by Hannula and colleagues to measure SFON tendency in adults (Hannula et al., 2009); thus far, it has yet to be used with children. Two other possible tasks are: (1) a change-detection task in which children are asked to determine what aspects of a visual scene have been transformed (this may be number, colour, shape, size, etc.), and (2) a card sorting task in which children are asked to sort a collection of cards in any way that they wish (again this could be number, colour, shape, size, etc.).

These sorts of experimental tasks offer at least two potential advantages over the current SFON measures. Not only do they allow SFON to be assessed independent of counting accuracy, but they provide several competing dimensions on which one can choose to focus. On a picture description task, for example, children may focus on the number of items in the visual scene, the colour, shape or size of these items, or even the emotional valence of the picture. This contrasts with the Posting and Model Tasks on which there is little information to focus on other than the number of letters posted or the number of "nodes"/"spikes" stamped.

2.4.1   Limitations of Study 1

It is important to acknowledge the limitations of this study when making generalisations from these findings. First and foremost, there were some unforeseen differences between the SFON tasks administered in the current study and the SFON tasks administered by Hannula and Lehtinen (2005). Although the procedures reported in Hannula and Lehtinen (2005) were followed directly, it transpired that some aspects of the tasks had been interpreted in a different way to what had been intended. Details of these differences are noted throughout the chapter (see, for example, Section 2.2.3.1). Given these differences between the tasks, conclusions regarding reliability and validity can only be drawn in relation to the current versions of the tasks, and not necessarily to the original versions.

Moreover, conclusions do not necessarily generalise to other samples of children beyond the age range (3 to 6 years) studied in the current study. The children in this study were younger than the children in previous SFON studies (e.g. Hannula & Lehtinen, 2005) and therefore it is possible that they lacked sufficient skills needed to perform the tasks. Arguably this is not the case because children's performance (based on 'pure accuracy' scores) was not at floor – scores were normally distributed on all of the tasks in both preschool and school-aged samples (see Figures 2.4 to 2.6). Nevertheless, future studies would benefit from using GFON tasks to demonstrate that children's enumeration and cognitive skills are at a sufficient level (see Hannula & Lehtinen, 2005).

2.4.2   Summary of Findings

The present study examined the current measures used to assess children's spontaneous focusing on numerosity (SFON). Using a test-retest design it was shown that two out of three tasks designed to measure children's SFON have moderate test-retest reliability in both preschool and school-aged children. These two tasks were significantly correlated with each other and also with children's numerical skills. In sum, these results make two key contributions: (1) theoretically, they replicate and extend previous findings of a relationship between children's SFON and their counting and arithmetic skills (note that previously SFON has only been measured in preschool samples), and (2) methodologically, they demonstrate that current SFON mea-

sures vary in their psychometric properties and they highlight the need to develop a new SFON measure that is independent of children's counting skills. If we can replicate the relationship between children's SFON and their numerical skills with a SFON measure that is less driven by counting accuracy, then we will be better able to separate the attentional processes of SFON from children's counting skills per se.

Chapter 3

Developing a New SFON Measure (Studies 2 & 3)

In the previous chapter I used a test-retest methodology to examine the psychometric properties of tasks currently used to assess children's spontaneous focusing on numerosity (SFON). A key finding that emerged from this study was that children's scores on these SFON tasks were driven largely by accuracy (i.e. producing the correct numerosity), as opposed to verbal and nonverbal quantitative acts (i.e. counting gestures and utterances). As such, it is difficult to disentangle the attentional aspects of SFON from children's procedural counting skills. To address this issue, in the current chapter I introduce a new picture-based task that generates SFON scores from verbal number references rather than counting accuracy. Three studies are presented: First, a pilot study to develop and test the Picture Task procedure (Study 2: Pilot). Second, a test-retest study to investigate the reliability of the Picture Task (Study 2: Main). Third, an observational experimental study to examine the predictive validity of the Picture Task by correlating it with an ecologically valid measure (Study 3). This chapter follows on directly from Chapter 2, thus the section below provides only a brief review of the relevant background.

3.1 Introduction

We know from recent research conducted by Hannula and colleagues that children vary in the extent to which they spontaneously focus on the numeri-

cal aspects of their environment. These individual differences in SFON have been shown to predict a range of mathematical outcomes, from early number word sequence production and object counting skills (Hannula & Lehtinen, 2005), to later arithmetical problem-solving performance (Hannula et al., 2010). In view of these findings SFON is emerging as an important factor for explaining individual differences in the development of children's numerical skills.

As the theoretical significance of SFON increases, there is a need to further our methodological understanding of the way in which children's SFON behaviour is assessed. Chapter 2 highlighted the fact that we know little about the psychometric properties of the experimental tasks designed to measure children's SFON. Consequently, Study 1 used a test-retest design to examine the reliability and validity of current SFON tasks.

The results of Study 1 demonstrated that the three SFON measures developed by Hannula and Lehtinen (2005) varied in their test-retest reliability. The Posting Task and the Model Task showed moderate SFON score stability while the Finding Task yielded no significant correlation between SFON scores at Time 1 and Time 2. These findings suggest that the Finding Task does not provide an reliable index of children's SFON tendency, at least not in children aged $3 - 6$ years.

While the Posting Task and the Model Task showed adequate test-retest reliability (and convergent validity), these measures were not unproblematic. An important finding that emerged from Study 1 was that children's SFON scores on these tasks were driven primarily by counting accuracy, as opposed to verbal and nonverbal quantitative acts.[1] This makes it difficult to separate the attentional aspects of SFON from children's counting skills per se. It is possible that the low SFON children were focusing on numerosity but that they were making counting errors. As such, one could argue that SFON, as measured by these tasks, is merely a proxy for counting skills.

One way to refute this claim is to provide evidence from guided focusing on numerosity (GFON) tasks to show that low SFON children are capable of recognising the numbers of items involved (see, for example, Hannula, 2005).

---

[1]The majority of children's SFON scores were based on them producing the correct numerosity (i.e. posting the correct number of letters on the Posting Task or stamping the correct number of "nodes" or "spikes" on the Model Task) rather than them eliciting any counting gestures and/or utterances.

Alternatively, we can separate the attentional aspects of SFON from children's counting accuracy by developing a new SFON measure that generates SFON scores from verbal number references rather than counting accuracy. One possible measure is a picture description task in which children are asked to describe a visual scene containing numerical and non-numerical aspects. This type of task has been used previously to measure SFON tendency in adults (Hannula et al., 2009).

In a study which sought to explore the neural correlates of SFON, Hannula et al. (2009) showed adult participants a series of photographs and asked them to complete two tasks. First, they were asked to indicate, after each photograph, whether it contained a mammal or not. Second, after a run of photographs, they were asked to recall as much as they could about each of the photographs they had seen. The descriptions provided on this secondary task were recorded and analysed in terms of whether or not they contained any references to number. Importantly, this picture-based task can provide a measure of SFON that is independent of counting accuracy. Participants can be scored as focusing on numerosity if their description contains any reference to number, regardless of whether they have enumerated the items in the picture correctly. This enables SFON scores to be separated from counting skills per se.

Up until now, this type of picture-based task has only been used to measure SFON in adults. Therefore, Study 2 sought to explore whether it could be adapted for use with children.

3.2   Study 2: Pilot

The aim of this pilot study was to develop and test the new Picture Task procedure with a large sample of children aged $4 - 9$ years.

3.2.1   Method

3.2.1.1   Participants

One-hundred and thirty-two children (68 girls) aged $4.0 - 9.8$ years ($M = 6.3$ years, $SD = 1.5$ years) were recruited through the University of Notting-

ham's 'Summer Scientist Week' scheme[2]. 'Summer Scientist Week' is annual event, run during the school summer holidays, which brings together children and parents with researchers in developmental psychology. Children (and their parents) are invited into the university to take part in a half-day (3 hour) session of games and research studies. All studies are approved by the University of Nottingham School of Psychology Ethics Committee and all parents provide written consent for their child to take part. At the end of the session children receive a 'Summer Scientist Week' certificate and a small gift-bag to thank them for taking part.

Four children were excluded from the analyses because they were identified by their parents as having learning difficulties. A further three children had missing data, leaving a total of 125 complete datasets. The final sample with complete data consisted of 25 four-year-olds (16 girls), 27 five-year-olds (9 girls), 33 six-year-olds (15 girls), 20 seven-year-olds (11 girls), 8 eight-year-olds (7 girls) and 12 nine-year-olds (7 girls).

3.2.1.2   Design

This pilot study was carried out in conjunction with another study looking at children's (and parent's) mathematics anxiety (see Study 6 presented in Chapter 7). Prior to completing the tasks associated with the mathematics anxiety study, children were presented with two (pilot) trials of a newly developed SFON Picture Task. These SFON trials were presented first to ensure that children (and their parents) were not aware of the numerical nature of the study.

As part of the wider 'Summer Scientist Week' event children also completed the British Picture Vocabulary Scale (BPVS), a standardised measure of receptive vocabulary. This was presented by a different researcher as a separate activity.

3.2.1.3   Materials & Procedure

This section describes the experimental procedure for the new SFON Picture Task. All other tasks presented as part of the mathematics anxiety study are described in Chapter 7 (Section 7.2.3).

---

[2]http://www.summerscientist.org/
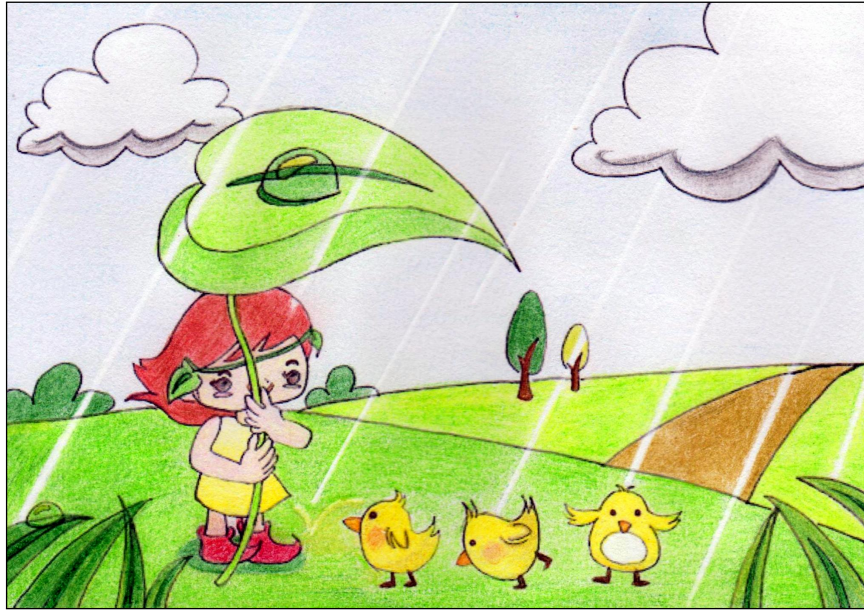
(a) Trial 1

(b) Trial 2

Figure 3.1: The pictures used in the first and second trials of the Picture Task.

Picture Task   This SFON task was programmed using E-Prime software 2.0 and was presented on a 15 inch touchscreen monitor attached to a laptop computer. Children were tested individually in a cordoned-off section of a university seminar room. Each child sat at a child-sized desk within comfortable reaching distance of the touchscreen monitor. The researcher sat to the left of the child (in front of the laptop computer) and parents, if present, sat behind the researcher so as not to distract their child.

The task began with the following instructions, presented on the computer screen and read aloud by the researcher: "This game is all about pictures. You will see a picture appear on the screen in front of you. After a short time this picture will disappear and you will see a question mark. When you see the question mark, your task is to tell me everything that you can remember about the picture." The researcher then checked that the child understood and prepared them for the first trial: "Ready? Steady? Go!"

On each of two trials a cartoon picture was presented for 7000ms, followed by a question mark until the researcher pressed the spacebar (on the laptop keyboard) to proceed. When the question mark appeared the researcher prompted the child to describe the picture: "Can you tell me what you saw?" The researcher then wrote down (with pen and paper) everything the child said. If the child was reluctant to speak, the researcher repeated her request: "Can you remember what you saw in the picture?" If the child spoke too quietly, the researcher asked them to speak a little louder. There was no time limit for children to respond. When the child finished the researcher asked: "Is that everything?" When the child was ready to move on the researcher proceeded to the next trial, or, after both trials, to a feedback screen with a smiley-faced gold star.

The pictures were presented in the same order for each child. Picture 1 (Figure 3.1a) showed a girl standing in the rain with a leaf umbrella and baby chicks. Picture 2 (Figure 3.1b) showed a boy and a girl in a hot air balloon with houses and trees below. Importantly, both pictures contained several small arrays (of objects, people or animals) that could be enumerated, for example: "three chicks" (Picture 1); "two children" (Picture 2). The set sizes of these arrays ranged $1-9$.

For each of the two trials children received a score of 0 or 1 depending on whether or not they spontaneously focused on numerosity. Children were

scored as spontaneously focusing on numerosity if their description contained any symbolic number words[3], regardless of whether they had enumerated the objects correctly. For example, if a child accurately described "three chicks" in Picture 1 they received a SFON score of 1. Likewise, if a child miscounted and described "four chicks" they too received a SFON score of 1. However, if a child described "some chicks" and made no other reference to number in their description then they received a SFON score of 0. Note that because SFON scores for each trial were binary, a child who mentioned number several times and a child who mentioned number only once both received the same score of 1. Each child received a total SFON score out of two.

3.2.2   Results

Descriptive statistics are first presented for children's SFON scores on the Picture Task. Next, group-level comparisons are carried out to examine the effects of age and gender on children's SFON scores on the Picture Task. Finally, further analyses are conducted to investigate the effect of children's verbal skills on their Picture Task performance.

Figure 3.2 shows the percentage of children in each age group spontaneously focusing on numerosity from zero to two times on the Picture Task. Visual inspection of this data shows that the distribution of scores varies across the different age-groups. SFON scores for the youngest children are at floor with 80% of the four-year-olds, and 59% of the five-year-olds, not focusing on numerosity on either of the two trials. In contrast, SFON scores for the older children show a slight tendency towards ceiling with 58% of the nine-year-olds focusing on numerosity on both possible occasions. These findings suggest that the Picture Task, in its current form, may vary in its suitability for use with children of different ages.

To formally examine the effect of age on Picture Task performance, SFON scores were compared across the age-groups with a nonparametric Kruskall-Wallis Analysis of Variance (ANOVA) on ranks. This revealed a significant difference in SFON scores across the age-groups, $\chi^2(5) = 24.55, p < .001$, reflecting the fact that older children demonstrated higher levels of

---

[3]These number words needed to be about the items in the picture and not about items, objects or events that were external to the task.
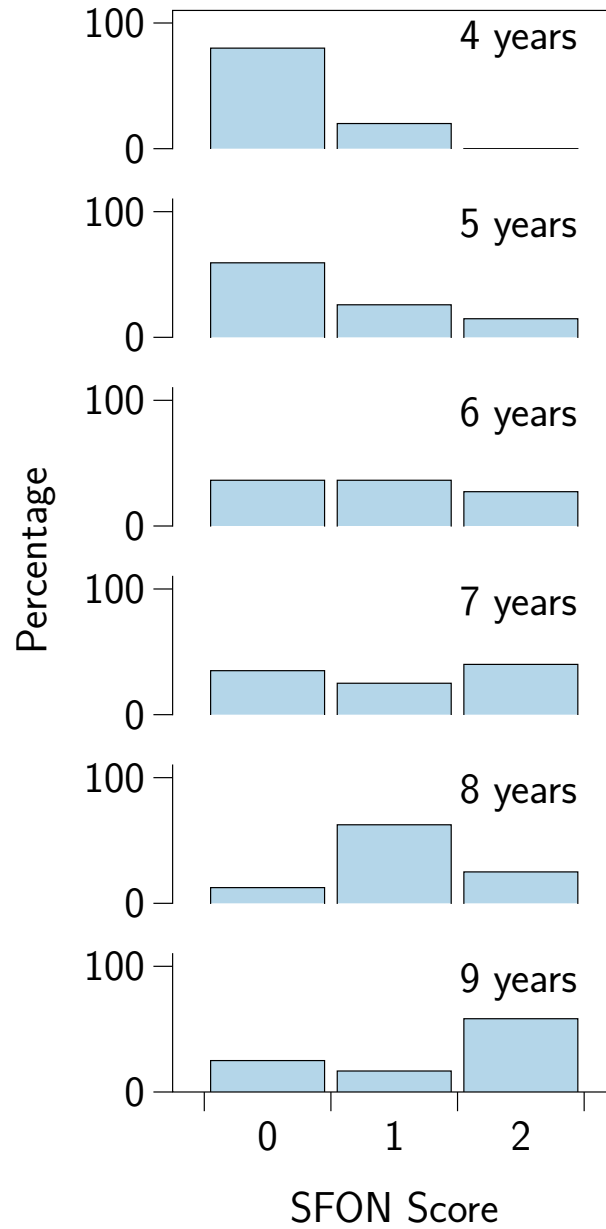
Figure 3.2: The percentage of children in each age group obtaining SFON scores from $0 - 2$ on the Picture Task.

SFON than the younger children.

Group-level comparisons were also made between boys and girls to examine the effect of gender on children's Picture Task performance. Collapsing across all age-groups, a nonparametric Mann-Whitney U test indicated that boys and girls showed no significant differences in their SFON scores on the Picture Task, $z = -0.32$, $p = .749$.[4] This suggests that boys and girls do not differ in their tendency to focus on the numerical aspects of their environment, at least not with this picture-based task.

The final set of analyses looked at the effect of children's verbal skills on their Picture Task performance. Since the Picture Task requires children to produce a verbal response, it is possible that SFON scores on this task are dependent on children's verbal abilities. To test this possibility two measures of verbal skills were examined: (1) children's scores on the BPVS, obtained separately as part of the wider 'Summer Scientist Week' event, and (2) the number of words children uttered on the Picture Task (indexed as an average across the two trials).[5] These measures were subjected to separate analyses, each presented in turn below.

A one-way Analysis of Covariance (ANCOVA) was conducted to compare BPVS scores in low, middle and high SFON children, while controlling for the (significant) effect of age that was previously demonstrated. SFON groups were determined by children's scores on the Picture Task; either 0, 1, or 2, respectively. The analysis used children's raw BPVS scores rather than standardised scores because children's scores on the Picture Task were raw. The results revealed a significant effect of age, $F(1, 119)=188.47$, $p < .001$, reflecting the fact that older children scored higher on the BPVS. There were no significant differences in BPVS scores between the low, middle and high SFON groups after controlling for this effect of age, $F(2, 119) = 1.05$, $p = .353$, suggesting that children's SFON scores on the Picture Task are independent of their receptive vocabulary skills.

A further one-way ANCOVA was performed to compare the average

---

[4]Note that separate analyses per age group also revealed no significant gender differences: four- and five-year-olds ($p = 1.000$); six- and seven-year-olds ($p = .734$); eight- and nine-year-olds ($p = .859$).

[5]There was a small significant correlation between children's BPVS (raw) scores and the average number of words they uttered on the Picture task, $pr = .20$, $p = .031$ (controlling for age).

number of words uttered on the Picture Task by low, middle and high SFON children, while controlling for the effect of age. This yielded a significant effect of age, $F(1, 119)$=32.13, $p < .001$, reflecting the fact that older children gave longer descriptions of the pictures. There were no significant differences in the number of words uttered by low, middle and high SFON children after controlling for this effect of age, $F(2, 119)$=0.23, $p = .796$. This indicates that children's SFON scores on the Picture Task are independent of the length of their picture descriptions.

3.2.3   Discussion

The aim of this pilot study was to assess the suitability of a newly developed picture description task for measuring children's SFON (Aim 1). The results suggest that this new Picture Task varies in its suitability for use with children of different ages. The four- and five-year-olds demonstrated large floor effects, therefore in its current form, the task does not appear to be appropriate for use with very young children. It is likely that the floor effects examined in this age-group were due to the memory demands of the task. Recall that children had to remember what they had seen in the picture rather than describing it whilst it was presented in front of them. In view of this, the experimental procedure of the Picture Task should be adapted for young children by eliminating the memory component.

Whilst the pilot results revealed significant differences in children's SFON scores on the Picture Task across different age-groups, there were no significant differences across gender. This suggests, in line with the findings from Study 1 (assessing current SFON measures), that there are no gender differences in terms of the extent to which children spontaneously attend to the numerical aspects of their environment.

Finally, the pilot results showed that children's verbal skills did not affect their SFON scores on the Picture Task. After controlling for the effects of age, low, middle and high SFON children showed no significant differences in receptive vocabulary (as measured by the BPVS) or average number of words uttered on the Picture Task. This finding is important because it demonstrates that the Picture Task provides a measure of children's SFON that is independent of their verbal skills.

3.2.3.1   Limitations of Study 2 (Pilot)

Given that this was a pilot study, the results should be treated with a degree of caution. The study used a convenience sample and whilst the overall sample size was relatively large ($N = 125$), there were different numbers of children in each age group. For example, there were over twice as many four- and five-year-olds ($N = 52$) as there were eight- and nine-year-olds ($N = 20$). These small and unequal groups make it difficult to carry out post-hoc comparisons across the individual age-groups. However, note that the purpose of this study was to develop and test the Picture Task procedure and not to test any theoretical predictions regarding age-related trends.

3.2.3.2   Summary

In sum, the results of this pilot study suggest that the Picture Task may provide an advantageous way of measuring children's SFON. This task generates SFON scores that are based on verbal number references, rather than counting accuracy, therefore it allows us to separate the attentional processes of SFON from children's procedural counting skills. As it stands, the experimental procedure of the task is only suitable for children aged 6 years and above, but with some minor modifications it may be used with younger children as well. In the main study that follows, the procedure is adapted for children aged $4 - 5$ years and the psychometric properties of the task are examined.

3.3   Study 2: Main

The following study had two aims:

1. To investigate the test-retest reliability and convergent validity of the Picture Task.

2. To examine the relationship between children's SFON scores on the Picture Task and their early numerical skills.

To address these aims, a test-retest experiment was carried out using a similar design to Study 1. Children aged $4 - 5$ years completed two SFON tasks; the Posting Task, developed by Hannula and Lehtinen (2005), and the newly developed Picture Task. These tasks were administered twice,

exactly one week apart. In a separate testing session children also completed a numerical task which tapped into the same numeracy skills measured by the teacher-based assessment used in Study 1.

3.3.1 Method

3.3.1.1 Participants

Thirty-eight children (22 boys) aged $4.2 - 5.3$ years ($M = 4.8$ years, $SD = 0.3$ years) were recruited from a primary school in Nottingham, UK. The school was of mid to high SES with fewer pupils receiving free school meals than the national average. All children were in the first term of their first year of school. Classes followed the EYFS framework described briefly in Section 2.2.1.

As with Study 1, participation was voluntary and the children received stickers to thank them for taking part. Written consent was obtained from all parents prior to the start of the study and children's assent was monitored closely during each testing session. Study procedures were approved by the Loughborough University Ethics Approvals (Human Participants) Subcommittee (see Section 2.1.4 for full details on ethical considerations).

Of the 38 children who took part in the study, four children were absent at the second time of testing, leaving a total of 34 complete datasets.

3.3.1.2 Design

In line with Study 1, a test-retest design was employed with an exact time interval of one week (± 60 minutes). During each of two sessions children completed two SFON tasks, the order of which was counterbalanced across participants. Children were presented with the tasks in the same order in each session. Following the SFON tasks, in Session 1 only, children completed the Block Design subtest of the WPPSI.

In addition to these measures children completed a short numeracy task. This data was collected in a third testing session in order to keep the numerical aspect of the study concealed (see Section 2.1.4 on ethical considerations). This session took place approximately one month later.

3.3.1.3   Materials & Procedure

SFON Tasks

Children were tested individually, in a corridor outside their classroom. The researcher was careful to ensure that the testing area was free from any numerical displays that may have prompted the children to focus on number. During the testing the children were not told that the tasks were in anyway numerical or quantitative. Likewise, the children's parents and teachers were not informed of the numerical aspect of the study; as in Study 1, they were told that the study was looking at general thinking skills.

In each of two testing sessions children completed two SFON measures: The Posting Task developed by Hannula and Lehtinen (2005) and the newly developed Picture Task adapted from Hannula et al. (2009). Throughout each of these tasks children received general praise (e.g. "You're watching really nicely") but no specific feedback was given. At the end of each task children were allowed to choose a sticker.

Posting Task   Children completed the same Posting Task used in Study 1. The experimental procedure was identical to that described in Section 2.2.3.1. The experimental materials and set-up of the task can be seen in Figures 2.1 and 3.3.

In accordance with Study 1 (and the procedure of Hannula & Lehtinen, 2005), there were three trials of the Posting Task. On each trial children received a score of 0 or 1 depending on whether they focused on numerosity or not. Children were scored as spontaneously focusing on numerosity if they posted the correct number of letters and/or if they presented any verbal or nonverbal quantitative acts (see Section 2.2.3.1). Each child received a total score out of three.

Picture Task   Children completed a modified version of the Picture Task developed and piloted in Section 3.2. In this modified version of the task the picture stimuli were presented on cards (16.0cm × 12.0cm) rather than on a computer screen and, crucially, the stimuli remained in front of the child so as to eliminate the memory demands.

The researcher introduced the task by saying: "This game is all about pictures. I'm going to show you a picture, but I'm not going to see the

Figure 3.3: The experimental set-up of the SFON Posting Task.

Figure 3.4: The picture used in the third trial of the Picture Task.

Note. The pictures used in the first and second trials are presented in Figure 3.1.

picture. Only you get to see the picture. This means I need your help to tell me what's in the picture." On each of three trials, the researcher held up a picture in front of the child and said: "What can you see in this picture?" The researcher wrote down everything the child said. If the child was reluctant to speak, the researcher repeated her request: "Can you tell me what you can see?" If the child spoke too quietly, the researcher prompted them to speak a little louder. There was no time limit for children to respond. When the child finished the researcher asked: "Is that everything?" When the child was ready to move on the researcher introduced the next trial: "Let's look at another picture. Ready? Steady? Go!"

There were three pictures in total and these were presented in the same order for each child. Picture 1 and Picture 2 were the same two stimuli presented in the pilot study (see Figure 3.1 in Section 3.2). Picture 3 showed a girl with a hat on holding a basket of flowers near the sea (see Figure 3.4). As in the pilot study all pictures contained several small arrays (of objects, people or animals) that could be enumerated, for example: "three chicks" (Picture 1); "two children" (Picture 2); "two butterflies" (Picture 3). The set sizes of these arrays ranged $1 - 9$.

For each of the three trials children received a score of 0 or 1 depending on whether they spontaneously focused on numerosity or not. As in the pilot study, children were scored as spontaneously focusing on numerosity if their description contained any mention of symbolic number, regardless of whether they had enumerated the array correctly. Each child received a total SFON score out of three.

Block Design

In order to control for individual differences in nonverbal reasoning ability children completed the Block Design subtest of the WPPSI. This was administered in accordance with the standardised procedure described in Section 2.2.3.2. Each child received a score out of 20.

Numeracy Skills

Children completed a 12-item (paper and pencil) numeracy task designed to measure the following skills: number word sequence production, numerical ordering, cardinality understanding and simple addition and subtraction.

Can you put the socks in order on the line?
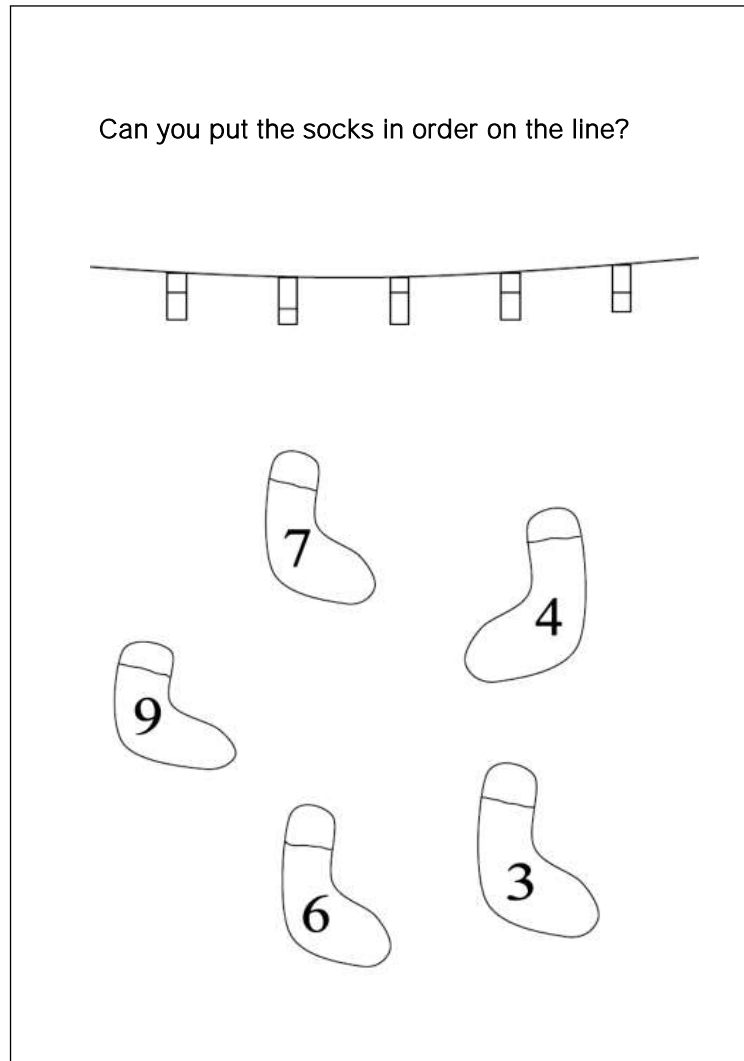
Figure 3.5: Example item from the numeracy task.

Figure 3.5 presents an example item assessing numerical ordering skills. A full list of items (the order of which was counterblanced across participants) is presented in Appendix A. This task was administered in small groups $(3 - 4$ children) in a corridor outside the classroom.

To allow comparison across studies, this task was designed to tap into the same numeracy skills measured by the teacher-based assessment used in Study 1 (see Section 2.2.3.3). Note that teacher ratings were not obtained in this study because children were new to the school (they were in the first half-term of their first year) and thus teachers were less able to provide reliable judgements. In accordance with Study 1, each child received a numeracy score $1 - 5$ (where $1 =$ the lowest numeracy score and $5 =$ the highest numeracy score) based on performance on the task. The criteria for each numeracy score can be seen in Chapter 2, Table 2.1.

3.3.2   Results

The results reported in this section are organised as follows: First, descriptive statistics are presented for children's performance on each of the SFON tasks. Next, preliminary analyses are conducted to check (i) whether there were any gender differences in children's SFON, (ii) whether children's SFON scores on the Picture Task differed depending on the length of their picture descriptions, and (iii) whether children's SFON scores changed between Time 1 and Time 2. Finally, in the main analyses section, correlations are performed to investigate the test-retest reliability and convergent validity of the newly developed SFON Picture Task (Aim 1), and the relationship between children's SFON and numerical skills (Aim 2).

3.3.2.1   Descriptive Statistics

Figure 3.6 presents the frequency of children spontaneously focusing on numerosity from zero to three times on each SFON task at Time 1. This demonstrates that children showed individual differences in SFON on both the Posting Task and the newly developed Picture Task. Importantly, children's scores on the Picture Task showed a relatively flat distribution, thus the modified procedure can be seen to have eliminated the floor effects found in the pilot study.
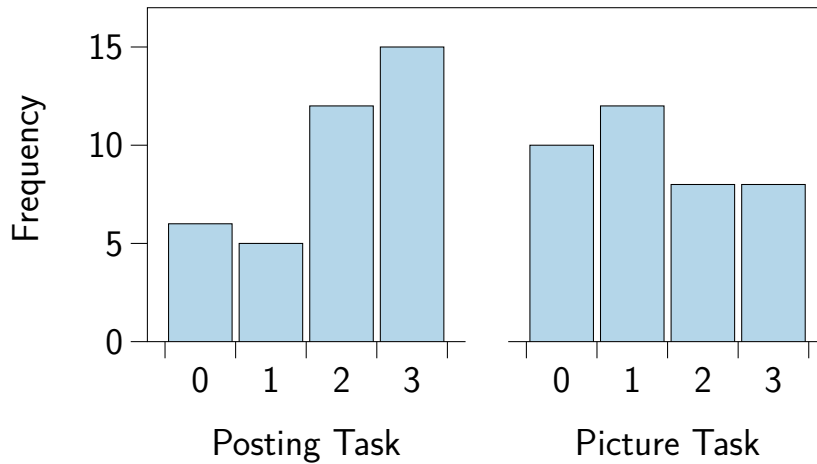
Figure 3.6: The frequency of children obtaining SFON scores from $0 - 3$ on the Posting Task and the Picture Task at Time 1.

3.3.2.2 Preliminary Analyses

Do boys and girls show different levels of SFON?

A series of Bonferroni-corrected Mann-Whitney U tests were performed to compare boys' and girls' SFON scores on each SFON task at Time 1 and Time 2. The results revealed no significant gender differences on the Posting Task or the Picture Task at either timepoint (all $ps > .219$). In line with the findings from Study 1 (and the pilot study), this indicates that boys and girls do not differ in their tendency to focus on the numerical aspects of their environment.

Do children's SFON scores on the Picture Task differ depending on the length of their picture descriptions?

A one-way ANOVA was performed to compare the average number of words uttered on the Picture Task (at Time 1) by children with different SFON tendencies (low, low-mid, mid-high or high). SFON groups were determined by children's scores on the Picture Task (at Time 1); either 0, 1, 2, or 3, respectively. The results yielded no significant differences in the number of words uttered by low, low-mid, mid-high or high SFON children, $F(3, 34) = 2.32$, $p = .092$, despite a trend towards high SFON children giving longer

descriptions. In accordance with the findings from the pilot study, this suggests that children's SFON scores on the Picture Task are independent of their verbal skills.

Do children's SFON scores differ between Time 1 and Time 2?

Wilcoxon signed ranks tests were performed to examine whether there were any changes in children's SFON scores across the two testing time-points. The results demonstrated no significant differences in children's SFON scores on the Posting Task or the Picture Task between Time 1 and Time 2: Posting Task, $z = -0.92$, $p = .358$; Picture Task, $z = -0.63$, $p = .528$. This indicates that there were no systematic practice (or boredom) effects.

3.3.2.3 Main Analyses

In a similar manner to Study 1, Spearman correlational analyses were performed to assess the psychometric properties of the SFON tasks and the relationship between children's SFON and numerical skills.

What is the test-retest reliability and convergent validity of the newly developed Picture Task?

The (one-week) test-retest reliabilities of the Posting Task and the Picture Task were calculated using Spearman rank-order correlation coefficients. Figure 3.7 presents bubbleplots of children's SFON scores at Time 1 and Time 2, on each SFON task. Results demonstrated significant test-retest correlations for the Posting Task ($r_s = .659$, $p < .001$) and the Picture Task ($r_s = .654$, $p < .001$), thus both tasks showed high SFON score stability.

The convergent validity between the Posting Task and the Picture Task was next examined. Figure 3.8 presents a bubbleplot depicting the relationship between children's SFON scores on Posting Task and the Picture Task (at Time 1). Results yielded a significant correlation between the tasks at Time 1 ($r_s = .329$, $p = .044$) and a marginally significant correlation at Time 2 ($r_s = .327$, $p = .059$). After taking into account the test-retest reliability of each task (by correcting for attenuation) these correlations were somewhat stronger: Time 1, $r_s = .501$; Time 2, $r_s = .498$. This suggests that the Posting Task and the Picture Task were, certainly to some extent, measuring the same SFON construct.
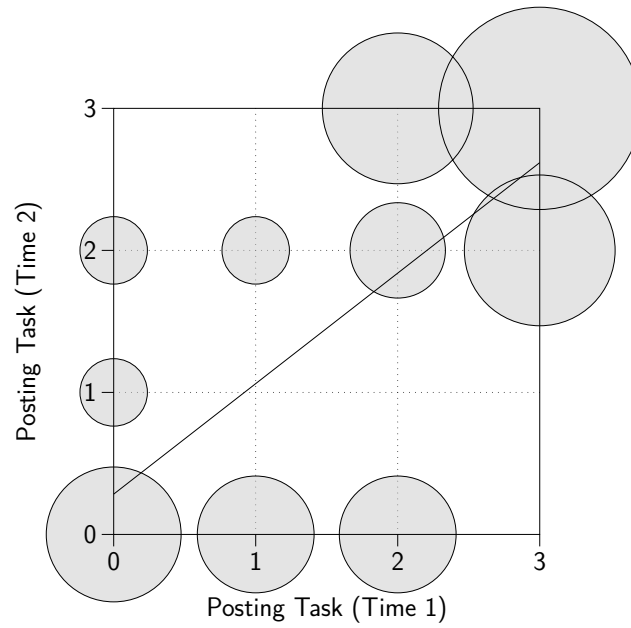
(a)
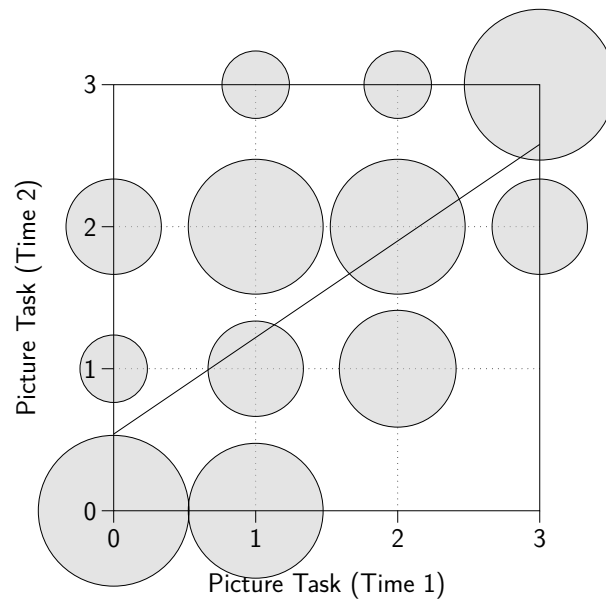
(b)

Figure 3.7: Bubbleplots of children's SFON scores at Time 1 and Time 2, for (a) the Posting Task and (b) the Picture Task.

Note. The area of each point indicates the relative number of children for each score.
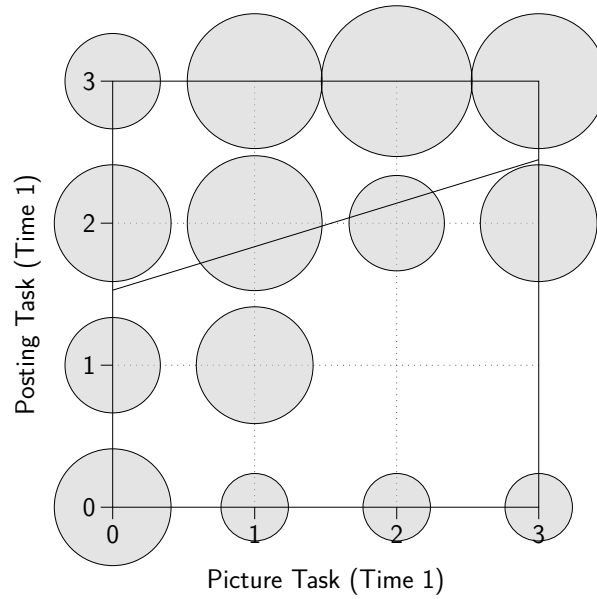
Figure 3.8: Bubbleplot depicting the relationship between SFON scores on the Posting Task and the Picture Task at Time 1.

Note. The area of each point indicates the relative number of children for each score.

What is the relationship between children's SFON and numerical skills?

Figure 3.9 presents bubbleplots depicting the relationship between children's numeracy scores and SFON scores (on the Posting & Picture Tasks) at Time 1. Spearman correlations revealed a significant positive correlation between children's numeracy score and SFON score on the Posting Task ($r_s = .344$, $p = .034$), thus replicating the results of Study 1. There was also a significant positive correlation between children's numeracy score and SFON score on the newly developed Picture Task ($r_s = .462$, $p = .004$). Together these findings indicate that the more children spontaneously focus on numerosity, the better their numerical skills.

Next, Spearman partial rank-order correlations were conducted to investigate whether these relationships remained significant after controlling for age, nonverbal IQ (Block Design score) and verbal skills (average number of words uttered on the Picture Task). Results demonstrated a significant correlation between children's numeracy and SFON on the Picture Task
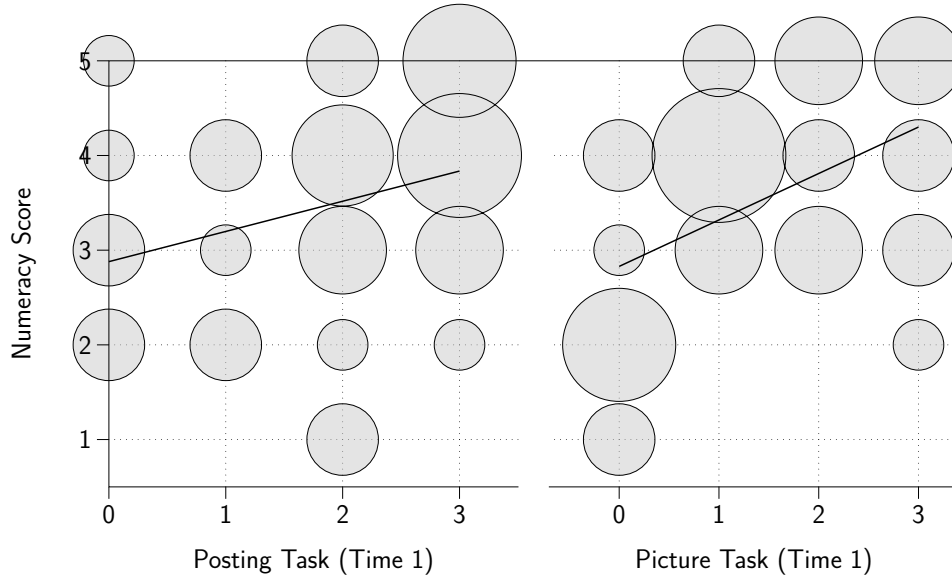
Figure 3.9: Bubbleplots depicting the relationship between numeracy scores and SFON scores (on the Posting & Picture Tasks) at Time 1.

Note. The area of each point indicates the relative number of children for each score.

$(pr_s = .400, p = .021)$, but the correlation between numeracy and Posting Task performance was no longer significant $(pr_s = .284, p = .109)$[6]. Therefore, the relationship between children's SFON and numerical skills is not accounted for, at least not for the Picture Task, by individual differences in age, nonverbal IQ and verbal skills.

3.3.3   Discussion

The present study used a similar test-retest design to Study 1 (Chapter 2) to examine the psychometric properties of a new picture-based task for measuring children's SFON. In each of two testing sessions, administered one week apart, children completed the Posting Task (developed by Hannula & Lehtinen, 2005) and a modified version of the Picture Task (developed and piloted in Section 3.2). This allowed me to compare the new picture-based

---

[6]Block Design was found to be the key covariate contributing to the relationship between numeracy and SFON on the Posting Task.

measure of SFON with the most reliable of the previously developed SFON measures (examined in Study 1).

The results showed that children's scores on the Picture Task were highly correlated between Time 1 and Time 2. This correlation was significant (at $p < .001$) and similar in magnitude to that of the Posting Task ($r_s = .654$ and .659 respectively). Therefore, the new picture-based task for measuring children's SFON demonstrates good test-retest reliability, comparable to that of the Posting Task measure.

The results also revealed some convergent validity between the Picture Task and the Posting Task. There was a small but nonetheless significant correlation between children's SFON scores on each task at Time 1 ($r_s = .329$) and a marginally significant correlation at Time 2 ($r_s = .327$). Note that after correcting for attenuation these correlations were slightly stronger. This indicates that the Posting Task and the Picture Task are, to some extent, measuring the same SFON construct.

The relatively small magnitude of these cross-task correlations may stem from three sources. First, they may be the result of different extraneous task demands. For example, the Posting Task required working memory resources (as it involved quantities that were invisible) and the Picture Task required verbal skills (as it involved verbal responses). Second, they may stem from the different SFON score criteria. SFON scores on the Posting Task were based primarily on counting accuracy whereas SFON scores on the Picture Task were based on verbal number references, regardless of accuracy. Third, they may be due to the small (four-point) scale of each of the measures and the relatively small sample size (at Time 1 $N = 38$; at Time 2 $N = 34$). As noted in Study 1, correlation coefficients are affected by the number of items per measure as well as the number of data points.

Future research may allow us to distinguish between these accounts by employing larger sample sizes and further controls for extraneous factors such as working memory. Recall that the present study did control for nonverbal IQ and verbal skills and, importantly, it was found that verbal skills did not moderate children's SFON scores on the Picture Task. In terms of the number of items per measure, these need to be kept to a minimum because the tasks seek to capture behaviour that is spontaneous.

As well as assessing the psychometric properties of the new Picture Task, the present study looked at the relationship between children's SFON and

their numerical skills. In accordance with Study 1, children received a numeracy score based on their number word sequence production, numerical ordering, cardinality understanding and simple addition and subtraction skills. The results showed a significant positive correlation between children's SFON scores on the Picture Task and their numerical skills, even after controlling for individual differences in age, nonverbal IQ and verbal skills. The magnitude of this relationship was similar to that found for the Posting Task (note, however, that the partial correlation for the Posting Task was not significant) and it aligned with the findings reported in previous SFON studies ($r_s \approx .4$). Importantly, this confirms that high SFON children show a numerical advantage, the theoretical implications of which are reviewed in the general discussion section of this chapter (Section 3.5).

3.3.3.1 Limitations of Study 2 (Main)

The results of this study should be interpreted within the limitations of the methodology. First, note that the sample is relatively small ($N = 38$) and it only includes children aged 4 to 5 years. As such, conclusions regarding the reliability and validity of the SFON tasks can only be generalised to this age group. In addition to this, note that the SFON tasks comprised only a small number of trials (three Posting Task trials and three Picture Task trials). This is a limitation not only in the current study but in SFON studies more generally. As stated previously, tasks designed to assess SFON seek to capture behaviour that is spontaneous (or unprompted) therefore trial numbers have to be kept to a minimum. The small number of trials limits the variance in SFON and is likely to reduce the effect sizes observed. One possible way of dealing with this is to administer multiple SFON tasks across a wide range of settings. However, before we can do this we need to establish that the tasks are reliable, and that thy are all measuring the same SFON construct.

3.3.3.2 Summary

Overall, the results of this study suggest that the newly developed Picture Task may be an effective tool for measuring children's SFON. It showed high test-retest reliability and moderate convergent validity with the previously developed Posting Task measure. Moreover it demonstrated predictive

validity because, in line with the literature, children's SFON scores were positively associated with their numerical skills. To further validate the Picture Task, it would be useful to establish whether it is predictive of children's SFON behaviour in their natural everyday environment. The next study (Study 3) addresses this issue by comparing children's SFON on the Picture Task with their SFON during child-parent play.

3.4   Study 3

3.4.1   Introduction

In order for an experimental task to be suitable for research purposes it needs to show robust psychometric properties. As discussed in Chapter 2, Section 2.1.1, it should demonstrate both measurement reliability and validity. In addition, where research is concerned with generalising findings to real-world contexts, it is important that the measures we use provide a good indicator of how participants would behave outside of an experimental setting. Behaviour on experimental tasks should predict behaviour on ecologically valid everyday activities. (This ecological validity is often defined as a type of external validity; see, for example, Bracht & Glass, 1968).

Issues of ecological validity are particularly relevant to the study of SFON because researchers are interested in the extent to which an individual spontaneously engages with numbers in their natural everyday environment. Importantly, the experimental tasks designed to assess SFON should correlate with (ecologically valid) measures of real-life 'spontaneous' behaviour. Therefore, the aim of the present study was to examine whether children's SFON scores on the newly-developed Picture Task relate to their SFON behaviour during child-parent play.

Thus far, only one study in the literature has directly compared children's SFON across experimental and play-based settings. This study, conducted by Edens and Potter (2013), used observational methods to examine preschool children's self-selected activities during free-time in the classroom. Children's activity choices were coded into several categories both mathematics-related (e.g. jigsaw puzzles and block construction) and non-mathematics-related (e.g. reading and painting). The findings from this study revealed no relationship between children's activity choices and their SFON tendency (measured using the pretend play activities developed by

Hannula & Lehtinen, 2005). Children who scored highly on the experimental SFON tasks were no more likely to select mathematics-related activities than their low SFON peers.

These findings question the assumption that high SFON children are more likely to engage with numbers in their natural environment. However, the results are not conclusive. There are at least three reasons to interpret them with caution. Firstly, children's overt choices of 'mathematics-related' versus 'non-mathematics-related' activities do not necessarily provide a proxy for their SFON during day-to-day routines. Numbers are not confined to specific mathematics-related activities; rather, they surround us everywhere. A child engaged in a non-mathematics-related activity may focus on number just as much as, or even more than, a child engaged in an activity that is considered to be mathematics-related. It is not clear from the findings presented by Edens and Potter (2013) how children's number talk varied across the different activities.

Further to this, it is important to note that Edens and Potter (2013) used a variety of pretend-play tasks (based on those of Hannula & Lehtinen, 2005) to measure children's SFON. We know from Study 1 (Chapter 2) that these task vary in terms of their reliability and validity, therefore the lack of correlation between children's task-based SFON and activity choices may stem from the psychometric properties of the tasks used. Alternatively, it may be due to the small sample size ($N = 14$). In view of these limitations there is a need for further research examining the relationship between children's task-based and play-based SFON. The present study sought to address this need in relation to the Picture Task developed in Study 2.

3.4.1.1 Aim of Study 3

As detailed above, the aim of this study was to examine the predictive validity of the new Picture Task by correlating it with behaviour on a task that has high ecological validity. To achieve this aim, experimental and observational methods were used to compare children's SFON on the Picture Task with their SFON during child-parent play.

Note that as well as examining the predictive validity of the Picture Task, this study aimed to investigate theoretical issues regarding the relationship between children's and parents' SFON and the role of home numeracy practices. These theoretical aspects are reported separately in Chapter 4.

3.4.2   Method

3.4.2.1   Participants

Fifty-six child-parent dyads were recruited through the University of Nottingham's 'Summer Scientist Week' scheme. Children (30 boys) were aged $4.0 - 5.9$ years ($M = 4.9$ years, $SD = 0.6$ years) and they came from a range of socioeconomic backgrounds (as measured by postcode). Roughly half of the children had just finished their first year of school ($N = 27$) and the other half were of preschool age ($N = 29$).

All parents provided written consent prior to the start of the study. Study procedures were approved by the University of Nottingham Ethics Committee (see Section 2.1.4 for full details on ethical considerations).

Four children were excluded from the analyses because English was not their native language ($N = 3$) or they were identified by their parents as having learning difficulties ($N = 1$). The final sample comprised 52 child-parent dyads, the gender distribution of which is reported in Table 3.1. Of these, 7 had missing data and were excluded from the individual analyses that included those data. Ns are reported for each variable in the results section.

	Daughter	Son
Mother	21 (40%)	23 (44%)
Father	4 (8%)	4 (8%)

Note. Of these 52 child-parent dyads, 49 parents were the child's primary caregiver.

Table 3.1: The gender distribution (frequency and percentage) of the child-parent dyads.

3.4.2.2   Design

This study used a mixed-method design with experimental tasks and structured observations. Children and parents took part in a single testing session lasting approximately 20 minutes. In the first half of the session, they were observed (video-recorded) whilst they played together with three different

sets of toys. In the second half of the session, they each (individually) completed a picture-based SFON task. Parents also filled in a home numeracy questionnaire.

As part of the wider 'Summer Scientist Week' event, 39 of the children also completed the BPVS, a standardised measure of receptive vocabulary. This was presented by a separate researcher as a separate activity during the half-day event.

3.4.2.3   Materials & Procedure

Children and parents were invited to take part in a study looking at children's play behaviour and general thinking skills. As with the previous SFON studies, they were not informed of the numerical aspect of the study (see Section 2.1.4.2). The study had two phases: first, an observational phase in which children and parents played together (Play SFON), and second, a task phase in which children and parents completed separate experimental tasks (Task SFON). Each of these phases are described in turn below.

Play SFON

Children and parents were observed (video-recorded) whilst they played three games/activities together:

1. Hungry Hippos. Participants were asked to play Hungry Hippos, a tabletop game in which players munch as many marbles as they can with their toy hippo/s. There were four hippos and 19 marbles (12 × red, 6 × silver and 1 × gold).

2. Lego Duplo. Participants were given a box of Lego Duplo and they were asked to construct one of three objects, either a boat, a car, or a house. Each object was presented on its own instruction card.

3. Picture Printing. Participants were presented with a collection of pictures that had missing pieces (e.g. a cat with no facial features). They were asked to choose a picture and fill in the missing pieces using different geometric-shaped stamps and coloured inks.

These games were chosen on the basis that they would be familiar and suitable for children and parents to play together. The researcher presented

each game sequentially, in the same order for each child-parent dyad. She remained in the room at all times but away from the area of testing (setting up later experimental tasks). Each game lasted approximately 3 minutes (Hungry Hippos $M = 2.41$ minutes, $SD = 0.95$ minutes; Lego Duplo $M = 3.40$ minutes, $SD = 1.29$ minutes; Picture Printing $M = 3.42$ minutes, $SD = 1.30$ minutes). Timings were variable to increase the naturalistic nature of the observations.

All play sessions were recorded on a Samsung Nexus smartphone placed discreetly on a windowsill so as not to distract participants. Video data was later coded in terms of the frequency of child-initiated SFON and parent-initiated SFON. Full details of this coding are presented in Section 3.4.3.1.

Task SFON

Children completed a SFON task with the researcher whilst parents completed a SFON task on a laptop computer. Parents wore headphones playing classical music so that they could not hear what their child was doing.

Child Picture Task   Children completed the Picture Task developed earlier in this chapter. The experimental procedure was identical to that used in Study 2 (Main) described in Section 3.3.1.3. On each of three experimental trials, children saw a cartoon picture and they were asked to verbally describe what they saw in the picture. The picture stimuli are presented in Figures 3.1 and 3.4. For each trial, children received a score of 0 or 1 depending on whether their description contained any reference to symbolic number. Each child received a total score out of three.

Parent Picture Task   Parents completed an adult version of the Picture Task. This was administered to address theoretical questions which are beyond the scope of the current chapter. Details of this task are reported in Chapter 4, Section 4.2 where these theoretical issues are explored.

Home Numeracy Questionnaire

Parents also completed a paper-based questionnaire designed to assess home numeracy (and literacy) experiences. As with the Parent Picture Task, this was administered to address wider theoretical questions and thus it is detailed separately in Chapter 4, Section 4.2.

3.4.3   Results

Before presenting the data analysis, I outline the data coding procedures for the video data.

3.4.3.1   Data Preparation & Reduction

All observations (video-recordings) of child-parent play were analysed to examine children's and parents' SFON behaviour. For each of the three games (Hungry Hippos, Lego Duplo & Picture Printing), play was coded in terms of child-initiated symbolic number references and parent-initiated symbolic number references.[7] A symbolic number reference was any mention of symbolic number (e.g. "I've got five") that was spontaneous or unprompted.

Table 3.2 illustrates the raw data coding for an example child-parent dyad. This tally (frequency) data was further reduced by giving children and parents a binary score of 0 or 1 for each game, depending on whether or not they focused on numerosity. The resulting binary scores were then summed across the three games. Children and parents received a total Play SFON score out of three. An example of this reduced data is presented in Table 3.3.

All video-recordings were coded by a single observer (myself). A second independent observer coded a random subset of 10% ($N = 5$) to establish inter-rater reliability. Overall, there were high levels of agreement between the coders. There was 100% concordance in the binary scores and a Pearson r correlation of .82 for the total (continuous) scores.

3.4.3.2   Data Analysis

The results are presented in three sections. To begin, descriptive statistics are reported for the Play SFON and Task SFON measures. Next, preliminary analyses are carried out to check for possible effects of age, gender and verbal skills on SFON. Finally, in the main analysis section, the correlation between children's Task SFON and Play SFON is examined to assess the relationship between the Picture Task and an ecologically valid measure.

---

[7]Play was also coded in terms of nonsymbolic number references (e.g. "You've got more") and counting references (e.g. "How many have you got?"), but these were found to be very infrequent and therefore were not analysed.

	Hippos	Duplo	Printing
Child-initiated SFON	‖‖	‖	
Parent-initiated SFON					‖

Table 3.2: An example coding table for one child-parent dyad.

	Hippos	Duplo	Printing	Total
Child-initiated SFON	1	1	0	2
Parent-initiated SFON	1	1	1	3

Note. This data corresponds with the example data presented in Table 3.2.

Table 3.3: An example of the reduced data for one child-parent dyad.

Descriptive Statistics

Figure 3.10 presents the frequency of children obtaining SFON scores from 0 to 3 during child-parent play (Play SFON) and on the Picture Task (Task SFON). This demonstrates that children showed individual differences in their play-based and task-based SFON. The frequencies of parents' SFON scores are presented and analysed separately in Chapter 4 where I examine the relationship between children's and parents' SFON.

Preliminary Analyses

● Do preschool and school-aged children show different levels of SFON?

Group-level comparisons were made between preschool and school-aged children to examine the effects of (age-related) development and formal schooling on children's SFON. Mann-Whitney U tests were carried out with the alpha level Bonferroni-corrected to .025 (.05/2). Results indicated no significant difference between preschool and school-aged children's play-based SFON ($z = -0.52$, $p = .601$). There was also no significant difference in terms of task-based SFON ($z = -1.82$, $p = .069$), despite a trend towards school-aged children obtaining higher scores on the Picture Task than
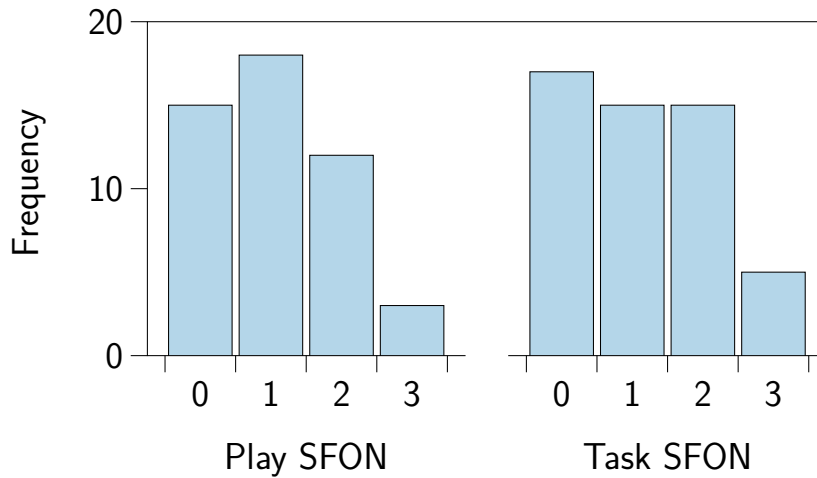
Figure 3.10: The frequency of children obtaining SFON scores from $0 - 3$ during child-parent play (Play SFON) and on the Picture Task (Task SFON).

Note. Play SFON, $N = 48$; Task SFON $N = 52$.

preschoolers. Overall there was, therefore, little evidence to suggest that SFON differed between preschool and school-aged children.

● **Do boys and girls show different levels of SFON?**

Further Bonferroni-corrected Mann-Whitney U tests were conducted to check for any gender differences in children's SFON. Results revealed no significant differences between boys' and girls' play-based ($z = -0.13$, $p = .896$) or task-based ($z = -0.43$, $p = .667$) SFON. In line with the findings from Studies 1 and 2, this indicates that boys and girls did not differ in their tendency to spontaneously focus on numerosity.

● **Do children's SFON scores differ depending on their verbal skills?**

Next, one-way ANCOVAs were carried out to examine the effect of children's verbal skills on their play-based and task-based SFON (while controlling for the effect of age). Children were categorised into SFON groups (low, low-mid, mid-high, high) based on their scores (0, 1, 2, 3) on the play-based and task-based measures respectively. Children's verbal skills were indexed by

their raw scores on the BPVS, a standardised measure of receptive vocabulary.

Results revealed no significant differences in BPVS scores between low, low-mid, mid-high and high SFON groups (after controlling for the effect of age): play-based SFON, $F(3, 31) = 1.71$, $p = .186$; task-based SFON, $F(3, 34) = 0.27$, $p = .850$. These findings align with the results of Study 2, providing further evidence that children's SFON is independent of their verbal abilities.

Main Analysis

● **How does task SFON relate to play SFON?**

A Spearman correlation analysis was performed to investigate the relationship between children's task-based and play-based SFON. Results revealed a significant positive correlation between SFON scores on the Picture Task and SFON behaviour during child-parent play, $r_s = .638$, $p < .001$. This correlation is depicted in Figure 3.11. The more children spontaneously focused on numerosity on the Picture Task the more (self-initiated) symbolic number references they made during child-parent play. This suggests that the Picture Task relates to an ecologically valid task.

3.4.4   Discussion

This study used a mixed methods design to examine the relationship between the Picture Task and an ecologically valid task. Children (and parents) took part in two phases of research: first an observational phase in which they played with different sets of toys, and second, a task phase in which they completed experimental SFON tasks. This allowed a comparison of children's SFON across experimental and play-based settings.

The results demonstrated that children's SFON scores on the Picture Task were highly correlated with their SFON during child-parent play ($r_s = .638$). This suggests that the Picture Task relates to an ecologically valid measure of children's (self-initiated) tendency to focus on numerosity in their natural everyday environment. The findings add to our limited understanding of children's SFON in play-based settings. Recall that to date, only one study has directly observed children's SFON during free-play and this
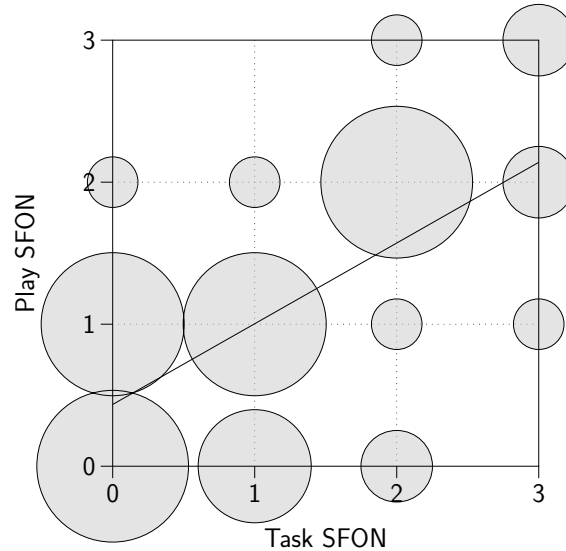
Figure 3.11: Bubbleplot depicting the relationship between children's SFON on the Picture Task and SFON during play.

Note. The area of each point indicates the relative number of children for each score.

study was limited because it measured activity choices (numerical versus non-numerical) rather than number talk (Edens & Potter, 2013).

3.4.4.1   Limitations of Study 3

In terms of the limitations of this study there are two key points to raise. First, it is worth mentioning that SFON on the Picture Task (Task SFON) and SFON during child-parent play (Play SFON) were both verbally-based measures. As such, the correlation between children's Task SFON and Play SFON may be confounded by children's verbal skills. To examine this possibility a partial Spearman's correlation analysis was performed with BPVS score as a covariate measure of verbal skills. This showed that the correlation between children's Task SFON and Play SFON remained highly significant, even after controlling for BPVS scores ($pr_s = .702$, $p < .001$). This suggests that the relationship between Task SFON and Play SFON is not confounded by verbal skills. Nevertheless, future studies should be aware of, and control for, this possible confound.

The second point to consider is the selected scoring criteria of the SFON measures. Recall that for each of the three trials of the Picture Task (and for each of the three games during child-parent play) children received a binary score of 0 or 1 depending on whether or not they focused on numerosity. Because SFON scores were binary, a child who mentioned number several times and a child who mentioned number only once both received the same score of 1. Each child received a total SFON score out of 3 for each of the SFON measures.

It is important to acknowledge the effect that this binary scoring might have on the results. Binary scores reduce the variance in the data and this may reduce the estimates of the effect sizes found. Crucially, if the analyses are re-run using children's total SFON scores (total number of times a child focuses on number across all trials) then the pattern of results is more or less the same. Note, however, that the distribution of these total scores is not normal and thus I chose to present only the binary scores.

3.4.4.2   Summary

Together with the results of Study 2 (Main) the findings from Study 3 indicate that the Picture Task is a reliable and valid tool for investigating children's SFON. The methodological advantages of this task are reviewed in the general discussion section that follows.

3.5   General Discussion

The research presented in this chapter advances (i) our methodological tools for examining children's SFON and, (ii) our theoretical understanding of the relationship between children's SFON and their early numerical skills. Below I first discuss the methodological contributions of this work before reviewing the key theoretical conclusions. Finally, I conclude the chapter with a summary of the findings and directions for future research.

3.5.1   Methodological Contributions

The present research introduced a new picture-based task for measuring children's SFON. Children were shown a cartoon picture and they were asked to describe what was in the picture. For each trial they were scored as focusing

on numerosity if their description contained any reference to exact symbolic number. Importantly, this Picture Task demonstrated high test-retest reliability and moderate convergent validity with the Posting Task measure (Study 2: Main). It also showed high predictive validity with respect to an ecologically valid measure of children's SFON during child-parent play (Study 3).

In contrast to the Posting Task and other previously developed SFON measures, the Picture Task generates SFON scores that are independent of children's counting accuracy. Children are deemed to be focusing on numerosity if they enumerate something in the picture, regardless of whether this enumeration is correct. Crucially, this allows us to tease apart the attentional aspects of SFON from children's counting skills per se; the theoretical significance of which is elaborated on in Section 3.5.2.

As well as providing a means to disentangle these cognitive processes, the Picture Task offers a number of potential advantages. One particular advantage is that there are several competing dimensions on which a child can choose to focus. Some children may focus on the number of items in the picture ("three houses") while others may focus on the colours in the picture ("bright blue sky") or the emotional content ("they look happy"). This differs from the Posting Task on which children can focus on little information other than the number of letters posted. To illustrate the diversity of children's responses on the Picture Task, three example quotations are presented below. Note that each of these examples are descriptions of Picture 1 presented in Figure 3.1a.

Example 1:

"A little girl and a leaf. There's two bushes and there's ducks. Three ducks. There's two ducks and one duck so that's how I know there's three ducks."

(Female, 4.6 years)

Example 2:

"It's a countryside. It's raining and there's a girl holding a leaf and a duck is looking at her. There are white clouds and there's a leaf not an umbrella."

(Male, 5.3 years)

Example 3:

"There's a little girl with a leaf to try and get warm because it's
raining. There's chicks in a line under the umbrella and a leaf
with a drop on."

(Female, 4.6 years)

In the first example the child is focusing almost entirely on the number of
objects while in the second and third examples the children are focusing on
the colour and form of the objects as well as the actions involved. Note that
in both Studies 2 and 3 there was no evidence for any gender differences
in children's SFON behaviour, a finding that aligned with the results of
Study 1. Given the range of visual aspects that children can spontaneously
choose to focus on here, the Picture Task may provide a better representation
of children's 'spontaneous focusing' in their everyday environment. For a
child to demonstrate SFON in a natural everyday setting, it is not just a
case of focusing on 'number' versus 'not number', rather they must choose
to focus on number over many other salient dimensions.

A further advantage of the Picture Task is that it is suitable for partic-
ipants of a wide age range. Unlike the pretend play activities of the previ-
ously developed SFON tasks, it may be adapted for use with older children
and adults. For example, the simple cartoon pictures could be substituted
for more complex visual scenes (see Chapter 4, page 105, for an example
picture-based task with adults). Importantly, this means that we can study
SFON throughout development in a simple and consistent manner.

Finally, the Picture Task benefits from being quick and easy to run.
While the Posting Task needs to be administered on a one-to-one basis, the
Picture Task may be flexibly administered in small or large whole group set-
tings. Here participants would be required to write down their descriptions
(rather than verbally responding), thus they would need to be of a certain
age.

Despite these potential advantages, the Picture Task is not without its
limitations. In view of the verbal requirements of the task it is only appropri-
ate for children who have developed verbal communication skills. It would

not be suitable for measuring SFON in infants or children with speech and language difficulties, and it may need to be used cautiously with bilinguals. That said, note that children's SFON scores on this Picture Task did not differ depending on their verbal abilities as indexed by (i) the average length of their picture descriptions, and (ii) their BPVS receptive vocabulary skills.

3.5.2   Theoretical Conclusions

The results of Study 2 (Main) replicated previous findings of a relationship between children's SFON and numerical skills. The more children spontaneously focused on numerosity on the newly developed Picture Task, the higher they scored on a measure of early numeracy. This replication is theoretically important because prior research employed tasks that were less able to tease apart the attentional aspects of SFON from children's counting skills per se. Note that here with the new Picture Task, children's SFON scores are based on verbal number references rather than counting accuracy. Therefore, we are better able to conclude that SFON is a distinct attentional process and not just a proxy for children's counting skills.

Given that the evidence for SFON is now more compelling, there is a need for further research to investigate the precise mechanisms of SFON. Why do high SFON children show a numerical advantage? And furthermore, where does children's SFON tendency come from? If we can address these questions then we can increase our understanding of how to support children's early numerical development.

In addition to these questions, it would be valuable for future studies to examine the developmental trajectory of children's SFON. The current literature suggests that SFON is an important factor in the preschool and early school years, but we know little about its role in later childhood – a time when children are exposed to more formal mathematics education. One interesting issue is whether we start to see any gender differences emerge. This is particularly relevant given recent concerns over women's participation in mathematics. We know from the vast literature on stereotype threat that societal $math = male$ biases can undermine women's mathematics performance (see Steele, Spencer, & Aronson, 2002 for a review); thus it is possible that as children grow older they become more aware of these societal stereotypes, and they may start to affect their SFON.

3.5.3   Summary of Findings

The present research developed and tested a new picture-based task for measuring children's SFON. This Picture Task offers an advantage over previously developed SFON measures because it generates SFON scores that are independent of children's counting accuracy; therefore it enables us to disentangle the attentional processes of SFON from children's counting skills per se. There were two key findings from the studies reported in this chapter: (1) the Picture Task showed good psychometric properties and was related to an ecologically valid task, and (2) children's SFON scores on the Picture Task were positively associated with their numerical skills. These findings strengthen the literature on SFON, verifying the need for further research into its developmental roots and causal mechanisms.

Chapter 4

The Developmental Roots of SFON (Study 3 continued)

The research presented in Chapters 2 and 3 demonstrated that SFON can be reliably and validly measured in preschool and school-aged children with two experimental tasks; namely, the Posting Task and the Picture Task. Now that the psychometric properties of these tasks have been established, I move on to focus on theoretical questions concerning the causes, or developmental roots, of SFON. In the current chapter I report further results from Study 3 which examine the relationship between children's and parents' SFON and the role of home numeracy practices.

4.1   Introduction

To recap, we know that children show individual differences in their tendency to spontaneously focus on the numerical aspects of their environment (SFON). These individual differences in SFON are relatively stable over time and they predict both current numerical skills and later arithmetical success. A natural question that arises from these findings is what causes some children to spontaneously focus on numerosity and others not?

One possibility is that differences in children's SFON tendency stem from differences in early home numeracy environment. High SFON children may be exposed to more numerical input at home than their low SFON peers. This possibility is theoretically plausible given the evidence that early environmental input is strongly associated with later cognitive development

(see, for example, Melhuish et al., 2008).

As highlighted in the Literature Review, Section 1.2.3.4, studies have demonstrated large variations in children's early home numeracy experiences. A recent observational study looking at parents' "math talk" found that whilst some parents produced as little as four number words in over 7.5 hours of child-parent interaction, other parents produced as many as 257 (Levine et al., 2010). Furthermore, intervention studies investigating the effects of playing number-board games have shown that children from low income backgrounds report only half as much experience playing number board games as their age-matched peers from higher income families. Many of these children report never having played a board game at home (Ramani & Siegler, 2008, 2011; Siegler & Ramani, 2008, 2009).

These variations in home numeracy experiences appear to have significant and lasting effects on children's mathematical outcomes. Levine et al. (2010) found that the frequency of parents' "math talk" (observed between 14 and 30 months) was positively related to children's later understanding of cardinality (at 46 months). Moreover, Ramani and Siegler (2008) showed that the more children played number board games at home, the more advanced their numerical knowledge; in particular, digit recognition, magnitude comparison, counting and number line linearity.

Research conducted by LeFevre and colleagues has emphasised the need to distinguish between different types of home numeracy practices. LeFevre et al. (2009) gave parents a home numeracy questionnaire in which they were asked to recall how often they had engaged with their child in various number-related activities in the past month. The findings revealed a distinction between parents' formal activities (directly teaching their child specific numerical skills, e.g. counting objects) and informal activities (indirectly involving their child in activities with numerical content, e.g. measuring ingredients whilst cooking). Overall parents reported higher levels of informal activities than they did formal activities and there was a strong positive relationship between the frequency of informal activities and children's mathematical skills. In a more recent study, Skwarchuk et al. (2014) found that parents' formal numeracy activities predicted children's symbolic number knowledge whilst informal numeracy activities predicted nonsymbolic numerical abilities.

The distinction between informal and formal activities may help to ex-

plain why some studies have found a relationship between home numeracy practices and numerical development (Blevins-Knabe & Musun-Miller, 1996; Huntsinger, Jose, & Larson, 1998; Huntsinger, Jose, Larson, Balsink Krieg, & Shaligram, 2000; LeFevre, Clarke, & Stringer, 2002; Saxe et al., 1987; Starkey et al., 1999) and others haven't (Blevins-Knabe, Austin, Musun, Eddy, & Jones, 2000; LeFevre, Polyzoi, Skwarchuk, Fast, & Sowinski, 2010). These mixed findings may also be due to the scarcity of parental numeracy-related activities reported in some studies (Blevins-Knabe et al., 2000). Note that researchers have found the frequency of home numeracy activities to be consistently lower than the frequency of home literacy activities (Anders et al., 2012; Blevins-Knabe & Musun-Miller, 1996; LeFevre et al., 2002, 2009; Skwarchuk, 2009).

In sum, there is increasing evidence to suggest that children's early home numeracy experiences are related to later mathematical achievement, and thus it seems plausible that children's SFON may stem from parental influences. Study 3 used a mixed methods design (with experimental tasks, structured observations and questionnaires) to test this possibility.

4.1.1   Aims of Study 3

In addition to examining the ecological validity of the Picture Task (see Chapter 3, Section 3.4) the aims of this study were as follows:

1. To investigate the relationship between children's SFON and their parent's SFON.

2. To explore the relationships between SFON and home numeracy experiences.

Based on previous findings showing a relationship between children's early home numeracy experiences and later numerical skills, it was hypothesised that children's SFON would be positively associated with (i) their parent's SFON (Prediction 1) and (ii) their home numeracy experiences (Prediction 2).

4.2   Method

The method of this study was reported previously in Chapter 3, Section 3.4.2, therefore in this section I provide only a brief overview (with full details of

the measures that were not previously described).

Fifty-six child-parent dyads (children aged $4 - 5$ years) took part in two phases of research: first, an observational phase in which they played together with different toys (Play SFON), and second, an experimental phase in which they completed individual SFON tasks (Task SFON). Parents also completed a home numeracy questionnaire.

The final sample comprised 52 child-parent dyads after excluding children with learning difficulties or English as a second language. Of these child-parent dyads, 85% of parents were mothers and 94% of parents were the child's primary caregiver. Moreover, 50% of the children had just finished their first year of school whilst the other 50% were of preschool age. The gender distribution of the sample is presented in Table 3.1 (page 89).

Play SFON

Children and parents were observed (video-recorded) whilst they played three games together (approximately 3 minutes each). Play was coded in terms of the frequency of child-initiated SFON and parent-initiated SFON. Full details of this coding are presented in Chapter 3, Section 3.4.3.1 (page 92).

Task SFON

Children completed a SFON task with the researcher whilst parents completed a SFON task on a laptop computer. Parents wore headphones playing classical music so that they could not hear what their child was doing.

Child Picture Task   Children completed the Picture Task developed in Chapter 3. The experimental procedure was identical to that used in Study 2 (Main) described in Section 3.3.1.3.

Parent Picture Task   Parents completed an adult version of the Picture Task (programmed using E-Prime software 2.0 and presented on a 15 inch LCD laptop screen). The experimental procedure was similar to that used in the pilot study of Study 2.

On each of three experimental trials, parents were presented with a photograph for 3500ms, followed by a question mark until response. When they saw the question mark their task was to write down (with pen and paper)

(a) Trial 1

(b) Trial 2

Figure 4.1: The photographs used in the first and second trials of the Parent Picture Task.

(c) Trial 3

Figure 4.1: The photograph used in the third trial of the Parent Picture Task.

a description of the photograph they had just seen. Once they had written their description they were instructed to press the spacebar on the laptop keyboard to continue.

The photographs were presented in the same order for each parent. Photograph 1 (Figure 4.1a) showed a student bar with women playing a game of pool. Photograph 2 (Figure 4.1b) showed a group of men running the marathon. Photograph 3 (Figure 4.1c) showed chefs preparing food in a kitchen. As with the child Picture Task, all photographs comprised several sets of objects or people that could be enumerated, for example: "five balls" (Photograph 1); "three cones" (Photograph 2); "four chefs" (Photograph 3).

In line with the child Picture Task, parents were scored as focusing on numerosity if their descriptions contained any reference to symbolic number. They received a score of 0 or 1 for each trial and thus a total score out of three.

Home Numeracy Questionnaire

Parents also completed a paper-based questionnaire designed to assess home numeracy (and literacy) experiences; this was adapted from LeFevre et al. (2009). This questionnaire comprised the following sections:

A. Background Information
   (E.g. Child's day-care provision)

B. Literacy Questions
   (E.g. Estimate the number of children's books in your household)

C. Numeracy Questions
   (E.g. How high can your child count?)

D. Toys & Games
   (E.g. Estimate the number of word-based games in your household)

E. Benchmarks
   How important is it for children to reach the following benchmarks before starting school? Five-point rating scale.
   (E.g. Count to 100)

F. Home Activities
   In the past month how often did you and your child engage in the following activities? Five-point rating scale.
   (E.g. Measuring ingredients when cooking)

G. Caregiver's Attitudes Towards Mathematics & Literacy
   Indicate the degree to which agree/disagree with the following statements. Five-point rating scale.
   (E.g. When I was in school I was good at mathematics)

A full list of items is presented in Appendix B and details on data coding and reduction are provided in Section 4.3.1 below.

4.3   Results

Before presenting the data analysis I outline the data coding procedures for the Home Numeracy Questionnaire.

4.3.1   Data Preparation & Reduction

The Home Numeracy Questionnaire included seven sections outlined in Section 4.2 (see also Appendix B). Items in the first sections (A – D) were scored individually while those in the later sections (E – G) were aggregated to form composite numeracy and literacy variables. The procedures for aggregating the data are described below for each of the relevant sections.

Section E - Benchmarks   Questions in this section measured parents' academic expectations for children's early numeracy and literacy achievement. Parents were asked to rate the importance (0 = not important to 4 = very important) of children reaching 8 benchmarks (4 numeracy and 4 literacy) before starting school. Two composite variables were formed:

- Parents' numeracy expectations (an average of the 4 numeracy items)

- Parents' literacy expectations (an average of the 4 literacy items)

The internal reliability for each composite variable was high: parents' numeracy expectations, Cronbach's $\alpha = .85$; parents' literacy expectations, Cronbach's $\alpha = .90$.

Section F - Home Activities   This section comprised a list of 27 home activities; 18 numeracy-related, 3 literacy-related and 6 general activities unrelated to numeracy or literacy. Parents were asked to indicate how often they had engaged in each of the activities with their child in the past month (0 = did not occur to 4 = almost daily).

First, frequencies were examined to identify whether there were any activities that rarely took place. Three activities were found to occur infrequently: wearing a watch, playing with calculators and "paint-by-number" activities. The percentages of parents who indicated that these activities never occurred were 69.5%, 60.0% and 61.0%, respectively. These items were deemed to be not applicable to the sample and thus were removed from data aggregation. The general activities, included as a control to increase the range of activities, were also omitted from data aggregation. In accordance with LeFevre et al.'s (2009) analysis, five composite variables were formed:

- Number Skills (an average of 7 items)

  - Identifying names of written numbers; counting objects; sorting things by size, colour or shape; counting down; learning simple sums; playing with number fridge magnets; printing/writing numbers.

- Number Books (an average of 3 items)

  - "Connect-the-dot" activities; using number activity books; reading number storybooks.

- Number Games (an average of 3 items)

  - Playing card games; being timed; playing board games with die or spinner.

- Number Applications (an average of 3 items)

  - Measuring ingredients when cooking; using calendars and dates; talking about money when shopping.

- Letter Activities (an average of 3 items)

  - Identifying names of written alphabet letters; identifying sounds of alphabet letters; printing/writing letters.

The four numeracy composites separated parents' formal activities (directly teaching their child specific numerical skills) from their informal activities (indirectly involving their child in activities with numerical content). The number skills and number books composites can be considered indices of children's direct exposure to specific numerical skills. Meanwhile, the number games and number applications composites can be thought to index children's indirect exposure to general numerical content.

The following composite variables showed moderate to good internal reliability: number skills, Cronbach's $\alpha = .74$; number books, Cronbach's $\alpha = .68$; number applications, Cronbach's $\alpha = .62$; letter activities, Cronbach's $\alpha = .87$. The number games composite demonstrated poor internal reliability, Cronbach's $\alpha = .46$, therefore it was not included in the analyses.

Section G - Caregiver's Attitudes Towards Mathematics & Literacy In this section parents were asked to indicate the extent to which they

agreed (1 = strongly disagree to 5 = strongly agree) with 9 statements relating to their attitudes towards mathematics and literacy. Two composite variables were formed:

- Parents' attitudes towards mathematics (an average of the 5 mathematics items)

- Parents' attitudes towards literacy (an average of the 4 literacy items)

Each composite variable showed good internal reliability: parents' attitudes towards mathematics, Cronbach's $\alpha = .77$; parents' attitudes towards literacy, Cronbach's $\alpha = .80$.

4.3.2   Data Analysis

The results are organised as follows: First, descriptive statistics are reported for the parental SFON measures and the Home Numeracy Questionnaire. Next, preliminary analyses are carried out to check for (i) differences between home numeracy and literacy variables, and (ii) differences in home numeracy and literacy practices for sons and daughters. Finally, in the main analyses section, correlations are performed to investigate the relationships among child and parent SFON (Aim 1), and SFON and home numeracy factors (Aim 2).

4.3.2.1   Descriptive Statistics

Figure 4.2 presents the frequencies of parents' SFON scores during child-parent play (Play SFON) and on the Picture Task (Task SFON). The frequencies of children's SFON scores were presented previously in Chapter 3 (page 94). Recall that for each SFON measure both parents and children received a total score from 0 to 3. For the Play SFON measure this depended on the number of games (out of three) on which they initiated symbolic number talk. For the Task SFON measure this depended on the number of picture trials (out of three) on which they referred to symbolic number.

It is clear from Figure 4.2 that parents showed individual differences in their (self-initiated) tendency to focus on numerosity during child-parent play. Parents also showed individual differences in their tendency to focus on numerosity on the Picture Task. Here there was a trend towards ceiling (48% of parents focused on numerosity on all trials). Note, however, that if
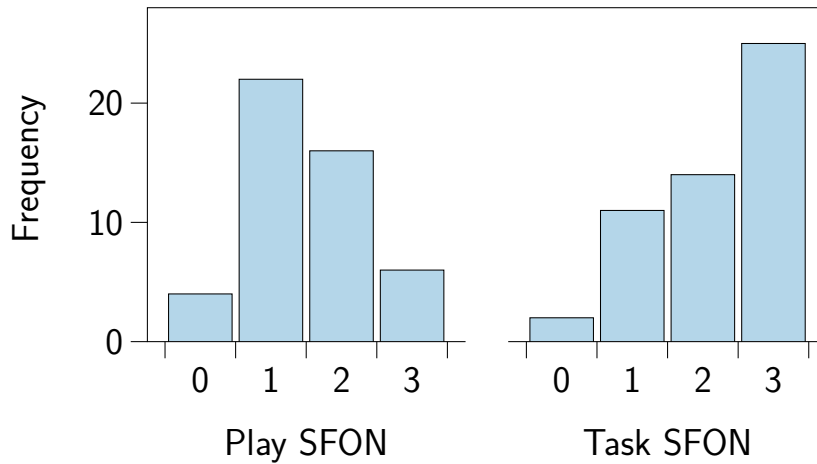
Figure 4.2: The frequency of parents obtaining SFON scores from $0-3$ during child-parent play (Play SFON) and on the Picture Task (Task SFON).

Note. Play SFON, $N = 48$; Task SFON $N = 52$.

we look at the total number of references to number, rather than assigning a binary score of 0 or 1 for each trial, then parents' scores show a normal distribution ($M = 4.46$, $SD = 2.93$). Therefore, the results yielded from the parent analyses are checked using the total (continuous) SFON data in addition to the binary data for each trial.

The descriptive statistics for variables from the Home Numeracy Questionnaire are presented in Table 4.1. Parents' estimates of the number of books and games in the home varied widely, in keeping with the broad SES range of the sample. Parents' academic expectations and home activities also showed large individual differences, with responses covering the entire range on the five-point ($0-4$) scale. In terms of parents' attitudes towards mathematics and literacy, responses varied more widely for mathematics than they did for literacy. Attitudes towards literacy were highly positive ($M = 4.40$, $SD = 0.64$) with a range from $3-5$ on the five-point ($1-5$) scale.

	N	Mean	SD	Range
Literacy questions:				
• Child books	51	136	111	$25 - 500$
• Adult books	51	183	232	$0 - 1000$
Numeracy questions:				
• Child's counting (how high?)	51	137	226	$5 - 1000$
Toys & games:				
• Word-related games	51	6.14	4.60	$0 - 20$
• Number-related games	51	9.59	6.34	$1 - 30$
Parents' academic expectations:				
• Numeracy	51	2.28	0.99	$0.00 - 4.00$
• Literacy	50	2.61	1.05	$0.00 - 4.00$
Home activities:				
• Number skills	49	2.17	0.76	$0.29 - 3.43$
• Number books	50	1.61	0.84	$0.00 - 3.33$
• Number applications	50	1.74	0.86	$0.00 - 3.33$
• Letter activities	51	3.08	1.01	$0.67 - 4.00$
Parents' attitudes towards:				
• Mathematics	51	3.51	0.82	$1.40 - 5.00$
• Literacy	51	4.40	0.64	$3.00 - 5.00$

Note. Nominal variables are not included in the table.

Table 4.1: Descriptive statistics for variables from the Home Numeracy Questionnaire.

4.3.2.2   Preliminary Analyses

Do parents' academic expectations, home activities and attitudes differ for numeracy and literacy?

Previous studies have shown that home literacy practices are more prevalent than home numeracy practices, thus preliminary analyses were carried out to examine whether there were any differences between the home numeracy and literacy variables. First, a paired-samples t-test compared parents' academic expectations for children's early numeracy and literacy achievement. This revealed a significant difference between numeracy and literacy expectations, $t(49) = -5.36$, $p < .001$, reflecting the fact that parents had higher expectations for children's literacy achievement ($M = 2.61$) than they did for children's numeracy achievement ($M = 2.28$).

Further paired-samples t-tests yielded significant differences in the frequency with which parents engaged in numeracy and literacy activities with their child: identifying written letters ($M = 3.10$) occurred more often than identifying written numbers ($M = 2.54$), $t(49) = 3.06$, $p < .004$, and printing/writing letters ($M = 2.84$) occurred more often than printing/writing numbers ($M = 2.26$), $t(49) = 3.97$, $p < .001$.

In addition to these differences in academic expectations and home numeracy activities, parents' attitudes towards literacy ($M = 4.40$) were found to be more positive than their attitudes towards mathematics ($M = 3.51$), $t(50) = -6.20$, $p < .001$. Together these findings align with previous studies providing consistent evidence to suggest that home literacy environments may be richer than home numeracy environments.

Do parents' academic expectations and home activities differ for sons and daughters?

In an exploratory manner, a series of independent samples t-tests were conducted to investigate whether there were any differences in home numeracy and literacy practices reported by parents of sons and daughters. Results revealed no significant differences in the numeracy or literacy expectations of parents of sons and daughters ($ps > .263$). Likewise, there were no significant differences in the frequency with which parents of sons and daughters engaged in numeracy and literacy activities ($ps > .148$), apart from a marginally significant difference in parental activities involving direct teach-

ing of number skills, $t(47) = 1.96$, $p = .057$. This marginal difference reflects the fact that parents of sons engaged in more number skills activities ($M = 2.38$) than parents of daughters ($M = 1.96$).

4.3.2.3 Main Analyses

How does parental SFON relate to child SFON?

Spearman correlation analyses examined whether children's SFON behaviour was related to their parent's SFON behaviour (Prediction 1). Table 4.2 presents the correlation coefficients between children's and parents' task-based and play-based SFON. Results revealed no significant correlations which indicates that children's tendency to focus on numerosity is not related to their parents' tendency to focus on numerosity. The 95% confidence intervals for these correlations were calculated based on the Fisher r-to-z transformation (see Table 4.2). All confidence intervals included zero, thus providing stronger support for the null hypothesis that children's SFON is not related to their parents' SFON.

	Parent Task SFON	Parent Play SFON
Child Task SFON	.073 [-.203, .339]	−.145 [-.412, .145]
Child Play SFON	−.036 [-.316, .250]	−.242 [-.492, .045]

Note. Correlation coefficients are presented for the binary SFON scores. Results are similar using the total scores.

Table 4.2: Spearman correlation coefficients [and 95% confidence intervals] between children's and parents' task-based and play-based SFON.

It is important to note that of the 52 child-parent dyads in the sample, only 3 parents were not the child's primary caregiver and if these parents are excluded from the analyses then the results are more or less the same. The results are also similar if we look separately at (i) sons and daughters, and (ii) preschool-aged children and school-aged children. This suggests that there is no relationship between children's SFON and parents' SFON regardless of the child's gender and exposure to formal numerical instruction.

How do home experiences relate to SFON?

Further Spearman correlations were performed to investigate the relationship between SFON and home numeracy (and literacy) factors. Table 4.3 presents a correlation matrix for all variables and composite indices.

It was hypothesised that children's SFON would be positively related to their home numeracy experiences (Prediction 2). Contrary to this hypothesis, there were no significant correlations between children's SFON (task-based or play-based) and parents' reports of engagement in home numeracy activities, apart from a negative correlation between children's task-based SFON and the frequency of number book activities, $r_s = -.340$, $p = .016$.[1] This negative correlation shows that the more parents and children engaged with number book activities (e.g. reading number storybooks), the less likely children were to spontaneously focus on numerosity. There was also a negative correlation between children's play-based SFON and parents' reports of the frequency of letter activities, $r_s = -.332$, $p = .023$; the higher the frequency of letter activities at home (e.g. printing/writing letters), the lower children's SFON.[1]

The results also demonstrated no significant associations between children's SFON and parents' numeracy expectations, attitudes towards mathematics and the amount of number-related games in the household. Likewise, parents' SFON (task-based and play-based) was not related to any of the home numeracy variables, apart from a positive correlation between parents' task-based SFON and their estimate of how high their child could count, $r_s = -.335$, $p = .016$.[1] This positive correlation reflects the fact that the more parents spontaneously focused on numerosity, the higher they estimated their child's counting skills.

Whilst there was little evidence for a relationship between children's SFON and home numeracy factors, there was a small positive correlation between children's task-based SFON and their SES, $r_s = .279$, $p = .050$.[1] Children from higher socioeconomic backgrounds were more likely to focus on numerosity on the Picture Task.

---

[1]This correlation did not survive Bonferroni-correction for multiple tests, but showed a trend towards significance.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Background variables																			
1 SES[a]	–																		
2 BPVS	.34*	–																	
Task-based SFON:																			
3 Child	.28*	.19	–																
4 Parent	.24†	.26	.07	–															
Play-based SFON:																			
5 Child	.20	−.12	.64***	−.04	–														
6 Parent	.03	−.11	−.15	.03	−.24†	–													
Literacy questions:																			
7 Child books	.24†	.24	.25†	.23	.20	−.04	–												
8 Adult books	.01	.47**	.15	.17	.03	.01	.46**	–											
Numeracy questions:																			
9 Child's counting (how high?)	.08	.31†	.23	.34*	.27†	−.13	.10	.08	–										
Toys & games:																			
10 Word-related games	.23	−.01	.18	.06	.22	−.22	.37**	.25†	.17	–									
11 Number-related games	.00	−.10	.21	.15	.14	−.15	.40**	.04	.24†	.42**	–								
Parents' academic expectations:																			
12 Numeracy	−.16	−.10	−.14	.07	.11	−.10	−.05	−.17	.15	.13	−.02	–							
13 Literacy	−.13	−.02	−.22	.10	.01	−.06	.04	−.01	.05	.11	−.03	.86***	–						
Home activities:																			
14 Number skills	−.06	.03	−.14	.07	−.08	−.07	.12	.07	.15	−.08	.11	.18	.16	–					
15 Number books	−.04	.09	−.34*	.06	−.28†	.17	−.17	−.02	.01	.02	.09	.05	.00	.42**	–				
16 Number applications	.00	.06	−.07	−.25†	−.05	−.18	−.09	.09	−.10	.02	.05	−.03	−.13	.49***	.43**	–			
17 Letter activities	−.15	.34*	−.25†	.23†	−.33*	.02	.18	.27†	.21	−.07	−.13	.07	.16	.60***	.35*	.15	–		
Parents' attitudes towards																			
18 Mathematics	.07	.25	−.07	.20	−.01	.01	−.05	.09	.38**	.12	.11	.06	.02	.16	−.03	.15	.26†	–	
19 Literacy	−.24†	.03	−.18	−.03	−.17	.29*	.19	.12	−.06	.05	−.08	.12	.12	.12	.14	−.12	.29*	.09	–

Note. †$p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$.
a. Higher numbers represent higher SES.

Table 4.3: Spearman correlation coefficients between all variables and composite indices.

In terms of the home literacy variables and children's literacy outcomes, two relationships were uncovered. Children's receptive vocabulary (BPVS) scores were positively associated with (i) the number of adult books in the home, $r_s = .469$, $p = .003$, and (ii) the frequency with which parents and children engaged in letter activities, $r_s = .340$, $p = .037$.[1] The more adult books in the home and the higher the frequency of letter activities, the higher children's BPVS scores.

Finally, it is important to point out that there were no significant correlations between parents' academic expectations and their reports of home numeracy and literacy practices. This contrasts with previous findings showing a positive relationship between parents' expectations in numeracy and literacy and the frequency with which they reported formal (numeracy and literacy) home learning activities (Skwarchuk et al., 2014). Nevertheless, in line with previous findings, there was a significant correlation between parents' attitudes towards literacy and the frequency of letter activities, $r_s = .293$, $p = .037$.[1] The more positive parents' attitudes towards literacy, the more they engaged in letter activities with their child. Note that there was no association between parents' attitudes towards mathematics and their home numeracy activities, a finding which also aligns with previous results.

4.4   Discussion

The study presented in this chapter used a mixed methods design to examine the relationship between children's and parents' SFON, together with the role of home numeracy practices. It was hypothesised that children's tendency to engage with numbers may be shaped by parental influences and thus children's SFON scores may be positively related to their parent's SFON scores (Prediction 1) and also to home numeracy factors (Prediction 2). The results yielded little support for either of these predictions. Below I consider the possible reasons for these null results.

In terms of Prediction 1, there were no significant correlations between children's SFON (task-based or play-based) and parents' SFON (task-based or play-based). This means that either the null hypothesis is true (there is no real relationship between children's and parents' SFON), or, the null hypothesis is false (there is a real relationship between children's and par-

ents' SFON) but there is not enough power to reject it. The 95% confidence intervals for the correlation coefficients were constructed to try and distinguish between these possibilities. These suggest that the true relationship lies relatively close to zero, and therefore is not of substantial theoretical importance.

One could argue that the SFON measures used may have lacked the sensitivity to detect a relationship between children's and parents' SFON. We know that correlation coefficients are affected by the number of items on a measure and both children's and parents' task-based and play-based SFON were scored on a small $0 - 3$ scale. (Recall that children and parents received a binary SFON score for each trial or game and these binary scores were summed to give a total score out of three). There are at least two reasons to question this account of the null results.

Firstly, if we analyse children's and parents' total continuous SFON scores rather than their summed binary scores then the correlations yielded are of similar magnitude. To give an example, the correlations between children's and parents' Task SFON using summed binary scores and total continuous scores are $r_s = .073$ [95% CI $= -.203, .339$] and $r_s = .038$ [95% CI $= -.237, .307$], respectively. This indicates that the low correlations are not solely due to the small 4-point scale of the measures.

Further to this, we know from the research presented in Chapter 3 that the Picture Task shows good psychometric properties including test-retest reliability and predictive validity of children's numerical skills. This implies that it should have been sensitive enough to detect any effects should they have been present. That said, the Picture Task has only been validated in children, thus we cannot rule out the possibility that it lacks measurement sensitivity in adults.

It is also important to consider theoretical reasons why the present study found no relationship between children's and parents' SFON. One possibility is that when children start school their SFON becomes less influenced by parental input and more influenced by teachers and peers. If this is true then we might expect children's SFON to be related to parents' SFON before they start school but not after. To explore this possibility, the correlations were re-run separately for the school-aged children and the preschool-aged children in the sample. These analyses yielded no significant correlations between children's and parents' SFON for either group of children. As such,

it seems that it is not the case that parental influence is reduced when children start school. Note, however, that many of the preschool-aged children attended nursery or preschool (42% part-time and 23% full-time) and this preschool attendance may have altered parental influence as well.

Another possibility to consider is that perhaps children's SFON is more heavily shaped by their father's SFON than their mother's SFON. Mathematics is stereotypically seen as a male domain and thus fathers may be more likely than mothers to engage with their child in number-based activities. If this is the case then fathers may have a more influential effect on children's SFON. Given the low number of fathers who took part in this study ($N = 8$) it is difficult to test this possibility with the current data. There is some evidence from previous studies to suggest that mothers and fathers do not differ in the frequency of their mathematics activities with children (see, for example, Jacobs & Bleeker, 2004) but this research has focused largely on older children with an emphasis on formal instruction (e.g. help with homework) rather than informal numerical practices.

As well as finding no relationship between children's and parents' SFON, the present study revealed little association between children's SFON and home numeracy factors. Contrary to Prediction 2, there were no significant correlations between children's SFON and parents' academic expectations and attitudes towards mathematics. Likewise, children's SFON was not correlated with parents' self-reports of home numeracy activities, apart from a small negative correlation between children's task-based SFON and the frequency of number book activities. Somewhat counterintuitively, the more parents reported engaging in number book activities with their child, the less likely their child was to focus on numerosity on the Picture Task. There was also a small negative correlation between the frequency of letter activities and children's play-based SFON which suggests that the more focus there is on literacy at home, the less likely children are to engage with numbers.

As discussed in relation to Prediction 1, it is important to ascertain whether the null hypothesis is true (that there is no relationship between SFON and home numeracy factors) or whether there was not enough power to reject the null hypothesis. It may be the case the Home Numeracy Questionnaire lacked measurement sensitivity. It relied on parents' retrospective recall and, as with all self-report measures, it may have been biased by social desirability effects. There are two sources of evidence which question this

account of the null results.

Firstly, the current study found significant associations between home literacy factors and children's scores on the BPVS, a measure of receptive vocabulary. Specifically, children's (BPVS) scores were positively correlated with the number of adult books in the home and the frequency with which parents and children engaged in letter activities. Moreover, parents' attitudes towards literacy were positively correlated with the frequency of letter activities. The fact that these relationships were found for the literacy variables suggests that issues of retrospective recall and social desirability did not hinder the measurement sensitivity of the questionnaire.

Further to this, previous studies using the Home Numeracy Questionnaire (developed by LeFevre et al., 2009) have found positive relationships between home numeracy factors and children's numerical skills (LeFevre et al., 2009; Skwarchuk et al., 2014). This implies that the questionnaire is sensitive enough to have detected a relationship with children's SFON should this relationship have been present.

Despite these sources of evidence, it is important to point out that there was a small positive correlation between children's task-based SFON and their SES (an index of the family's economic and social position). This correlation suggests that SFON is not entirely independent of parental factors. There appears to be something about the home environment that promotes SFON but it is not clear from the Home Numeracy Questionnaire what this something is. It may likely be something subtle, for example, to do with the quality of the home numeracy activities rather than their frequency. Further studies using more detailed self-reports or observational methods may allow us to unpick such subtleties.

4.4.1 Summary of Findings

The research presented in this chapter sought to further our understanding of where children's SFON tendency comes from by examining the role of parental influences. The findings revealed no significant relationship between children's SFON and parents' SFON. Moreover, there was little relationship between children's SFON and home numeracy factors, such as parents' academic expectations, home activities and attitudes relating to mathematics. These null effects suggest that parental influences on children's SFON may be too small to be of theoretical significance. Future research needs

to examine other possible causes or influences on children's SFON as well as increasing our understanding of the nature of the relationship between children's SFON and numerical skills.

Chapter 5

The Mechanisms of SFON (Studies 4 & 5)

SFON is emerging as a key factor for explaining variations in children's numerical development. However, the mechanisms behind this relationship are not yet clear. In the current chapter I address this issue with two cross-sectional studies. First, I investigate the relationship between children's SFON and their basic nonsymbolic and symbolic numerical processing skills (Study 4). Next, I explore whether the relationship between SFON and arithmetical skills can be accounted for by individual differences in fluency with nonsymbolic and symbolic representations of number (Study 5).

5.1   Introduction

We know that SFON is associated with a numerical advantage. High SFON children show more advanced counting and arithmetical skills than their low SFON peers. What we don't know yet is why SFON is associated with this numerical advantage. What are the mechanisms underlying this relationship?

Hannula et al. (2007) proposed that the more children focus on the numerical aspects of their environment, the more practice they acquire with enumeration and thus, the better their counting skills become. To explore this possibility, they looked at the relationships between children's subitizing-based enumeration (i.e. the rapid perception of the numerosity of small sets, without counting), object counting and SFON. Regression analy-

ses revealed a direct relationship between children's SFON and their number sequence production skills. In contrast, there was an indirect relationship between SFON and object counting that was explained by individual differences in subitizing-based enumeration skills. This provides some evidence to suggest that SFON promotes perceptual subitizing skills which in turn supports the development of children's counting skills.

Other research has explored motivational factors in the development of children's SFON and early numerical skills. In one of the first SFON studies to be conducted outside of Finland, Edens and Potter (2013) measured SFON and counting skills in 4-year-old children in US preschools. They also obtained teacher reports of children's motivation, attentional self-regulation, persistence and interest in mathematics. Results showed a positive correlation between preschoolers' SFON and counting skills, thus replicating the results of Hannula and colleagues (e.g. Hannula & Lehtinen, 2005). Furthermore motivational measures were significantly correlated with children's counting skills, but not with SFON. As such, SFON does not appear to reflect motivational factors such as an interest in mathematics.

Together these studies suggest that the factors underpinning the relationship between SFON and children's numerical development are cognitive rather than affective. However, the precise mechanisms involved need further investigation. The current literature is sparse and somewhat limited in scope. Thus far, studies exploring the mechanisms of SFON have focused solely on its relationship with early counting skills. We do not know why SFON is related to children's later arithmetical development. We also do not know how SFON relates to more basic (low-level) numerical competencies such as nonsymbolic processing skills or 'number sense' (Dehaene, 2001). The studies presented in this chapter sought to further our understanding of both of these issues.

5.2   Study 4

The aim of this study was to investigate the effect of differences in SFON on children's nonsymbolic and symbolic numerical processing skills. To address this aim children completed the SFON Picture Task followed by two numerical processing tasks (nonsymbolic and symbolic magnitude comparison). Given the lack of previous research in this area, no prior predictions

were made.

Note that this study also aimed to investigate the effect of differences in mathematics anxiety on children's numerical magnitude processing. As reported in Chapter 1, Section 1.3, there was no relationship between children's SFON and their mathematics anxiety therefore, the anxiety aspect of this Study is reported separately in Part III (Chapter 8).

5.2.1   Method

5.2.1.1   Participants

One-hundred and nine children (58 boys) aged $8.1-9.8$ years ($M = 8.9$ years, $SD = 0.4$ years) were recruited from a primary school in Nottingham, UK. The school was of mid to high SES with fewer pupils receiving free school meals than the national average. All children were in Year 4, Key Stage 2 of the UK National Curriculum.

All parents received opt-out consent letters at least 10 days before the start of the study. Parents who did not wish their child to take part were asked to sign and return an attached form to their child's class teacher. (There were no children whose parents declined permission for them to take part). Children gave verbal assent prior to the start of the study and they received stickers to thank them for their participation. Study procedures were approved by the Loughborough University Ethics Approvals (Human Participants) Subcommittee (see Section 2.1.4 for full details on ethical considerations).

5.2.1.2   Design

Children took part in two testing sessions (one group and one individual) separated by approximately $1 - 2$ weeks. In the first (group) testing session, children completed a mathematics anxiety questionnaire administered by their teacher as part of normal classroom activity. In the second (individual) testing session children completed the SFON Picture Task followed by two numerical processing tasks (nonsymbolic and symbolic magnitude comparison). These tasks were administered by the researcher in a quiet corridor outside the classroom. The tasks were presented in the same order for every child.

5.2.1.3   Materials & Procedure

This section describes the experimental procedure for the SFON and numerical processing tasks only. The mathematics anxiety questionnaire is described separately in Chapter 8 (Section 8.2.2) where the anxiety aspect of the study is presented.

SFON Picture Task

Children completed the Picture Task developed in Chapter 3. This Picture task (programmed using E-Prime software 2.0) was presented on a 15 inch LCD laptop screen. The experimental procedure was the same as that used in the pilot study of Study 2 (described in Section 3.2.1.3, page 67). The picture stimuli can be seen in Figure 3.1. On each of two experimental trials, children were presented with a cartoon picture for 7000ms and immediately afterwards they were asked to verbally describe what they had seen. If the child's description contained any reference to symbolic number then they received a score of 1. Each child thus received a total SFON score out of 2.

Numerical Processing Tasks

Children completed two computer-based numerical processing tasks designed to assess (i) nonsymbolic magnitude comparison skills, and (ii) symbolic magnitude comparison skills. Both tasks were programmed using E-Prime software 2.0 and presented on a 15 inch LCD laptop screen. The researcher was present at all times throughout the tasks.

Nonsymbolic Comparison Task   This task measured children's ability to compare nonsymbolic numerical stimuli. Children were presented with two arrays of dots (simultaneously side-by-side) and they were asked to indicate which of the arrays was more numerous. The task was incorporated into a game in which the children saw two fictional characters and were asked to quickly decide, without counting, who had the most marbles.

Numerosities ranged from 5 to 30 and the numerical ratios between the two arrays were 0.5, 0.6, 0.7, 0.8 and their inverses. Small numerosities were excluded to prevent children from subitizing. Dot arrays were generated following the method of Gebuis and Reynvoet (2011) which controls for dot size and envelope area to prevent participants reliably using strategies

based on continuous quantity variables. The dot arrays (one red and one blue) were presented on a grey background as shown in Figure 5.1a. The side of the correct array was counterbalanced.

Each of 80 experimental trials began with a fixation cross for 1000ms, followed by the two dots arrays (red left and blue right) for 600ms, followed by a question mark until response. Stimuli presentation times were chosen based on previous numerosity experiments (e.g. Gilmore, Attridge, De Smedt, & Inglis, 2014) and pilot testing with children of the same age. The dots were presented for a very limited time to prevent children from counting. Children responded by pressing red (left bigger) and blue (right bigger) stickers in place of the 'c' and 'm' keys on a standard keyboard. The order of the trials was randomised and children were prompted to take a break after every 20 trials.

The experimental trials were preceded by two blocks of 4 practice trials. In the first practice block children received no time limit; they were presented with a fixation cross followed by the two dot arrays until response. In the second practice block, the researcher introduced the experimental time limit of 600ms. Note that this was designed to prevent the children from counting. The researcher instructed children to respond as quickly and accurately as possible, and children were encouraged to have a guess if they were not sure. Throughout the task children received general praise but no specific feedback was given. Each child received an accuracy score based on the proportion of items they answered correctly.

Symbolic Comparison Task   This task measured children's ability to compare symbolic numerical stimuli. Children were presented with two Arabic digits (simultaneously side-by-side) and they were asked to indicate the numerically larger of the two. Numerosities ranged from 5 to 30. The problems were identical to the nonsymbolic problems, except the numerosities were presented as Arabic digits rather than dot arrays. Symbolic stimuli were black digits on a grey background. The left digit had a red border and the right digit had a blue border as shown in Figure 5.1b.

Each of 80 experimental trials began with a fixation cross for 1000ms, followed by the two Arabic digits for 300ms, followed by a question mark until response. Stimuli presentation times were chosen based on previous numerosity experiments (e.g. Gilmore et al., 2014) and pilot testing with
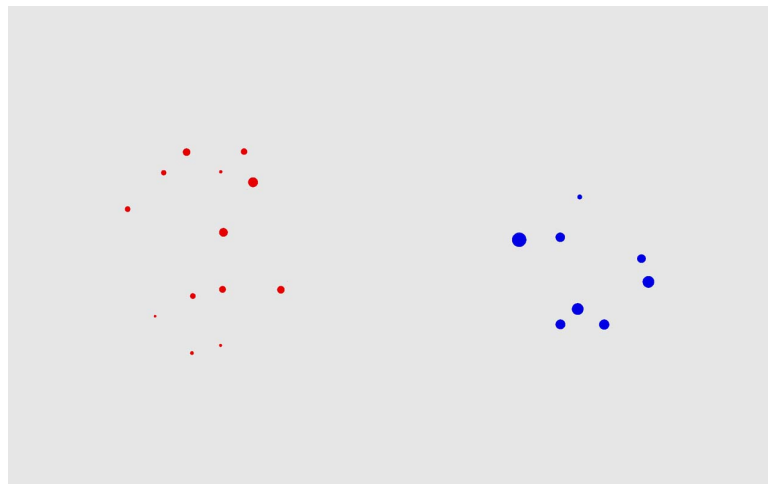
(a)
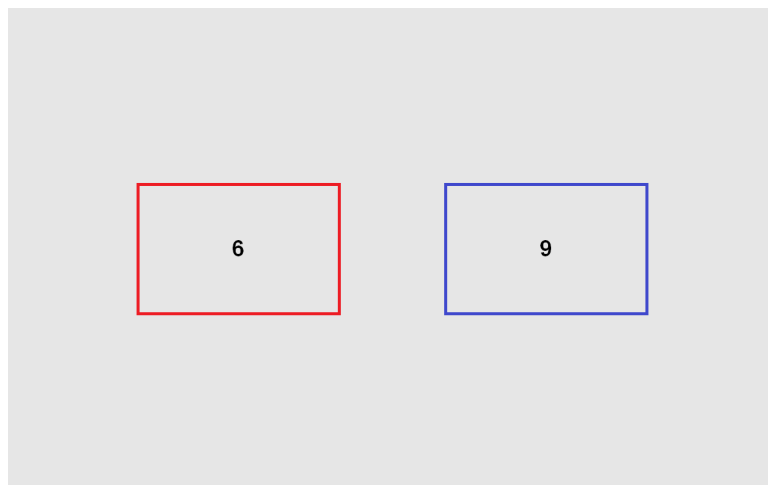
(b)

Figure 5.1: Example stimuli used in (a) the nonsymbolic comparison task, and (b) the symbolic comparison task.

children of the same age. They varied across the nonsymbolic and symbolic versions of the task to avoid floor and/or ceiling effects. In line with the nonsymbolic version of the task, children responded by pressing red (left bigger) and blue (right bigger) stickers in place of the 'c' and 'm' keys on a standard keyboard. The experimental trials were preceded by two blocks of 4 practice trials. The first practice block had no time limit and the second practice block introduced the experimental time limit of 300ms. All trials were presented in a random order and children were prompted to take a break after every 20 trials. Each child received an accuracy score based on the proportion of items they answered correctly.

5.2.2   Results

First, descriptive statistics are presented for children's performance on each of the experimental tasks. Next, a mixed ANOVA is conducted to investigate the effect of differences in SFON on children's nonsymbolic and symbolic numerical processing skills.

Children showed individual differences in their SFON tendency on the Picture Task. Of the 109 children who took part, 34 (31%) did not focus on numerosity on either of the two picture trials, 27 (25%) focused on numerosity on one of the two trials, and 48 (44%) focused on numerosity on both of the two trials. In line with Studies 1, 2 and 3, a Mann-Whitney U test indicated no gender differences in children's SFON, $z = -0.33$, $p = .742$.

Children also demonstrated a range of performance on the numerical processing tasks. Both nonsymbolic comparison accuracy, $M = .76$, $SD = .08$, and symbolic comparison accuracy, $M = .81$, $SD = .14$. were significantly above the 50% chance level ($ps < .001$). Independent samples t-tests demonstrated no gender differences in children's nonsymbolic comparison performance, $t(107) = 1.07$, $p = .289$, but boys performed marginally better than girls on the symbolic version of the task, $t(107) = 2.02$, $p = .045$. Issues surrounding gender differences are discussed further in relation to the mathematics anxiety aspect of this study (presented in Part III, Chapter 8).

To investigate the effect of differences in SFON on children's numerical comparison performance a 2×4×2 ANOVA was conducted with two within-subjects factors: Comparison Task (nonsymbolic, symbolic) and Numerical Ratio (0.5, 0.6, 0.7, 0.8), and one between-subjects factor: SFON (low, high). Low SFON children were those who attended to numerosity on nei-

ther of the two trials (score of 0) and high SFON children were those who attended to numerosity on both trials (score of 2). This revealed no significant main effect of SFON ($p = .537$), but there was a significant interaction between SFON and comparison task, $F(1, 80) = 5.73$, $p = .019$, $\eta_p^2 = .07$. High SFON children performed better on the symbolic version of the comparison task than the nonsymbolic version. Low SFON children showed no advantage on the symbolic version (see Figure 5.2).[1]
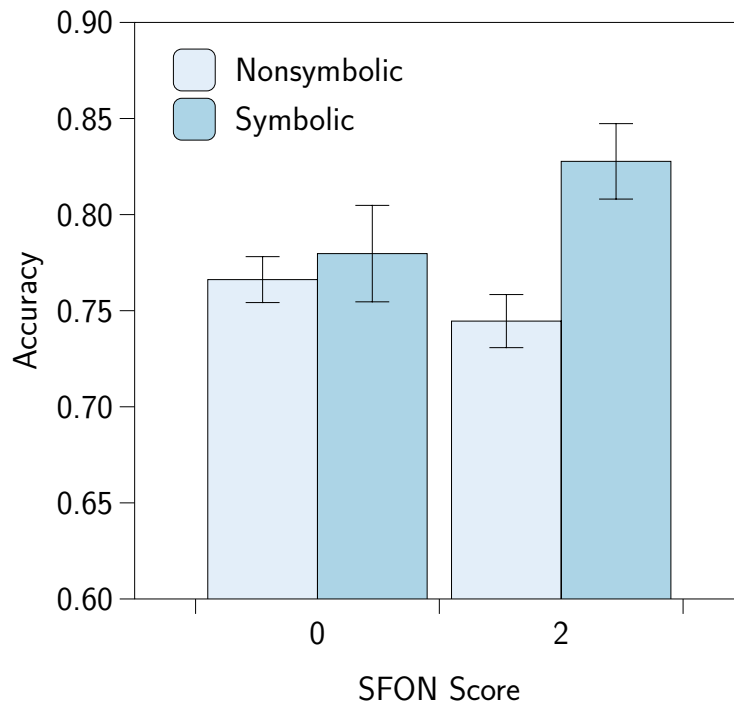
Figure 5.2: Interaction between Task and SFON on comparison accuracy (error bars show ± 1 standard error of the mean).

Note that as expected there was a highly significant main effect of ratio, $F(3, 240) = 114.22$, $p < .001$, $\eta_p^2 = .59$. Children's accuracy decreased as

---

[1]If we also include mid SFON children (those scoring 1 out of 2) in this analysis, then the pattern of results is the same but to a lesser degree. The interaction between SFON and comparison task is marginally significant, $F(2, 106) = 2.72$, $p = .071$, $\eta_p^2 = .05$.

the ratio between the numerosities to be compared approached 1, reflecting the approximate nature of their underlying magnitude representations. There was no interaction between ratio and SFON ($p = .828$) indicating that low and high SFON children showed similar ratio effects.

5.2.3   Discussion

These results advance our understanding of how SFON relates to children's basic (low-level) numerical processing skills. Specifically, they demonstrate that SFON is associated with improved symbolic skills. High SFON children were more accurate on the symbolic comparison task than the nonsymbolic comparison task, whereas low SFON children showed no symbolic advantage. Note that the children in this study (aged $8 - 9$ years) were older than the children in previous SFON studies, thus these results also add our understanding of the developmental trajectory of children's SFON. I return to this issue in the general discussion of this chapter (Section 5.4) but for now I focus on the possible mechanisms through which SFON might work.

In view of the current findings, it is possible that SFON promotes mathematical development by increasing children's fluency with number symbols. High SFON children may get more practice mapping between their newly-acquired symbolic representations of number and pre-existing nonsymbolic representations. As children get practice with and improve the precision of these mappings, their counting and arithmetic skills may develop. This possibility is theoretically likely because we know from previous research that mapping ability is related to mathematics achievement (Booth & Siegler, 2008; Brankaer, Ghesquière, & De Smedt, 2014; Holloway & Ansari, 2009; Mundy & Gilmore, 2009; see De Smedt et al., 2013 for a review). For example, Mundy and Gilmore (2009) found that children aged $6 - 8$ years showed individual differences in their ability to map between nonsymbolic representations (dot arrays) and symbolic representations (Arabic digits). These individual differences explained a significant amount of variation in children's school mathematics achievement.

Some initial support for this possibility comes from two recent studies. Firstly, Bull (2013) found that high SFON children (aged $5 - 7$ years) performed better than their low SFON peers on a numerical estimation task, in which they had to assign a symbolic number to a nonsymbolic array of dots. In other words, children who consistently focused on numerosity were better

able to map between nonsymbolic and symbolic representations of number. Secondly, research exploring the transition from informal to formal mathematics knowledge has highlighted the role of mapping ability. Purpura, Baroody, and Lonigan (2013) demonstrated that the link between children's informal and formal mathematics knowledge was fully explained by individual differences in symbolic number identification and the understanding of symbol to quantity relations.

To directly test this hypothesis the next study (Study 5) investigates whether the relationship between children's SFON and mathematical skills can be accounted for by individual differences in fluency with nonsymbolic and symbolic representations of number.

5.3   Study 5

This study sought to investigate the mechanism through which SFON exerts its positive influence on children's mathematical skills. As stated above, it aimed to test whether the relationship between SFON and mathematical skills is explained by variations in children's fluency with nonsymbolic and symbolic representations of number.

To achieve this objective, children aged $4 - 5$ years were given a battery of tasks designed to assess SFON, nonsymbolic magnitude comparison, symbolic magnitude comparison, nonsymbolic-to-symbolic mapping and arithmetic skills. Children were also given a digit recognition task to determine their knowledge of number symbols and a visuospatial working memory task to control for individual differences in domain-general cognitive skills. This control task was chosen because visuospatial working memory is a specific predictor of early mathematics achievement (see, for example, Bull et al., 2008). Note that previous studies investigating SFON have not controlled for executive function skills.

Two hypotheses were made. First, it was predicted that SFON would show a significant positive correlation with children's mathematical skills, thus confirming the results of previous studies (Prediction 1). Second, it was predicted that this relationship would be largely explained by individual differences in children's ability to map between nonsymbolic and symbolic representations of number (Prediction 2).

5.3.1 Method

5.3.1.1 Participants

One-hundred and thirty children (66 boys) aged $4.5 - 5.6$ years ($M = 5.0$ years, $SD = 0.3$ years) were recruited from three primary schools in Nottinghamshire and Leicestershire, UK. The schools were of varying SES: one low ($N = 32$), one medium ($N = 58$) and one high ($N = 40$); based on the proportion of children eligible for free school meals compared to the national average. All children were in the second term of their first year of school. At this stage classes are very informal, following the EYFS framework outlined in Section 2.2.1.

As with Study 4, parental consent was obtained on an opt-out basis. Parents received an information sheet (at least 10 days before the start of the study) and those parents who did not wish their child to take part were asked to sign and return an attached form to their child's class teacher. The number of children whose parents declined permission for them to take part was low ($N = 4$). Children gave verbal assent before each testing session and they received stickers to thank them for taking part. Study procedures were approved by the Loughborough University Ethics Approvals (Human Participants) Subcommittee (see Section 2.1.4 for full details on ethical considerations).

Nine children were excluded from all of the analyses for the following reasons: English was not their native language ($N = 2$), speech and language difficulties and/or selective mutism ($N = 2$), other special educational needs ($N = 3$), failure to identify numerical digits beyond 1 ($N = 2$). A further two children did not complete all of the measures during Session 2, leaving a total of 119 complete datasets.

5.3.1.2 Design

Children took part in two testing sessions, approximately 20 minutes each, scheduled one week apart. At Time 1 they completed two SFON tasks and a visuospatial working memory task. At Time 2 they completed a series of computer-based numerical processing tasks (nonsymbolic comparison, symbolic comparison, digit recognition and nonsymbolic-to-symbolic mapping) followed by a standardised measure of arithmetic. The tasks were presented in the same order for every child (however, note that the two SFON tasks

were counterbalanced). Each task is described in turn below, in the order in which it was presented.

5.3.1.3   Materials & Procedure

Children were tested individually in a quiet room or corridor outside their classroom. The researcher ensured that the testing area was free from any numerical displays that might have prompted the children to focus on number (during Session 1) or helped them to solve a numerical problem (during Session 2). During testing Session 1, children were not told that the tasks were in anyway numerical or quantitative. Likewise, the children's parents and teachers were not informed of the numerical focus of the study; as with Studies 1 and 2 they were told that the study was looking at children's general thinking skills (see Section 2.1.4 on ethical considerations).

Throughout all tasks children received general praise but no specific feedback was given. At the end of each task children were allowed to choose a sticker.

SFON

SFON was measured using the Posting Task developed by Hannula and Lehtinen (2005) and the Picture Task developed in Chapter 3. The order of these tasks was counterbalanced.

Posting Task   Children were presented with the same Posting Task used in Studies 1 and 2. The experimental procedure was identical to that described in Section 2.2.3.1 (page 37). The experimental materials and set-up of the task can be seen in Figures 2.1 and 3.3. As in the previous studies there were three trials, on each of which children received a score of 0 or 1 depending on whether they focused on numerosity or not. Children were scored as focusing on numerosity if they posted the correct number of letters and/or if they presented any verbal or nonverbal quantitative acts (see Section 2.2.3.1). Each child received a total score out of three.

Picture Task   Children completed the Picture Task developed and used in Studies 2, 3 and 4. The task was presented with the same procedure used in Study 2 (Main) and Study 3 (see Section 3.3.1.3, page 75). On each

of three experimental trials, children saw a cartoon picture and they were asked to verbally describe what they saw in the picture. The picture stimuli can be seen in Figures 3.1 and 3.4. For each trial, children received a score of 0 or 1 depending on whether their description contained any reference to symbolic number. As with the Posting Task, each child received a total score out of three.

Working Memory

Visuospatial working memory skills were measured using a visual search ('Spin-the-Pots') task adapted from Hughes and Ensor (2005).

Spin-the-Pots Task   The materials were a circular silver tray (diameter = 39.5cm), 11 different coloured paper cups (diameter = 7.0cm, height = 9.5cm), 9 stickers (2.0cm × 2.0cm) and an A3 piece of card.

The researcher randomly positioned each cup upside down around the rim of the circular tray. She then introduced the task to the child by saying: "Now we're going to play a finding game. Here are some cups. They are all different colours. Can you tell me what colours they are?" This question intended to check whether the child could distinguish between all of the different colours. The researcher then placed each sticker on top of a cup, pointing out to the child that there were not enough stickers for all of the cups and that two cups would not have stickers. Next, she instructed the child: "Watch carefully whilst I hide the stickers under the cups. Later, you can have a go at finding them." The researcher hid all of the stickers and then covered the cups with a piece of card. She told the child: "Now, I'm going to spin the cups. Then you can choose one cup and see if there's a sticker inside." The researcher spun the cups and then removed the card for the child to choose a cup. If they found a sticker then they took it out and kept it beside them. The researcher continued by covering up the cups again and spinning them round before allowing the child to choose another. This continued until the child found all 9 stickers, or, until the maximum number of spins (18) was reached. Each child received a score out of 18 depending on the number of errors they made.

Mathematical Skills

Children completed four computer-based numerical processing tasks (all programmed using E-Prime software 2.0 and all presented on a 15 inch LCD laptop screen) followed by a standardised measure of arithmetic. The researcher was present at all times throughout each of the tasks.

Nonsymbolic Comparison Task   This task measured children's ability to compare nonsymbolic numerical stimuli. As with the nonsymbolic comparison task used in Study 4 (Section 5.2.1.3), children were presented with two arrays of dots and they were asked to select the more numerous of the two arrays. The task was incorporated into a game in which the children saw two fictional characters and were asked to quickly decide, without counting, who had the most marbles.

Numerosities ranged from 4 to 9 and the numerical distance between the two numbers being compared was either small (a distance of 1 or 2) or large (a distance of 3 or 4). Numerosities $1-3$ were excluded because they are in the subitizing range. Dot arrays were generated randomly in accordance with previous numerosity experiments, such that no two dot arrays for the same quantity were the same. Stimuli were created using the method used by Dehaene, Izard, and Piazza (2005) to control for continuous quantity variables such as dot size and envelope area.[2] All dot arrays were black dots on a white circular background as shown in Figure 5.3a. The side of the correct array was counterbalanced.

Each of 40 experimental trials began with a fixation cross for 1000ms, followed by the two dot arrays (side-by-side) for 1250ms, followed by a question mark until response. Stimuli presentation times were chosen based on previous numerosity experiments (e.g. Gilmore et al., 2014) and pilot testing with children of the same age. The dots were presented for a very limited time to prevent children from counting. Children responded by pointing to the character with the most marbles. The researcher recorded these responses via the 'c' (left bigger) and 'm' (right bigger) keys on a standard keyboard.

---

[2]This method of dot stimuli creation was chosen over that used in Study 4 because it is better suited to generating small numerosities $(1-9)$. Importantly, it was able to produce physically larger dots which were visually more appropriate for the younger age group of children used in this study.

The order of the trials was randomised and children were prompted to take a break after 20 trials.

The experimental trials were preceded by two blocks of 4 practice trials. In the first practice block children received no time limit; they were presented with a fixation cross followed by the two dot arrays until response. In the second practice block, the researcher introduced the experimental time limit of 1250ms to prevent the children from counting. The researcher emphasised that it was a speeded game, and children were encouraged to have a guess if they were not sure. Each child received an accuracy score based on the proportion of items they answered correctly.

Symbolic Comparison Task   This task measured children's ability to compare symbolic numerical stimuli. As with the symbolic comparison task used in Study 4 (Section 5.2.1.3), children were presented with two Arabic digits and they were asked to select the numerically larger of the two. Numerosities ranged from 4 to 9. The problems were identical to the non-symbolic problems, except the numerosities were presented as Arabic digits instead of dot arrays. Symbolic stimuli were black digits on a white circular background as shown in Figure 5.3b.

Each of 40 experimental trials began with a fixation cross for 1000ms, followed by the two Arabic digits (side-by-side) for 750ms, followed by a question mark until response. Stimuli presentation times were chosen based on previous numerosity experiments (e.g. Gilmore et al., 2014) and pilot testing with children of the same age. They varied across tasks to avoid floor and/or ceiling effects. In line with the nonsymbolic version of the task, children responded by pointing to the character with the larger number of marbles and the researcher recorded these responses on the computer.

The experimental trials were preceded by two blocks of 4 practice trials. The first practice block had no time limit and the second practice block introduced the experimental time limit of 750ms. All trials were presented in a random order and children were prompted to take a break half-way through. Each child received an accuracy score based on the proportion of items they answered correctly.

Digit Recognition Task   This task measured children's knowledge of Arabic digit stimuli. Children were asked to read aloud a series of Ara-

bic digits (ranging from 1 to 9) presented one by one in a random order on the laptop screen. Children scored one point for each correct identification giving a total score out of 9.

Numerical Mapping Task   This task measured children's ability to map nonsymbolic numerical stimuli onto symbolic numerical stimuli. Children were presented with an array of dots and they were asked to quickly (without counting) decide which of two Arabic digits matched the numerosity of the dots. The task was adapted from Mundy and Gilmore (2009).

Numerosities ranged from 2 to 9 and the numerical distance between the two symbolic choices was either small (a distance of 1 or 2) or large (a distance of 3 or 4). The number range included small numerosities within the subitizing range because pilot testing revealed some children to be performing at chance with the larger numerosities. Stimuli were presented simultaneously with the dot array centred at the top and the symbolic stimuli at the bottom left and right hand sides of the screen (as shown in Figure 5.3c).

Each of 40 experimental trials began with a fixation cross for 1000ms, followed by the numerical stimuli for 2000ms, followed by a question mark until response. Stimuli presentation times were chosen based on previous numerosity experiments (e.g. Gilmore et al., 2014) and pilot testing with children of the same age. The dot array disappeared when the question mark appeared to prevent children from counting. Children responded by pointing to the digit that matched the numerosity of the dots. The researcher recorded these responses via the 'c' (left matching) and 'm' (right matching) keys on a standard keyboard.

As with the comparison tasks, the experimental trials were preceded by two blocks of 4 practice trials (first with no time limit, then with the experimental time limit of 2000ms). Again the researcher emphasised that it was a speeded game and children were encouraged to have a guess if they were not sure. Each child received an accuracy score based on the proportion of items they answered correctly.

Arithmetic Task   The arithmetic subtest of the WPPSI (Wechsler, 1967) was administered in accordance with the standard procedure. There were 20 questions in total. Questions $1 - 4$ required children to make nonsymbolic judgments about size or quantity (e.g. "Here are some balls. Point to the one

(a)   (b)

(c)
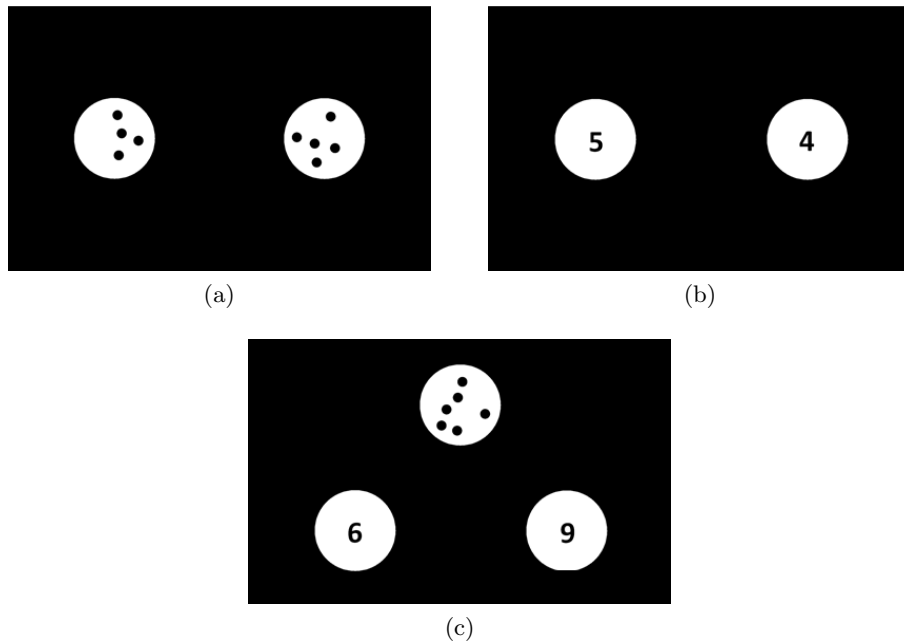
Figure 5.3: Example stimuli used in (a) the nonsymbolic comparison task, (b) the symbolic comparison task and (c) the numerical mapping task.

that is the biggest"), questions $5 - 8$ required children to perform counting tasks with blocks (e.g. "Can you give me all of the blocks except four") and questions $9 - 20$ required children to mentally solve arithmetic word problems (e.g. "Mary had five dolls. She lost two. How many did she have left?"). Children continued until they had answered four consecutive questions incorrectly. They received a raw score out of 20.

5.3.2   Results

The results reported in this section are organised as follows: First, descriptive statistics (and preliminary analyses) are presented for children's performance on each of the experimental tasks. Next, the correlations among children's performance on each task are explored. Finally, the nature of these relationships is examined with a series of hierarchical regression models.

5.3.2.1   Descriptive Statistics & Preliminary Analyses

Figure 5.4 shows the number of children focusing on numerosity from zero to three times on each SFON task. Children's performance on all other
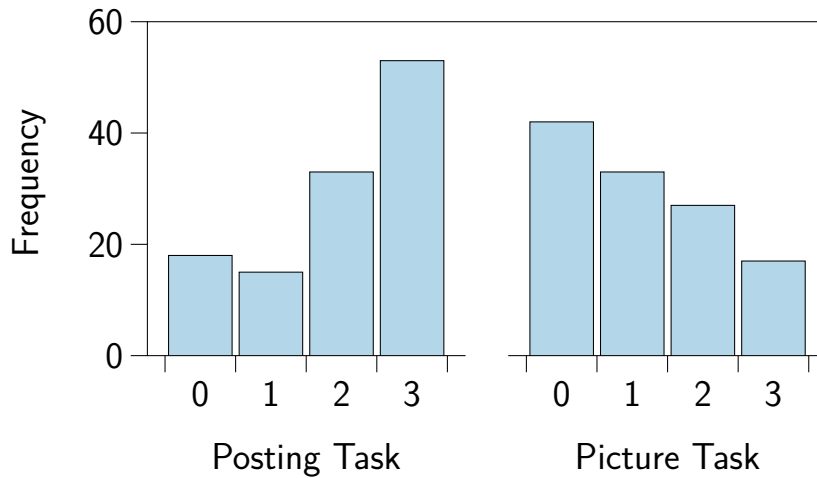
Figure 5.4: The frequency of children obtaining SFON scores from $0 - 3$ on the Posting Task and the Picture Task.

tasks is presented in Table 5.1. Together these demonstrate that children showed individual differences in SFON and a range of performance on the working memory and mathematical tasks. The two SFON tasks varied in terms of difficulty. Scores on the Posting Task were negatively skewed (44% of children obtained the maximum score of 3) and scores on the Picture Task were positively skewed (35% of children obtained the minimum score of 0). As a likely result of this, performance on these two tasks was not significantly correlated ($r_s = .06$, $p = .533$).

A series of group-level comparisons were carried out to check for any gender differences in children's SFON or mathematical skills. In accordance with previous studies, Mann-Whitney U tests indicated no significant differences between boys' and girls' levels of SFON on the Posting Task ($p = .238$) or the Picture Task ($p = .649$). Likewise, independent samples t-tests demonstrated no significant differences between boys' and girls' performance on the numerical processing tasks ($ps > .183$) or the arithmetic test ($p = .825$).

	M	SD	Range
SFON Posting	2.02	1.09	$0-3$
SFON Picture	1.16	1.07	$0-3$
Working memory (errors)	6.30	3.53	$0-11$
Nonsymbolic comparison	0.74	0.12	$0.43-0.95$
Symbolic comparison	0.75	0.17	$0.38-1.00$
Digit recognition	8.22	1.31	$4-9$
Numerical mapping	0.73	0.16	$0.38-1.00$
Arithmetic (WPPSI raw score)	10.95	2.63	$5-17$

Note. Mean accuracy on the nonsymbolic and symbolic comparison tasks and the numerical mapping tasks was significantly above the 50% chance level ($ps < .001$). $N = 119$.

Table 5.1: Descriptive statistics for performance on each of the eight tasks.

5.3.2.2   Correlational Analyses

Correlations between all variables are reported in Table 5.2. These show that the SFON tasks (while not significantly correlated with each other) were both positively related to performance on the mathematical tasks. The correlation between SFON and arithmetic was .30 for the Posting Task and .47 for the Picture Task. These correlations are similar in magnitude to those found in previous SFON studies (e.g. Hannula et al., 2007). Importantly, they remain significant even after controlling for age, working memory skills, verbal skills (average word count on the SFON Picture Task) and Arabic digit recognition: Children's arithmetic scores were positively correlated with SFON scores on the Posting Task ($pr_s = .29$, $p = .002$) and the Picture Task ($pr_s = .42$, $p < .001$).

These correlations confirm previous findings that high SFON children show more advanced mathematical skills than their low SFON peers (thus supporting Prediction 1). To explore the nature of these relationships a series of hierarchical regression models were constructed.

Variable	1	2	3	4	5	6	7	8
1 SFON Posting	–							
2 SFON Picture	.06	–						
3 Working memory	−.04	−.08	–					
4 Nonsymbolic comparison	.22*	.25**	−.17	–				
5 Symbolic comparison	.22*	.31**	−.25**	.55***	–			
6 Digit recognition	.11	.15	−.25**	.23*	.50***	–		
7 Numerical mapping	.25**	.37***	−.25**	.51***	.72***	.47***	–	
8 Arithmetic	.30**	.47***	−.15	.43***	.65***	.42***	.57***	–

Note. Spearman's r_s coefficients are reported for the correlation between the two SFON tasks. All other coefficients are Pearson's r. $N = 119$, *$p < .05$, **$p < .01$, ***$p < .001$.

Table 5.2: Zero-order correlations between all variables.

5.3.2.3   Regression Analyses

Hierarchical regression models were computed to test whether nonsymbolic skills and the mapping between nonsymbolic and symbolic representations could account for the relationship between SFON and mathematics achievement (Prediction 2). Two mathematical outcome measures were used: symbolic comparison performance and arithmetic performance. These were found to be highly correlated, $r = .65$, $p < .001$. For each dependent variable two models were built. In the first model baseline variables were entered in Step 1, followed by both SFON measures in Step 2, and nonsymbolic comparison and mapping performance in Step 3. In the second model, the order of steps 2 and 3 were reversed.

As shown in Table 5.3, SFON was a significant predictor of symbolic comparison performance when entered in Step 2, before the nonsymbolic comparison and mapping tasks, but not when it was entered after these variables in Step 3. In other words, SFON did not explain significant variance in symbolic comparison performance once nonsymbolic comparison and mapping performance had been taken into account. This demonstrates that the relationship between SFON and symbolic processing skills can be accounted for by individual differences in nonsymbolic skills and mapping skills. With arithmetic performance as the dependent variable, we see a different pattern of results. SFON was a significant predictor of arithmetic when entered into the model at Step 2 and also at Step 3. This shows that SFON explains additional variance in arithmetic performance over that explained by nonsymbolic skills and mapping skills. Therefore, the relationship between SFON and arithmetic skills is only partly accounted for by individual differences in nonsymbolic skills and mapping skills (lending partial support for Prediction 2).

5.3.3   Discussion

These results add to our limited understanding of how children's informal (spontaneous) interactions with number relate to their early mathematical skills. First, they replicate previous studies showing that SFON is associated with an arithmetic advantage (e.g. Hannula et al., 2010). Second, they extend previous findings by providing evidence that this association persists even after controlling for individual differences in symbolic number identi-

fication, verbal skills and working memory. Third, and most importantly, they advance our theoretical understanding of how SFON exerts its positive influence on arithmetic skills. Specifically, the findings suggest that SFON leads to increased practice mapping between nonsymbolic and symbolic representations of number which improves symbolic fluency and, in part, leads to better counting and arithmetic skills. In the general discussion section below I review the conclusions that can be drawn from these findings and consider directions for future research.

5.4   General Discussion

The results of the studies presented in this chapter generate two main conclusions:

1. SFON is associated with a symbolic number processing advantage.

2. The relationship between SFON and arithmetic skills can be explained, in part, by individual differences in children's ability to map between nonsymbolic and symbolic representations of number.

In addition to these conclusions regarding the mechanisms of SFON, the findings provide evidence on developmental issues. The children in Study 4 (aged $8 - 9$ years) were older than the children in previous SFON studies, thus it is possible to conclude that the relationship between SFON and symbolic number skills is present not only in the initial stages of schooling, but in later years as well.

Following on from this research there are some important questions for future studies to address. In particular, since mapping ability only partly accounted for the relationship between SFON and arithmetic skills, there is a need to explore the additional mechanisms involved. Here I highlight two possible mechanisms. One possibility is that SFON improves the precision with which children execute arithmetic procedures. High SFON children may get more practice counting and as a result they may develop more mature counting strategies which lead to more accurate arithmetic calculations. We know that as children become more proficient at counting they become less reliant on finger counting and they start to use more mature counting strategies, e.g. 'counting on' as opposed to 'counting all'. Numerous studies

			Symbolic comparison		Arithmetic	
Model	Step	Independent variable	β	ΔR^2	β	ΔR^2
1	1	Digit recognition	.46***	.27***	.40***	.20***
		Working memory	−.14		−.06	
		Average word count	.09		.15	
	2	Digit recognition	.42***	.07**	.33***	.19***
		Working memory	−.12		−.03	
		Average word count	.01		.02	
		SFON Posting	.15		.23**	
		SFON Picture	.22**		.39***	
	3	Digit recognition	.21**	.26***	.21**	.09***
		Working memory	−.03		.02	
		Average word count	.00		.01	
		SFON Posting	.02		.15*	
		SFON Picture	.04		.29***	
		Nonsymbolic comparison	.25**		.16	
		Numerical mapping	.47***		.25**	
2	1	Digit recognition	.46***	.27***	.40***	.20***
		Working memory	−.14		−.06	
		Average word count	.09		.15	
	2	Digit recognition	.20**	.33***	.21*	.19***
		Working memory	−.03		.02	
		Average word count	.01		.10	
		Nonsymbolic comparison	.26***		.19*	
		Numerical mapping	.49***		.37***	
	3	Digit recognition	.21**	.002	.21**	.08***
		Working memory	−.03		.02	
		Average word count	.00		.01	
		Nonsymbolic comparison	.25**		.16	
		Numerical mapping	.47***		.25**	
		SFON Posting	.02		.15*	
		SFON Picture	.04		.29***	

Note. $N = 119$, *$p < .05$, **$p < .01$, ***$p < .001$.

Table 5.3: Linear regression models predicting symbolic comparison and arithmetic performance.

have related these advanced counting and arithmetic strategies to improved performance on arithmetic tasks (e.g. Geary, Hoard, Byrd-Craven, & DeSoto, 2004). Therefore, if SFON supports the acquisition of more mature counting strategies then it may also advance children's arithmetic skills, over and above the advantage provided by high SFON children's mapping ability.

A second possibility is that SFON provides an arithmetic advantage because it makes children better at extracting and modelling numerical information from real-world contexts. We know that being able to construct a mental representation of an arithmetic problem (i.e. understanding the quantitative relations between, and the actions upon different numerical sets in a problem) is an important process in numerical problem solving (Kintsch & Greeno, 1985; Thevenot, 2010; Verschaffel & De Corte, 1993). Children with high SFON tendency may not necessarily have more advanced computational skills; rather, they may be better at working out when (and which of) these computational skills need to be used. Note that the standardised arithmetic task used in the present study comprised several word-based problems in which children needed to extract and model numerical information from a real-word story context, e.g. buying apples, sharing sweets and losing toys.

As well as testing these possibilities, it would be valuable for future studies to examine issues surrounding causality. Data presented here are cross-sectional, thus we can only tentatively specify the causal nature of SFON based on prior longitudinal research. Hannula and Lehtinen (2005) showed that children's SFON predicted their counting skills, and likewise their counting skills predicted their SFON. This suggests that SFON and arithmetic skills are likely to develop together in a cumulative cycle. Further longitudinal work (and training studies) will allow us to determine whether SFON increases symbolic fluency, and therefore arithmetic skills, and/or vice versa. Importantly, this longitudinal work will also help to build a clearer picture of the developmental trajectory of children's SFON.

5.4.1   Limitations of Studies 4 & 5

As with all studies there are limitations to consider. First, as highlighted above, the data from Studies 4 and 5 is cross-sectional. It is thus important to recognise that the relationship between SFON and numerical skills is not

necessarily a causal one. Also, note that this relationship is likely to be explained by several factors, not just the ones measured in Study 5. One important factor that has not yet been investigated in SFON studies is inhibition. As noted by Hannula et al. (2010), in order to focus on a particular aspect (such as number), one has to simultaneously inhibit the processing of other aspects of a task or situation. Several studies have shown that children's inhibitory control skills are related to their mathematics achievement (e.g. Bull & Scerif, 2001), thus future studies should look at the link between SFON, inhibition and mathematical skills.

There is also a methodological limitation regarding the convergent validity of the SFON tasks. In Study 5, SFON scores on the Posting Task and the Picture Task were not significantly correlated. This lack of correlation may stem from the oppositely skewed distributions of children's performance on the two tasks. Scores on the Posting Task were negatively skewed while scores on the Picture Task were positively skewed. Alternatively, the lack of correlation between the tasks may be due to the different task demands and response modes. The Posting task required a nonverbal response whereas the Picture Task required a verbal response. The tasks can also be seen to differ in the extent to which the researcher focuses on (or points towards) number. On the Posting Task the researcher focuses on number in a way that they do not on the Picture task. The researcher does a "one by one posting action" rather than posting a "bunch of letters". Further research is needed to untangle the subtle differences between the tasks and the SFON constructs that they are measuring.

5.4.2   Summary of Findings

The research presented in this chapter aimed to increase our understanding of the mechanisms of SFON. The results from two studies were reported. The first study (Study 4) revealed that SFON is associated with a symbolic number processing advantage: The more children spontaneously attend to the numerical aspects of their environment, the more fluent they are at comparing number symbols. The second study (Study 5) demonstrated that the relationship between SFON and formal symbolic arithmetic can be explained, in part, by individual differences in children's nonsymbolic numerical processing skills and their ability to map between nonsymbolic and symbolic representations of number. Further research will allow us to

explore the additional mechanisms involved.

Chapter 6

Discussion of SFON Studies

The aim of this chapter is to bring together all of the empirical findings from the SFON studies presented in Chapters 2 – 5. First, I provide an overview of the findings. Next, I review the methodological and theoretical conclusions that can be drawn from these findings. Finally, I consider the educational implications of this work, and directions for future research, before concluding Part II of this thesis.

6.1 Overview of Findings

The research presented in Chapters 2 – 5 (Part II) of this thesis sought to further our understanding of children's spontaneous focusing on numerosity (SFON), a recently-discovered predictor of early counting and arithmetic skills. Five studies were conducted to examine both methodological and theoretical questions within this emerging body of research. The main findings from each study are summarised below.

Study 1, reported in Chapter 2, used a one-week test-retest design to investigate the test-retest reliability and convergent validity of the three tasks (Posting, Model & Finding) developed by Hannula and Lehtinen (2005) to assess children's SFON. The results showed that the tasks varied in their psychometric properties. The Posting Task and the Model Task showed moderate test-retest reliability and convergent validity. They also showed predictive validity as children's SFON scores were positively correlated with their numerical skills. A difficulty with these tasks is that children's SFON scores were driven largely by accuracy (i.e. producing the correct numeros-

ity), as opposed to verbal and nonverbal quantitative acts (i.e. counting gestures and utterances). This makes it difficult to separate the attentional aspects of SFON from children's procedural counting skills.

In view of these results, the studies presented in Chapter 3 developed and tested a new picture-based task for measuring children's SFON. This Picture Task, in which children are asked to describe a series of cartoon images, generates SFON scores from verbal number references rather than counting accuracy. In Study 2 (Main), the Picture Task was found to show high test-retest reliability and moderate convergent validity with the Posting Task measure. It also demonstrated predictive validity, in terms of children's numerical skills. In Study 3, children's SFON scores on the Picture Task were found to be highly correlated with their SFON during child-parent play. Together these findings indicate that the Picture Task has good psychometric properties and is related to an ecologically valid task.

Chapter 4 shifted away from methodological issues to focus on theoretical questions regarding the causes, or developmental roots, of SFON. Further data was presented from Study 3 to examine the role of parental influences. The findings revealed no significant relationship between children's SFON and their parents' SFON. Also, there was little relationship between children's SFON and home numeracy factors (including parents' academic expectations, home activities and attitudes relating to mathematics).

Finally, Chapter 5 investigated the mechanisms of SFON. In Study 4 it was found that SFON was associated with a symbolic numerical processing advantage. Children who spontaneously focused on numerosity on the Picture Task were more fluent at comparing number symbols than children who did not focus on numerosity. In Study 5 it was found that the relationship between SFON and arithmetic skills can be explained, in part, by individual differences in children's ability to map between nonsymbolic and symbolic representations of number.

6.2   Methodological Conclusions

The studies reported in Chapters 2 and 3 give rise to some novel methodological conclusions. It is evident from Study 1 that SFON can be reliably measured in preschool and school-aged children (aged $3-6$ years) with two out of the three tasks developed by Hannula and Lehtinen (2005). However,

these tasks are limited because children's SFON scores are driven largely by counting accuracy rather than verbal or nonverbal quantitative acts. If we wish to disentangle the attentional processes of SFON from children's numerical competencies, then we either need to use GFON tasks to demonstrate that all children are capable of recognising the numbers of items involved (see, for example, Hannula & Lehtinen, 2005), or we need a measure of SFON that is independent of counting skills. The new Picture Task developed in Study 2 can provide such a measure.

Two key conclusions are drawn in relation to the Picture Task: First, it shows measurement reliability and validity; and second, it relates to an ecologically valid measure of children's SFON. It is, therefore, highly suitable for research purposes. As stated above, the Picture Task offers an advantage over existing measures because it generates SFON scores that are independent of children's counting accuracy. Children are deemed to be focusing on numerosity if they enumerate something in the picture, regardless of whether this enumeration is correct. Importantly, this allows us to tease apart children's attention to numbers from their counting skills per se.

The Picture Task also affords some other advantages. As discussed in Chapter 3, Section 3.5.1, it provides several competing dimensions on which children can choose to focus (e.g. number, colour, emotional content). This contrasts with existing measures on which children can focus on little information other than number. Also, the picture stimuli can be adapted for use with children of all ages. This allows us to study SFON in a consistent manner throughout development. Moreover, the task is quick and easy to run. It can be administered flexibly with either paper-based or computer-based presentation, and it has the potential to be adapted for group (as well as individual) testing sessions.

6.3   Theoretical Conclusions

The findings speak to a number of theoretical issues within the emerging literature on SFON. First, they replicate previous studies showing a relationship between children's SFON and their numerical skills (Hannula & Lehtinen, 2005; Hannula et al., 2010). This replication is significant because previous research employed tasks that were unable to tease apart the attentional aspects of SFON from children's counting skills per se. Data from

the new Picture Task helps to confirm Hannula and Lehtinen's (2005) assertion that SFON is an "intentional separate sub-process in enumeration" (p. 237). Further to this, nearly all of the previous research has been carried out in Finland, with preschoolers who have yet to receive formal symbolic number instruction. By replicating these results with children in England (who start school 3 years earlier than children in Finland), it is possible to conclude that SFON is an important factor for explaining variations in numerical skills both in preschool and school-aged children.

As well as supporting previous findings, the present research yields some new conclusions. In terms of the causes of SFON, the results suggest that it does not stem from either (i) parental influences, or (ii) mathematics anxiety[1]. So where then does SFON come from? What makes some children more inclined to engage with numbers than others? It may be the case that SFON is fostered by something subtle in the home environment, something that was not captured by the Home Numeracy Questionnaire. This possibility is theoretically likely for two reasons. Firstly, children's SFON was positively correlated with their family's social and economic position, i.e. SES. Secondly, the questionnaire did not tell us anything about the quality of the home numeracy practices, only the frequency with which certain activities occurred. Follow-up studies using a broader range of home numeracy measures may allow us to test this possibility.

In terms of the consequences of SFON, the results are more conclusive. They show that SFON is positively associated with children's low-level symbolic processing skills as well as their higher-level arithmetic. This association persists even after controlling for other cognitive factors (e.g. verbal skills, working memory and symbolic number identification) and it is present throughout the primary school years. Finally, and perhaps most importantly, conclusions are drawn with regards to the mechanisms of SFON. The results suggest that its association with arithmetic is partly explained by individual differences in children's nonsymbolic skills and their ability to map between nonsymbolic and symbolic representations of number.

---

[1]Recall from Section 1.3 that there were no differences in the mathematics anxiety levels reported by low, middle and high SFON children. Mathematics anxiety is to be investigated separately in Part III.

6.4   Educational Implications

The findings show that SFON is an important factor in the development
of children's early numerical skills. This raises interesting questions as to
whether SFON is something that can be trained. Can we increase children's
tendency to recognise and use numbers in informal everyday contexts? If
so, do increases in SFON lead to better mathematical outcomes?

Researchers have started to explore these issues. Hannula et al. (2005)
conducted a preliminary small-scale intervention with 34 preschoolers in
Finnish day-care settings. In this study, an experimental group ($N = 17$)
received 4 weeks of day-care aimed at enhancing SFON whilst a control
group ($N = 17$) participated in regular day-care with no intervention. The
results revealed some positive effects of SFON enhancement. Those children
who started off with some initial SFON tendency showed greater SFON
during the follow-up than control children (matched for initial SFON). Im-
portantly, this SFON enhancement was associated with improved cardinality
skills, suggesting that SFON-based interventions may help to support chil-
dren's early counting development.

In addition to SFON training itself, low SFON children may benefit
from intervention activities that help them to map between nonsymbolic
and symbolic representations of number. Recall that the relationship be-
tween SFON and arithmetic is explained, in part, by the precision of these
nonsymbolic-to-symbolic mappings. One way of targeting these mappings is
through number board games in which children match quantities and sym-
bols (e.g. Siegler & Ramani, 2008). With more practice making the con-
nections between quantities and symbols, low SFON children may develop
more precise mappings between nonsymbolic and symbolic representations
of number. This in turn may lead to a better, more flexible understanding
of what numbers mean and the relationships between them.

Overall, the present studies highlight the role of children's informal
(spontaneous) interactions with number in the development of formal sym-
bolic number skills. This corresponds with previous research showing that
the more children engage with and enjoy informal numerical activities before
school, the more they engage with formal mathematics throughout school
and higher education (Linder, Powers-Costello, & Stegelin, 2011; Seefeldt
& Galper, 2008; Van de Walle & Lovin, 2006). The findings have impor-

tant implications for current debates over early years numeracy practices. In particular, they question recent educational policies, such as the British Government's EYFS Framework, which calls for more formal mathematical content in the preschool curriculum (Department for Education, 2012).

6.5   Future Research

In view of the findings from this research, there are a number of specific questions for future studies to address. These are discussed in the relevant sections below. Note that broader research directions are reviewed in the general discussion of this thesis (Part IV, Chapter 12).

6.5.1   Mechanisms of SFON

It was shown that children's mapping ability only partly accounted for the relationship between SFON and arithmetic so there is a need to investigate the additional mechanisms involved. How else might SFON exert its positive influence on children's early numerical skills?

As proposed in Chapter 5 (page 144), one possibility is that SFON improves the precision with which children execute arithmetic procedures. High SFON children may get more practice counting and as a result they may develop more mature counting strategies (e.g. 'counting on' as opposed to 'counting all'). An alternative possibility is that SFON provides an arithmetic advantage because it makes children better at extracting and modelling numerical information from real-world contexts. Children with high SFON tendency may not necessarily have more advanced computational skills; rather, they may be better at working out when (and which of) these computational skills need to be used.

One way of testing these possibilities would be to give low and high SFON children an arithmetic task in which problems are systematically varied for (i) computational complexity, and (ii) difficulty extracting and modelling numerical information. If SFON works by improving children's computational skills, then we would expect high SFON children to show a greater advantage (over low SFON children) for the problems with high computational complexity. On the other hand, if SFON works by improving children's ability to extract numerical information, then we would expect the performance advantage for high SFON children to be greater for the

arithmetic problems that require the extraction and modelling of numerical information. Note that it is also possible that SFON works by improving both computational skills and extraction skills, in which case we would expect to see a performance advantage for high SFON children on both sets of problems (those with high computational complexity and high extraction difficulty).

In addition to testing these potential mechanisms, it is important for future studies to examine issues surrounding causality. Data presented here are cross-sectional, therefore we cannot determine the direction of the relationships between SFON, mapping ability and arithmetic skills. As noted in the discussion of Chapter 5 (Section 5.4), it may be the case that these skills develop together in a cumulative cycle. This is theoretically likely because previous longitudinal research found SFON to be reciprocally related to children's counting skills (Hannula & Lehtinen, 2005). Further longitudinal work (and training studies) will extend our causal understanding of SFON. Crucially, if we can establish its causal mechanisms then we will have a firm basis to start developing interventions for low SFON children (see Section 6.5.3 for further discussion on interventions).

6.5.2   Developmental Roots of SFON

Recall that the present findings suggest that children's SFON does not stem from either (i) parental influences, or (ii) mathematics anxiety. Therefore, the causes (or developmental roots) of SFON remain a question for further investigation.

As stated in the theoretical conclusions (Section 6.3) it is possible that SFON is influenced by something subtle in the home environment; for example, something to do with the quality, rather than the frequency, of home numeracy activities. In order to test this possibility, there is a need to obtain more detailed measures of the specific kinds of number-related interactions that occur between parents and children. This would perhaps be best achieved through observations within the home environment; see Hur (2010) for an example observation study of mother-child "math talk" during a home cooking activity. So far, very few studies have collected this type of home numeracy data.

In future studies looking at the relationship between SFON and home numeracy factors researchers should also try to recruit more fathers. This

will allow us to examine whether there are any gender differences in the provision of home numeracy activities (and thus the promotion of SFON). There is some evidence to suggest that parents provide more mathematics input to their sons than their daughters (e.g. Chang, Sandhofer, & Brown, 2011)[2] but we don't know how mothers' and fathers' "math talk" compares. The limited number of studies published in this area have focused on parents' formal mathematics activities with older children (e.g. mathematics homework) rather than informal activities in the early years.

As well as examining parental influences on children's SFON, it would be valuable for future studies to look at the role of educational factors. Studies 1 and 3 investigated whether there were any differences in overall SFON tendency between school-aged children and preschoolers. The results were mixed (school-aged children showed higher levels of SFON than preschoolers on some measures but not others), therefore it is not clear whether children's SFON tendency increases as they get older and start to receive mathematics education in school. Note that even if these results had been conclusive, the experimental design would not have allowed us to separate the effects of formal education from the effects of age-related development.

One way of disentangling the effects of age and formal education on children's SFON would be to compare SFON tendencies in children in different countries, who start school at different ages. Specifically, researchers could run a cross-cultural experiment comparing SFON tendencies in children in Northern Ireland, England, Belgium and Finland (who start school, and start to receive symbolic number instruction, at 4, 5, 6, and 7 years, respectively). If SFON is enhanced by formal mathematics education then we would expect to see different patterns of SFON development across countries. Levels of SFON would depend on an interaction between children's age and country of origin. In particular, the $4-5$ year old children in Northern Ireland and England (receiving symbolic number instruction) would be expected to show significantly greater levels of SFON than the same aged children in Belgium and Finland (not receiving symbolic number instruction). If SFON develops independently of formal mathematics education then we would expect to see no significantly different levels of SFON when

---

[2]Note that these gender differences are discussed in more detail in relation to mathematics anxiety in Part III of this thesis (see page 168).

comparing the same aged children across countries.

Overall, further research into the developmental roots of SFON may help us to identify low SFON children who may be at risk from mathematical difficulties. Together with research into the mechanisms of SFON, this may provide a strong basis for developing SFON-based interventions (see Section 6.5.3 below).

6.5.3 SFON-based Interventions

From an applied perspective, future studies should examine the potential of SFON-based interventions. Can we support low SFON children's numerical development by encouraging, or training, them to focus on the numerical aspects of their environment? And, what are the best ways to direct children's attention to numerosity?

As discussed in the educational implications (Section 6.4), the results from an initial small-scale intervention were promising. Hannula et al. (2005) found that preschool children's SFON was enhanced through social interaction in day-care settings and, importantly, this enhancement was associated with improved cardinality skills. Note, however, that the effects of this intervention varied depending on children's SFON tendency to begin with. SFON was only enhanced in those children who started out with some initial SFON, therefore further studies are needed to investigate whether SFON can also be increased in children with very low initial SFON tendency.

There are at least three reasons why Hannula et al.'s (2005) intervention may have been less effective in low SFON children. Firstly, the one-month intervention may not have been not long enough to see effects in this group. Secondly, day-care practitioners may have found it difficult to direct low SFON children's attention to numerosity, and thus they may have engaged in fewer SFON promoting activities with these children. Thirdly, these children may have needed to experience SFON promoting activities in a wider range of contexts (both at home and in preschool). Follow-up studies could test these possibilities with (i) longer intervention periods, (ii) frequent monitoring of intervention delivery, and (iii) two or more intervention groups (e.g. preschool intervention/home intervention/preschool and home intervention).

6.6   Summary

To summarise, the empirical work presented in Part II of this thesis focused on a recently-discovered predictor of numerical skills, namely the extent to which children spontaneously focus on numerosity (SFON). The findings from five studies were reported. These findings add to the current literature on SFON both methodologically and theoretically. They advance our tools for measuring children's SFON in a reliable and valid manner. They also further our understanding of why SFON is associated with an arithmetic advantage. Specifically, they show that the relationship between SFON and mathematical skills can be explained, in part, by individual differences in children's ability to map between nonsymbolic and symbolic representations of number. This suggests that SFON works by increasing children's fluency with number symbols.

Part III

Mathematics Anxiety

Chapter 7

Parents' and Children's Mathematics Anxiety (Study 6)

Mathematics anxiety refers to the syndrome of negative emotions that many individuals experience when engaging in tasks demanding numerical or mathematical skills. It has long been recognised by educators and researchers and it has been shown to have a range of negative consequences, including poorer performance on mathematical tasks and avoidance of mathematics-related activities. Up until recently, research into mathematics anxiety has focused on older children and adults. As outlined in the literature review, little is known about the emergence of mathematics anxiety in early childhood. It is not clear how, or why, mathematics anxiety develops. One possibility is that parents play a role in shaping their children's attitudes (and anxieties) towards mathematics. Parents may transmit negative feelings towards the subject with comments such as "I've always been hopeless with numbers". In the current chapter I test this possibility by investigating the relationship between children's and parents' mathematics anxiety.

7.1 Introduction

While the causes of mathematics anxiety are not yet well understood (Eden et al., 2013), it is typically reported to have multiple origins (Godbey, 1997; Jain & Dowson, 2009; Norwood, 1994). These origins can be broadly clas-

sified into environmental (e.g. negative experiences at home or school), personal (e.g. lack of confidence) and cognitive (e.g. poor working memory) factors (Rubinsten & Tannock, 2010). Here I focus specifically on the environmental influence of parents.

As highlighted by Gunderson et al. (2012), there are two key ways in which parents may affect their children's developing attitudes (such as anxiety, self-efficacy and self-concept) towards mathematics. One way is through their expectations and beliefs about their child's competence in mathematics. Another way is through their own attitudes towards mathematics. In other words, parents may function both as "expectancy socializers" and as "role models" (Parsons, Adler, & Kaczala, 1982).

In support of the "expectancy socializers" account, Parsons et al. (1982) showed that parents' expectations of their child's mathematics ability predicted children's expectations and self-concepts, more so than children's previous mathematics performance (see also Jacobs, 1991, and Jayaratne, 1987). More recently, in a study with younger children, Vukovic, Roberts, and Green Wright (2013) demonstrated a relationship between parents' home support and expectations and children's mathematics anxiety. Interestingly, they found that parents' expectations had a positive effect on children's mathematical problem solving performance by reducing children's levels of anxiety.

It is important to note that within this body of research there has been a large emphasis on gender issues. Studies have shown that parents expectations of children's mathematics ability are often higher for boys than they are for girls, even when boys' and girls' levels of mathematics achievement don't differ (Eccles, Jacobs, & Harold, 1990; Yee & Eccles, 1988). These gender-biased expectations have been found for preschoolers as well as school-aged children (Blevins-Knabe & Musun-Miller, 1991) and they are greater for parents who hold stronger "math = male" stereotypes (Jacobs, 1991).

In terms of the "role models" account, there is little research examining the relationship between parents' attitudes towards mathematics and their children's attitudes towards mathematics (Gunderson et al., 2012). Thus far, studies looking at the adult-child transmission of mathematics attitudes have tended to focus on teachers (Beilock et al., 2010; Midgley, Feldlaufer, & Eccles, 1989). One study which did look at the concordance of parent-

child attitudes found some evidence for role modelling processes in $5^{\text{th}}-12^{\text{th}}$ grade students (Jayaratne, 1987). Jayaratne showed that parents' attitudes (including present and past ability, difficulty, enjoyment and effort in mathematics) were positively correlated with children's attitudes; however, these correlations were only significant between mothers and daughters. Similarly a study looking at the concordance of parent-child general anxiety (not specific to mathematics) found that the strongest relationship was between girls' anxiety and their mothers' anxiety (Adams & Sarason, 1963).

These gender differences in the transmission of attitudes and anxieties have also been observed in studies looking at the influence of teachers. Beilock et al. (2010) demonstrated that female teachers' mathematics anxiety was related to girls' (but not boys') mathematics achievement and endorsement of math–gender stereotypes. This may reflect the fact that children are more likely to emulate same-sex models than opposite-sex ones (Bussey & Bandura, 1984; Perry & Bussey, 1979). Further research is needed with male teachers and fathers to determine whether this is the case.

To summarise, parents are often cited as a possible cause of children's mathematics anxiety, yet few studies have directly tested this hypothesis. There is some evidence to suggest that parents act as "expectancy socializers", while less is known about their function as "role models". Study 6 sought to address this gap by investigating the link between children's mathematics anxiety and their parent's mathematics anxiety. Given previous reports of gender differences in this domain, the findings are reported separately for boys and girls.

7.1.1   Aim of Study 6

The aim of this study was to investigate the relationship between mathematics anxiety in primary school children and their parents. To achieve this aim, children (aged $6-9$ years) and parents each completed a self-report measure of mathematics anxiety.

7.2 Method

7.2.1 Participants

Thirty-eight child-parent dyads were recruited through the University of Nottingham's 'Summer Scientist Week' scheme.[1] Children (26 girls) were aged $6.0 - 9.8$ years ($M = 7.6$ years, $SD = 1.1$ years) and they came from a range of socioeconomic backgrounds (as measured by postcode). The gender distribution of child-parent dyads is presented in Table 7.1.

All parents provided written consent prior to the start of the study. Study procedures were approved by the University of Nottingham Ethics Committee (see Section 2.1.4 for full details on ethical considerations).

	Daughter	Son
Mother	25 (65%)	11 (29%)
Father	1 (3%)	1 (3%)

Table 7.1: The gender distribution (frequency and percentage) of the child-parent dyads.

7.2.2 Design

Children and parents took part in a single testing session in which they each (individually) completed a previously validated mathematics anxiety questionnaire. Note that children also completed other measures as part of the SFON strand of research (Study 2: Pilot) and the wider 'Summer Scientist Week' event.

7.2.3 Materials & Procedure

This section describes the experimental procedure for the child and parent mathematics anxiety measures only. The SFON data was collected as a separate pilot study presented previously in Part II, Chapter 3.

---

[1]A further 23 children were recruited and tested but their parents did not take part and thus their data is not analysed.

7.2.3.1   Child Mathematics Anxiety Questionnaire

Children were presented with the Child Mathematics Anxiety Questionnaire (CMAQ) developed by Ramirez et al. (2013) for children aged $5 - 9$ years. Two adaptations were made. First, any American words were replaced with British equivalents (e.g. candy was replaced with sweets). Second, a five-point scale was used instead of a sixteen-point sliding scale.

For each of 8 items children were asked to rate how they feel (or would feel) in various mathematics-related situations, on a five-point smiley-face scale. Four of the items involved specific mathematical calculations (e.g. "How do you feel when you have to solve $34 - 17$?"). The other four items involved specific scenarios in which a child might be confronted with mathematics (e.g. "How do you feel when getting your mathematics book and seeing all of the numbers in it?"). An example item is presented in Figure 7.1 and the full questionnaire can be found in Appendix C.

This questionnaire was administered by the researcher on a laptop computer. The researcher read aloud each item (presented sequentially) and then the child responded by pointing to the appropriate face (and the researcher recorded their response). For each item children received a score from $1 - 5$ (where $1 =$ not anxious and $5 =$ very anxious). Total mathematics anxiety scores were computed by averaging across all items. The questionnaire showed good internal reliability, Cronbach's $\alpha = .81$.
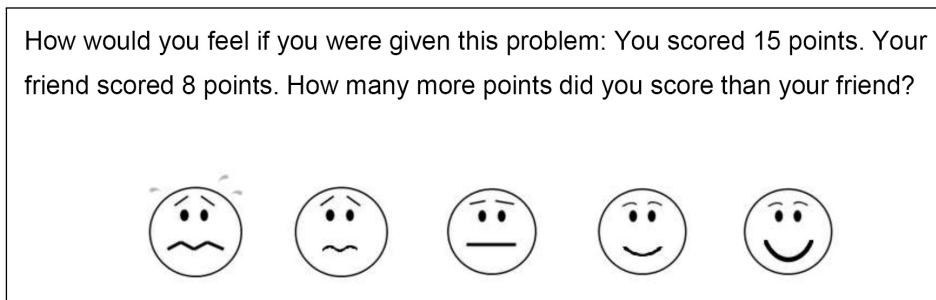
Figure 7.1: An example item from the CMAQ.

7.2.3.2   Parent Mathematics Anxiety Questionnaire

Parents were presented with the Mathematics Anxiety Scale–UK (MAS-UK) developed by Hunt, Clark-Carter, and Sheffied (2011). This 23-item questionnaire was originally designed and validated on British undergraduate students therefore some of the items were not relevant to this parent population ($N = 5$) and were removed from the scale.

For each of 18 items, parents were asked to rate how anxious they would feel in various mathematics-related situations on a five-point scale (ranging from not at all to very much). There were items measuring the following three factors: mathematics evaluation anxiety (e.g. "Having someone watch you multiply 12×23 on paper"), everyday/social mathematics anxiety (e.g. "Working out how much your shopping bill comes to") and mathematics observation anxiety (e.g. "Listening to someone talk about mathematics"). The full questionnaire can be found in Appendix D.

Parents completed this paper and pencil questionnaire individually, in a separate area to their child and the researcher. For each item they received a score from $1-5$ (where $1 =$ not anxious and $5 =$ very anxious). As with the child measure, total mathematics anxiety scores were computed by averaging across all items. The questionnaire showed good internal reliability overall, Cronbach's $\alpha = .89$, and for each of the three subscales: mathematics evaluation anxiety, Cronbach's $\alpha = .89$; everyday/social mathematics anxiety, Cronbach's $\alpha = .81$; and mathematics observation anxiety, Cronbach's $\alpha = .77$.

7.3   Results

The results reported in this section are organised as follows: First, descriptive statistics are presented for children's and parents' mathematics anxiety. Next, group-level comparisons are carried out to check for any age and gender differences in anxiety. Finally, correlations are performed to investigate the relationship between parents' mathematics anxiety and sons' and daughters' mathematics anxiety.

7.3.1   Descriptive Statistics & Group Comparisons

Overall, children showed individual differences in their reports of mathematics anxiety on the CMAQ ($M = 2.45$, $SD = 0.87$). Likewise, parents showed individual differences in their mathematics anxiety on the MAS-UK ($M = 1.78$, $SD = 0.59$). As shown in the histograms in Figure 7.2, children's mathematics anxiety scores were normally distributed whilst parents' scores were more positively skewed.

A one-way ANOVA was conducted to examine the effect of age on children's mathematics anxiety. This revealed no significant differences between the six-, seven-, eight- and nine-year-olds, $F(23, 34)=0.36$, $p = .780$. Further to this, an independent samples t-test was carried out to see whether there were any gender differences in children's mathematics anxiety. This demonstrated no significant differences between the boys' ($M = 2.56$, $SD = 1.07$) and girls' ($M = 2.40$, $SD = 0.78$) scores, $t(36) = 0.53$, $p = .598$. Given the lack of fathers who took part in the study ($N = 2$), it was not possible to test for any gender differences in parents' mathematics anxiety.

7.3.2   Correlational Analyses

Pearson correlations were calculated between parents' mathematics anxiety and sons' and daughters' mathematics anxiety. These revealed a significant positive association between the anxiety levels of parents and sons ($r = .76$, $p = .004$). In contrast there was no association between the anxiety levels of parents and daughters ($r = .05$, $p = .811$). As depicted in Figure 7.3, the higher the mathematics anxiety in parents the higher the mathematics anxiety in sons, but not daughters. A Fisher's z test revealed that these correlations were significantly different, $z = 2.39$, $p = .017$. The 95% confidence intervals were $[.33, .93]$ and $[-.34, .43]$ for sons and daughters, respectively.

As stated previously, there were only two fathers in this sample therefore it was not possible to test for differences in terms of parents' gender. However, if the fathers are excluded from the analyses then the results remained similar.[2]

---

[2]There was a significant positive correlation between the anxiety levels of mothers and sons ($r = .69$, $p = .020$), but not mothers and daughters ($r = -.02$, $p = .915$). These correlations were significantly different, $z = 2.09$, $p = .037$.
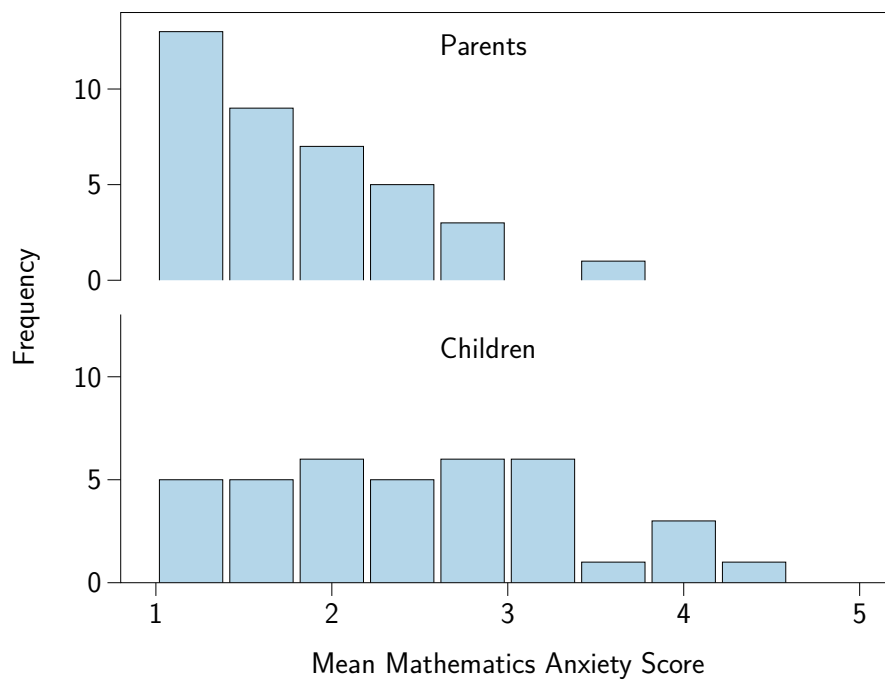
Figure 7.2: Distribution of children's and parents' mathematics anxiety scores.

7.4   Discussion

This study yielded three main results. First, in line with recent research, it was shown that mathematics anxiety is present in primary school children as young as 6 years (Ramirez et al., 2013; Wu et al., 2012; Young et al., 2012). Second, it was demonstrated that boys and girls show no significant differences in their anxiety. This adds to the mixed findings of gender differences in the literature (Devine, Fawcett, Szucs, & Dowker, 2012). Third, and most interestingly, it was revealed that parents' mathematics anxiety is related to sons' mathematics anxiety, but not daughters' mathematics anxiety. Below I interpret this result in view of the existing literature, before considering possible reasons for this gender difference.

As reviewed in the introduction to this chapter, only a few studies have examined the adult-child transmission of mathematics attitudes (and anxieties). These studies – conducted with female teachers and mothers with older children – found that role modelling processes were stronger in girls (Beilock et al., 2010; Jayaratne, 1987). This contrasts with the current finding that mathematics anxiety was related in parents and sons, but not parents and daughters. Given that the majority of parents in this study were mothers, these results cannot be reconciled by the fact that children are more apt to imitate same-sex models than opposite-sex ones (Bussey & Bandura, 1984).

How then can we account for these results? One possibility is that initially (in the early preschool and primary school years) parents engage in more mathematics-related activities with their sons than with their daughters. As a result, young boys may have more opportunities to pick up on their parents' mathematics anxiety than young girls. Later, as children get older and start to learn more complex mathematics in school, parents may become more involved with girls' mathematics (e.g. helping with homework) and the reverse may be true.

In accordance with this hypothesis, there is evidence for varying gender differences in the provision of parents' mathematics-related activities (Chang et al., 2011; Jacobs & Bleeker, 2004; Simpkins, Davis-Kean, & Eccles, 2005). Chang et al. (2011) demonstrated that mothers' number talk with their preschool children was significantly greater for boys than it was for girls; e.g. mothers used number words with boys in approximately 10%
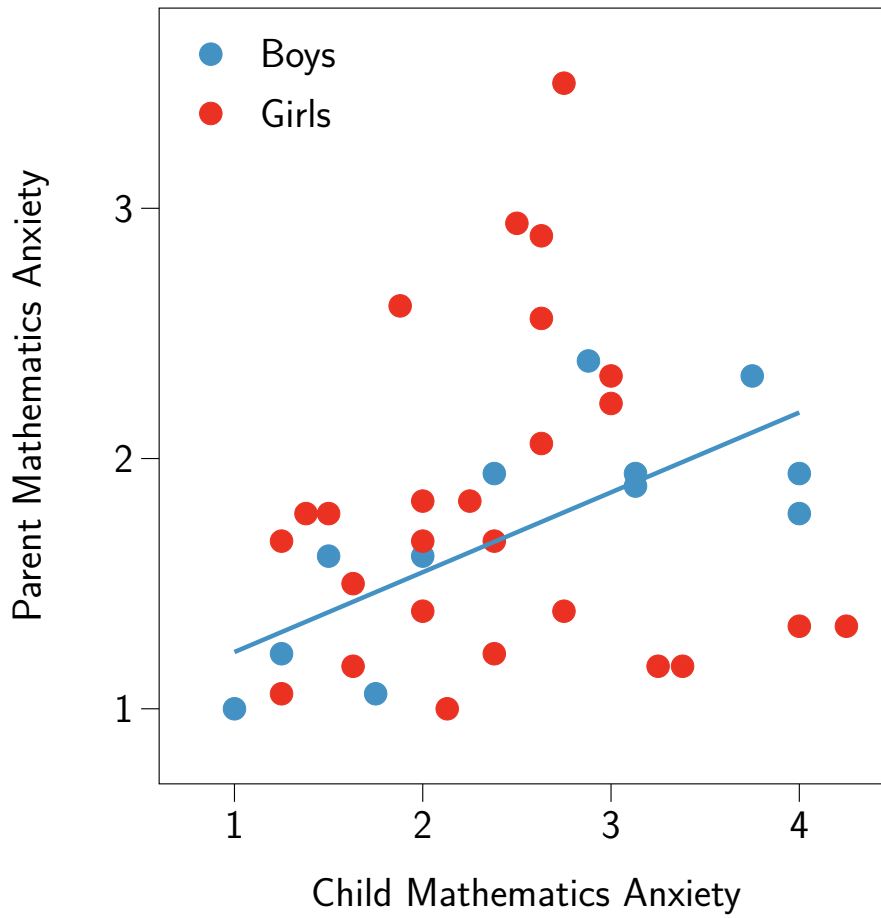
Figure 7.3: Scatterplot depicting the relationship between children's and parents' mathematics anxiety scores.

of utterances compared to 5% of utterances with girls. In a study with older elementary school children, Jacobs and Bleeker (2004) found that whilst parents provided more mathematics-related toys to their sons, they were more involved with their daughters' mathematics homework. Further to these findings, recall that in Study 3 of this thesis, parents of four- and five-year-olds reported more number skills activities with sons than daughters (see Part II, Chapter 4, page 115).

As an alternative to this account, the gender differences in the current study may stem from the way in which mathematics anxiety was measured. It is possible that the CMAQ was measuring different constructs in boys and girls. Girls may have been responding based on how they thought they should feel (e.g. in line with societal $math = male$ stereotypes), whilst boys were responding based on how they actually feel. Arguably, this would add more noise to the girls' mathematics anxiety scores, thus weakening the correlation between girls' mathematics anxiety and their parent's mathematics anxiety. If this is the case, then we would also expect to see gender differences in other correlates of mathematics anxiety (such as mathematics achievement). Boys' scores on the CMAQ should show more predictive validity than girls' scores.

7.4.1   Limitations of Study 6

The main limitation of this study is its small size and narrow scope, both of which should be taken into account when making generalisations from the findings. The results are based on a small sample of child-parent dyads ($N = 38$; children aged 6 to 9 years) who completed only a single measure of mathematics anxiety at a single timepoint. It is not clear whether the gender differences observed are stable over time, or robust across different measures of mathematics anxiety. In terms of the scope of the study, it is important to recognise that mathematics anxiety is inextricably linked to other cognitive, motivational and affective factors. It is situated within a much larger frame of children's motivational behaviour (Bandura, 1993; Berhenke, Miller, Brown, Seifer, & Dickstein, 2011) and it is hard to separate from other constructs such as perceived self-efficacy.[3]

---

[3]Self-efficacy can be defined as a person's beliefs about their ability to exercise control over a partiular course of action (Bandura, 1993).

7.4.2   Summary of Findings

The present study investigated the role of parents in the development of children's anxiety towards mathematics. Interestingly, the findings indicate that gender impacts on parental influence. Parents' mathematics anxiety was significantly associated with sons' mathematics anxiety, but not daughters' mathematics anxiety. This gender difference may reflect the fact that parents do more mathematics activities with boys, and thus are more likely to transmit their mathematics anxiety to their sons than their daughters. Alternatively, it may be a result of measurement issues. In the chapters that follow, further investigations are conducted to distinguish between these possibilities.

Chapter 8

Mathematics Anxiety and Numerical Processing Skills (Study 4 continued)

In the previous chapter it was found that mathematics anxiety was present in children as young as 6 years. This mathematics anxiety did not differ between boys and girls; however, there was a gender difference in the relationship between children's mathematics anxiety and their parents' mathematics anxiety. In the current chapter I examine whether this gender difference is specific to the relationship between parent-child anxiety, or whether it can also be found for the relationship between anxiety and performance. Here I present further data from Study 4, which investigates the association between children's mathematics anxiety and their basic (low-level) nonsymbolic and symbolic numerical processing skills.

8.1 Introduction

As reviewed in Chapter 1 (Section 1.2.3.3), research has demonstrated that mathematics anxiety is associated with poorer performance on a range of mathematical tasks. High mathematics-anxious individuals show impairments in basic, low-level, numerical processing – such as dot enumeration and Arabic digit comparison (Maloney et al., 2010, 2011) – as well as more complex, higher-level mathematics (see, Hembree, 1990 and Ma, 1999 for reviews).

Thus far, this research has focused largely on older children and adults (Eden et al., 2013). It is only recently that mathematics anxiety has started to be measured in younger children (Aarnos & Perkkilä, 2012; Krinzinger et al., 2009; Ramirez et al., 2013; Thomas & Dowker, 2000; Wu et al., 2012; Young et al., 2012), and it is not yet clear how this anxiety relates to the development of numerical skills. Some studies have found a relationship between primary school children's mathematics anxiety and their mathematics performance (Ramirez et al., 2013; Wu et al., 2012), whilst others haven't (Krinzinger et al., 2009; Thomas & Dowker, 2000). Note that these studies all used mathematics reasoning and calculation as measures of performance rather than tasks assessing basic numerical processing.

Following on from these studies (and the study reported in the previous chapter), the present research asked two questions: Is there a relationship between primary school children's mathematics anxiety and their low-level numerical processing skills? And, are there any gender differences in this relationship? Having found in Study 6 that boys' (but not girls') mathematics anxiety was related to their parents' mathematics anxiety, it was questioned whether boys' anxiety would more strongly predict their numerical skills than girls'. Recall that if the gender difference in the relationship between parent-child anxiety stems from the way in which mathematics anxiety is measured (i.e. we are measuring different constructs in boys and girls), then we would also expect to observe gender differences in the relationship between anxiety and other factors.

Interestingly, previous studies with adult and student populations have found some evidence for gender differences in the relationship between mathematics anxiety and performance (Betz, 1978; Devine et al., 2012; Hembree, 1990; Miller & Bichsel, 2004). In a meta-analysis of 151 studies, Hembree (1990) demonstrated that mathematics anxiety was a stronger predictor of mathematics performance in males than females. Similarly, Miller and Bichsel (2004) showed that mathematics anxiety explained significantly more variance in basic mathematics performance (but not applied mathematics performance) in men than women. In contrast to these findings, a recent study revealed that after controlling for test anxiety, the negative association between mathematics anxiety and mathematics performance remained for female students only (Devine et al., 2012).

Whilst these gender differences have been reported, there has been little

research into why they might exist (and why there are some inconsistencies). As noted by Devine et al. (2012), the number of studies that have looked at gender differences in the relationship between mathematics anxiety and performance is relatively low, especially when compared to the number of studies comparing males' and females' overall levels of mathematics anxiety. Given this gap, there is a need for further research looking at the nature of the relationships among mathematics anxiety, gender and mathematics performance.

8.1.1   Aim of Study 4

The aim of this study was to investigate the relationship between mathematics anxiety and numerical processing skills in primary school-aged boys and girls. To achieve this objective, children aged $8 - 9$ years completed a mathematics anxiety questionnaire (the CMAQ) and two numerical processing tasks (nonsymbolic and symbolic magnitude comparison). Given the mixed evidence in terms of the relationship between children's mathematics anxiety and their mathematics achievement, no prior predictions were made.

8.2   Method

The method of this study was reported previously in Chapter 5, Section 5.2.1, thus in this section I provide only a brief overview (with full details of the mathematics anxiety measure that was not previously described).

8.2.1   Participants & Design

One-hundred and nine children (58 boys) aged $8.1 - 9.8$ years took part in two testing sessions (approximately $1 - 2$ weeks apart). In the first session children completed a mathematics anxiety questionnaire administered by their teacher as a whole class activity. In the second session children completed a SFON task followed by two numerical processing tasks (nonsymbolic and symbolic magnitude comparison). These were administered individually by the researcher in a quiet corridor outside the classroom. Seven children had missing data on the mathematics anxiety questionnaire leaving a total of 102 complete datasets.

8.2.2   Materials & Procedure

This section details the experimental procedure for the mathematics anxiety questionnaire and the numerical processing tasks only. As reported in Chapter 1, there was no relationship between children's SFON and their mathematics anxiety, therefore the SFON aspect of the study was presented separately in Part II (Chapter 5).

Child Mathematics Anxiety Questionnaire

Children were presented with a paper and pencil version of the CMAQ used in Study 6. This was administered by teachers as part of normal classroom activity in order to minimise its association with the experimental tasks. The teacher read aloud each questionnaire item to the whole class allowing time for children to respond. An example item is presented in Figure 7.1 and the full questionnaire can be found in Appendix C.

For each of the 8 items children received a score from $1 - 5$ (where $1 =$ not anxious and $5 =$ very anxious). Total mathematics anxiety scores were computed by averaging across all items. The questionnaire showed good internal reliability, Cronbach's $\alpha = .87$.

Numerical Processing Tasks

Children completed a nonsymbolic comparison task followed by a symbolic comparison task (both presented on a 15 inch LCD laptop screen). On the nonsymbolic comparison task children had to decide which of two arrays of dots was more numerous (see page 126 and Figure 5.1a). On the symbolic comparison task children had to decide which of two Arabic digits was numerically larger (see page 127 and Figure 5.1b). Each task comprised 80 trials and the problems were identical across both tasks. Numerosities ranged from 5 to 30 and the numerical ratios between the two arrays were 0.5, 0.6, 0.7, 0.8 and their inverses. For each task children received an accuracy score based on the proportion of items they answered correctly.

8.3   Results

Descriptive statistics (and group comparisons) are first presented for children's mathematics anxiety scores. In the sections that follow, correlation

and regression analyses are performed to examine the relationships among children's mathematics anxiety, gender and numerical processing skills.

8.3.1 Descriptive Statistics & Group Comparisons

Overall, children showed individual differences in their reports of mathematics anxiety on the CMAQ ($M = 1.90$, $SD = 0.68$). Figure 8.1 presents the distribution of mathematics anxiety scores for boys and girls. This shows that boys' scores ($M = 1.84$, $SD = 0.73$) were more positively skewed than girls' scores ($M = 1.97$, $SD = 0.62$), yet note that there were no significant gender differences in mathematics anxiety, $t(100) = -1.01$, $p = .314$. This aligns with the findings from Study 6.

As reported in Chapter 5, Section 5.2.2, children also showed a range of performance on the nonsymbolic and symbolic comparison tasks. There were no gender differences in nonsymbolic comparison performance, but boys performed marginally better than girls on the symbolic version of the task.
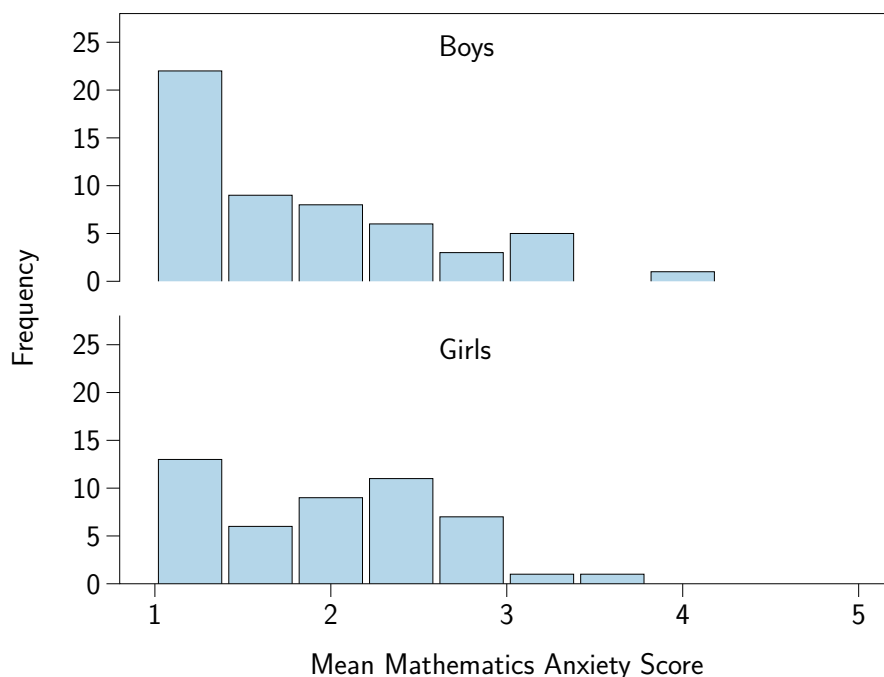
Figure 8.1: Distribution of mathematics anxiety scores on the CMAQ for boys and girls.

8.3.2 Correlational Analyses

To examine the relationship between mathematics anxiety and numerical comparison performance a series of Pearson correlations were conducted. Table 8.1 presents these correlations separately for boys and girls. Interestingly, there appeared to be some gender differences. For boys, there was a significant negative correlation between mathematics anxiety and symbolic comparison performance. As depicted in Figure 8.2, boys who reported high levels of mathematics anxiety performed worse on the symbolic comparison task than their low mathematics-anxious male peers. They also showed a tendency to perform worse on the nonsymbolic comparison task; however, note that there was one outlier and if this outlier is removed then the correlation is no longer significant ($r = -.22$, $p = .110$). In contrast to these results for boys, there were no significant correlations between mathematics anxiety and numerical comparison performance in girls.

To assess the significance of these gender differences a Fisher's z test was used to compare the correlations. This demonstrated that the relationship between mathematics anxiety and symbolic comparison performance was significantly stronger for boys than it was for girls, $z = -3.28$, $p = .001$. Conversely, the relationship between anxiety and nonsymbolic comparison performance did not differ significantly by gender, $z = -1.87$, $p = .062$, although there was a trend in the same direction.

As shown in Table 8.1, there also appeared to be a gender difference in the relationship between symbolic comparison performance and nonsymbolic comparison performance. In particular, there was a significant correlation between symbolic comparison and nonsymbolic comparison for boys, but not for girls. However, these correlations did not differ significantly, $z = 1.59$, $p = .112$.

8.3.3 Regression Analyses

Next, linear regression models were used to investigate the effects of mathematics anxiety and gender on children's numerical processing skills. Separate models were computed with nonsymbolic comparison accuracy and symbolic comparison accuracy as dependent variables. In both models the independent variables were gender, mathematics anxiety and the gender by mathematics anxiety interaction.

		1	2	3
Boys				
1	Mathematics anxiety	–		
2	Nonsymbolic comparison	$-.34$*	–	
3	Symbolic comparison	$-.55$***	$.46$**	–
Girls				
1	Mathematics anxiety	–		
2	Nonsymbolic comparison	$-.03$	–	
3	Symbolic comparison	$.05$	$.17$	–

Note. Pearson's r coefficients, *$p < .05$, **$p < .01$, ***$p < .001$.

Table 8.1: Zero-order correlations between all variables for boys and girls.

As reported in Table 8.2, both models were significant (accounting for approximately 9% of the variance in nonsymbolic comparison accuracy and 21% of the variance in symbolic comparison accuracy). Children's nonsymbolic comparison accuracy was significantly predicted by mathematics anxiety but not by gender or the gender by anxiety interaction. This indicates that high mathematics-anxious children showed poorer nonsymbolic comparison skills than low mathematics-anxious children, irrespective of gender. In contrast, children's symbolic comparison accuracy was significantly predicted by mathematics anxiety, gender and the gender by anxiety interaction. The significant gender by anxiety interaction ($\beta = 0.90$, $p = .003$) shows that mathematics anxiety was a stronger predictor of symbolic comparison accuracy in boys than girls.

8.4 Discussion

The results from this study suggest that there are gender differences in the relationship between children's mathematics anxiety and their (low-level) numerical processing skills. Specifically, it was shown that mathematics anxiety was more negatively associated with symbolic comparison performance in boys than it was in girls. Meanwhile, there were no significant gender differences in the relationship between mathematics anxiety and nonsymbolic comparison performance. Note that overall, mathematics anxiety accounted for more variance in symbolic comparison performance than nonsymbolic

Figure 8.2: Scatterplots depicting the relationship between mathematics anxiety and numerical comparison performance in boys and girls.

Independent variable	Nonsymbolic comparison $R^2 = .09$, $F(3, 98) = 3.06$, $p = .032$			Symbolic comparison $R^2 = .21$, $F(3, 98) = 8.56$, $p < .001$		
	b	$SE\ b$	β	b	$SE\ b$	β
Gender^a	-0.09	0.05	-0.54	-0.27	0.08	-0.99**
Mathematics anxiety	-0.04	0.02	-0.34**	-0.10	0.02	-0.51***
Gender × mathematics anxiety	0.04	0.02	0.49	0.11	0.04	0.90**

Note. $N = 102$, *$p < .05$, **$p < .01$, ***$p < .001$
a. 0 = male, 1 = female.

Table 8.2: Linear regression models predicting nonsymbolic and symbolic comparison performance by gender and mathematics anxiety.

comparison performance.

These gender differences may help to explain the inconsistent findings in the literature on primary school children's mathematics anxiety. It is possible that previous studies which found a relationship between children's mathematics anxiety and their mathematics performance (Ramirez et al., 2013; Wu et al., 2012) had a greater ratio of boys to girls than those which didn't (Krinzinger et al., 2009; Thomas & Dowker, 2000). Indeed, there is evidence that this was the case: For example, Wu et al.'s (2012) sample comprised more boys than girls (55.6% and 44.4% respectively), whilst Krinzinger et al.'s (2009) sample comprised more girls than boys (57.1% and 42.9% respectively).

Despite this evidence, it is not clear whether the gender differences found in the current study generalise to the different measures of mathematics anxiety and mathematics performance used in previous studies. The current study employed low-level numerical processing tasks as opposed to the higher-level mathematics reasoning and calculation measures used previously. To explore this issue, the next section will analyse some existing data from the Premature Infants' Skills in Mathematics (PRISM) study (Simms, 2014; Simms, Cragg, Gilmore, Marlow, & Johnson, 2013).

8.4.1   Analysis of Data from the PRISM Study

The PRISM study[1] (conducted in the UK between $2011-2013$) investigated differences between the mathematical skills of children born very premature with those born full-term. This section will analyse the mathematics anxiety data collected from the control children (those born full-term) as part of this study (Simms, 2014). In particular, it will examine whether there are gender differences in the relationship between mathematics anxiety and mathematics reasoning performance.

8.4.1.1   Method & Results

Participants were 76 children (39 boys) aged $7.8-10.8$ years ($M = 9.5$ years, $SD = 0.7$ years). All children participated in the PRISM study as controls. As part of a large battery of tasks they completed the Mathematics Anx-

---

[1]http://www.prismstudy.org.uk/

iety Questionnaire (MAQ), developed by Thomas and Dowker (2000), and the mathematical reasoning subtest of the Wechsler Individual Achievement Test (WIAT).

The MAQ comprised a total of 16 items. Children were presented with four mathematics-related situations (mathematics in general, written mathematics, mental mathematics and mathematics outside of school). For each situation they were asked four different types of questions ("How much do you like...?" "How worried are you if you have problems with...?" "How good are you at...?" and "How happy or unhappy are you if you have problems with...?"). The answer to each question was rated on a five-point pictorial scale (see Krinzinger et al., 2009). The questionnaire showed good internal reliability, Cronbach's $\alpha = .88$.

Children showed individual differences in their self-reported anxiety towards mathematics; total mathematics anxiety scores ($M = 29.86$ $SD = 13.27$) ranged from 4 to 63 out of a possible 64 (where low scores = low anxiety and high scores = high anxiety). In line with the results from Studies 4 and 6, there was no significant difference between boys' ($M = 27.31$ $SD = 12.11$) and girls' ($M = 32.54$ $SD = 14.06$) mathematics anxiety scores, $t(74) = -1.74$, $p = .086$. Likewise, there was no significant gender difference in mathematics reasoning scores, $t(74) = 0.04$, $p = .967$.

Pearson correlations were performed to examine the relationship between mathematics anxiety and mathematics reasoning performance in boys and girls. These revealed a significant negative correlation for boys ($r = -.51$, $p = .001$), but not for girls ($r = -.15$, $p = .367$). As shown in the scatterplot in Figure 8.3, high mathematics-anxious boys performed worse on the WIAT mathematics reasoning subtest than their low mathematics-anxious male peers. In contrast, there was no significant relationship between mathematics anxiety and mathematics reasoning performance in girls.

To investigate the effects of mathematics anxiety and gender on children's mathematics reasoning performance a linear regression model was computed. The results of this model are reported in Table 8.3. Overall the model was significant, accounting for approximately 18% of the variance in mathematics reasoning scores. Children's mathematics reasoning scores were significantly predicted by mathematics anxiety and the gender by mathematics anxiety interaction. The interaction between gender and mathematics anxiety ($\beta = 0.76$, $p = .018$) demonstrates that mathematics

anxiety was a stronger predictor of mathematics reasoning performance in boys than girls.

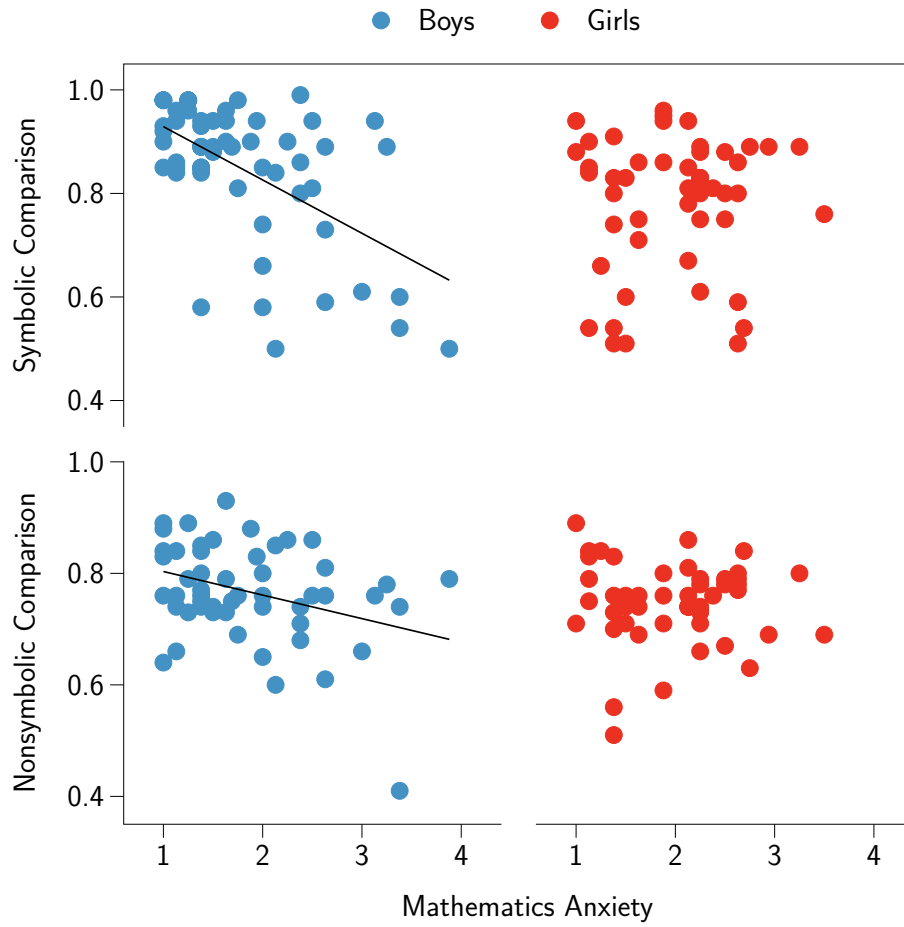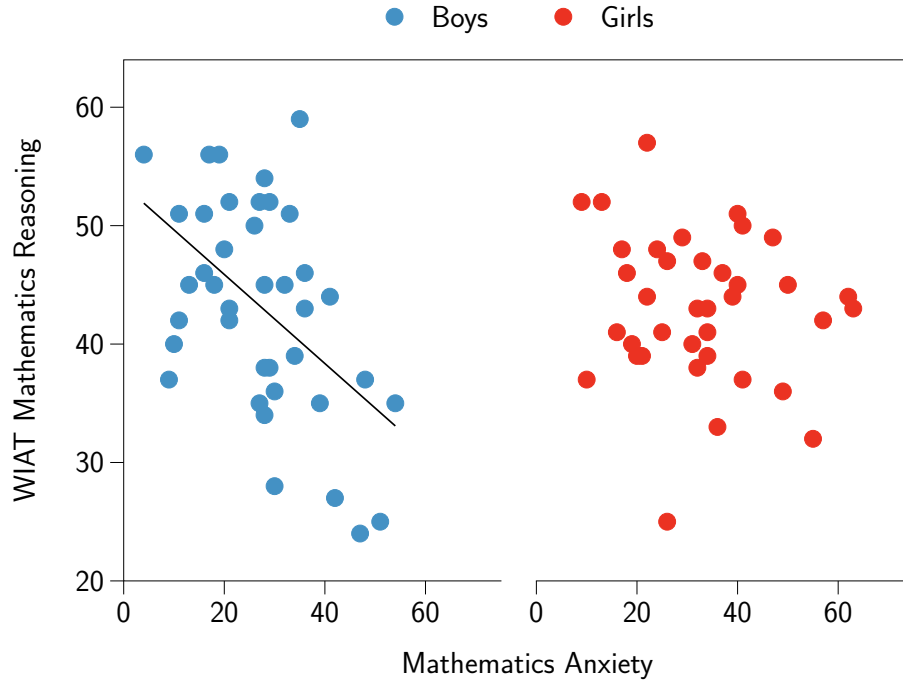Figure 8.3: Scatterplots depicting the relationship between mathematics anxiety and mathematics reasoning performance in boys and girls.

	Mathematics reasoning $R^2 = .18$, $F(3, 72) = 5.37$, $p = .002$		
Independent variable	b	$SE\ b$	β
Gender[a]	−8.12	4.13	−0.53
Mathematics anxiety	−0.38	0.10	−0.65**
Gender × mathematics anxiety	0.31	0.13	0.76*

Note. $N = 76$, *$p < .05$, **$p < .001$.
a. $0 =$ male, $1 =$ female.

Table 8.3: Linear regression model predicting mathematics reasoning performance by gender and mathematics anxiety.

8.4.2   Implications & Possible Explanations

The results yielded from the PRISM study data indicate that there are gender differences in the relationship between mathematics anxiety and mathematics reasoning performance. In particular, it was shown that mathematics anxiety was more negatively associated with mathematics reasoning performance in boys than it was in girls. This aligns with the findings from Study 4 and with previous studies conducted in older children and adults (Hembree, 1990; Miller & Bichsel, 2004).

Importantly, the PRISM study used different measures of mathematics anxiety and mathematics performance to those employed in Study 4. As a result, the gender differences observed can be seen to generalise across tasks. Mathematics anxiety appears to be a stronger predictor of mathematics performance in boys than it is in girls, and this holds for low-level numerical processing as well as higher-level mathematics. Given the cross-sectional nature of these studies, further longitudinal research is needed to determine the direction of causality between mathematics anxiety and performance.

These findings add to, and are consistent with, the gender differences found between children's and parents' mathematics anxiety in Study 6 (presented in Chapter 7). Recall that boys' (but not girls') mathematics anxiety was associated with their parents' mathematics anxiety. The consistency of these gender differences suggests that the scales used to assess mathematics anxiety may be measuring different constructs in boys and girls. As proposed in Chapter 7, it is possible that girls respond to mathematics anxiety questionnaires based on how anxious they think they should feel (e.g. in line with societal $math = male$ stereotypes), whilst boys respond based on how anxious they actually feel.

In contrast to this possibility, it may be the case that we are measuring the same mathematics anxiety construct in boys and girls, but that boys' performance is, for some reason, more negatively affected by their anxiety. As noted by Devine et al. (2012), mathematics is seen as a male domain therefore, boys may be less likely to communicate and receive help in dealing with their mathematics anxiety. This account does not explain the gender difference in the relationship between children's and parents' anxiety, but there could be an entirely separate explanation for this result.

Further investigations are needed to distinguish between these possibilities. In particular, one avenue for future research is to examine whether

(self-report) measures of mathematics anxiety vary in their psychometric properties for boys and girls. The current findings suggest that they have more predictive validity in boys, but it is not clear whether there are gender differences in their reliability. Crucially, with a better understanding of reliability issues we will be better able to assess the theoretical significance of the gender differences observed.

8.4.3   Limitations of Study 4

One limitation of this study (and other studies using the CMAQ) is the potential confound between children's self-reported mathematics anxiety and their mathematical skills. The CMAQ comprises eight items, four of which involve specific mathematical calculations (e.g. "How do you feel when you have to solve $34 - 17$?"). These calculation-related items may not only be measuring children's mathematics anxiety, but their perceived self-efficacy and their mathematical skills. Indeed, many children worked out the answers to these calculations and recorded the answers on their questionnaire.

This possible confound between mathematics anxiety and mathematical skills is less evident in other mathematics anxiety measures such as the MAQ (used in the PRISM Study reported in Section 8.4.1). Importantly, the findings from the PRISM study align with those from Study 4. This suggests that the relationship between boys' (but not girls') mathematics anxiety and mathematics performance is robust across different measures and is not just an artefact of this confound.

8.4.4   Summary of Findings

The research presented in this chapter demonstrates that there are gender differences in the relationship between primary school children's mathematics anxiety and their mathematics performance. Mathematics anxiety was found to be more negatively associated with mathematical outcomes in boys than it was in girls. This was the case for low-level symbolic comparison skills as well as higher-level mathematics. Interestingly, these gender differences are consistent with the results from Study 6 showing that boys' (but not girls') mathematics anxiety was related to their parents' mathematics anxiety. Together these findings call into question whether the self-report scales used to assess mathematics anxiety are measuring different constructs

in boys and girls.

Chapter 9

Measuring Mathematics Anxiety in Boys and Girls (Study 7)

The research presented in Chapters 7 and 8 revealed some consistent gender differences in the correlates of primary school children's mathematics anxiety. First, it was shown that boys' (but not girls') mathematics anxiety was associated with their parents' mathematics anxiety. Second, it was demonstrated that mathematics anxiety was more negatively associated with mathematical outcomes in boys than it was in girls. It is not clear whether these observed gender differences reflect actual structural differences between boys and girls, or whether they stem from measurement issues. To address this question, in the current chapter I examine possible gender differences in the reliability of the measures used to assess children's mathematics anxiety.

9.1   Introduction

There has been a considerable emphasis on gender differences in mathematics anxiety, yet little research has compared the psychometric properties of the measures used to assess mathematics anxiety in males and females. This is both surprising and problematic. As stated by Becker (2000), "gender differences in reliability may be the cause of apparent differences in correlations obtained in men and women when no real differences exist" (p. 372).

To help interpret the findings reported in Chapters 7 and 8, it is therefore important to consider whether there are gender differences in the reliability of the mathematics anxiety measures used. It is possible, for example, that boys' mathematics anxiety is more strongly correlated with mathematical outcomes because boys' reports of mathematics anxiety are more reliable than girls'. To investigate this possibility, the next section will examine whether there were any gender differences in the internal reliability of the CMAQ (used in Study 6 and Study 4) and the MAQ (used in the PRISM study).

9.1.1   Internal Reliability Analyses

As outlined in Chapter 2, Section 2.1.1, internal reliability refers to the consistency of the items within a measure, and it is typically estimated using Cronbach's alpha. If the Cronbach's alpha coefficient is high (≥ 0.7) then the items within a measure can be seen to be reliably measuring the same construct (Field, 2005).

9.1.1.1   Child Mathematics Anxiety Questionnaire (CMAQ)

The CMAQ was used to measure children's mathematics anxiety both in Study 6 and Study 4 (presented in Chapters 7 and 8). Recall that children were asked to rate how they feel (or would feel) in various mathematics-related situations, on a five-point smiley-face scale (see Appendix C). Overall, the CMAQ showed high internal reliability: Cronbach's $\alpha = .81$ and $.87$, for Studies 6 and 4 respectively. Table 9.1 presents the Cronbach's alpha coefficients separately for boys and girls in each study.

	Boys		Girls	
	N	Cronbach's α	N	Cronbach's α
Study 6	12	.83	26	.79
Study 4	54	.84	48	.76

Note. 8 items on the CMAQ.

Table 9.1: Cronbach's alpha reliability coefficients for boys and girls on the CMAQ used in Studies 6 and 4.

As shown in Table 9.1, the internal reliability of the CMAQ was slightly higher for boys than it was for girls in both studies. To examine whether there were any significant differences between these coefficients an F test was performed using the procedure outlined by Feldt, Woodruff, and Salih (1987).[1] Results revealed no significant differences between the Cronbach's alpha coefficients for boys and girls in Study 6, $F(11, 25) = 1.50$, $p = .313$. Moreover, there were no significant differences between the coefficients for boys and girls in Study 4, $F(53, 47) = 1.24$, $p = .080$, despite a trend towards higher Cronbach's alpha coefficients in boys. There is, therefore, no evidence for any gender differences in the internal reliability of the CMAQ.

9.1.1.2   Mathematics Anxiety Questionnaire (MAQ)

The MAQ was used to measure children's mathematics anxiety in the PRISM study, of which data were analysed in Chapter 8, Section 8.4.1. Children were presented with various mathematics-related situations and they were asked to rate how they feel on four subscales (like, worried, good and happy). As with the CMAQ, the answer to each question was rated on a five-point scale.

Overall, the internal reliability of the MAQ was high, Cronbach's $\alpha = .88$. Table 9.2 presents the Cronbach's alpha coefficients separately for boys and girls, both for the whole scale and the four subscales. A series of F tests were performed to examine whether there were any gender differences in these coefficients.[1]. In line with the findings from the CMAQ, results yielded no significant differences between the coefficients for boys and girls, either for the whole scale, $F(36, 38) = 0.79$, $p = .760$, or the four subscales ($ps > .172$). These results therefore suggest that there are no gender differences in the internal reliability of the MAQ.

9.1.2   Summary

The internal reliability analyses for the CMAQ and the MAQ showed a consistent pattern of results. The Cronbach's alpha coefficients were high

---

[1]The ratio of $(1 - \text{smaller } \alpha)$ over $(1 - \text{larger } \alpha)$ was compared to an F distribution with degrees of freedom equal to (N for larger $\alpha - 1$) and (N for smaller $\alpha - 1$) for the numerator and denominator, respectively.

	Cronbach's α	
	Boys ($N = 39$)	Girls ($N = 37$)
Like	.75	.81
Worried	.65	.74
Good	.73	.63
Happy	.63	.68
Total MAQ	.86	.89

Note. 16 items on the MAQ, 4 per subscale.

Table 9.2: Cronbach's alpha reliability coefficients for boys and girls on the MAQ (subscales and total scale) used in the PRISM study.

for both mathematics anxiety scales, thus suggesting that the items within each scale were reliably measuring the same construct. Interestingly, the coefficients did not differ significantly for boys and girls. In view of this, there is no evidence that the observed gender differences in the correlates of children's mathematics anxiety (found in Study 6, Study 4 and the PRISM study) were due to gender differences in the internal consistency of the scales. Still, it may be the case that boys and girls show differences in the temporal stability of their self-reported mathematics anxiety. This would suggest different amounts of noise associated with boys' and girls' mathematics anxiety scores, which therefore, may explain differences in their correlates. Study 7 sought to investigate this possibility using the test-retest method.

9.1.3   Aim of Study 7

The aim of this study was to investigate possible gender differences in the test-retest reliability of children's self-reported mathematics anxiety. To address this aim, children (aged $7 - 9$ years) completed the CMAQ at two timepoints, exactly 9 weeks apart.

9.2   Method

One-hundred and seventeen children (60 boys) aged $7.7 - 9.1$ years ($M = 8.3$ years, $SD = 0.4$ years) were recruited from a primary school in Nottinghamshire, UK. All children were in Years 3 and 4, Key Stage 2 of the National Curriculum. Five children were excluded from the study because they

were identified by their teacher as having learning difficulties. A further 13 children had missing data leaving a total of 99 complete datasets.

A test-retest design was employed with an exact time interval of 9 weeks. This time interval took place during term-time, with a one-week half-term holiday part way through. During each of two sessions children completed the CMAQ. As in Study 4, this was administered by teachers as a whole class (paper and pencil) activity; see Chapter 8, Section 8.2.2, for a description of the procedure. An example item is presented in Figure 7.1 and the full questionnaire can be found in Appendix C. For each of 8 items children received a score from 1 to 5 (where 1 = not anxious and 5 = very anxious). Total mathematics anxiety scores were computed by averaging across all items.

9.3   Results

First, descriptive statistics are presented for girls' and boys' mathematics anxiety at each timepoint. Next, preliminary analyses are conducted to check for (i) gender differences in mathematics anxiety, and (ii) changes in mathematics anxiety over time. Finally, in the main analyses section, reliability statistics are computed and compared across genders.

9.3.1   Descriptive Statistics & Preliminary Analyses

Table 9.3 shows the mean mathematics anxiety scores for boys and girls at each timepoint. In line with the previous studies, there were no significant gender differences in mathematics anxiety, either at Time 1, $t(103) = -0.67$, $p = .508$, or Time 2, $t(103) = -0.26$, $p = .797$. In terms of changes in mathematics anxiety over time, children's scores were lower at Time 2 than they were at Time 1. These differences over time were significant for girls, $t(49) = 2.35$, $p = .023$, and marginally significant for boys, $t(48) = 2.00$, $p = .052$. This demonstrates that children (both girls and boys) reported lower levels of mathematics anxiety at the second time of testing.

	Boys			Girls		
	N	M	SD	N	M	SD
Time 1	51	2.06	0.96	54	2.17	0.73
Time 2	54	1.89	0.87	51	1.93	0.76

Note. Mathematics anxiety scores ranged from $1 - 5$ (where $1 =$ not anxious and $5 =$ very anxious).

Table 9.3: Descriptive statistics for boys' and girls' mathematics anxiety scores on the CMAQ at Time 1 and Time 2.

9.3.2   Main Analyses

9.3.2.1   Internal Reliability

First, internal reliability estimates were calculated for boys' and girls' mathematics anxiety scores. Results showed that at Time 1, the internal reliability of the CMAQ was slightly higher for boys (Cronbach's $\alpha = .89$) than it was for girls (Cronbach's $\alpha = .82$). This difference was found to be significant, $F(50, 53) = 1.64$, $p = .039$. Note that this contrasts with the findings from the previous studies (see Section 9.1.1.1); however, these previous studies had smaller sample sizes and thus had less power to detect a significant difference. In line with the previous results, at Time 2, there were no significant gender differences in the internal reliability of the CMAQ (Cronbach's $\alpha = .88$ and .87 for boys and girls, respectively), $F(53, 50) = 1.08$, $p = .393$.

9.3.2.2   Test-Retest Reliability

Next, the test-retest reliability of the CMAQ was computed using Pearson correlational analyses. Table 9.4 reports the correlation coefficients for boys and girls, together with their 95% confidence intervals. Figure 9.1 presents a scatterplot of boys' and girls' mathematics anxiety scores at Time 1 and Time 2.

Overall, children's scores on the CMAQ showed high test-retest reliability. As shown in Table 9.4, this test-retest reliability was greater for boys than it was for girls. A Fisher's z was used to compare the correlation coefficients and the results revealed a significant gender difference, $z = -2.20$, $p = .028$. These findings therefore suggest that boys' reports of mathematics anxiety are more stable over time than girls' reports of mathematics anxiety.

	N	r	95% CI
Boys	49	.87***	[.78, .92]
Girls	50	.71***	[.53, .82]

Note. ***$p < .001$.

Table 9.4: Pearson (nine-week) test-retest correlation coefficients for the CMAQ for boys and girls.
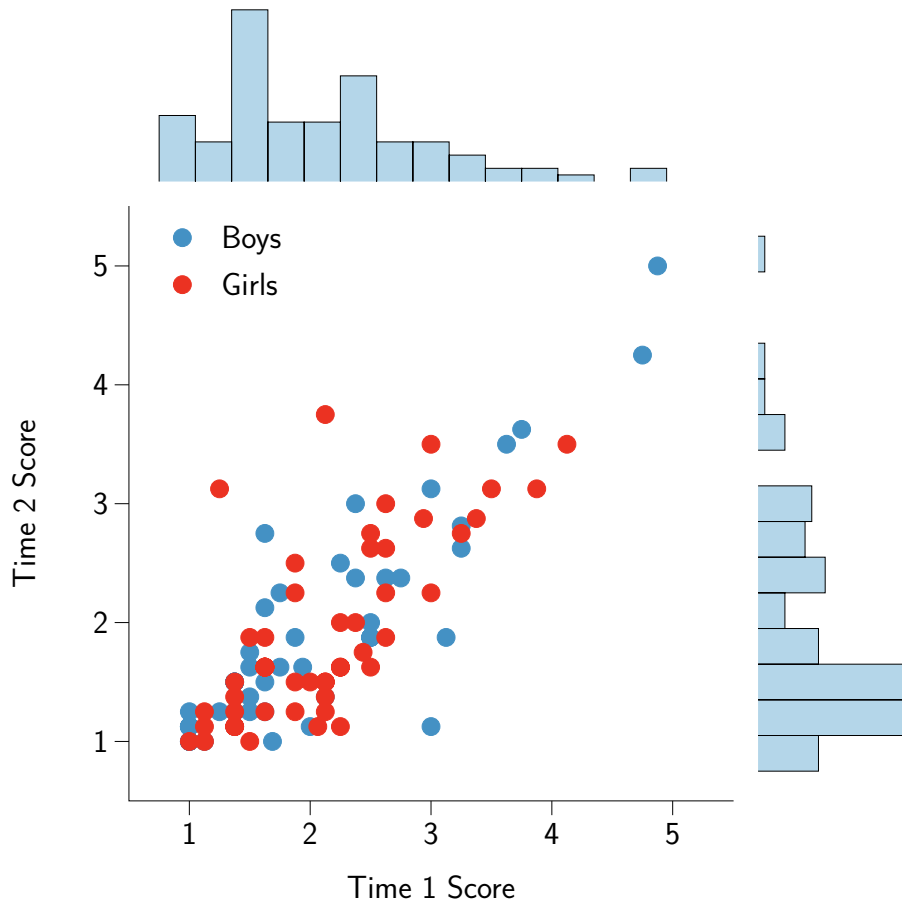

Figure 9.1: Scatterplot (with histogram) depicting the relationship between mathematics anxiety scores at Time 1 and Time 2 for boys and girls.

9.4   Discussion

This study used a nine-week test-retest design to investigate the temporal stability of children's self-reported mathematics anxiety on the CMAQ. In particular, it examined possible gender differences in the stability of children's mathematics anxiety scores. The results demonstrated that the CMAQ showed high test-retest reliability. Interestingly, this test-retest reliability was greater for boys than it was for girls.

These findings may help to interpret the gender differences in the correlates of children's mathematics anxiety reported in Chapters 7 and 8. Given that boys' scores on the CMAQ were found to be more reliable than girls' scores on the CMAQ, there is some evidence to suggest that there is more noise associated with girls' self-reported mathematics anxiety. This noise may be contributing to the lower correlations obtained in girls than boys (e.g. between mathematics anxiety and mathematical outcomes).

Despite the gender differences observed in this test-retest study, it is important to note that the reliability of the CMAQ was nevertheless still high for girls. The correlation between girls' mathematics anxiety scores at Time 1 and Time 2 was 0.71, which is above the level typically deemed to be suitable for research purposes (Field, 2005). It seems, therefore, that there may be more noise associated with girls' mathematics anxiety scores than boys' mathematics anxiety scores; however, girls' reports of mathematics anxiety are still stable over time.

Taken together with the internal reliability results, these findings suggest that the gender differences in the correlates of mathematics anxiety cannot be solely explained by psychometric issues. So how else might they be explained? One possibility is that there are actual structural differences in the causes and consequences of mathematics anxiety in boys and girls. For example, boys might get their mathematics anxiety from judgements of their own mathematics performance, whilst girls get their mathematics anxiety from societal influences such as math–gender stereotypes. If this is the case, then boys' mathematics anxiety should be more strongly correlated with mathematical outcomes than girls' mathematics anxiety (as found in Study 4 and the PRISM study). Meanwhile, girls' mathematics anxiety should be more strongly correlated with their implicit math–gender stereotypes. The research presented in the next chapter investigates this possibility.

9.4.1   Limitations of Study 7

There are two limitations worth noting. First, this study used a relatively narrow sample. Although the sample size was large ($N = 117$), children were all aged $7 - 9$ years and they were all recruited from a single primary school of mid to high SES. Second, there was only one measure used to assess children's mathematics anxiety (the CMAQ). As such, conclusions regarding the test-retest reliability of children's mathematics anxiety might not necessarily extend to other mathematics anxiety measures, or to children of different ages and socio-economic backgrounds.

In addition to these limitations it is important to note that both boys and girls reported higher levels of mathematics anxiety at Time 1 than they did (nine weeks later) at Time 2. This systematic difference suggests that there were some confounding practice or familiarity effects. Children may have felt generally more relaxed and less anxious at Time 2 because they had seen the questionnaire before and they knew what to expect.

9.4.2   Summary of Findings

The present chapter examined possible gender differences in the reliability of the measures used to assess children's mathematics anxiety. First, internal reliability analyses were conducted for the CMAQ and the MAQ used in the previous studies. Next, the temporal stability of the CMAQ was assessed using a nine-week test-retest design. The results provide some evidence to suggest that boys' reports of mathematics anxiety are more reliable than girls' reports of mathematics anxiety. Specifically, the test-retest reliability of the CMAQ was found to be greater for boys than it was for girls. Despite this evidence, the reliability of the CMAQ was still high for girls, thus suggesting that the gender differences in the correlates of mathematics anxiety (reported previously in Chapters 7 and 8) are not entirely due to measurement issues.

Chapter 10

Mathematics Anxiety, Gender and Implicit Stereotypes (Study 8)

In the current chapter I continue to investigate the causal factors underlying the gender differences in the correlates of children's mathematics anxiety (reported in Chapters 7 and 8). Recall that boys' mathematics anxiety was more negatively associated with their mathematical performance than girls' mathematics anxiety. Moreover, boys' (but not girls') mathematics anxiety was related to their parents' mathematics anxiety. These gender differences may reflect the fact that boys' and girls' mathematics anxiety stems from different sources. In particular, it may be the case that boys get their mathematics anxiety from judgements of their own mathematics performance, whilst girls get their mathematics anxiety from societal math–gender stereotypes. To test this possibility, here I examine possible gender differences in the relationships among children's mathematics anxiety, mathematics performance and implicit math–gender stereotypes.

10.1   Introduction

Research suggests that the gender gap in mathematics attainment has narrowed considerably over recent decades. Studies have shown that gender

differences in mathematics achievement are often not found and when they are found the effects are very small (Friedman, 1989; Hyde, Fennema, & Lamon, 1990).[1] According to a recent review, girls are now performing at similar levels to boys on standardised tests of mathematics achievement (Hyde, Lindberg, Linn, Ellis, & Williams, 2008).

Whilst the gender gap in mathematics performance is closing, mathematics is still stereotypically seen as a male domain, and women are still underrepresented in mathematics-intensive fields. In the UK, 44% of undergraduate mathematics qualifications were obtained by females in 2010/11 but only 25% of mathematics graduates from doctoral programmes were female. In higher academic positions, women occupied only 21% of UK mathematics lecturer and researcher posts and 6% of professorships (Higher Education Statistics Agency, 2011). A similar drop-off in women's participation persists in the US. Although nearly half of undergraduate mathematics degrees were obtained by women in US universities in 2010, only a quarter of faculty positions in the mathematical sciences were occupied by women (National Science Foundation, 2013).

Concerns over women's participation in mathematics have sparked an increase in research into the role of societal math–gender stereotypes. Numerous studies have shown that girls underperform on mathematics tests when their gender identity is made salient (see Steele et al., 2002, for a review). These stereotype-threat effects are robust across multiple settings and have been shown to emerge in children as young as $5 - 6$ years (Ambady, Shih, Kim, & Pittinsky, 2001; Galdi, Cadinu, & Tomasetto, 2014; Tomasetto, Alparone, & Cadinu, 2011).

In a recent study conducted by Galdi et al. (2014), 6-year-old children were asked to colour in one of three pictures: either a stereotype-consistent picture (a boy correctly solving a mathematics calculation), a stereotype-inconsistent picture (a girl correctly solving a mathematics calculation), or a control picture (a landscape). Results demonstrated that girls in the stereotype-consistent condition performed worse on a subsequent mathematics test than girls in the stereotype-inconsistent condition. Meanwhile, for boys, there was no effect of condition on mathematics performance. In

---

[1] A meta-analysis of 100 studies demonstrated that all effect sizes were less than $d = 0.10$ (Hyde et al., 1990)

line with these stereotype-threat effects, cross-national studies have demonstrated that gender differences in mathematics achievement are more likely to prevail in countries where implicit math–gender stereotypes are stronger (Nosek et al., 2009) and where there is less gender equity in terms of education and job opportunities (Else-Quest, Hyde, & Linn, 2010).

Interestingly, there is also growing evidence to suggest that societal stereotypes may undermine women's interest and motivation in mathematical domains as well as their test performance (e.g. Davies, Spencer, Quinn, & Gerhardstein, 2002). Research has shown that implicit math–gender stereotypes predict girls' attitudes and identification with mathematics (Nosek, Banaji, & Greenwald, 2002) together with their school enrolment preferences (Steffens, Jelenec, & Noack, 2010) and later career choices (Kiefer & Sekaquaptewa, 2007; Lane, Goh, & Driver-Linn, 2012; Smeding, 2012). Girls with stronger implicit $math = male$ stereotypes show more negative attitudes towards mathematics and avoid mathematics-related activities, regardless of whether they explicitly endorse these beliefs.

In the last few years, implicit math–gender stereotypes have been demonstrated in primary school children, both in Western (Cvencek, Meltzoff, & Greenwald, 2011; Cvencek, Meltzoff, & Kapur, 2014; Galdi et al., 2014) and Asian (Cvencek et al., 2014) cultures. Cvencek et al. (2014), for example, measured 7- to 11-year-old Singaporean children's automatic associations between mathematics and male using a child version of the Implicit Association Task (IAT). In accordance with standard IAT procedures, implicit stereotypes were based on differences in the speed with which children categorised mathematics with male (and reading with female) compared to mathematics with female (and reading with male). Note that such tasks have been used in many domains to measure implicit beliefs (see Gawronski & Payne, 2010, for a review).

Overall, research suggests that societal ($math = male$) stereotypes start to influence children from a very early age. These math–gender stereotypes have been shown to negatively affect girls' mathematics performance together with their attitudes towards mathematics and participation in mathematics-related activities. In view of these findings, it is thus theoretically plausible that girls' mathematics anxiety is moderated by their math–gender stereotypes. Girls with stronger implicit math–gender stereotypes may be more anxious towards mathematics than girls with weaker implicit

stereotypes (e.g. for fear that they might confirm the negative stereotype). In contrast, boys' mathematics anxiety may be unrelated to their math–gender stereotypes. If this is the case, then implicit stereotypes may be a key factor for explaining the gender differences in the correlates of children's mathematics anxiety observed in the previous studies (reported in Chapters 7 and 8). Study 8 sought to test this hypothesis.

10.1.1   Aim of Study 8

The aim of this study was to investigate possible gender differences in the relationships among primary school children's mathematics anxiety, mathematics performance and implicit math–gender stereotypes. To address this aim, children aged $7-9$ years completed the CMAQ, the mathematics reasoning subtest of the WIAT and a math–gender IAT. Children also completed a mathematics confidence questionnaire. This was included as a means to separate subtleties of mathematics affect which are not distinguished in the literature.

Two predictions were made. First, in line with the results from Study 4 and the PRISM study, it was hypothesised that boys' mathematics anxiety would be more strongly related to their mathematics performance than girls' mathematics anxiety (Prediction 1). Second, it was hypothesised that girls' mathematics anxiety would be more strongly related to their implicit math–gender stereotypes than boys' mathematics anxiety (Prediction 2).

10.2   Method

10.2.1   Participants

Two hundred children (101 girls) aged $7.8-9.8$ years ($M = 8.7$ years, $SD = 0.6$ years) were recruited from three primary schools in Nottinghamshire, UK. The schools were of mid to high SES with fewer pupils receiving free school meals than the national average. All children were in Years 3 and 4, Key Stage 2 of the UK National Curriculum.

Participation was voluntary and the children received stickers to thank them for taking part. All parents and guardians were informed of the nature of the study and were given the opportunity for their child to opt out. The number of children whose parents declined permission for them to take part

was low ($N = 2$). Children gave verbal assent to take part prior to the individual testing sessions. The study and consent procedures were approved by the Loughborough University Ethics Approvals (Human Participants) Subcommittee (see Section 2.1.4 for full details on ethical considerations).

Nineteen children were excluded from the analyses for having special educational needs. This included children with dyslexia ($N = 4$) who reported difficulty with the IAT. The final sample comprised 181 children (96 girls). Of these, 17 children had missing data and were excluded from the individual analyses that included those data. Ns are reported for each measure in the results section.

10.2.2   Design

Children completed four measures in four separate testing sessions. First, they completed a mathematics anxiety questionnaire administered by their teacher as a whole class activity. One to two weeks later they completed a mathematics reasoning task and an implicit association task. These were administered by a researcher in two individual testing sessions that took place in a quiet room or corridor outside the child's classroom. The order of these tasks was counterbalanced by gender. Children completed them in separate sessions, one in the morning and one in the afternoon (approximately 3 hours apart) to reduce crossover effects. Finally, one to two weeks later children completed a mathematics confidence questionnaire administered by their teacher as another whole class activity. Questionnaires were administered by teachers as part of normal classroom activity to minimise their association with the experimental tasks. Each task is described in the section below in the order in which it was presented.

10.2.3   Materials & Procedure

Child Mathematics Anxiety Questionnaire

Children were presented with the same 8-item CMAQ used in the previous studies. This questionnaire was administered by teachers as a whole class (paper and pencil) activity; see Chapter 8, Section 8.2.2, for a description of the procedure. An example item is presented in Figure 7.1 and the full questionnaire can be found in Appendix C. Note that teachers and children

in some classes used the term numeracy rather than mathematics, so for those children all use of the word mathematics was replaced with numeracy.

For each item children received a score from 1 to 5 (where 1 = not anxious and 5 = very anxious). Total mathematics anxiety scores were computed by averaging across all items. In line with previous studies the questionnaire showed good internal reliability, Cronbach's $\alpha = .87$. The internal reliability was higher for boys (Cronbach's $\alpha = .89$) than it was for girls (Cronbach's $\alpha = .84$) and this difference was significant, $F(84, 95) = 1.44$, $p = .042$.

Mathematical Reasoning Task

The mathematical reasoning subtest of the WIAT was administered in accordance with the standard procedure. Children completed a series of word problems read aloud by the researcher and presented in writing or with illustrations. They continued until they had answered six consecutive questions incorrectly. Each child received a raw score out of 63.

10.2.3.1   Child Implicit Association Task (Child IAT)

A child version of the math–gender stereotype IAT (adapted from Cvencek et al., 2011) was programmed using E-Prime software 2.0 and presented on a 15 inch LCD laptop screen. This task measured the relative strength of children's automatic associations between 'mathematics–boy' and 'literacy–girl' as compared to 'mathematics–girl' and 'literacy–boy'.

Children were asked to quickly sort words into target categories (boy, girl) and attribute categories (mathematics[2], literacy). There were seven blocks of trials: three single categorisation (practice) blocks and four double categorisation (experimental) blocks. The experimental blocks are depicted in Figure 10.1 and the IAT procedure is summarised in Table 10.1.

In Block 1 children were asked to sort boys' and girls' names by pressing left (boy) and right (girl) arrows in place of the 'c' and 'm' keys on a standard keyboard. The boys' and girls' names appeared one by one in the middle of the screen and were presented orally via headphones (at the onset of the written word on screen). Stimuli remained on the screen until the child

---

[2]Note that the category 'mathematics' was replaced with 'numeracy' for those children who were more familiar with the term numeracy.

responded. In Block 2 children practised sorting mathematics words and literacy words using the same left and right arrows. Each of these blocks comprised 16 trials with an inter-trial interval of 250ms. Incorrect trials were followed by a red question mark (for 250ms) and the trial was repeated until corrected.

Next, children completed two double categorisation blocks (Blocks 3 & 4) in which they were asked to sort target words and attribute words together. They responded to boys' names and mathematics words by pressing the left arrow and to girls' names and literacy words by pressing the right arrow. Target categories and words were presented in green and attribute categories and words were presented in purple to simplify the task.

Following these double categorisation blocks the target categories were switched so that girls' names were positioned on the left and boys' names on the right. In Block 5 children practised sorting target words into these new positions. Children then completed two more double categorisation blocks (Blocks 6 & 7), this time sorting girls' names and mathematics words with the left arrow and boys' names and literacy words with the right arrow.

For each condition, stereotype-congruent ('boy–mathematics', 'girl–literacy') and stereotype-incongruent ('girl–mathematics', 'boy–literacy'), the first double categorisation blocks (Blocks 3 & 6) comprised 16 trials and the second blocks (Blocks 4 & 7) comprised 40 trials.

The target and attribute words are listed in Table 10.2. Target words were chosen based on the most popular boys' and girls' names in the UK between 2004 and 2006 (the years in which the children were born). Attribute words were chosen based on their frequency in the Key Stage 2 National Curriculum (UK). This ensured that all words were familiar to the children. The order of the blocks and the order of the trials within each block were both fixed to ensure that the switching demands of the task were constant across participants.

IAT scores were calculated based on the D-algorithm (Greenwald, Nosek, & Banaji, 2003) which computes the difference in mean reaction time (RT) between the conditions divided by the standard deviation of all RTs for both conditions. This produces a bounded score between -2 and +2. In-group stereotype scores were derived so that positive scores reflected an association between mathematics and one's own gender.
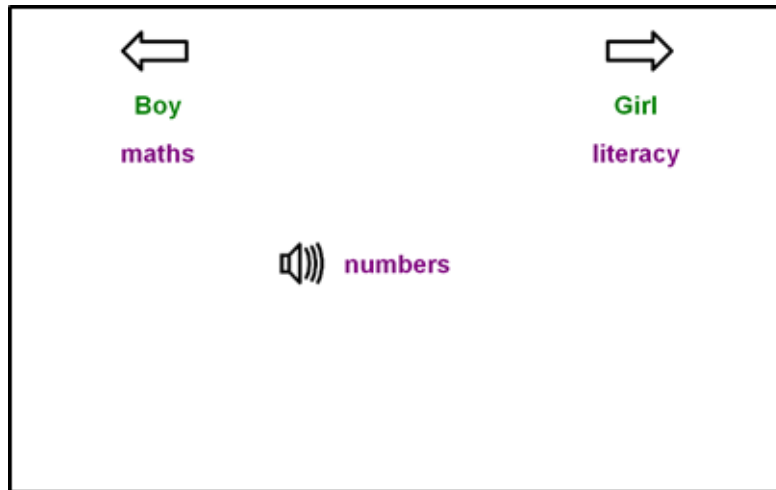
(a) stereotype-congruent
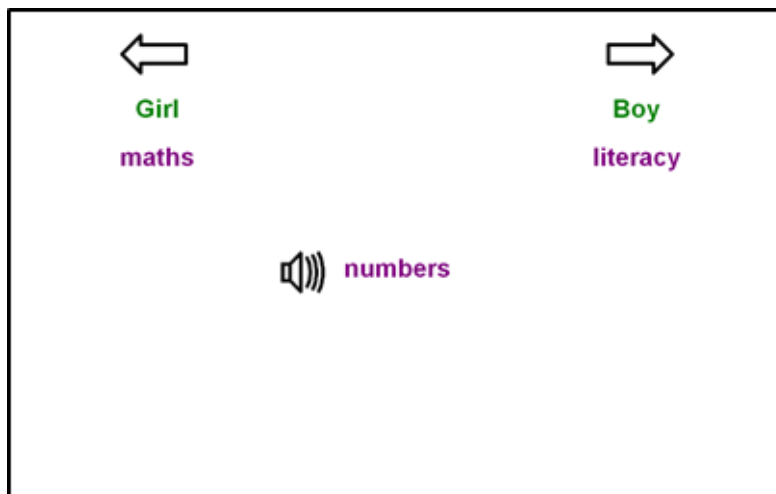
(b) stereotype-incongruent

Figure 10.1: A pictorial representation of the two experimental conditions of the Child IAT.

Note. In the stereotype-congruent condition boys' names and mathematics words share the same (left arrow) response key and girls' names and literacy words share the same (right arrow) response key. In the stereotype-incongruent condition the pairings are switched. Children with the math–gender stereotype ('mathematics is for boys') should respond faster to the stereotype-congruent condition than the stereotype-incongruent condition.

Block	Block Type	No. Trials	Categorisation Type	Description
1	Practice	16	Single (target) categorisation	Boy (left), Girl (right)
2	Practice	16	Single (attribute) categoriation	mathematics (left), literacy (right)
3	Experimental	16	Double categoriation	Boy/mathematics (left), Girl/literacy (right)
4	Experimental	40	Double categoriation	Boy/mathematics (left), Girl/literacy (right)
5	Practice	16	Single (target) categorisation	Girl (left), Boy (right)
6	Experimental	16	Double categoriation	Girl/mathematics (left), Boy/literacy (right)
7	Experimental	40	Double categoriation	Girl/mathematics (left), Boy/literacy (right)

Table 10.1: Implicit Association Task procedure.

Category			
Boy	Girl	Mathematics	Literacy
Jack	Sophie	numbers	letters
Thomas	Emily	counting	spelling
William	Jessica	subtract	reading
Joshua	Ellie	multiply	writing

Table 10.2: Implicit Association Task word stimuli.

10.2.3.2 Mathematics Confidence Questionnaire

Children were presented with an 8-item (paper and pencil) questionnaire designed to assess their confidence in mathematics. This questionnaire was adapted from the self-rating subscale of the Mathematics Attitudes and Anxiety Questionnaire (Thomas & Dowker, 2000; Krinzinger et al., 2007).

Children were asked to rate their ability at various aspects of mathematics, on a five-point scale of ticks and crosses (e.g. "How good are you at working out mathematics problems on paper?").[3] There were six items focusing on the following areas: mathematics in general, mental sums, written sums, easy mathematics, difficult mathematics and learning new things in mathematics. In addition there were two control items which asked children to rate their ability at reading and writing (Literacy Confidence). An example item is presented in Figure 10.2 and the full questionnaire can be found in Appendix E.

As with the mathematics anxiety questionnaire, the teacher read aloud each questionnaire item to the whole class allowing time for children to respond. For each item children received a score from 1 to 5 (where 1 = not confident and 5 = very confident). Total mathematics confidence scores were computed by averaging across all items. The questionnaire showed good internal reliability, Cronbach's $\alpha = .82$ for the six mathematics items. The internal reliability did not differ for boys and girls.
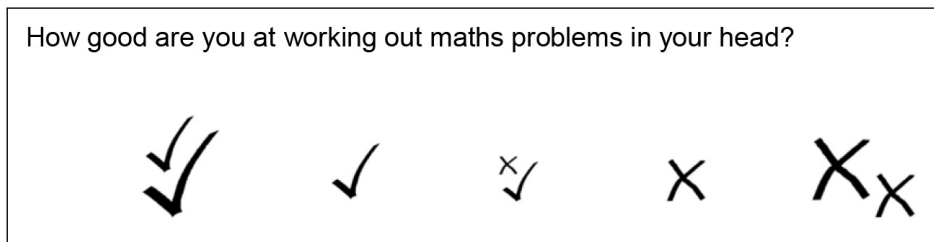

Figure 10.2: An example item from the Mathematics Confidence Questionnaire.

---

[3]As with the mathematics anxiety questionnaire and the IAT, use of the word mathematics vs. numeracy varied across participants, depending on that which children were most familiar with using in school.

10.3   Results

The results are reported in three sections. First, descriptive statistics (and group comparisons) are presented for boys' and girls' scores on each of the measures (mathematics anxiety, mathematics reasoning performance, implicit in-group stereotypes and mathematics confidence). Next, the correlations between these measures are explored separately for boys and girls. Finally, to test the predictions outlined in Section 10.1.1, the nature of these relationships is examined with a series of regression analyses.

10.3.1   Descriptive Statistics & Group Comparisons

Descriptive statistics for all measures are presented separately for boys and girls in Table 10.3. A series of independent samples t-tests were conducted to examine whether there were any gender differences in children's scores on each of these measures. The alpha level was Bonferroni corrected to .01 (.05/5) to control for the effect of multiple group comparisons. Results from the IAT showed that boys associated mathematics more strongly with their own gender than did girls, $t(176) = 7.74$, $p < .001$. There were also trends towards gender differences in children's ratings of confidence. Boys reported higher levels of mathematics confidence than girls, $t(162) = 2.20$, $p = .029$, whereas girls reported higher levels of literacy confidence than boys, $t(162) = 2.64$, $p = .009$. There were no gender differences in terms of mathematics anxiety or mathematics reasoning performance ($ps > .29$).

10.3.2   Correlational Analyses

Correlations between all variables are reported separately for boys and girls in Table 10.4. These show that in-group math–gender stereotypes were significantly correlated with all mathematical outcome measures in boys. As depicted in the scatterplots in Figure 10.3, the more boys associated mathematics with their own gender the higher their mathematics reasoning performance, the lower their mathematics anxiety and the higher their mathematics confidence. For girls, in-group math–gender stereotypes were significantly correlated with mathematics confidence, but not with mathematics performance or mathematics anxiety. A Fisher's z test revealed that the differences between these correlations for boys and girls were not

	Boys			Girls		
Measure	*N*	*M*	*SD*	*N*	*M*	*SD*
Mathematics anxiety	85	1.98	0.94	93	2.06	0.78
Mathematics reasoning	85	41.08	8.86	93	39.82	6.71
In-group stereotypes	85	0.28	0.29	93	−0.12	0.40
Mathematics confidence	78	4.10	0.83	86	3.84	0.73
Literacy confidence	78	3.94	0.77	86	4.22	0.62

Table 10.3: Descriptive statistics for all measures for boys and girls.

significant. Nevertheless, the was a trend towards boys' in-group stereotypes being more strongly associated with their mathematics reasoning performance than girls' in-group stereotypes, $z = 1.32$, $p = .093$.

Both boys and girls showed a strong negative correlation between the two affective measures mathematics anxiety and mathematics confidence ($r = -.715$ and $-.681$ respectively), reflecting the fact that children with higher levels of mathematics anxiety reported lower levels of mathematics confidence. These affective measures were significantly related to mathematics reasoning performance in boys and girls. The control measure, literacy confidence, was not related to mathematics performance in either gender.

		1	2	3	4	5
Boys						
1	Mathematics anxiety	–				
2	Mathematics reasoning	−.64***	–			
3	In-group stereotypes	−.28*	.32**	–		
4	Mathematics confidence	−.68***	.66***	.32**	–	
5	Literacy confidence	−.17	.09	−.05	.37**	–
Girls						
1	Mathematics anxiety	–				
2	Mathematics reasoning	−.35**	–			
3	In-group stereotypes	−.19	.13	–		
4	Mathematics confidence	−.72***	.36**	.24*	–	
5	Literacy confidence	−.26*	.16	−.03	.53**	–

Note. Pearson's r coefficients, *$p < .05$, **$p < .01$, ***$p < .001$.

Table 10.4: Zero-order correlations between all variables for boys and girls.

10.3.3   Regression Analyses

Two linear regression models were computed to test the predictions outlined in Section 10.1.1. The first model investigated whether boys' mathematics anxiety was more strongly related to their mathematics performance than girls' mathematics anxiety (Prediction 1). To test this hypothesis, mathematics reasoning performance was predicted by mathematics anxiety, gender and the mathematics anxiety by gender interaction. The second model
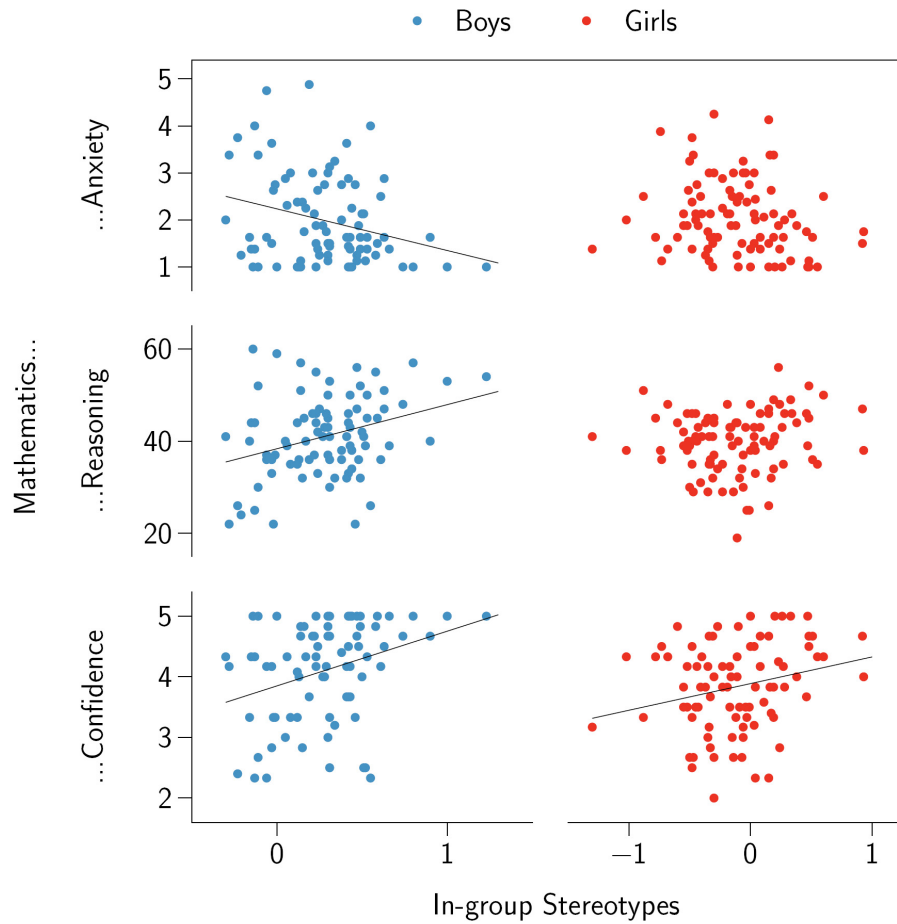
Figure 10.3: Scatterplots depicting the relationship between in-group stereotypes and each of the three mathematical outcome measures (mathematics anxiety, mathematics reasoning and mathematics confidence).

Independent variable	Mathematics reasoning $R^2 = .30$, $F(3, 174) = 25.05$, $p < .001$			In-group stereotypes $R^2 = .29$, $F(3, 174) = 23.79$, $p < .001$		
	b	$SE\ b$	β	b	$SE\ b$	β
Gender^a	-7.06	2.56	-0.45**	-0.38	0.14	-0.47**
Mathematics anxiety	-6.03	0.77	-0.66***	-0.09	0.04	-0.18*
Gender $\times$ mathematics anxiety	3.04	1.17	0.46*	-0.01	0.06	-0.04

Note. $N = 178$, *$p < .05$, **$p < .01$, ***$p < .001$
a. 0 = male, 1 = female.

Table 10.5: Linear regression models predicting mathematics reasoning performance and in-group stereotypes by gender and mathematics anxiety.

investigated whether girls' mathematics anxiety was more strongly related to their implicit math–gender stereotypes than boys' mathematics anxiety (Prediction 2). To test this hypothesis, in-group stereotypes were predicted by mathematics anxiety, gender and the mathematics anxiety by gender interaction. The results of both models are reported in Table 10.5.

Overall, both models were highly significant (accounting for approximately 30% of the variance in mathematics reasoning performance and 29% of the variance in implicit in-group stereotypes). The results from the first model demonstrate that children's mathematics reasoning was significantly predicted by mathematics anxiety, gender and the gender by anxiety interaction. The significant gender by anxiety interaction ($\beta = 0.46$, $p = .010$) shows that mathematics anxiety was a stronger predictor of mathematics reasoning performance in boys than it was in girls. This aligns with the findings from the PRISM study (reported in Chapter 8) and lends support for Prediction 1.

The results from the second model show that children's implicit in-group stereotypes were significantly predicted by gender and mathematics anxiety. The main effect of gender demonstrates that boys were more likely to associate mathematics with their own gender than girls. The main effect of mathematics anxiety shows that high mathematics-anxious children were less likely to associate mathematics with their own gender than low mathematics-anxious children. The interaction between gender and anxiety was not significant which indicates that there were no gender differences in the extent to which mathematics anxiety predicted in-group stereotypes. This therefore provides no support for Prediction 2.

10.4   Discussion

The study presented in this chapter sought to examine possible gender differences in the relationships among primary school children's mathematics anxiety, mathematics performance and implicit math–gender stereotypes. It was hypothesised (i) that boys' mathematics anxiety would be more strongly related to their mathematics performance than girls' mathematics anxiety (Prediction 1) and (ii) that girls' mathematics anxiety would be more strongly related to their implicit math–gender stereotypes than boys' mathematics anxiety (Prediction 2). The results yielded support for Predic-

tion 1, but not Prediction 2. Below I discuss these findings in relation to the earlier mathematics anxiety studies and the wider literature on stereotype threat.

In line with the findings from the previous studies, these results suggest that there are no significant gender differences in children's overall levels of mathematics anxiety; however, there are gender differences in terms of the strength of the relationship between mathematics anxiety and mathematics performance. In particular, the association between mathematics anxiety and mathematics performance is stronger for boys than it is for girls (thus supporting Prediction 1). Interestingly, this gender difference does not appear to be explained by gender differences in the relationship between children's mathematics anxiety and implicit math–gender stereotypes. Contrary to Prediction 2, the current results indicate that the association between mathematics anxiety and implicit stereotypes does not differ between boys and girls.

In view of these findings, the reasons behind the gender differences in the correlates of children's mathematics anxiety (reported in Chapters 7 and 8) remain a question for further research. It is evident from the reliability analyses presented in Chapter 9 that they cannot be solely explained by psychometric issues. As such, it seems that there may be actual structural differences in the causes and consequences of mathematics anxiety in boys and girls. Further possible explanations for the observed gender differences are considered in the next chapter (Section 11.4) which reviews all of the findings from the mathematics anxiety part of this thesis.

Whilst these findings do not provide an explanation for the gender differences in the previous studies, they do add to the growing body of research on primary school children's implicit math–gender stereotypes. First, they replicate recent studies showing that implicit $math = male$ stereotypes are acquired early in the absence of gender differences in mathematics achievement (e.g. Cvencek et al., 2011). Second, they extend previous findings by demonstrating that these implicit stereotypes are related to affective factors, such as mathematics anxiety, as well as mathematics test performance. Thus far, there has been some evidence to suggest that math–gender biases are related to motivational variables (e.g. school enrolment preferences and later career choices), but these studies have tended to be conducted with older children and university students (Kiefer & Sekaquaptewa, 2007; Lane

et al., 2012; Smeding, 2012).

Further to this, previous studies have focused primarily on the impact of societal stereotypes on girls. Research has emphasised that math–gender biases undermine female students' (but not male students') performance and participation in mathematics (e.g. Spencer, Steele, & Quinn, 1999). In the current study it was found that implicit stereotypes were related to mathematical outcomes (mathematics anxiety, mathematics reasoning and mathematics confidence) in boys as well as girls. In fact, as shown in Figure 10.3, there was a trend towards in-group stereotypes being more strongly associated with boys' mathematics reasoning performance ($r = .32$) than girls' mathematics reasoning performance ($r = .13$). The difference between these correlations was marginally significant. These results therefore question the traditional assumption that societal $math = male$ stereotypes are harmful for girls whilst neutral for boys.

It is important to consider the causal mechanisms that might underlie the relationship between children's in-group stereotypes and their mathematical outcomes. One possibility is that high mathematics-anxious children struggle with mathematics and avoid mathematics-related activities, and as a result they may dissociate mathematics from their own gender. Conversely, low mathematics-anxious children who perform well in mathematics may be more motivated to engage with mathematics-related activities, thus strengthening the association between mathematics and their in-group gender. A difficulty with this account is that implicit stereotypes have been shown to precede explicit beliefs (Galdi et al., 2014). Nevertheless, there is some evidence to suggest that these implicit stereotypes may be malleable (Dasgupta & Greenwald, 2001; Galdi et al., 2014).

A second possibility is that stronger in-group biases make children more motivated to enjoy (and succeed in) mathematics, increasing their engagement with mathematics-related activities and fostering more positive attitudes towards mathematics. On the other hand, weaker in-group biases may make children less inclined to enjoy (and less driven to achieve in) mathematics, reducing their engagement with mathematics-related activities and their subsequent mathematics affect. A third possibility is that the relationship between children's math–gender stereotypes and their mathematics anxiety and mathematics achievement is reciprocal. In-group biases may increase children's motivation, achievement and attitudes towards mathematics and,

in turn, higher motivation, achievement, and more positive attitudes may serve to reinforce these implicit biases.

In order to distinguish between these possibilities, there is a need for longitudinal research tracking the development of children's motivation, achievement and anxiety towards mathematics together with their math–gender stereotypes. Training studies may also help us to untangle the developmental trajectory of these factors. Some initial research with adults has shown that women who are re-trained to like mathematics exert more effort on a subsequent mathematics task. Furthermore, women who are re-trained to associate mathematics with their own gender show greater working memory capacity and increased mathematics performance under stereotype threat conditions (Forbes & Schmader, 2010). These findings suggest that attitudes may play a casual role in later mathematical outcomes for women who experience stereotype threat. Interestingly, the results from this study were not replicated in men, therefore the casual links between these factors may vary according to gender. Interventions designed to shape implicit math–gender associations in children may further our understanding of developmental (and gender) issues surrounding causality.

10.4.1   Limitations of Study 8

It is important to acknowledge that the data from this study is cross-sectional, thus the relationships between mathematics anxiety, mathematics achievement and in-groups stereotypes should not be interpreted as causal. Also, note that the percentage of the variance explained in the models is quite low. Mathematics anxiety, gender and the gender by anxiety interaction account for only 30% of the variance in mathematics reasoning performance and 29% of the variance in implicit in-group stereotypes. This means that there are other factors not included in the model.

Other limitations of this study overlap with those associated with the earlier mathematics anxiety studies presented. For example, as stated in Chapter 8 (Section 8.4.3), children's self-reported mathematics anxiety on the CMAQ may be confounded by their mathematical skills. Half of the items on the CMAQ involve specific calculations (e.g. "How do you feel when you have to solve $34 - 17$?"). These calculation-related items may not only be measuring children's mathematics anxiety, but their perceived self-efficacy and their mathematical skills. We know that mathematics anx-

iety is inextricably linked to a range of cognitive and motivational variables
(e.g. Bandura, 1993). Future studies should try to tease apart these con-
structs.

10.4.2 Summary of Findings

This study sought to investigate whether the gender differences in the cor-
relates of children's mathematics anxiety observed in the previous studies
(reported in Chapters 7 and 8) can be explained by gender differences in the
association between mathematics anxiety and math–gender stereotypes. It
was hypothesised that boys might get their mathematics anxiety from judge-
ments of their own mathematics performance (which are strongly related to
their actual mathematics performance), whilst girls get their mathemat-
ics anxiety from societal stereotypes (which are less related to their actual
mathematics performance). The results revealed little evidence for this hy-
pothesis. Specifically, there were no gender differences in the strength of the
relationship between mathematics anxiety and in-group stereotypes. Over-
all, the findings add to the literature on children's math–gender biases in
two ways. First, they demonstrate that implicit in-group stereotypes are re-
lated to affective factors as well as mathematics performance. Second, they
challenge the widespread assumption that societal $math = male$ stereotypes
are associated with mathematical outcomes in girls but not boys.

Chapter 11

Discussion of Mathematics Anxiety Studies

The aim of this chapter is to review all of the empirical findings from the mathematics anxiety studies reported in Chapters 7 – 10. First, I present an overview of the findings. Next, I discuss the theoretical conclusions that can be drawn from these findings, together with the educational implications, and directions for future research. Finally, I provide a summary to conclude Part III of this thesis.

11.1 Overview of Findings

The research presented in Chapters 7 – 10 (Part III) of this thesis sought to further our understanding of the causes and consequences of mathematics anxiety in the primary school years. Thus far, mathematics anxiety has been widely researched in adolescent and adult populations but little is known about how it develops in young children. Data from five studies were presented with a particular focus on gender issues. The key findings from each study are summarised in turn below.

To begin, Study 6 (reported in Chapter 7), investigated the role of parents in the development of children's anxiety towards mathematics. The results showed that mathematics anxiety was present in children as young as 6 years. This mathematics anxiety did not differ between boys and girls; however, gender had an impact on parental influence. Boys' (but not girls') mathematics anxiety was significantly associated with their parents' math-

ematics anxiety.

Following on from this result, the studies presented in Chapter 8 examined possible gender differences in the relationship between children's mathematics anxiety and their mathematics performance. The findings revealed that mathematics anxiety was more negatively associated with mathematical outcomes in boys than it was in girls. This was the case for low-level symbolic processing skills (Study 4) as well as higher-level mathematics reasoning performance (the PRISM study). Once again there were no significant gender differences in children's overall levels of mathematics anxiety.

In view of these consistent gender differences in the correlates of children's mathematics anxiety, it was questioned whether the self-report scales used to assess mathematics anxiety vary in their psychometric properties for boys and girls. To investigate this possibility, Study 7 (reported in Chapter 9) used a nine-week test-retest design to examine possible gender differences in the temporal stability of children's mathematics anxiety scores. Internal reliability analyses were also conducted. The results demonstrated that boys' reports of mathematics anxiety were more reliable over time than girls' reports of mathematics anxiety. Nevertheless, the test-retest coefficient for girls was still high (above the 0.7 level typically deemed to be suitable for research purposes).

Finally, Study 8, reported in Chapter 10, tested the hypothesis that boys get their mathematics anxiety from judgements of their own mathematics performance, whilst girls get their mathematics anxiety from societal math–gender stereotypes. The results yielded little support for this hypothesis. Although the relationship between mathematics anxiety and mathematics performance was stronger for boys than it was for girls, there were no gender differences in the relationship between mathematics anxiety and implicit in-group stereotypes. For both girls and boys, stronger in-group biases were associated with lower mathematics anxiety.

11.2   Theoretical Conclusions

The findings add to the growing literature on mathematics anxiety in the primary school years. In particular, they further our theoretical understanding of gender differences in the correlates of children's mathematics anxiety. Overall, it is possible to draw three main conclusions. Firstly,

that mathematics anxiety is present in the initial stages of symbolic number acquisition as well as later mathematics development. Secondly, that mathematics anxiety is an important factor for explaining individual differences in mathematical outcomes, particularly in boys. Thirdly, that parental mathematics anxiety influences sons' mathematics anxiety more than daughters' mathematics anxiety.

The first of these conclusions aligns with recent findings showing mathematics anxiety can emerge in children in the early school years. The second and third conclusions are novel because previous studies have focused on gender differences in overall levels of mathematics anxiety rather than gender differences in the correlates of mathematics anxiety. Interestingly, the findings may help to explain why some studies have found a relationship between primary school children's mathematics anxiety and mathematics performance (Ramirez et al., 2013; Wu et al., 2012) and others haven't (Krinzinger et al., 2009; Thomas & Dowker, 2000).

In terms of explaining the observed gender differences, the present research helps to rule out two possible accounts. It can be concluded that the gender differences in the correlates of mathematics anxiety are not simply an artefact of gender differences in the reliability of the measures used to assess children's mathematics anxiety. Moreover, they are not explained by gender differences in the relationship between mathematics anxiety and implicit math–gender stereotypes. Further possible accounts are considered in Section 11.4 within the discussion of future research directions.

As well as contributing to the literature on mathematics anxiety, the findings extend our knowledge of children's math–gender stereotypes. In accordance with previous studies it can be concluded that implicit $math = male$ biases arise early, despite no evidence for any gender differences in mathematics achievement (Cvencek et al., 2011, 2014; Galdi et al., 2014). These biases are associated not only with children's mathematics performance, but with affective factors such as mathematics anxiety and mathematics confidence. The implications of these findings are discussed in the section that follows.

11.3 Educational Implications

The findings demonstrate that mathematics anxiety is associated with a range of mathematical outcomes from basic (low-level) numerical processing to higher-level mathematics reasoning. This association appears to be stronger in boys than it is in girls. As such, contrary to what media reports suggest (e.g. Oullette, 2010), there is a need to help boys overcome mathematics anxiety, not just girls.

Importantly, research into the antecedents and causal mechanisms of mathematics anxiety may help to inform educators of ways to prevent its occurrence and reduce its negative effects. As outlined in the literature review, Section 1.2.3.3, studies with older children and adults have suggested that mathematics anxiety results in worrisome thoughts and ruminations which load working memory and disrupt information processing during mathematical tasks (Ashcraft & Kirk, 2001; Ashcraft & Krause, 2007; Faust, 1996; Hopko et al., 1998). Based on these studies, researchers have started to develop interventions which encourage high mathematics-anxious individuals to offload their worries (e.g. through expressive writing) prior to performing a mathematics-related activity. The idea here is that this helps to free up working memory resources, thus reducing the detrimental effects of mathematics anxiety. So far, these interventions have shown promising results in secondary school-aged students (Park, Ramirez, & Beilock, 2014; Ramirez & Beilock, 2011). Further research is needed to explore possible interventions for younger children.

Given the findings that mathematics anxiety is associated with low-level numerical processing as well as higher-level mathematics, high mathematics-anxious individuals may also benefit from interventions that target basic symbolic number skills. These basic skills include number symbol identification, magnitude comparison and symbol-to-quantity matching. Further to this, interventions that target early attitudes towards mathematics, and in particular math–gender stereotypes, may also help to remediate mathematics anxiety. The findings from the present research suggest that stereotypes are acquired young and are related to affective factors as well as mathematics performance. Importantly, longitudinal research and training studies will help to establish the causal nature of these relationships and test these possible interventions.

11.4   Future Research

Following on from this research, there are some specific questions for future studies to address. These are detailed in the relevant sections below. Note that broader research directions are reviewed in the general discussion of this thesis (Part IV, Chapter 12).

11.4.1   Theoretical Issues

As discussed previously, the findings from this research revealed some significant gender differences in the correlates of children's mathematics anxiety. These gender differences were consistent across studies and were not accounted for by gender differences in (i) the reliability of the measures used to assess mathematics anxiety, or (ii) the relationship between mathematics anxiety and implicit math–gender stereotypes. It is thus important for future studies to continue to investigate their causal roots.

One possibility to consider is that the gender differences observed may be better explained by variations in personality. This is theoretically plausible because recent findings suggest that the effects of gender and personality are confounded in explaining mathematical outcomes in undergraduate students (Alcock, Attridge, Kenny, & Inglis, 2014). Alcock et al. (2014) found that gender was a significant predictor of students' achievement and behaviour in mathematics when entered into a regression model on its own; however, it was no longer significant when personality profiles were taken into account. Interestingly, the same could not be said the other way round. Gender did not explain variations in achievement and behaviour over and above that accounted for by personality.

A simple way of testing this personality hypothesis would be to give primary school children measures of mathematics anxiety and mathematics reasoning, together with a Big 5 personality questionnaire. If the gender differences observed in the present research are better explained by individual differences in personality, then we would expect the interaction between mathematics anxiety and personality to predict mathematics performance over and above the interaction between mathematics anxiety and gender. Conversely, if the gender differences are not better explained by personality, then we would expect the mathematics anxiety by gender interaction to remain significant even after accounting for variations in personality.

Another factor that may be important to consider is independent control or uncontrollability. According to behavioural and neuroscientific research, anxiety and controllability are both important determinants of task performance (Amat et al., 2005; Zirk-Sadowski, Lamptey, Devine, Haggard, & Szũcs, 2014). In a recent study, Zirk-Sadowski et al. (2014) demonstrated that gender differences in overall levels of mathematics anxiety were mediated by variations in children's perception of their control in mathematics. It could therefore be the case that the gender differences in the strength of the relationship between mathematics anxiety and mathematics performance are also explained by gender differences in independent control. Arguably, perceptions of control can be seen as a personality construct (Rotter, 1975), thus this explanation ties in with the first hypothesis.

In addition to testing these possibilities, it would be valuable for future studies to address the key question of causality. It is evident that there are relationships between mathematics anxiety, mathematics performance and implicit math–gender stereotypes. But are these relationships causal ones? For example, will a decrease in math–gender stereotypes lead to a reduction in mathematics anxiety and an improvement in mathematics performance? Crucially, to test causality, there is a need for longitudinal research and training studies. This will help to further not only our theoretical knowledge but our practical tools for supporting children's mathematical development.

11.4.2   Methodological Issues

There are also methodological questions for future studies to investigate. The scales used to assess children's mathematics anxiety have demonstrated good reliability; however, their validity is limited by the subjective nature of self-reports. Children's responses may be influenced by a number of extraneous factors such as the way in which the questionnaire is administered (e.g. individual vs. group, teacher vs. researcher). Indeed, it is interesting to note that in Study 6, where the CMAQ was administered by the researcher in individual testing sessions, children reported higher levels of mathematics anxiety ($M = 2.45$, $SD = 0.87$) than in Study 4, where the CMAQ was administered by teachers as a whole class activity ($M = 1.90$, $SD = 0.68$).[1]

---

[1]This difference which can be seen in in Figure 1.1 (page 21) was found to be significant, $t(138) = 3.93$, $p < .001$.

Whilst this difference may reflect actual variations between the samples, it may also stem from the different testing conditions. It is possible that children feel less able to report feelings of mathematics anxiety in the presence of their teacher and their peers than in the presence of an unknown researcher. Alternatively, the researcher or the unfamiliar research situation may increase children's anxiety, so they are more able to access a sense of themselves being anxious.

This raises the issue of how to obtain more objective measures of children's mathematics anxiety. One way may be to record physiological responses or implicit associations between mathematics and anxiety (Faust, 1992; Egloff & Schmukle, 2002). A second way may be to observe a child's emotional reactions to a mathematics task by interpreting facial, vocal and behavioral cues (Berhenke et al., 2011). A third way may be to gather real-time reports of (state) anxiety experienced during a mathematics task. There is evidence to suggest that measures of state anxiety are less influenced by subjective beliefs than measures of trait anxiety (Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013; Robinson & Clore, 2002).

In a recent study, Goetz et al. (2013) assessed children's trait anxiety, using a standard mathematics anxiety questionnaire, together with their state anxiety, using a single self-report question ("I am anxious") administered at multiple time points during a mathematics test. Interestingly, the results showed that girls experienced more trait anxiety than boys, but there were no gender differences in terms of state anxiety. This trait-state discrepancy was explained by lower levels of perceived competence or self-efficacy in girls. The findings therefore suggest that children's subjective beliefs impact their self-reported trait anxiety more than their self-reported state anxiety.

In view of these findings, it is possible that further studies employing real-time (state-based) measures of mathematics anxiety may help to explain the gender differences in the correlates of mathematics anxiety observed in the present research. More generally, state-based measures may provide a useful tool for investigating the causal mechanisms of children's mathematics anxiety.

11.5   Summary

In sum, the empirical work presented in Part III of this thesis examined the causes and consequences of mathematics anxiety in the primary school years. The findings from five studies were reported. These findings reveal some consistent gender differences in the correlates of children's mathematics anxiety. In particular, they show that boys' (but not girls') mathematics anxiety is related to their parents' mathematics anxiety. Moreover, boys' mathematics anxiety is more negatively associated with their mathematics performance than girls' mathematics anxiety. There is little evidence to suggest that these gender differences stem from psychometric issues or from gender differences in the relationship between mathematics anxiety and implicit math–gender stereotypes. Further research is needed to uncover their causal roots. Overall, it can be concluded that mathematics anxiety is an important factor for explaining variations in children's early numerical development as well as later mathematics success.

Part IV

General Discussion

Chapter 12

Conclusions

The aim of this final chapter is to bring together the findings from the SFON and mathematics anxiety studies. First, I review the thesis aims and background. Next, I consider the key findings and emerging questions. Finally, I revisit the problem statement that motivated this research.

12.1 Overview of Thesis Aims

Current research highlights a range of cognitive and non-cognitive factors that may affect children's early symbolic number development. Of these factors, little is known about the role of dispositional predictors such as attitudes, beliefs and motivations. This thesis sought to address this gap by investigating the causes and consequences of individual differences in children's SFON and mathematics anxiety.

The specific aims with regards to SFON were threefold. Firstly, to assess the psychometric properties of the tasks used to measure children's SFON. Secondly, to develop and validate a new task that measures SFON independently of children's counting accuracy. Thirdly, to investigate the developmental roots and possible mechanisms of SFON.

In terms of mathematics anxiety, there were two primary aims. Firstly, to examine the relationship between children's SFON and their parent's SFON. Secondly, to investigate the effects of mathematics anxiety on children's numerical processing skills. Further to these objectives, questions regarding gender issues were addressed.

12.2   Research Findings and Emerging Questions

Overall, the findings from both strands of research suggest that children's dispositions towards mathematics are related to their early numerical outcomes. High SFON children, who are more inclined to focus on the numerical aspects of their environment, show more advanced counting and arithmetic skills than their low SFON peers. On the other hand, children who are anxious towards mathematics show poorer performance on numerical processing and mathematics reasoning tasks. There was no relationship between children's SFON and mathematics anxiety. This indicates that SFON and mathematics anxiety are independent constructs.

Given the fact that SFON and mathematics anxiety were unrelated, the findings were presented in two separate parts and different issues were explored in relation to each construct (see Chapters 6 and 11 for a general discussion of SFON and mathematics anxiety, respectively). Despite these differences, there are some commonalities that can be drawn. In particular, both SFON and mathematics anxiety were associated with children's basic (low-level) numerical processing. As such, whilst they may not be directly related, they may both interact with domain-specific cognitive factors to influence children's mathematical skills. Further research needs to examine the interactions among the different cognitive and non-cognitive predictors of early mathematics.

As well as examining these interactions, it is important for future studies to explore issues surrounding causality. We know that SFON and mathematics anxiety are related to children's mathematics achievement but are these relationships causal? If we can establish the causal relationships (and interactions) between predictors then researchers can start to build a comprehensive model of early numeracy.

12.3   Statement of the Problem Revisited

Numerical skills are important both for the individual negotiating life's daily demands and for modern society as a whole. As emphasised by the UK's National Numeracy organisation[1]:

---

[1]A charity set up in 2012 to help improve the nation's numeracy levels.

> "Numeracy is a life skill. Being numerate goes beyond simply 'doing sums'; it means having the confidence and competence to use numbers and think mathematically in everyday life..." (National Numeracy, 2015).

Importantly, numeracy is not just a cognitive process. An individual's success in dealing with numbers depends not only on their numerical skills, but the dispositions (such as attitudes, beliefs and motivations) that they bring with them to the numerical task or situation. Thus, if we are to improve the nation's numeracy, then we need to understand not only the specific cognitive skills involved in learning and performing mathematics, but the non-cognitive 'dispositional' factors that play a role as well.

As highlighted in Chapter 1, recent years have seen widespread efforts to raise the numerical skills of young people worldwide. In particular, these efforts have focused on the early years (Munn, 2006; Williams, 2008). Educational policies such as the British Government's EYFS Framework have called for more formal mathematics in the preschool curriculum (Department for Education, 2012). This follows recent findings that gaps in numeracy start to emerge before children start school (Ginsberg et al., 2008), and that early interventions are more effective and economically efficient (Heckman & Masterov, 2007).

Given this shift in attention to the early years, research into the factors affecting young children's numerical development has come to the forefront. Here, the current thesis highlights the role of non-cognitive 'dispositional' factors. Two dispositional factors – SFON and mathematics anxiety – were found to be independently related to a range of numerical outcomes in children aged 3 to 9 years. Overall, these findings suggest that we need to foster children's emerging attitudes and motivations towards mathematics, as well as intervening at the cognitive level.

Broadly speaking, this interplay between cognitive and dispositional factors is central not only to our understanding of children's early numeracy, but to other areas of mathematics and to learning and development more generally. If we are to succeed in a particular discipline then we need to acquire subject relevant knowledge and skills, together with a productive disposition to use those skills.

Part V

References & Appendices

References

Aarnos, E., & Perkkilä, P. (2012). Early signs of mathematics anxiety? *Procedia-Social and Behavioral Sciences*, *46*, 1495-1499.

Adams, E. B., & Sarason, I. G. (1963). Relation between anxiety in children and their parents. *Child Development*, *34*, 237-246.

Alcock, L., Attridge, N., Kenny, S., & Inglis, M. (2014). Achievement and behaviour in undergraduate mathematics: Personality is a better predictor than gender. *Research in Mathematics Education*, *16*, 1-17.

Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, *21*, 133-137.

Amat, J., Baratta, M. V., Paul, E., Bland, S. T., Watkins, L. R., & Maier, S. F. (2005). Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nature Neuroscience*, *8*, 365-371.

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, *12*, 385-390.

Anders, Y., Rossbach, H. G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & von Maurice, J. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly*, *27*, 231-244.

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, *130*, 224-237.

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, *14*, 243-248.

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology*, *96*, 699-713.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*, 117-148.

Baroody, A. J. (2000). Does mathematics instruction for three- to five-year-olds really make sense? *Young Children*, *55*, 61-67.

Becker, G. (2000). How important is transient error in estimating reliability? Going beyond simulation studies. *Psychological Methods*, *5*, 370-379.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*, 1860-1863.

Berhenke, A., Miller, A. L., Brown, E., Seifer, R., & Dickstein, S. (2011). Observed emotional and behavioral indicators of motivation predict school readiness in Head Start graduates. *Early Childhood Research Quarterly*, *26*, 430-441.

Betz, N. E. (1978). Prevalence, distribution, and correlates of math anxiety in college students. *Journal of Counseling Psychology*, *25*, 441-448.

Blevins-Knabe, B., Austin, A. B., Musun, L., Eddy, A., & Jones, R. M. (2000). Family home care providers' and parents' beliefs and practices concerning mathematics with young children. *Early Child Development and Care*, *165*, 41-58.

Blevins-Knabe, B., & Musun-Miller, L. (1991). *Parental beliefs about the development of preschool children's number skills.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Seattle.

Blevins-Knabe, B., & Musun-Miller, L. (1996). Number use at home by children and their parents and its relationship to early mathematical performance. *Early Development & Parenting*, *5*, 35-45.

Bojorque, G., Torbeyns, J., Hannula-Sormunen, M. M., Van Nijlen, D., & Verschaffel, L. (2014). *The development of spontaneous focusing on numerosity and early mathematical skills in Ecuadorian Kindergartners: The role of school type and quality of early math education.* Paper presented at the Biennial EARLI Conference of SIG 5 – Learning and Development in Early Childhood, Jyväskylä.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations

influence arithmetic learning. *Child Development*, *79*, 1016-1031.

Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. *American Educational Research Journal*, *5*, 437-474.

Brankaer, C., Ghesquière, P., & De Smedt, B. (2014). Children's mapping between non-symbolic and symbolic numerical magnitudes and its association with timed and untimed tests of mathematics achievement. *PloS one*, *9*, e93565.

Briars, D., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology*, *20*, 607-618.

Bryman, A. (2008). *Social research methods* (3rd ed.). Oxford University Press.

Buckingham, B. R., & MacLatchy, J. (1930). *The number abilities of children when they enter grade one. Report of the Society's Committee on Arithmetic, 29th Yearbook of the National Society for the Study of Education.* Bloomington, Illinois: Public School Publishing.

Buckley, S., & Reid, K. (2013). *Learning and fearing mathematics: Insights from psychology and neuroscience.* Paper presented at the Australian Council for Educational Research Conference, Melbourne. Retrieved from `http://research.acer.edu.au/cgi/viewcontent.cgi?article=1182&context=research_conference`

Bull, R. (2013). *Examining sources of individual differences in acuity of the Approximate Number System.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Seattle.

Bull, R., Espy, K. A., & Wiebe, S. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*, 205-228.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*, 273-293.

Bussey, K., & Bandura, A. (1984). Influence of gender constancy and social power on sex-linked modeling. *Journal of Personality and Social Psychology*, *47*, 1292-1302.

Bynner, J., & Parsons, S. (2006). *New light on literacy and numeracy.* National Research and Development Centre for Adult Literacy and Numeracy.

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology*, *37*, 231-242.

Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus*, 59-68.

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS One*, *7*, e33832.

Chang, A., Sandhofer, C. M., & Brown, C. S. (2011). Gender biases in early number exposure to preschool-aged children. *Journal of Language and Social Psychology*, *30*, 440-450.

Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of Cognition and Development*, *15*, 2-11.

Clark, C. A. C., Pritchard, V. E., & Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology*, *46*, 1176-1191.

Clearfield, M., & Mix, K. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*, 408-411.

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks Project. *Journal for Research in Mathematics Education*, *38*, 136-163.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Cohen, R. J., & Swerdlik, M. E. (1999). *Psychological testing and assessment: An introduction to tests and measurement*. Mountain View, CA: Mayfield Publishing.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 96-104.

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math–gender stereotypes in elementary school children. *Child Development*, *82*, 766-779.

Cvencek, D., Meltzoff, A. N., & Kapur, M. (2014). Cognitive consistency and math–gender stereotypes in Singaporean children. *Journal of Experimental Child Psychology*, *117*, 73-91.

D'Amico, A., & Guarnela, M. (2005). Exploring working memory in chil-

dren with low arithmetical achievement. *Learning and Individual Differences*, *15*, 189-202.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800-814.

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, *28*, 1615-1628.

Davis, B. S. (2009). *Preschool children's acquisition of mathematical competencies.* Unpublished doctoral dissertation, Northern Michigan University.

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

Dehaene, S. (2001). Précis of the number sense. *Mind and Language*, *16*, 16-36.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the human and animal brain. *Trends in Neurosciences*, *21*, 355-361.

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital: Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 626-641.

Dehaene, S., Izard, V., & Piazza, M. (2005). *Control over non-numerical parameters in numerosity experiments.* Retrieved from `www.unicog` `.org`

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

Department for Education. (2012). *Statutory Framework for the Early Years Foundation Stage.* Runcorn: Department for Education.

De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*, 48-55.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*, 469-479.

Devine, A., Fawcett, K., Szucs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and Brain Functions*, *8*, 2-9.

Dew, K. H., & Galassi, J. P. (1983). Mathematics anxiety: Some basic issues. *Journal of Counseling Psychology*, *30*, 443-446.

Douglass, H. R. (1925). The development of number concept in children of preschool and kindergarten ages. *Journal of Experimental Psychology*, *8*, 443-470.

Dreger, R. M., & Aiken Jr, L. R. (1957). The identification of number anxiety in a college population. *Journal of Educational Psychology*, *48*, 344-351.

Duncan, G. J., Claessens, A., Huston, A. C., Pagani, L. S., Engel, M., Sexton, H., . . . Duckworth, K. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428-1446.

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues*, *46*, 183-201.

Eden, C., Heine, A., & Jacobs, A. M. (2013). Mathematics anxiety and its development in the course of formal schooling–a review. *Psychology*, *4*, 27-35.

Edens, K. M., & Potter, E. F. (2013). An exploratory look at the relationships among math skills, motivational factors and activity choice. *Early Childhood Education Journal*, *41*, 235-243.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology*, *83*, 1441-1455.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*, 103-127.

Epsy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions

to emergent mathematical skills in preschool children. *Developmental Neuropsychology*, *26*, 465-486.

Esping-Andersen, G. (2004). Untying the Gordian knot of social inheritance. *Research in Social Stratification and Mobility*, *21*, 115-138.

Faust, M. W. (1992). *Analysis of physiological reactivity in mathematics anxiety*. Unpublished doctoral dissertation, Bowling Green State University; Ohio.

Faust, M. W. (1996). Mathematics anxiety effects in simple and complex addition. *Mathematical Cognition*, *2*, 25-62.

Feigenson, L. (2005). A double-dissociation in infants' representations of object arrays. *Cognition*, *95*, B37-B48.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science*, *13*, 150-156.

Feigenson, L., Carey, S., & Spelke, E. S. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*, *44*, 33-66.

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307-314.

Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, *11*, 93-103.

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

Forbes, C. E., & Schmader, T. (2010). Retraining attitudes and stereotypes to affect motivation and cognitive capacity under stereotype threat. *Journal of Personality and Social Psychology*, *99*, 740-754.

Freeman, F. N. (1912). Grouped objects as a concrete basis for the number idea. *Elementary School Teacher*, *8*, 30014.

Friedman, L. (1989). Mathematics and the gender gap: A met-analysis of recent studies on sex differences in mathematical tasks. *Review of Educational Research*, *59*, 185-213.

Frye, D., Braisby, N., Lowe, J., Maroudas, C., & Nicholls, J. (1989). Young children's understanding of counting and cardinality. *Child Development*, *60*, 1158-1171.

Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer-Verlag.

Fuson, K. C., & Hall, J. W. (1983). The acquisition of early number word meanings: A conceptual analysis and review. In H. P. Ginsburg (Ed.),

The development of mathematical thinking (p. 49-107). New York: Academic Press.

Galdi, S., Cadinu, M., & Tomasetto, C. (2014). The roots of stereotype threat: When automatic associations disrupt girls' math performance. *Child Development*, *85*, 250-263.

Gallistel, C. R. (2007). Commentary on Le Corre & Carey. *Cognition*, *105*, 439-445.

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

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, *4*, 59-65.

Galloway, C., & Richards, B. (1994). *Input and interaction in language acquisition*. CUP, Cambridge.

Gawronski, B., & Payne, B. K. (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. New York: Guilford Press.

Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology*, *77*, 236-263.

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

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, *78*, 1343-1359.

Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods*, *43*, 981-986.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gelman, R., & Meck, E. (1983). Preschoolers' counting: Principles before skill. *Cognition*, *13*, 343-359.

Gelman, R., Meck, E., & Merkin, S. (1986). Young children's numerical competence. *Cognitive Development*, *1*, 1-29.

Gilmore, C. K., Attridge, N., De Smedt, B., & Inglis, M. (2014). Measuring

the approximate number system in children: Exploring the relationships among different tasks. *Learning and Individual Differences*, *29*, 50-58.

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

Ginsberg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report - Giving Child and Youth Development Knowledge Away*, *22*, 1-24.

Ginsburg, H., & Pappas, S. (2004). SES, ethnic, and gender differences in young children's informal addition and subtraction: A clinical interview investigation. *Journal of Applied Developmental Psychology*, *25*, 171-192.

Godbey, C. (1997). *Mathematics anxiety and the underprepared student* (Tech. Rep. No. JC 990 063). TN: Middle Tennessee State University.

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science*, *24*, 2079-2087.

Gray, S., & Reeve, R. (2013). *Core number abilities and preschool children's numerical cognition.* Poster presented at the Biennial Meeting of the Society for Research in Child Development, Seattle.

Greeno, J. G., Riley, M. S., & Gelman, R. (1984). Conceptual competence and children's counting. *Cognitive Psychology*, *16*, 94-143.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197-216.

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*, 153-166.

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*, 11116-11120.

Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with math achievement. *Nature*, *455*, 665-668.

Hannula, M. (2005). *Spontaneous focusing on numerosity in the development of early mathematical skills.* Annales Universitatis Turkuensis B, 282. Turku: Painosalama.

Hannula, M., Grabner, R. H., Lehtinen, E., Laine, T., Parkkola, R., & Ansari, D. (2009). *Neural correlates of spontaneous focusing on numerosity (SFON).* Poster presented at the Annual Meeting of the Organization for Human Brain Mapping, San Francisco.

Hannula, M., & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction*, *15*, 237–256.

Hannula, M., Lepola, J., & Lehtinen, E. (2010). Spontaneous focusing on numerosity as a domain-specific predictor of arithmetical skills. *Journal of Experimental Child Psychology*, *107*, 394–406.

Hannula, M., Mattinen, A., & Lehtinen, E. (2005). Does social interaction influence 3-year old children's tendency to focus on numerosity? A quasi-experimental study in day care. In L. Verschaffel, E. De Corte, G. Kanselaar, & M. Valcke (Eds.), *Powerful environemnts for promoting deep conceptual and strategic learning* (p. 63-80). Leuven University Press.

Hannula, M., Räsänen, P., & Lehtinen, E. (2007). Development of counting skills: Role of spontaneous focusing on numerosity and subitizing-based enumeration. *Mathematical Thinking and Learning*, *9*, 51–57.

Hannula-Sormunen, M. M. (2014). Spontaneous focusing on numerosity and its relation to counting and arithmetic. In R. Cohen-Kadosh & A. Dowker (Eds.), *Oxford handbook of mathematical cognition.* doi: 10.1093/oxfordhb/9780199642342.013.018.

Hart, S. A., Petrill, S. A., & Thompson, L. A. (2010). A factorial analysis of timed and untimed measures of mathematics and reading abilities in school aged twins. *Learning and Individual Differences*, *20*, 63-69.

Heckman, J. (2004). Invest in the very young. In R. E. Tremblay, R. G. Barr, & R. D. V. Peters (Eds.), *Encyclopedia on early childhood development* (p. 1-2). Montreal, Quebec: Centre of Excellence for Early Childhood Development.

Heckman, J., & Masterov, D. V. (2007). The productivity argument for investing in young children. *Applied Economic Perspectives and Policy*, *29*, 446-493.

Heckman, J., & Wax, A. (2004). Home alone. *Wall Street Journal*, *23*, A14.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, *21*, 33-46.

Higher Education Statistics Agency. (2011). *Statistics–Students and qualifiers at UK HE institutions.* Retrieved from `https://www.hesa.ac.uk/stats`

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, *103*, 17-29.

Hopko, D. R., Ashcraft, M. H., Gute, J., Ruggiero, K. J., & Lewis, C. (1998). Mathematics anxiety and working memory: Support for the existence of a deficient inhibition mechanism. *Journal of Anxiety Disorders*, *12*, 343-355.

Hughes, C., & Ensor, R. (2005). Theory of mind and executive function: A family affair. *Developmental Neuropsychology*, *28*, 645-668.

Hunt, T. E., Clark-Carter, D., & Sheffied, D. (2011). The development and part validation of a UK scale for mathematics anxiety. *Journal of Psychoeducational Assessment*, *29*, 455-466.

Huntsinger, C. S., Jose, P. E., & Larson, S. L. (1998). Do parent practices to encourage academic competence influence the social adjustment of young European American and Chinese American children? *Developmental Psychology*, *34*, 747-756.

Huntsinger, C. S., Jose, P. E., Larson, S. L., Balsink Krieg, D., & Shaligram, C. (2000). Mathematics, vocabulary, and reading development in Chinese American and European American children over the primary school years. *Journal of Educational Psychology*, *92*, 745-760.

Hur, J. H. (2010). *Math talk between preschoolers and their mothers during a home cooking activity.* Unpublished undergraduate thesis, Purdue University.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236-248.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*,

107, 139-155.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*, 494-495.

Iuculano, T., Tang, J., Hall, C. W., & Butterworth, B. (2008). Core information processing deficits in developmental dyscalculia and low numeracy. *Developmental Science*, *11*, 669-680.

Jacobs, J. E. (1991). Influence of gender stereotypes on parent and child mathematics attitudes. *Journal of Educational Psychology*, *83*, 518-527.

Jacobs, J. E., & Bleeker, M. M. (2004). Girls' and boys' developing interests in math and science: Do parents matter? *New Directions for Child and Adolescent Development*, *106*, 5-21.

Jain, S., & Dowson, M. (2009). Mathematics anxiety as a function of multidimensional self-regulation and self-efficacy. *Contemporary Educational Psychology*, *34*, 240-249.

Jameson, M. M. (2013). The development and validation of the Children's Anxiety in Math Scale. *Journal of Psychoeducational Assessment*, *31*, 391-395.

Jayaratne, T. E. (1987). *The impact of mothers' math experiences on their daughters' attitudes toward math.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Baltimore.

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes, gender identification, and math-related outcomes a prospective study of female college students. *Psychological Science*, *18*, 13-18.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, *92*, 109-129.

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk". *Developmental Psychology*, *42*, 59-69.

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

Krinzinger, H., Kaufmann, L., Dowker, A., Thomas, G., Graf, M., Nuerk, H. C., & Willmes, K. (2007). [German version of the math anxiety questionnaire (FRA) for 6-to 9-year-old children]. *Zeitschrift fur Kinder-und Jugendpsychiatrie und Psychotherapie*, *35*, 341-351.

Krinzinger, H., Kaufmann, L., & Willmes, L. (2009). Math anxiety and math ability in early primary school years. *Journal of Psychoeducational Assessment*, *27*, 206-225.

Kroesbergen, E. H., Van Luit, J. E. H., Van Lieshout, E. C. D. M., Van Loosbroek, E., & Van de Rijt, B. A. M. (2009). Individual differences in early numeracy: The role of executive functions and subitizing. *Journal of Psychoeducational Assessment*, *27*, 226-236.

Kucian, K., Kohn, J., Hannula-Sormunen, M. M., Richtmann, V., Grond, U., Käser, T., ... von Aster, M. (2012). Kinder mit Dyskalkulie fokussieren spontan weniger auf Anzahligkeit. *Lernen und Lernstörungen*, *1*, 241-253.

Kyttala, M., Aunio, P., Lehto, J. E., Van Luit, J., & Hautamaki, J. (2003). Visuospatial working memory and early numeracy. *Educational and Child Psychology*, *20*, 65-76.

Lane, K. A., Goh, J. X., & Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles*, *66*, 220-234.

Lang, P. J. (1968). Fear reduction and fear behavior: Problems in treating a construct. In J. M. Schlien (Ed.), *Research in psychotherapy. III* (p. 90-102). American Psychological Association; Washington, DC.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*, 395-438.

Le Corre, M., & Carey, S. (2008). Why the verbal counting principles are constructed out of representations of small sets of individuals: A reply to Gallistel. *Cognition*, *107*, 650-662.

Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Revisiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, *52*, 130-169.

LeFevre, J.-A. (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., Clarke, T., & Stringer, A. P. (2002). Influences of language and parental involvement on the development of counting skills: Comparisons of French-and English-speaking Canadian children. *Early Child Development and Care*, *172*, 283-300.

LeFevre, J.-A., Polyzoi, E., Skwarchuk, S. L., Fast, L., & Sowinski, C. (2010). Do home numeracy and literacy practices of Greek and Canadian parents predict the numeracy skills of kindergarten children? *International Journal of Early Years Education*, *18*, 55-70.

LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science*, *41*, 55-66.

Levine, S. C., Suriyakham, L. W., Rowe, M. L., Huttenlocher, J., & Gunderson, E. A. (2010). What counts in the development of young children's number knowledge? *Developmental Psychology*, *46*, 1309-1319.

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, *14*, 1292-1300.

Libertus, M. E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica*, *141*, 373-379.

Linder, S. M., Powers-Costello, B., & Stegelin, D. A. (2011). Mathematics in early childhood: Research-based rationale and practical strategies. *Early Childhood Education Journal*, *39*, 29-37.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large number discrimination in human infants. *Psychological Science*, *14*, 396-401.

Lipton, J. S., & Spelke, E. S. (2004). Discrimination of large and small numerosities by human infants. *Infancy*, *5*, 271-290.

Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development*, *76*, 978-988.

LoBello, S. G. (1991). A short form of the Wechsler Preschool and Primary Scale of Intelligence–Revised. *Journal of School Psychology*, *29*, 229-236.

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences*, *109*, 18737-18742.

Ma, X. (1999). A meta-analysis of the relationship between anxiety towards mathematics and achievement in mathematics. *Journal for Research*

in Mathematics Education, *30*, 520-540.

Maloney, E. A., Ansari, D., & Fugelsang, J. A. (2011). The effect of mathematics anxiety on the processing of numerical magnitude. *The Quarterly Journal of Experimental Psychology*, *64*, 10-16.

Maloney, E. A., & Beilock, S. L. (2012). Math anxiety: Who has it, why it develops, and how to guard against it. *Trends in Cognitive Sciences*, *16*, 404-406.

Maloney, E. A., Risko, E. F., Ansari, D., & Fugelsang, J. A. (2010). Mathematics anxiety affects counting but not subitizing during visual enumeration. *Cognition*, *114*, 293-297.

Marsh, M. (2011). *Numeracy counts: NIACE committee of inquiry on adult numeracy learning*. National Institute of Adult Continuing Education.

Mazzocco, M. M., Hanich, L. B., & Noeder, M. M. (2012). Primary school age students' spontaneous comments about math reveal emerging dispositions linked to later mathematics achievement. *Child Development Research*, *2012*, doi:10.1155/2012/170310.

McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, *15*, 776-781.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology*, *74*, 240-260.

Melhuish, E., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues*, *64*, 95-114.

Melhuish, E., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M., & Malin, A. (2008b). Preschool influences on mathematics achievement. *Science*, *321*, 1161-1162.

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self-and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, *81*, 247-258.

Miller, H., & Bichsel, J. (2004). Anxiety, working memory, gender, and math performance. *Personality and Individual Differences*, *37*, 591-606.

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1997). Numerical abstraction in infants: Another look. *Developmental Psychology*, *33*, 423-428.

Monette, S., Bigras, M., & Guay, M. C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology*, *109*, 158-173.

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

Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, *103*, 490-502.

Munn, P. (2006). Mathematics in early childhood – the early years maths curriculum in the UK and children's numerical development. *International Journal of Early Childhood*, *29*, 99-111.

Myers, L., & Sirois, M. J. (2006). Spearman correlation coefficients, differences between. In S. Kotz (Ed.), *Encyclopedia of statistical sciences* (p. 7901-7903). New York: John Wiley and Sons.

National Numeracy. (2014). *PISA 2012 UK country reports.* Retrieved from `http://www.nationalnumeracy.org.uk/resources/132/index.html`

National Numeracy. (2015). *What is numeracy?* Retrieved from `http://www.nationalnumeracy.org.uk/what-is-numeracy/index.html`

National Science Foundation. (2013). *Women, minorities, and persons with disabilities in science and engineering.* Retrieved from `http://www.nsf.gov/statistics/`

Newcombe, N. S. (2002). The nativist-empiricist controversy in the context of recent research on spatial and quantitative development. *Psychological Science*, *13*, 395-401.

Norris, E. (2012). *Solving the maths problem: International perspectives on mathematics education.* The Royal Society for the Encouragement of Arts, Manufactures and Commerce.

Norwood, K. S. (1994). The effect of instructional approach on mathematics anxiety and achievement. *School Science and Mathematics*, *94*, 248-254.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44-59.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., ... Greenwald, A. G. (2009). National differences in gender–

science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*, 10593-10597.

Nùñez-Peña, M. I., & Suàrez-Pellicioni, M. (2014). Less precise representation of numerical magnitude in high math-anxious individuals: An ERP study of the size and distance effects. *Biological Psychology*, *103*, 176-183.

Oullette, J. (2010, September). *Guest Blog: Helping girls overcome math anxiety.* Retrieved from `http://www.rachelsimmons.com/2010/09/guest-blog-helping-girls-overcome-math-anxiety/`

Park, D., Ramirez, G., & Beilock, S. L. (2014). The role of expressive writing in math anxiety. *Journal of Experimental Psychology: Applied*, *20*, 103-111.

Parsons, J. E., Adler, T. F., & Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: Parental influences. *Child Development*, *53*, 310-321.

Passolunghi, M. C., & Lanfranchi, S. (2012). Domain-specific and domain-general precursors of mathematical achievement: A longitudinal study from kindergarten to first grade. *British Journal of Educational Psychology*, *82*, 42-63.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, *80*, 44-57.

Perry, D.-G., & Bussey, K. (1979). The social learning theory of sex differences: Imitation is alive and well. *Journal of Personality and Social Psychology*, *37*, 1699-1712.

Piaget, J. (1952). *The child's conception of number.* Routledge & Kegan Paul Ltd.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*, 499-503.

Poltz, N., Wyschkon, A., Hannula-Sormunen, M. M., von Aster, M., & Esser, G. (2014). *Spontaneous focusing on numerosity and the development of mathematical skills: Stability and prediction in kindergarten.* Paper presented at the Biennial EARLI Conference of SIG 5 – Learning and Development in Early Childhood, Jyväskylä.

Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica*, *140*, 50-57.

Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The transition from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology*, *105*, 453-464.

Raban, B., & Ure, C. (2000). Early literacy – A government concern? *Early Years: An International Journal of Research and Development*, *20*, 47-56.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*, 110-122.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable inprovements in low-income children's numerical knowledge through playing number board games. *Child Development*, *79*, 375-394.

Ramani, G. B., & Siegler, R. S. (2011). Reducing the gap in numerical knowledge between low- and middle-income preschoolers. *Journal of Applied Developmental Psychology*, *32*, 146-159.

Ramirez, G., & Beilock, S. L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science*, *331*, 211-213.

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory and math achievement in early elementary school. *Journal of Cognition and Development*, *14*, 187-202.

Reeve, R., Reynolds, F., Humberstone, J., & Butterworth, B. (2012). Stability and change in markers of core numerical competencies. *Journal of Experimental Psychology: General*, *141*, 649-666.

Reynolds, A. J., & Temple, J. A. (1998). Extended early childhood intervention and school achievement: Age 13 findings from the Chicago Longitudinal Study. *Child Development*, *69*, 231-246.

Richardson, F. C., & Suinn, R. M. (1972). The Mathematics Anxiety Rating Scale. *Journal of Counseling Psychology*, *19*, 551554.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*, 934-960.

Rotter, J. B. (1975). Some problems and misconceptions related to the construct of internal versus external control of reinforcement. *Journal of Consulting and Clinical Psychology*, *43*, 56-67.

Rubinsten, O., & Tannock, R. (2010). Mathematics anxiety in children with developmental dyscalculia. *Behavioral and Brain Functions*, *6*, 1-13.

Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B., Toth, K., ... Smees, R. (2012). *Influences on students' attainment and progress in Key Stage 3: Academic outcomes in english, maths and science in Year 9*. Department for Education.

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. Routledge.

Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number–space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology*, *114*, 418-431.

Sasanguie, D., Van den Bussche, E., & Reynvoet, B. (2012). Predictors for mathematics achievement? Evidence from a longitudinal study. *Mind, Brain, and Education*, *6*, 119-128.

Saxe, G. B., Guberman, S. R., & Gearhart, M. (1987). Social processes in early number development. *Monographs of the Society for Research in Child Development*, *52*, 153-159.

Seefeldt, C., & Galper, A. (2008). *Active experiences for active children: Mathematics* (2nd ed.). Pearson Education.

Segal, D. L., & Coolidge, F. L. (2004). Objective assessment of personality and psychopathology: An overview. In M. Hersen (Ed.), *Comprehensive handbook of psychological assessment, personality assessment* (p. 3-13). John Wiley & Sons, Inc.

Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science*, *11*, 655-661.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games–but not circular ones–improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology*, *101*, 545-560.

Simms, V. (2014). *Mathematics anxiety predicts mathematical achievement in very preterm children*. Poster presented at the Pediatric Academic Society Annual Meeting, Vancouver.

Simms, V., Cragg, L., Gilmore, C., Marlow, N., & Johnson, S. (2013). Mathematics difficulties in children born very preterm: Current research and future directions. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *98*, F457-63. 10.1136/archdischild-2013-303777.

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2005). Parents' socializing behavior and children's participation in math, science, and computer out-of-school activities. *Applied Developmental Science*, *9*, 14-30.

Skwarchuk, S.-L. (2009). How do parents support preschoolers' numeracy learning experiences at home? *Early Childhood Education Journal*, *37*, 189-197.

Skwarchuk, S.-L., Sowinski, C., & LeFevre, J.-A. (2014). Formal and informal home learning activities in relation to children's early numeracy and literacy skills: The development of a home numeracy model. *Journal of Experimental Child Psychology*, *121*, 63-84.

Smeding, A. (2012). Women in science, technology, engineering, and mathematics (STEM): An investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. *Sex Roles*, *67*, 617-629.

Smith, A. (2004). *Making mathematics count.* Department for Education.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4-28.

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

Starkey, P., Klein, A., Chang, I., Dong, Q., Pang, L., & Zhou, Y. (1999). *Environmental supports for young children's mathematical development in China and the United States.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, Albuquerque.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, *34*, 379-440.

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math–gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology*, *102*, 947-963.

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

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, *104*, 743-762.

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences, U.S.A*, *95*, 7836-7841.

The British Psychological Society. (2010). *Code of human research ethics.* Retrieved from `http://www.bps.org.uk/sites/default/files/documents/code_of_human_research_ethics.pdf`

Thevenot, C. (2010). Arithmetic word problem solving: Evidence for the construction of a mental model. *Acta Psychologica*, *133*, 90-95.

Thomas, G., & Dowker, A. (2000). *Mathematics anxiety and related factors in young children.* Paper presented at the British Psychological Society Developmental Section Conference, Bristol.

Toll, S. W., Van der Ven, S. H., Kroesbergen, E. H., & Van Luit, J. E. (2011). Executive functions as predictors of math learning disabilities. *Journal of Learning Disabilities*, *44*, 521-532.

Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology*, *47*, 943-949.

Torbeyns, J., Rathé, S., & Verschaffel, L. (2014). *The relationship between kindergartners' SFON and their number-related utterances during picture book reading.* Paper presented at the Biennial EARLI Conference of SIG 5 – Learning and Development in Early Childhood, Jyväskylä.

Trochim, W. M. (2006, October). *The research methods knowledge base.* Retrieved from `http://www.socialresearchmethods.net/kb/`

Vanbinst, K., Ghesquiere, P., & De Smedt, B. (2012). Numerical magnitude representations and individual differences in children's arithmetic strategy use. *Mind, Brain, and Education*, *6*, 129-136.

Van de Walle, J., & Lovin, L. (2006). *Teaching student-centered mathematics: Grades K-3.* Pearson Education.

Van Loosbroek, E., & Smitsman, A. W. (1992). Visual perception of numerosity in infancy. *Developmental Psychology*, *26*, 916-922.

Verdine, B. N., Irwin, C. M., Golinkoff, R. M., & Hirsh-Pasek, K. (2014).

Contributions of executive function and spatial skills to preschool mathematics achievement. *Journal of Experimental Child Psychology*, *126*, 37-51.

Verschaffel, L., & De Corte, E. (1993). A decade of research on word problem solving in Leuven: Theoretical, methodological, and practical outcomes. *Educational Psychology Review*, *5*, 239-256.

Vorderman, C., Budd, C., Dunne, R., Hart, M., & Porkess, R. (2011). *A world-class mathematics education for all our young people* (P. Nicholson, Ed.). Conservative Party.

Vukovic, R. K., Roberts, S. O., & Green Wright, L. (2013). From parental involvement to children's mathematical performance: The role of mathematics anxiety. *Early Education and Development*, *24*, 446-467.

Walport, M. (2010). *Science and mathematics secondary education for the 21st century: Report of the Science and Learning Expert Group.*

Wechsler, D. (1967). Manual for Wechsler Preschool and Primary Scale of Intelligence. *San Antonio, TX The Psychological Corporation*.

Whyte, J. C., & Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Developmental Psychology*, *44*, 588-596.

Williams, P. (2008). *Independent review of mathematics teaching in early years settings and primary schools.* DCSF Publications.

Wu, S. S., Barth, M., Amin, H., Malcarne, V., & Menon, V. (2012). Math anxiety in second and third graders and its relation to mathematics achievement. *Frontiers in Psychology*, *3*, 1-11.

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

Wynn, K., Bloom, P., & Chiang, W.-C. (2002). Enumeration of collective entities by 5-month-old infants. *Cognition*, *83*, B55-B62.

Xenidou-Dervou, I., De Smedt, B., van der Schoot, M., & van Lieshout, E. C. (2013). Individual differences in kindergarten math achievement: The integrative roles of approximation skills and working memory. *Learning and Individual Differences*, *28*, 119-129.

Yee, D. K., & Eccles, J. S. (1988). Parent perceptions and attributions for children's math achievement. *Sex Roles*, *19*, 317-333.

Young, C. B., Wu, S. S., & Menon, V. (2012). The neurodevelopmental basis of math anxiety. *Psychological Science*, *23*, 492-501.

Zirk-Sadowski, J., Lamptey, C., Devine, A., Haggard, M., & Szűcs, D. (2014). Young-age gender differences in mathematics mediated by independent control or uncontrollability. *Developmental Science*, *17*, 366-375.

Appendix A

Numeracy Task

The following task was used to assess children's numerical skills in Study 2: Main (presented in Chapter 3). Items were designed to measure the following skills: number word sequence production, numerical ordering, cardinality understanding and simple addition and subtraction. The order of the items was counterbalanced across participants.
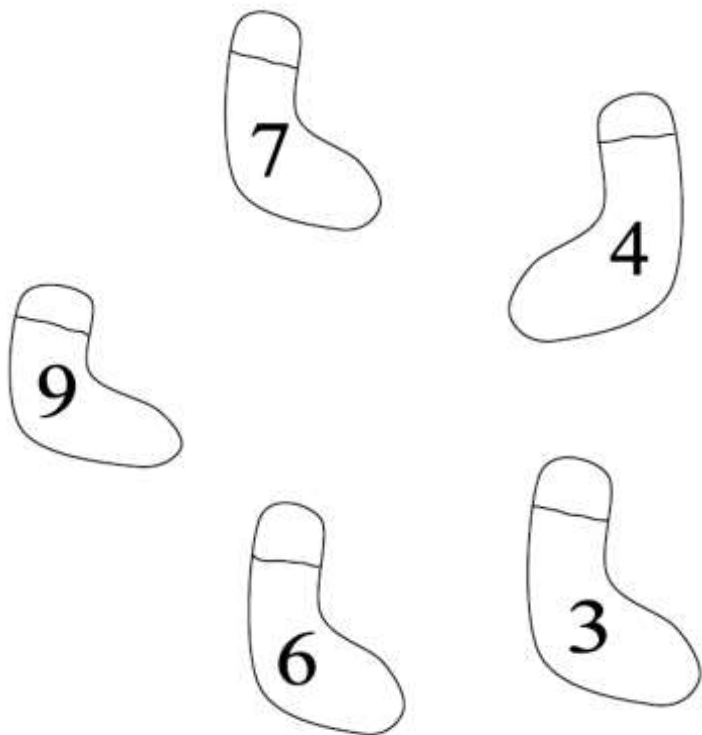
How many apples are on the tree? ____
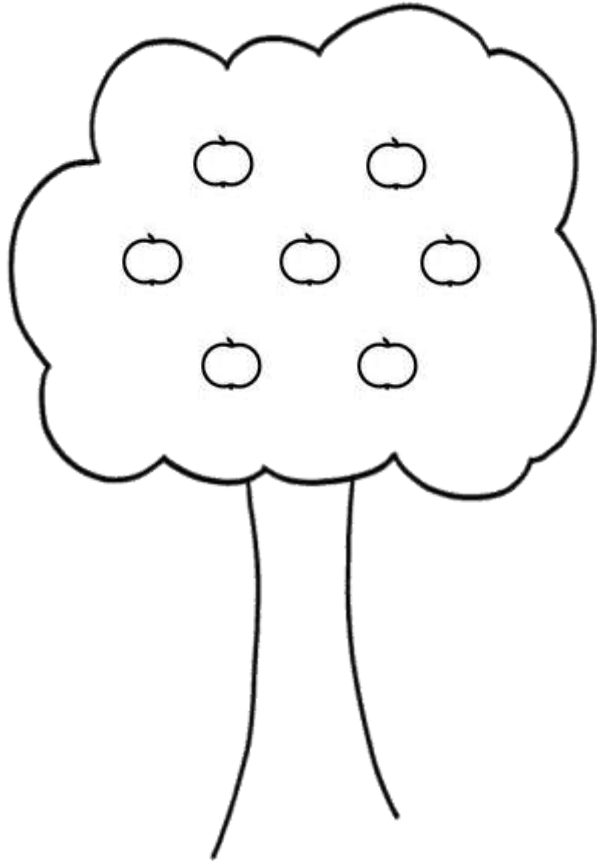
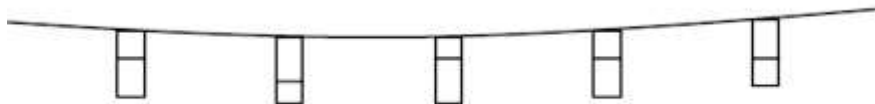**Two apples fall off the tree. How many apples are there left?** ____

Fill in the missing squares

1		3		5
6		8		10
11		13		15
16		18		20
21		23		25
26		28		30

Can you put the socks in order on the line?

Fill in the gaps in the clouds

7

4

9

6

3

One more than 10 = ____

Two more than 3 = ____

One less than 9 = ____

Two less than 6 = ____

Match the raindrops to the clouds

5	eight
10	seven
8	five
	ten
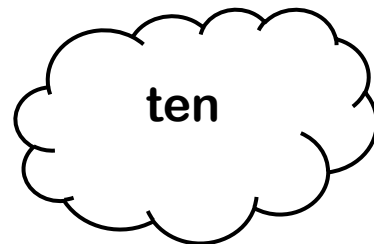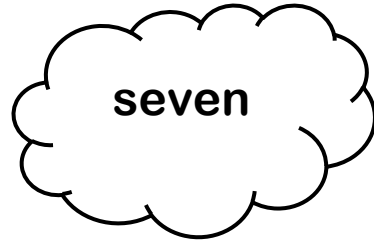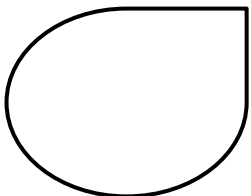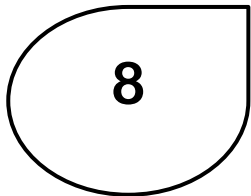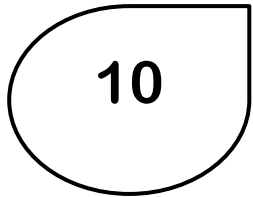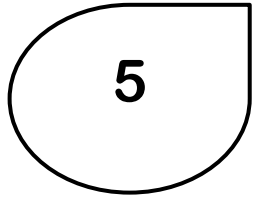
Fill in the missing raindrop

Draw two red apples and four green apples on the tree.
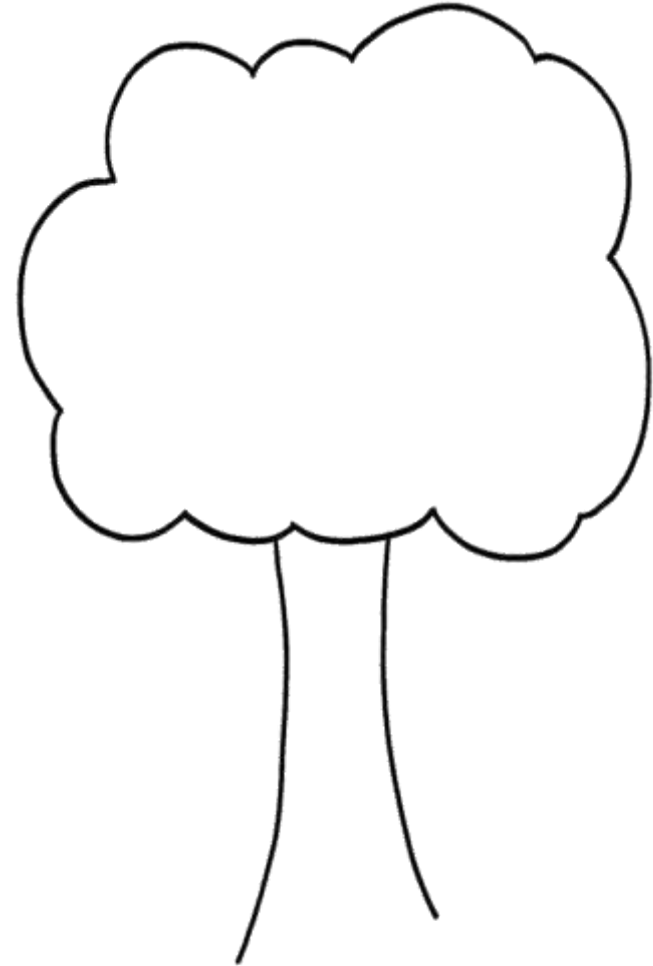
How many apples are there altogether? ___

Appendix B

Home Numeracy Questionnaire

The following questionnaire was given to parents in Study 3 (presented in Chapter 4) to measure children's home numeracy experiences. It was adapted from LeFevre et al. (2009).

Parent Questionnaire

SECTION A - Background Information

1) Relation to the child:

☐ Mother

☐ Father

☐ Other (please state) _____

2) Are you your child's primary caregiver?

☐ Yes

☐ No

3) Child's day care provision:

☐ School

☐ Preschool (☐Part-time or ☐full-time?)

☐ Nursery (☐Part-time or ☐full-time?)

☐ Child-minder (☐Part-time or ☐full-time?)

☐ Other (please state) _____

4) Do you have any other children? (If yes, how old are they?)

☐ Yes (Ages _____)

☐ No

SECTION B - Literacy Questions

1) Please estimate the number of children's books in your household: _____

2) Please estimate the number of adult's books in your household: _____

SECTION C – Numeracy Questions

1) **How high can your child count**: _____

2) **Did you ask your child to count to answer the above question:** _____

SECTION D – Toys & Games

1) **Does your child play computer games at home (including games on phones/i-pads etc)?**

   ☐ Yes   (What kind of games? _____)

   ☐ No

2) **Please estimate the number of word-based games in your household (e.g. "Guess Who?", "What's Up?"):** _____

3) **Please estimate the number of number-based games in your household (e.g. board games with die, dominoes)**: _____

SECTION E – Benchmarks

In your opinion, how important is it for children to reach the following benchmarks before starting school?

	Not important				Very important
Identify names of written alphabet letters	0	1	2	3	4
Identify sounds of alphabet letters	0	1	2	3	4
Print/Write name	0	1	2	3	4
Print/Write alphabet letters	0	1	2	3	4
Count to 10	0	1	2	3	4
Count to 100	0	1	2	3	4
Identify/recognise written numbers	0	1	2	3	4
Simple sums	0	1	2	3	4

SECTION F – Home Activities

In the past month, how often did you and your child engage in the following activities?

	Did not occur	1-3 times a month	Once a week	2-4 times a week	Almost daily
Identifying names of written numbers	0	1	2	3	4
Playing with number fridge magnets	0	1	2	3	4
Counting objects	0	1	2	3	4
Sorting things by size, colour or shape	0	1	2	3	4
Counting down (10, 9, 8, 7..)	0	1	2	3	4
Learning simple sums (e.g. 2+2=4)	0	1	2	3	4
Printing/writing numbers	0	1	2	3	4
Talking about money when shopping (e.g., "which costs more?")	0	1	2	3	4
Measuring ingredients when cooking	0	1	2	3	4
Being timed	0	1	2	3	4
Playing with calculators	0	1	2	3	4
"Connect-the-dot" activities	0	1	2	3	4
Using calendars and dates	0	1	2	3	4
Having your child wear a watch	0	1	2	3	4
Using number activity books	0	1	2	3	4
Reading number storybooks	0	1	2	3	4
Playing board games with die or spinner	0	1	2	3	4
Playing card games	0	1	2	3	4
Putting pegs in a board or shapes into holes, playing with puzzles	0	1	2	3	4
Playing with blocks	0	1	2	3	4
"Paint-by-number" activities	0	1	2	3	4
Watching educational TV shows	0	1	2	3	4

	Did not occur	1-3 times a month	Once a week	2-4 times a week	Almost daily
Using educational software	0	1	2	3	4
Building Lego or construction set (Duplo, Megablocks, etc.)	0	1	2	3	4
Identifying names of written alphabet letters	0	1	2	3	4
Identifying sounds of alphabet letters	0	1	2	3	4
Printing/writing letters	0	1	2	3	4

SECTION G – Caregiver's Attitudes Towards Maths and Literacy

Please read the following statements and indicate the degree to which you agree/disagree.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
When I was in school, I was good at maths.	1	2	3	4	5
When I was in school, I enjoyed maths.	1	2	3	4	5
The career path I have chosen is maths related.	1	2	3	4	5
When I was in school, I was good at language arts activities such as reading.	1	2	3	4	5
When I was in school, I enjoyed language arts activities such as reading.	1	2	3	4	5
I find maths activities enjoyable.	1	2	3	4	5
I find reading enjoyable.	1	2	3	4	5
It is important for my child to be exposed to mathematical concepts every day.	1	2	3	4	5
It is important for my child to be read to every day.	1	2	3	4	5

Thank you for your time.

Appendix C

Child Mathematics Anxiety Questionnaire

The following questionnaire was used to measure children's mathematics anxiety in all of the studies presented in Part III (Chapters 7 – 10). It was adapted from the CMAQ developed by Ramirez et al. (2013).

Attitudes to Maths Questionnaire

Name:

…………………………………………………………………………………………

Date:

…………………………………………………………………………………………

Date of birth:

…………………………………………………………………………………………

Please circle the face that matches how you feel.

1. How do you feel when taking a big test in your maths class?

2. How would you feel if you were given this problem: There are 13 ducks in the water. There are 6 ducks in the grass. How many ducks are there in all?

3. How would you feel if you were given this problem: You scored 15 points. Your friend scored 8 points. How many more points did you score than your friend?

4. How do you feel when getting your maths book and seeing all the numbers in it?

5. How do you feel when you have to solve 27 + 15?

6. How do you feel when figuring out if you have enough money to buy a chocolate bar and a soft drink?

7. How do you feel when you have to solve 34 – 17?

8. How do you feel when you get called on by the teacher to explain a maths problem on the board?

Appendix D

Parent Mathematics Anxiety Questionnaire

The following questionnaire was used to measure parents' mathematics anxiety in Study 6 (presented in Chapter 7). It comprised 18 items from the MAS-UK developed by Hunt et al. (2011). There were three subscales:

1. Mathematics evaluation anxiety (items 1, 5, 6, 9, 16, 17).

2. Everyday/social mathematics anxiety (items 2, 3, 4, 7, 10, 12, 13, 18).

3. Mathematics observation anxiety (items 8, 11, 14, 15).

ID Number:
Relationship to child: Mother/Father

		How anxious would you feel in the following situations? Please circle the appropriate numbers below				
		Not at all	Slightly	A fair amount	Much	Very much
1	Having someone watch you multiply 12 x 23 on paper	1	2	3	4	5
2	Adding up a pile of change	1	2	3	4	5
3	Being asked to add up the number of people in a room	1	2	3	4	5
4	Calculating how many days until a person's birthday	1	2	3	4	5
5	Taking a maths exam	1	2	3	4	5
6	Being asked to calculate £9.36 divided by 4 in front of several people	1	2	3	4	5
7	Being given a telephone number and having to remember it	1	2	3	4	5
8	Reading the word "algebra"	1	2	3	4	5
9	Calculating a series of multiplication problems on paper	1	2	3	4	5
10	Working out how much time you have left before you set off to work or place of study	1	2	3	4	5
11	Listening to someone talk about maths	1	2	3	4	5
12	Working out how much change a cashier should have given you in a shop after buying several items	1	2	3	4	5
13	Deciding how much each person should give you after you buy an object that you are all sharing the cost of	1	2	3	4	5
14	Reading a maths textbook	1	2	3	4	5
15	Watching someone work out an algebra problem	1	2	3	4	5
16	Being asked to memorize a multiplication table	1	2	3	4	5
17	Being asked to calculate three fifths as a percentage	1	2	3	4	5
18	Working out how much your shopping bill comes to	1	2	3	4	5

Thank you for your participation.

Appendix E

Child Mathematics Confidence Questionnaire

The following questionnaire was used in Study 8 (presented in Chapter 10) to measure children's confidence in mathematics. It was adapted from the self-rating subscale of the Mathematics Attitudes and Anxiety Questionnaire (Thomas & Dowker, 2000; Krinzinger et al., 2007). It comprised 6 mathematics items (questions $3 - 8$) and two literacy control items (questions $1 - 2$).

What am I like?

Name:

…………………………………………………………………………………………..

Date:

…………………………………………………………………………………………..

Please circle the ticks/crosses that match how good you think you are.

| Really good | Good | Ok | Not good | Really not good |

Remember that all children are different. Some children are really good at sports and games, while other children are better at writing stories and drawing pictures. There are no right answers and everyone will have different answers. Make sure that your answers show how you feel about yourself!

1. How good are you at writing?

2. How good are you at reading?

3. How good are you at maths?

✓✓   ✓   ˣ✓   ✗   ✗ˣ

4. How good are you at working out maths problems in your head?

✓✓   ✓   ˣ✓   ✗   ✗ˣ

5. How good are you at working out maths problems on paper?

✓✓   ✓   ˣ✓   ✗   ✗ˣ

6. How good are you at learning new things in maths?

✓✓   ✓   ˣ✓   ✗   ✗ˣ

7. How good are you at working out easy maths problems?

✓✓   ✓   ˣ✓   ✗   ✗ˣ

8. How good are you at working out hard maths problems?

✓✓   ✓   ˣ✓   ✗   ✗ˣ

List of Abbreviations

ANCOVA Analysis of Covariance.

ANOVA Analysis of Variance.

BPVS British Picture Vocabulary Scale.

CMAQ Child Mathematics Anxiety Questionnaire.

EYFS Early Years Foundation Stage.

GFON guided focusing on numerosity.

IAT Implicit Association Task.

MA mathematics anxiety.

MAQ Mathematics Anxiety Questionnaire.

MAS-UK Mathematics Anxiety Scale–UK.

PISA 'Programme for International Student Assessment'.

PRISM Premature Infants' Skills in Mathematics.

SES socioeconomic status.

SFOL spontaneous focusing on spatial locations.

SFON spontaneous focusing on numerosity.

WIAT Wechsler Individual Achievement Test.

WPPSI Wechsler Preschool and Primary Scale of Intelligence.